University of Strathclyde

Department of Management Science

Playing the Rankings Game

A game theoretic decision model using Data Envelopment Analysis evaluated for university league tables

Carolyn Booker

A thesis presented in fulfilment
of the requirements for the
Degree of Doctor of Philosophy

November 2013

Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed: Date:

Acknowledgements

The first thanks in this thesis must go to Stephen Magee at the University of St Andrews, who asked me the question which I am still trying to answer. Next I must give credit to my supervisors, Professors John Quigley and Lesley Walls, who have shown exceptional patience and humour in prodding me through this process; to Dr Kerem Akartunali, whose expertise in optimisation has saved my sanity; and to my examiners, Professors Emmanuel Thanassoulis and Valerie Belton, whose feedback has pushed me over the finish line.

I also owe a large thank you to the rest of the Management Science department: to those who have helped me to juggle admin and teaching along with research, who have pushed me into presentations and asked me difficult questions, who have shared tea breaks and lunch breaks through all our highs and lows. Thanks are due, too, to the friends and family who have persistently asked me “How’s your thesis going?” even though I have seldom felt very grateful at the time of asking. I apologise for the prolonged bad mood and I hope to make it up to you.

As I write these last words, though, I am chiefly aware of the people who will never read them. My beloved brother Matthew Eggleton (1970 – 2009), who died suddenly in the first few weeks of this endeavour; my sister-in-law, Dawn Wood (1965 – 2013), who patiently endured years of multiple sclerosis; and my friend and fellow research student, Bernadine Williams (1969 – 2014), who slipped away at the very end just when I thought I was home and dry.

I would not have survived these difficult years without the steadfast support of my husband, Stuart. He has made me porridge at 5:30 in the morning and picked me up from the train at eight o’clock at night. He has cajoled, comforted, reasoned and read. He has been patient beyond measure. Thank you.

Abstract

The purpose of this research is to examine trade-offs and conflicts between performance measures in the UK Higher Education sector. The performance measures under consideration are those which are imposed on a university from outside, such as statutory performance indicators and newspaper league tables, and which bring rewards in the form of either status or funding. The existing literature provides evidence that such measures are causing tension within institutions, but there has to date been no attempt to examine that tension using the tools of management science.

The main tool used here is Data Envelopment Analysis (DEA). A new DEA model has been developed which extends a trade-off model to incorporate a weighted preference structure. This model is used first to determine the production possibility set for a group of Higher Education Institutions and then to explore the options open to them. In many kinds of performance measurement system the reward achieved by an HEI, such as a "top ten" position or a share of a fixed amount of funding, depends not only on that institution's own decisions but also on the strategic decisions of others. Game Theory provides a range of structures which model such interactive decisions and can aid a decision-maker in determining optimal strategies. The results of the DEA model are therefore processed using a typical league table construction and then evaluated through the lens of Game Theory.

The analytical framework developed in this thesis has the potential for application in other educational and social contexts where external performance measures are known or suspected to have an influence on decision-making. The DEA model can additionally be used outwith this framework in any context where it is desirable to permit selective target-setting and to accommodate value judgments in the specification of the production possibility set.

Table of Contents

Table of Tables

Table of Figures

Notation

The following table has been compiled as a guide to the notation used in the DEA model.

$x_{ij} \geq 0 \quad i=1..m, j=1..n$	A single input *i* for a single DMU *j*; it must be non-negative
$y_{rj} \geq 0 \quad r=1..s, j=1..n$	A single output *r* for a single DMU *j*; it must be non-negative
(x,y)	The vectors *x* (of inputs) and *y* (of outputs) which describe a single DMU
X	An *m* x *n* matrix: one row of each of *m* inputs, one column for each of *n* DMUs
Y	An *s* x *n* matrix: one row of each of *s* outputs, one column for each of *n* DMUs
$w_i^- > 0 \quad i \in I_0$	A user-selected weight for a single input *i*
I_0	The subset of inputs *i* for which a user-selected weight has been defined
$\overline{I_0}$	The subset of inputs *i* for which no weight has been defined
$w_r^+ > 0 \quad r \in R_0$	A user-selected weight for a single output *r*
R_0	The subset of outputs *r* for which a user-selected weight has been defined
$\overline{R_0}$	The subset of outputs *r* for which no weight has been defined
$0 < a_i \leq 1 \quad i \in I_0$	A multiplier for the reduction of input $i \in I_0$
$z_r \geq 1 \quad r \in R_0$	A multiplier for the increase of output $r \in R_0$
$d_i \geq 0 \quad i \in \overline{I_0}$	A slack variable for the reduction of input $i \in \overline{I_0}$
$c_i \geq 0 \quad i \in \overline{I_0}$	A buffer variable to ensure that input *i* is not reduced to a negative quantity
$e_r \geq 0 \quad r \in \overline{R_0}$	A slack variable for the increase of output $r \in \overline{R_0}$
$\lambda_j \geq 0 \quad j=1..n$	Coefficient of a single DMU *j* in a convex combination of DMUs
$P_t \quad t=1..T$	A vector of *m* trade-off values, one for each input
$Q_t \quad t=1..T$	A vector of *s* trade-off values, one for each output
$\pi_t \geq 0 \quad t=1..T$	Coefficient of the t^{th} trade-off vector
$0 < \varepsilon \ll 1$	A very small positive number: it is assumed that $\varepsilon \ll w_r^+ \ \forall r \in R_0$ and $\varepsilon \ll w_i^- \ \forall i \in I_0$.

1 Introduction

This chapter introduces the topic of this thesis, league tables and ranking systems in UK higher education, and briefly establishes the context in which the present research is based. An overview of the objectives, tools and philosophy of the research follows and then the structure of the rest of the thesis is presented. The final section contains a guide to the notation used in this research.

1.1 Background

Performance measurement in the university sector has been around for at least a hundred years (Hazelkorn, 2011:29) but in the UK it came to prominence in the 1980s. Part of the government's agenda at that time was to increase the public sector's accountability through "scrutiny of the way in which the taxpayers' money [was] being used" (Johnes and Taylor, 1990:1) and the rise of mechanisms for evaluation went hand-in-hand with changes to the funding regime.

University league tables first appeared in the UK press in the 1990s. If performance measurement began as an exercise in accountability to the tax-payer, then league tables might be said to extend that accountability to the market-place (King, 2009). While the original intended audience for these tables was the population of potential students and their parents, this has since broadened to include "international postgraduate students and faculty, other HEIs and HE organizations, government and policymakers, employers, sponsors, foundations, private investors and industrial partners" (Hazelkorn, 2011:40).

Such a varied audience might be expected to have diverse interests and needs, but one of the most beguiling aspect of league tables appears to be their simplicity. Rather than capturing the complexities of the sector, the typical newspaper league table offers "as with restaurants, televisions or hotels … an easy guide to quality" (ibid).

Whatever the limitations of these ranking systems, in the last 20 years they have become increasingly visible. A number of major studies have been conducted which examine both their structure (e.g. Centre for Higher Education Research and Information et al., 2008a, Usher and Savino, 2006) and their impact (e.g. Hazelkorn, 2007, Institute for Higher Education Policy, 2009). These studies have identified tensions in the university sector resulting from the competing demands of different performance measures, including league tables.

1.2 Research Objectives and Tools

The aim of this research is to improve our understanding of these tensions through quantitative modelling of the league table environment. This is a new approach to the area as the research to date has been largely qualitative, with quantitative research focused on the structure of league tables and on individual measures (e.g. Lee and Buckthorpe, 2008, Tofallis, 2012) rather than on their impact.

One of the key features of the league table style of performance assessment is that one's score is affected not only by one's own performance but by the performance of everyone else. The number of universities jostling for a place in the "top ten" of any table is likely to be many more than the ten places available. This consideration has led us to test game theory as a suitable framework for this decision process. Described as "a model of interactive decisions" (Bennett et al., 1989: 287), game theory potentially provides the language we need for evaluating the effectiveness of university strategies in a league table context.

However, in order to apply the language of game theory to this problem, we first need to furnish the components of a game: 'players', 'strategies' and 'payoffs', for example (see section 4.2). In order to achieve this, we have identified Data Envelopment Analysis as an appropriate tool. The multi-dimensional nature of the Higher Education sector, with its many inputs and outputs, makes DEA a very practical choice. We will use the relatively conservative production space of a DEA model to determine targets which a university might aim for. By developing a new DEA model which incorporates the freedom for a DMU to set its own goals, we will be able to generate alternative targets which can serve as strategies in the game theoretic framework. We will then be in a position to evaluate the different outcomes resulting from the successful completion of the strategies.

We can therefore summarise the objectives of this research as follows:

1.2.1 Develop a quantitative framework for evaluating university decision making in relation to league table rankings

Drawing on the basic principles of game theory, this framework will allow us to consider the decisions of university managers in the context of competing performance measures and of the strategies of other institutions.

1.2.2 Develop a DEA model suitable for determining targets in the context of the university sector

To facilitate the use of the framework described above, a model will be developed which can furnish alternative strategies for university decision-makers.

These objectives will be considered in more detail as the research context is elaborated in chapters 2, 3 and 4.

1.3 Research Approach

In the introduction to the background and aims of this research we have emphasised that we are taking a quantitative view of the tensions surrounding league tables and ranking systems. Quantitative modelling is traditionally the dominant mode of management science, particularly under its earlier name of operational research. The bread-and-butter techniques of OR included linear programming, regression modelling, discrete event simulation. In some parts of the world, notably the US, this is still the case. However, in the UK the scope of management science has broadened out since it was strongly criticised in the 1960s and 70s by Churchman, Ackoff and others (e.g. Churchman, 1979, Ackoff, 1979). Their criticism was aimed not so much at the tools of management science but at the assumptions which underpinned their use or, in other words, the philosophy of management science. Assumptions about the nature of reality and our ability to apprehend it objectively were leading to inappropriate applications of these quantitative methods to situations where they might potentially do more harm than good (Ackoff, 1977). In the 1960s management science was apparently unable to tackle the kinds of unstructured problems which had become increasingly important in the world of organisations, but since that time new approaches have blossomed and flourished. These approaches are typically characterised as Problem Structuring Methods (PSMs) and designated ‘soft OR’ in contrast to the ‘hard OR’ of optimisation and computer simulation.

What is the philosophy, then, that lies behind this diverse subject? The labels ‘hard’ and ‘soft’ typically carry with them assumptions about the philosophy espoused by their practicioners. A researcher in ‘hard OR’ might be described as a positivist, one who asserts the existence of an objectively knowable reality, while a researcher in ‘soft OR’ might be characterised as an interpretivist, claiming a socially constructed reality which can only be experienced subjectively. The field is far more interesting, however, as there are a range of philosophical views informing management science research and a lively debate between them.

The philosophical approach taken in this research is realist, but it is not positivist. To aver a realist ontology is to say that there exists a reality which is independent of the scientist; it is not to say that the scientist can objectively observe and record that reality. Conflation of realist ontology and objective epistemology into an 'empirical realism'[1] which demands a correspondence theory of truth is, unfortunately, all too common in the literature of social research and its philosophy (e.g. Burrell and Morgan (1979), Jardine (1978:107)) but those who observe the distinction have developed rich ways of thinking about the world.

The first question such thinkers need to address is this: if the world is there but we cannot know it objectively, how can we know it at all? The realist responds that "our knowledge is always provisional, and historically and culturally relative" (Mingers, 2006:204). Further, we acknowledge that the objects of the social world differ from the objects of the natural world in that they are concept-dependent, i.e. they are affected by the meanings we ascribe to them (Sayer, 1992). There is an important distinction, however, between the notion of concept-dependence and complete subjectivity, and it is well-expressed by Sayer Sayer (1992:49): "Although social phenomena cannot exist independently of actors or subjects, they usually do exist independently of the particular individual who is studying them." It is possible for a realist to study a reality outside herself.

A second question quickly follows: what is the purpose of our partial, provisional knowledge? The realist, says Trigg (1980:ix), "will not be surprised if some portions of [reality] elude man's grasp forever," so we must have some other goal in mind than that of mirroring the real world. John Mingers noted that operational research is "above all, concerned with taking action, solving problems, improving situations, and so our interest is in methodologies and theories oriented towards action not simply description or analysis." (Mingers, 1992).

In the last twenty years critical realism has become a very popular philosophical approach among management scientists, including Mingers (2000, 2006), because it recognises the "ontological gap" between the world and our experience of it (Danermark et al., 2002:39) .

[1] for use of this term see Johnson and Duberley (2000) p 149-50

However, there is one aspect of critical realism which makes some researchers uncomfortable and that it is its approach to causal explanation. Critical realists assert that “by reasoning we can obtain knowledge of what properties are required for a phenomenon to exist” (Danermark et al., 2002:206), a process of transcendental argument which is designated retroduction. For some, including the present author, retroduction has a taste of gnosis (Jackson, 2006).

While it would be fascinating to give time and attention to the underpinnings of critical realism and the ways in which we might test ideas about the underlying causes of events, such a study is not, fortunately, a prerequisite for productive research in OR/MS. As the quotation from Mingers above demonstrates, there is a strong pragmatic streak in management science, where we are looking for knowledge that “helps us to deal, whether practically or intellectually, with either the reality or its belongings” (James, 1907:94). This type of knowledge meets the essential pragmatic criteria for truth. Johnson & Duberley maintain that the pragmatic paradigm occupies very similar ontological and epistemological ground to critical realism, but avoids the potential pitfall of attributing causality to “demons or witches’ spells” (Halfpenny, quoted in Johnson and Duberley, 2000:156).

The aim of this research is to improve our understanding of the tensions generated in the university sector by league tables and ranking systems; in other words to “help us to deal” with them. It is in the pragmatic paradigm, therefore, that the present research is located.

1.4 Structure of this thesis

Figure 1-1 illustrates the main tasks involved in the fulfilment of the research objectives outlined above and shows how they are addressed in the chapters of this thesis.

The problem context is explored in detail in **Chapter 2**, where we discuss the landscape of UK Higher Education, fill in the background to performance measurement in the sector and look in depth both at league tables and at research into their structure and impact.

We have outlined our reasons for selecting Data Envelopment Analysis and Game Theory as appropriate tools for this research. In **Chapters 3 and 4** we will present a thorough exploration of the literature on these tool. This will include identifying the models and approaches most suitable for this research; assessing the strengths and limitations of each method; evaluating how they have previously been applied to the study of higher education and looking at ways in which they have been jointly employed.

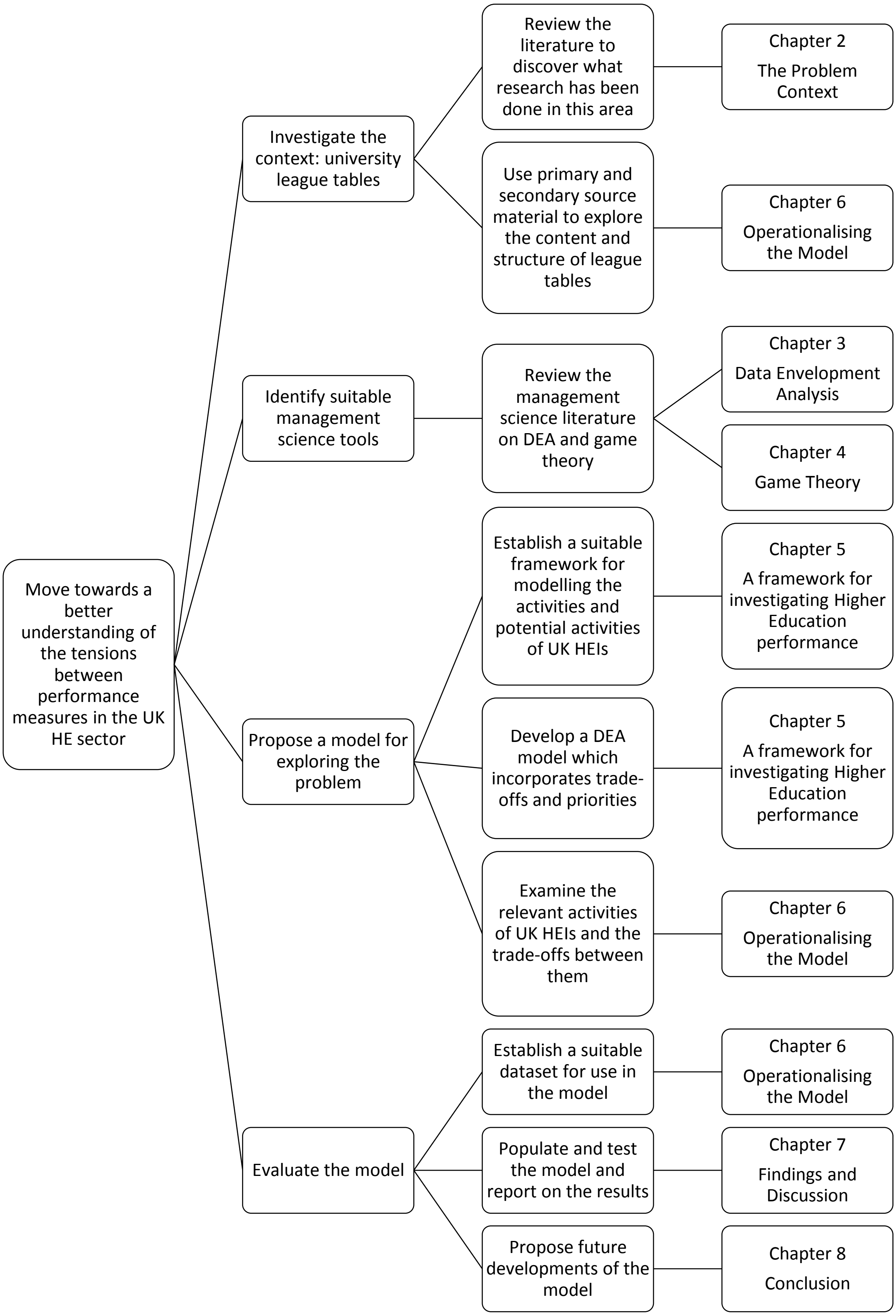
In **Chapter 5** we will develop the framework required for our first objective and the DEA model required for our second objective.

Our exploration of the problem context resumes in **Chapter 6**, where we examine the data required to populate a model of UK universities in a league table environment. We consider the available data sources in detail and establish a suitable dataset to illustrate the use of our model.

In **Chapter 7** we populate and test the framework and DEA model. Results are presented and discussed. We note the limitations of the model and identify insights arising from it.

Finally, we draw our conclusions in **Chapter 8**. We summarise what has been achieved in respect of our original research objectives and highlight other areas where the resulting framework could be applied. We then identify future work to develop further understanding of the problem context.

Figure 1-1: Structure of thesis



2 Higher Education and Performance Measurement

In this chapter we will establish the context of this research. We will begin with an overview of the Higher Education sector in the UK and then introduce some of the systems which are used to evaluate it. Then we will survey the literature on Higher Education and performance measurement, paying particular attention to studies of league tables and ranking systems.

2.1 Terminology

Before we proceed, it is important to clarify some of the terminology we will be using in this chapter and throughout this thesis.

2.1.1 Referring to Universities

Government agencies dealing with the university sector in the UK typically use the phrase **Higher Education Institution**, abbreviated to **HEI**, when referring to an organisation in their portfolio. This has the advantage of including universities, university colleges and other types of organisation under one umbrella term. However, in the academic literature it is much more common to see the word **university** used, and it is clearly intended in the inclusive sense described above. In this thesis, therefore, both these terms will be used and should be understood as encompassing all organisations which are accredited providers of higher education, not only those which are strictly designated universities.

It is worth noting at this juncture that an equivalent term in use in North America is **Institution of Higher Learning**, abbreviated to **IHL**, and this is used in some of the literature we will be reviewing. We will therefore need to refer to this term on occasion, but will not adopt it for our own use.

2.1.2 Referring to League Tables

This is a rather more complex area of terminology. There are many ways in which university performance is measured and we will be referring to a wide variety of literature. The terms **performance measurement** and **performance measurement system** (or **PMS**) will be used when we are discussing the general area of interest. This term is a neutral one in a management science context, but nonetheless has the power to raise hackles in the university environment where it is associated with a managerial worldview and therefore suspect (Broad et al., 2007:124).

Within this overall landscape, there are local measures of performance, devised and managed by the universities themselves, "official" measures of performance such as the indicators compiled by the funding councils on behalf of government and "unofficial" measures such as the ranking systems used by newspapers and other commercial actors. It is this last group which holds the greatest interest for this research. The term **league table** has been widely adopted, particularly in the UK, because the ordered lists of universities published in newspapers and magazines resemble a football league table (Usher and Savino, 2006). The term **rankings** is often used interchangeably with league table, even though there are ranking systems which are not presented in a league table format. Hazelkorn (2007, 2009) employs the abbreviation **LTRS** as shorthand for the composite term **league tables and ranking systems**. We will need to use all these terms as we work through the literature and, unless it is specifically stated, no particular properties are being attributed through the use of one term rather than another.

2.2 Higher Education in the UK

This chapter opened with the promise of an overview of the UK Higher Education sector. However, in the light of devolution and subsequent policy developments, it is fair to ask whether such a thing as "UK Higher Education" still exists. Indeed, a recent paper suggests that it never did: "It is dubious whether there ever was a British system of higher education given that the Scottish tradition has deviated in critical ways from the English model: wider social access, four-year degrees and broader degree programmes" (Filippakou et al., 2012:108). We will consider some of the evidence.

Data published by the Higher Education Statistics Agency (HESA) and Universities UK in 2011 give an overall view of the size and activities of the sector as shown in Table 2-1.

The UK picture is clearly dominated by the sector in England, which accounts for 80% of the institutions and 84% of the students. Nonetheless, from the post-war expansion of HE in the 1960s, through the funding cuts of the 1980s, to the second wave of expansion in the 1990s, government policy was applied across the UK as a whole. Keating (2005) reflects on the position of Scottish HEIs and notes that during this period "the policy community ... was UK-wide, with Scottish universities considering themselves part of a British system to be judged on the same criteria as those in England." (p 425)

Table 2-1: Data on Higher Education in the UK (HESA, 2011a, Universities UK, 2011b)

	UK	England	Scotland	Wales	N Ireland
Number of HEIs	165	131	19	11	4
Number of universities	115	89	14	10	2
Total income in 2009/10	£26.7 billion	£22.1 billion	£2.8 billion	£1.2 billion	£0.5 billion
Mean income in 2009/10	£162 million	£169 million	£147 million	£112 million	£130 million
Total no. of students in 2009/10	2,493,415	2,093,635	220,910	127,885	50,990
Total no. of undergraduates	1,914,710	1,608,300	166,985	99,570	39,855
Total no. of postgraduates	578,705	485,355	53,925	28,315	11,135
% of students from outside UK	16%	16%	19%	19%	11%
Mean no. of students in 2009/10	15,112	15,982	11,627	11,626	12,748

The administration of government funding was first devolved in 1992, when separate funding councils were set up for England, Wales and Scotland, and political devolution followed within the decade. However, many aspects of Higher Education funding and policy are still controlled at a UK-wide level, including the Research Councils, the Research Assessment Exercise and national pay bargaining. Significant challenges are also shared by institutions across the UK. Keating identifies

- the need to absorb many more students with limited resources
- the requirement to provide mass education and world-leading research at the same time
- the expectation that universities will contribute both to economic development and to overcoming social exclusion

Thus, although there are "no state-wide framework laws governing universities […] Scottish universities are part of a highly articulated UK policy community, with frequent contact among academics and managers, shared ideas and a determination not to be reduced to mere regional importance." (Keating, 2005:427)

This policy community can be seen reflected in the make-up of special interest groups such as the Russell Group, the 1994 Group and the Million+. These are shaped around perceptions of a similar mission rather than national groupings. The Russell Group, for example, identifies itself as representing “20 leading UK universities which are committed to maintaining the very best research, an outstanding teaching and learning experience and unrivalled links with business and the public sector.” (Russell Group, 2012) These 20 universities include Cardiff University (Wales), the Universities of Edinburgh and of Glasgow (Scotland) and Queen’s University, Belfast (Northern Ireland) as well as 16 institutions in England. Thus when Malcolm Tight refers to the “fiction of a single university system” (Tight, 2000:41) he is reflecting on the variety of missions which institutions espouse rather than on their geographical location.

However, since the introduction of tuition fee loans for students in England and Wales in 2004 there has been much discussion of a diverging sector (e.g. Keating (2005), Filippakou et al. (2012)). The present coalition government has adopted key proposals from Lord Browne’s review of higher education funding (Browne, 2010) and this will mean tuition fees for English students rising to as much as £9,000 per year in 2012/13.

It remains to be seen whether these developments have a significant impact on cross-border migration for the purposes of Higher Education. Looking back at 2009/10, the number of UK full-time first degree students who were studying outside their country of domicile was 73,300, comprising 7% of this population. This is a difficult figure to analyse statistically, because of the different sizes of the countries involved, but we can examine it from two perspectives (see Table 2-2).

One perspective, “Where do students go to study?” shows that English and Scottish students both have the strongest propensity to stay at home. In 2009/10 95% of English domiciled students choose to study for their first degree in England and the same proportion of Scottish domiciled students choose to study in Scotland. Students from Wales and Northern Ireland are much more likely to move: in 2009/10 30% of students from each country chose to study in another part of the UK.

However, the question “Where do the students in each country come from?” has a rather different answer. Institutions in Scotland are teaching a UK undergraduate population which is approximately 82% Scottish, 14% English and 4% Northern Irish while in Wales the population is 58% Welsh and 42% English. Only very small numbers of Welsh students choose to study in Scotland and vice versa. Notwithstanding the relatively large numbers − as proportions of their home populations − coming from other parts of the UK to study in

England, the UK undergraduate population in England remains 97% English; and in Northern Ireland 98% of UK-domiciled first degree students are from that country.

At the time of writing, early in 2012, preliminary figures released by the Universities and Colleges Admissions Service (UCAS) show an overall drop in applications to universities of 8.7% from 2011/12 to 2012/13. Applications from all the countries of the UK have fallen, with applications from English students falling by 9.9%. However, there remain more applicants then there are HE places available, and at this stage it would be difficult to forecast the composition of the actual student intake in 2012/13.

Table 2-2: Migration between the countries of the UK for the purposes of Higher Education in 2009/10 (Universities UK, 2011a)

Percentage of institutions' UK students coming from each domicile				
	Institution			
Domicile	England	Wales	Scotland	Northern Ireland
England	96.7	41.6	13.6	1.4
Wales	1.8	57.7	0.4	0.1
Scotland	0.6	0.2	82.4	0.1
Northern Ireland	1.0	0.5	3.6	98.4

Percentage of domiciled students going to institutions in each country				
	Institution			
Domicile	England	Wales	Scotland	Northern Ireland
England	95.3	2.9	1.7	0.0
Wales	30.0	69.1	0.8	0.0
Scotland	5.0	0.1	94.8	0.0
Northern Ireland	19.9	0.8	9.7	69.6

We are highlighting this mixed picture because it is often set aside. Quantitative studies of the Higher Education sector are sometimes headlined *United Kingdom* while the "fine print" tells another story. For instance, Athanassopoulos and Shale (1997) decided to exclude Scottish universities from their UK study because courses are structured differently north of the border. However, in the world of league tables and ranking systems, which we will consider in the next section, all UK HEIs are treated alike irrespective of their structural differences. This research will therefore also need to be inclusive rather than exclusive, but without forgetting the complexities of the current situation.

2.3 Performance Measurement in UK Higher Education

Performance measurement in the university sector has been around for at least a hundred years (Hazelkorn, 2011:29) but in the UK it came to prominence in the 1980s. Part of the government's agenda at that time was to increase the public sector's accountability through "scrutiny of the way in which the taxpayers' money [was] being used" (Johnes and Taylor, 1990:1) and the rise of mechanisms for evaluation went hand-in-hand with changes to the funding regime. This link between performance measurement and funding policy was quite explicit. A member of the Working Group tasked with developing the first set of Performance Indicators (PIs) was quoted in the *Times Higher Education Supplement* saying, "if universities wish to receive increased sums of public money — and they must receive more — they must exhibit evidence that what has been received has been well applied" (THES, 1987 quoted in Cave et al., 1988:34).

In their 1986 report, the Working Group, which consisted of representatives from the Committee of Vice-Chancellors and Principals (CVCP) and the University Grants Committee (UGC), adopted the well-known production model of inputs, processes and outputs. This is a standard model which is widely used in the analysis of organisations and their productivity across a whole spectrum of activities from "products" to "services" (Slack et al., 2010:11) and has been accepted without question by those accustomed to it, such as the economists Johnes and Taylor (1990:51). However, there is often a reluctance in the academic community to admit that "concepts borrowed from industrial management have a role in the university sector" (Cave et al., 1988:19). 20 years later, when Broad et al. (2007) examined the use of performance measurement through two case studies of universities and two of local authorities, they found that dissatisfaction with performance measurement was present in all four cases, but that in the HE sector "the dissatisfaction was more fundamental as PMs were generally seen as 'managerialist' and in conflict with the academic worldview" (p 125). However, they found in general that locally-determined performance measures were more readily accepted than those imposed from the outside. It is these externally imposed systems of measurement which are of particular interest in this research.

2.4 League Tables and Ranking Systems

In their report *A World of Difference*, Usher and Savino define university league tables and rankings as "lists of certain groupings of institutions … comparatively ranked according to a common set of indicators in descending order” (Usher and Savino, 2006:5). The key words to note in this definition are:

2.4.1 “lists”

A league table is always presented as a list of institutions, and in most cases the list is sorted so that the highest-scoring institution appears at the top and the lowest-scoring at the bottom.

2.4.2 “comparatively ranked”

Performance indicators used by governments, or by institutions themselves, may use comparison with other institutions or with targets for information or evaluation, but league tables “are designed specifically as a comparative measure, pitting institutions against each other” (Usher and Savino, 2006:5).

2.4.3 “common set of indicators”

Although a league table could be based on a single indicator, the overall score is usually an aggregation of several scores. A group of indicators is selected which is intended to represent the theme of the league table: this might be a general theme such as “teaching” or “teaching and research” (e.g. the Guardian, the Sunday Times) or it might be more specific such as “executive education” (the Financial Times) or “sustainability” (Times Higher Education). All institutions are scored on all the indicators. We will consider the scoring mechanisms in more detail below.

University league tables first appeared in the UK press in the 1990s. If performance measurement began as an exercise in accountability to the tax-payer, then league tables might be said to extend that accountability to the market-place (King, 2009). Roger King characterises league tables as an example of “increased ‘private authority’ being exercised over universities” (p 135-6) and notes that this private authority “appears to confound the more legitimated and democratic policies and processes associated with the state and public interest accountabilities” (p 137). In response to criticism of this kind, newspapers are wont to retort that their league tables “help to make institutions more accountable to the wider public” (p 138) although, given that in the UK this is exclusively a broadsheet phenomenon, the public they reach is unlikely to represent the population as whole.

Hazelkorn notes that the audience originally targeted — prospective students and their parents — has broadened to include “international postgraduate students and faculty, other HEIs and HE organizations, government and policymakers, employers, sponsors, foundations, private investors and industrial partners" (Hazelkorn, 2011:40) and suggests that the key to league table popularity lies in their simplicity. Where the initial impetus behind the introduction of PIs was to capture some of the complexities of Higher Education, the motivation behind league tables is to simplify: "as with restaurants, televisions or hotels, the ranking of universities appears to provide an easy guide to quality" (ibid).

In a direct comparison between university and football league tables, Malcolm Tight (2000) notes that they share some structural problems. In both cases the tables are dominated by a "self-perpetuating elite" which tends to consist of older institutions with a long-established monopoly of both resources and reputation. Newer entrants not only find it hard to compete with this elite for the top positions, but are met with "distaste and disbelief" if they succeed (p 38). This finding is echoed by King when he considers the problem from the compilers' perspective. He uses the phrase "league table rationality" to summarise the thinking behind them, namely that there is some perfect university ideal against which all institutions can be measured. However, he finds that the compilers undermine their own argument by their need to make their rankings appear "authoritative". To achieve this, they must ensure that the elite continue to be ranked highly, lest they withdraw their data and refuse to cooperate. This consideration drives both the selection of criteria and the relative weightings given to them (King, 2009:142).

2.5 League Table Design

In recent years there have been a few comprehensive surveys of university league tables, including the report by Usher and Savino (2006) which considers 19 league tables from around the world and the UK-focused *Counting what is measured or measuring what counts?* (Centre for Higher Education Research and Information et al., 2008a) which looks in depth at five different league tables. Their findings highlight several key points.

2.5.1 Types of data

It has been noted many times (e.g. Dill and Soo (2005)) that league table indicators are chosen from a mix of inputs, processes and outcomes with little regard for how these are inter-related, if indeed they are. Centre for Higher Education Research and Information et al. (2008a) develop an earlier analysis by Coates (2007) to give a 3 x 3 matrix of the types of data used in the tables they studied (see Table 2-3 below). The first dimension specificies whether the data relates to an input, process or outcome; the second whether it most closely concerns staff, students or the institutions themselves.

Table 2-3: The Inputs-Processes-Outcomes framework with examples of performance indicators given in Centre for Higher Education Research and Information et al. (2008a:10-11)

	Inputs	**Processes**	**Outcomes**
Institutions	Capital resources Income (research, donations, knowledge exchange etc) Reputation Expenditure on library, computing services etc Expenditure per student Bursaries and other student support Range of disciplines covered	Audit (quality and financial) Accreditation	Reputation (opinion surveys)
Staff	Qualifications (teaching and research) Research assessment (to inform teaching) Staff numbers (teaching, research, support etc) International staff Student:staff ratios	Quality assessment Professional accreditation	Awards Research assessment Research publications Citations Teaching/course materials
Students	Entry qualifications Entry requirements Demand/Selectivity: ratio of applicants to admissions Level of study Diversity/access/equity International students	Retention and progression Engagement processes Feedback on course and on graduation	Completion Qualifications and classifications/grades Value added (actual over predicted achievement) Graduate employment, further training etc Feedback on graduation Feedback some time after graduation Awards for alumni

The analysis makes clear that the population of indicators is very diverse, but it also highlights where there are gaps. Most of the institution-level information, for example, is focused on input measures such as expenditure and capital resources; while the only measure of outcome is reputation, assessed through opinion surveys. Reputation could in fact be considered as an input, as Tight's (2000) analysis suggests, and the analysis by CHERI et al includes it in both categories.

An alternative analysis is provided by Usher and Savino (2006), who classify teaching-related indicators into five groups which follow the student's progress through Higher Education:

> Beginning characteristics: attributes of students at the point of entry
>
> Learning inputs – resources: finance and facilities
>
> Learning inputs – staff: includes attributes of the teaching regime, such as contact hours, as well as of the people who provide it
>
> Learning outputs: attributes of students at the point of departure
>
> Final outcomes: the "ultimate ends to which the educational system may contribute" (p 15), such as employment, job satisfaction, citizenship

Together with measures of Research and Reputation, these categories provide a helpful framework for examining league tables from many different countries.

For example, Beginning Characteristics encompasses indicators such as A-level scores and other school-leaving performance measures, which will be familiar to users of UK league tables. However, the category also includes measures of diversity, e.g. the proportion of students who come from other countries, and measures of "study status", e.g. the proportion of students who are attending part-time. The authors point out that where these indicators are used in a league table it is because the compilers believe that they are proxies for measures of quality. A high proportion of international students is believed to indicate an institution which can attract a strong student body, while a high proportion of part-time students is taken as "evidence that an insitution is becoming less rigid in its timetabling" (p 17).

The connection between league table indicators and the aspects of performance they are intended to measure is often unsubstantiated. For example, CHERI et al found that the student:staff ratio "is purported to show the amount of contact time a student might expect" (Centre for Higher Education Research and Information et al., 2008b:39) although no information about how much time each staff member spends in teaching activities is included. For example, the London School of Economics typically scores well on this indicator, but in 2007-8 there was a flurry of activity in the press over the institution's heavy reliance on graduate teaching assistants, who taught about 75% of undergraduate classes. A writer for the student newspaper commented that students on some courses did not meet the senior figures in their departments until their second or third year of study (Newman, 2007).

At the time of the CHERI report, one league table had settled on the somewhat arbitrary student:staff ratio of 10:1 "as a benchmark for excellence" (Appendix C:25) although in practice this is seldom achieved even by the UK's elite institutions. Furthermore, as Hazelkorn explains, "a smaller ratio is viewed as equivalent to better teaching, but in reality this may say more about the funding or efficiency level of the institution and supporting HE system" (2011:62). The key advantage of this indicator is that it is readily available, and not just for UK institutions. It is "the only globally comparable and available indicator that has been identified to address the stated objective of evaluating teaching quality" (QS Intelligence Unit, 2011) even if it has not been proven to do this effectively.

Other typical proxy measures are shown in Table 2-4 below.

Table 2-4: Typical proxy measures, compiled from analysis in Usher and Savino (2006) and Hazelkorn (2011)

Indicator	**Is intended to measure**
Qualifications on entry Proportion of international students	Strength of the student body
Number of part-time students Number of mature students	Flexibility of institution
Staff:Student Ratio National Student Survey scores Number of first and upper second class degrees awarded	Quality of teaching
Citation scores Amount of money received in research grants Proportion of international staff	Quality of research

It is not the purpose of this research to examine such indicators in detail as this is already richly covered by the existing literature. The single most comprehensive contribution has been the book *Performance Indicators in Higher Education* (Johnes and Taylor, 1990) which looks in detail at the range of activities in HE and evaluates ways of measuring them. However, this volume is now over 20 years old and the HE sector has undergone considerable changes in that time. More recent studies which examine individual indicators include Bratti et al. (2004), Smith et al. (2000) who consider graduate destinations and Lee and Buckthorpe (2008) who investigates completion rates. In each case, the authors assess the suitability of an item as a performance measure and suggest an appropriate design.

Others have looked for a relationship between inputs and outputs, such as spend on HE and graduate earnings (Belfield and Fielding, 2001). The findings in this case are ambiguous: some correlation is initially apparent, but after controlling for certain factors the relationship is less clear. Of particular interest is the work by Cyrenne and Grant (2009) which proposes a "reputation function". Assuming that university managers will wish to maximise this function, they posit a number of variables – such as number of graduates, quality of research – which contribute to reputation as measured by a league table ranking. They use an Ordered Probit Model to investigate the relationship between changes in the variables and changes in reputation. As they are studying Canadian universities, they use the league table published by *Maclean's* magazine: a feature of this league table is that it divides universities into categories based on a classification of the institution's mission. Cyrenne and Grant's findings reflect this classification. For example, they find that the reputation of a research-intensive university is enhanced when more resources are focused on this area of activity, while teaching-oriented institutions improve their standing when student-related variables are increased.

Often, however, the conclusion of researchers is that "institutional ranking systems don't measure what their authors think they are measuring" (Usher and Savino, 2006:32). Rather than being effective proxies for attributes of an institution's quality, these indicators may be an expression of an underlying factor. This factor seems likely to be "nothing more than the resource available" concludes Michael (2005:381), which corresponds with Malcolm Tight's analysis (Tight, 2000) of the self-perpetuating nature of elite status.

2.5.2 Scores and Aggregation

Once a certain number of measurable attributes have been identified and the data collected, some manipulation is carried out in order to map the different variables onto a common scale. Methods for doing this include

- awarding a score of 100 to the highest performing institution and awarding a percentage score to the remaining institutions based on their performance relative to the highest (Sunday Times)
- choosing a fixed value as a maximum and awarding all institutions a percentage score based on their performance relative to this value (Sunday Times)
- calculating a z-score (Times) and then converting it to a “band” from 1 to 10 (Guardian).

In their survey, Usher and Savino found that the first of these methods was the most common (2006:8). However, methods are sometimes combined, as in the presentation of the Guardian’s subject tables. Here “the total S-scores [the Guardian’s term for z-scores] are rescaled so that the institution with the best S-score receives 100 points and all others get a lower (but positive) point score” (Hiely-Rayner, 2011).

In UK league tables it is almost always the case that an institution’s score is calculated based on a single year’s data, i.e. it constitutes a point estimate of the institution’s performance and no allowance is made for any fluctuation or uncertainty. Goldstein and Spiegelhalter (1996) have analysed in detail the weakness of this approach in relation to schools data, noting that ranking systems “are particularly sensitive to sampling variability” (p 391). The same concerns apply in the higher education context.

Once the individual scores have been determined they are aggregated into an overall score, typically a weighted sum (Usher and Savino, 2006:8). The weights ascribed to the component scores have a substantial impact on the final ranking and yet their selection often appears to be quite idiosyncratic.

For instance, Table 2-5 shows the relative weights of the indicators used by the Times and the Guardian in their respective indicators. The Guardian has elected to use only teaching-related indicators and so an institution’s RAE performance has no place in their table, while the Times uses it to generate 17% of an institution’s total score. On the other hand, the Guardian gives much greater weight to the National Student Survey than the Times. In general, since the Times has more indicators than the Guardian, each individual component is accorded less weight in the total.

A number of authors have taken issue with the additive approach to aggregation, and some, such as Tofallis (2012), have proposed a multiplicative approach instead. The advantages of multiplication are that the rescaling of variables would be unnecessary and that reverse indicators (such as student:staff ratio, where a smaller value is preferable to a larger one) would not need to be inverted. However, there is currently no sign that any league table is planning to adopt this method in place of the weighted sum.

Table 2-5: Weights ascribed to different indicators by the Times (in red) and the Guardian (in blue)

	Inputs	Processes	Outcomes
Institution	Service & Facilities spend (11% / 15%)		
Staff	Student:Staff ratio (11% / 15%)		RAE score (17%)
Students	Entry standards (11% / 15%)	NSS scores (17% / 25%)	Value Added (15%) Completion (11%) Good Honours (11%) Graduate Jobs (11% / 15%)

2.6 Impact of League Tables

Simon Marginson has said of university rankings that they function as a "meta-performance indicator". Their importance is such that the "criteria used to determine a university's position in the ranking system become meta-outputs that every university is duty bound to place on priority" (Marginson, 2007:2). It is no longer enough to carry out one's mission and then submit to measurement: the measurement has begun to define the mission.

Qualitative research into the impact of league tables ranges from the very local (e.g. (Dixon, 2006)) to the global (Hazelkorn, 2007, 2009, 2011) in scope. Hazelkorn's work is extensive and has already been referenced above. It is especially relevant to this research as she has undertaken some in-depth study of the impact which league tables have on strategic and operational decision-making in universities in Australia, Japan, Germany and Canada. Her analysis of the actions taken by university managers in response to league tables is helpfully mapped on to the indicators most affected by these actions, and she discusses their choices in terms which have a distinct ring of gaming about them (2009:16):

> "The most logical response is to identify indicators which are easiest to influence."

> "The simplest and most cost-neutral actions are those that affect brand and institutional data, and choice of publication or language."

Her quantitative analysis of survey data (Hazelkorn, 2007) is less convincing as the author is prone to citing percentages without making a clear statement of population size, but nonetheless the work does highlight issues of interest such as positive as well as negative perceptions of ranking systems.

A recent report from the Institute for Higher Education Policy (2009) draws heavily on the work of Ellen Hazelkorn (2007, 2009). The difference lies in the intention of the report, which uses the international case studies not only to understand the impact of league tables on institutional decision-making but also to apply this understanding "in ways that might benefit higher education in the United States" (Institute for Higher Education Policy, 2009:1). The report concludes with some high-level recommendations for institutions, e.g. to use rankings as a driver to improve data collection or as a spur to collaboration with international partners.

In the UK there are a number of organisations who have commissioned work on HE and performance measurement, including SCOP (Yorke and Longden, 2005) and HEFCE, the Higher Education Funding Council for England. In 2007 HEFCE charged a steering group with conducting a review of the performance indicators collated and published by HESA (HEFCE, 2007), and then in 2008 they turned their attention to newspaper league tables with the CHERI report (Centre for Higher Education Research and Information et al., 2008a).

The latter report is a very substantial piece of work which looks in detail both at the compilation of league tables (see discussion in section 2.5 above) and the impact on a number of different types of institution. The impact study is restricted to institutions in England, and includes a survey and six case studies. Overall the findings appear to confirm the widespread suspicion that league tables do not provide a comprehensive or well-founded picture of the sector (see, for example, West, 2009, Zemsky, 2008). The research also found a "tension between league table performance and institutions' and government policies and initiatives" (Centre for Higher Education Research and Information et al., 2008a:57) which was giving institution managers concern: it is this tension which motivates the present research.

2.7 Implications for this research

Our topic of concern is the tension between external performance measures, such as league tables and government statistics, in the university sector. We know, from studies such as Hazelkorn (2007) and Broadbent and Loughlin (2010), that such performance measurement systems have an influence on institutions' behaviour. We also know that the priorities arising from the influence of league tables may conflict with other external goals, such as government targets e.g. Centre for Higher Education Research and Information et al. (2008a). This leads us to ask whether we can quantify that tension: if a university pursues one goal, can the effect on another goal be measured?

The CHERI study, and further work by King (2009), also tell us that league tables create uncertainty for institutions as they do not know how others will respond. This leads us to ask whether we can quantify tension between institutions as well as between goals: if two universities pursue the same goal, such as a "top ten" position, what effect will that have on the ability of each to achieve it?

These questions provide a further refinement of the objectives set out in Chapter 1 and will help to focus our development of the proposed quantitative model.

2.8 Summary

In this chapter we have reviewed the history of performance measurement in the UK higher education sector and, in particular, the rise of newspaper league tables. We have considered both the design of such ranking systems and their impact on decision making in HEIs. This discussion has enabled us to identify some specific tensions which will provide a focus for the present research. As described in Chapter 1, the methods we have identified for this research are Data Envelopment Analysis and game theory, which we will introduce fully in the next two chapters.

3 Data Envelopment Analysis

The purpose of this chapter is to give an overview of Data Envelopment Analysis (DEA). DEA is a set of techniques for evaluating the relative efficiencies of a group of comparable decision making units (DMUs), which might be branches of a business, such as banks or factories, or public sector facilities, such as schools or hospitals.

First the basic features of DEA are introduced, then its application to Higher Education is considered through an in-depth discussion of this body of literature.

3.1 Origins

"The story of data envelopment analysis begins with Edwardo Rhodes's PhD dissertation research at Carnegie Mellon University's School of Urban and Public Affairs" (Charnes et al., 1994: 3).

In the 1970s, when Rhodes was seeking to evaluate and compare the performance of different initiatives for disadvantaged school pupils, the models available for measuring efficiency were econometric ones based essentially on price. In a public sector context, where price information was not readily available, the results he obtained were "unsatisfactory and even absurd" (Cooper et al., 2004: 4).

An alternative approach was found in "The Measurement of Productive Efficiency", a paper by M J Farrell (Farrell, 1957) in which he draws on the concepts of activity analysis (Koopmans, 1951) to estimate an empirical production function. Possibly because of computational limitations which were recognised at the time[2], Farrell's model appears to have been left relatively untouched for the next twenty years (Coelli et al., 1998). However, by the time Rhodes took it to his supervisor, William Cooper, computational power had increased and Cooper himself had done some key work in linear programming in collaboration with Abraham Charnes (e.g. Charnes and Cooper, 1962, 1973) which made it possible to put Farrell's idea into practice. The results were published in 1978 under the title "Measuring the Efficiency of Decision-Making Units" (Charnes et al., 1978).

[2] Farrell presented his paper at a meeting of the Royal Statistical Society and questions and comments from the audience are recorded. For instance, Professor M G Kendall remarked that Farrell's method "would strain the resources even of an electronic computer if many variables were involved"(Farrell, 1957).

The original motivation for developing DEA was "to provide a methodology whereby … those [DMUs] exhibiting best practice could be identified" (Cook and Seiford, 2009: 1) so that this best practice could be shared, and this is still the motivation for much DEA research today (Dyson and Shale, 2009). It is worth noting that this aim presupposes a shared management or a degree of cooperation between DMUs.

3.2 Key characteristics of DEA

3.2.1 Performance not based on price

As its origins suggest, DEA is particularly appropriate for and has been widely applied to public sector organisations where price is not the key consideration in evaluating performance.

3.2.2 Multiple inputs and outputs

Another feature which is useful in the public sector context is DEA's ability to handle multiple inputs and multiple outputs.

3.2.3 Non-parametric approach

DEA is a non-parametric approach, which means that efficiency measures are derived from the observed data without assuming an underlying functional form.

3.2.4 Efficient frontier

Unlike a regression model, which considers deviation from a central tendency, DEA builds a frontier "on top of" the observed data.

3.3 Efficiency in DEA

The following statements define what is meant by 'efficiency' in the context of DEA.

3.3.1 Extended Pareto-Koopmans Efficiency

"Full efficiency is attained by any DMU if and only if none of its inputs or outputs can be improved without worsening some of its other inputs or outputs."

3.3.2 Relative Efficiency

"A DMU is rated to be fully efficient on the basis of available evidence if and only if the performances of other DMUs does not show that some of its inputs or outputs can be improved without worsening some of its other inputs or outputs."

(Cooper et al., 2004: 3)

3.4 Basic model

Productivity is commonly defined to mean the ratio of output to input (Fried et al., 2008). Such a ratio is easily calculated where there is only one output and one input, so a case of this sort makes a good illustration of the principles of DEA.

3.4.1 Single input–single output case

For this illustration we adopt a very simple dataset with eight DMUs as shown in Table 3-1 below.

Table 3-1: Simple dataset for DEA illustration

DMU	Input	Output
A	1.0	8
B	2.0	10
C	2.5	17
D	2.0	8
E	4.0	20
F	0.8	4
G	3.0	14
H	3.0	7

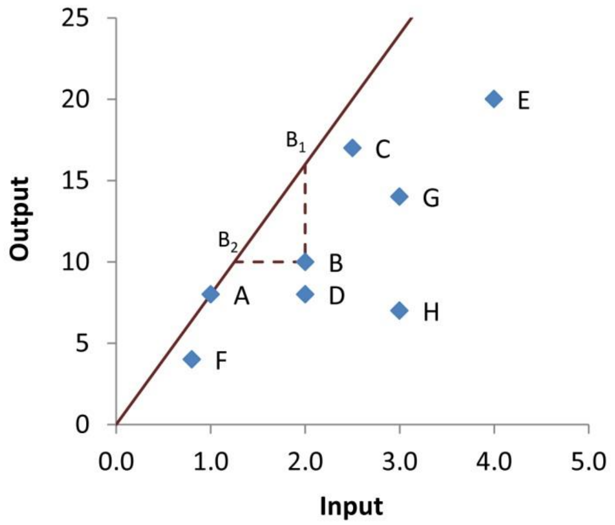


Figure 3-1: A single input—single output case with constant returns to scale

Table 3-1 and Figure 3-1 show eight DMUs with varying levels of input and output. The DMU with the highest ratio of output to input is DMU *A*. The line from the origin through *A* shows the **efficient frontier**, which can be used to determine either

- the level of output we would expect to obtain from a DMU with a given level of input (this approach is termed **output-oriented**), or
- the level of input we would expect a DMU to use to achieve a given level of output (this approach is termed **input-oriented**).

In this simple example we are assuming for the moment that **returns to scale** are constant, i.e. that the achievable ratio of output to input is the same no matter how large or small the DMU.

The efficiency of the other units is determined relative to *A*, so that for each DMU we calculate a ratio

$$0 \leq \frac{\text{output of unit } x \,/\, \text{input of unit } x}{\text{output of unit } A \,/\, \text{input of unit } A} = \frac{\text{productivity of } x}{\text{productivity of } A} \leq 1 \qquad (3.1)$$

(Cooper et al., 2006)

Figure 3-1 also illustrates how an inefficient unit such as *B* can be made efficient, either by increasing the level of output (to B_1) or decreasing the level of input (to B_2). Both of these new units have the same productivity as *A* and therefore an efficiency of 1.

3.4.2 Two input—single output case

A useful first step in generalising this simple example is to consider the case of two inputs and one output. For example, a university might be considered to have its academic staff and facilities as inputs contributing to the output of graduates.

We divide each input by the output to obtain the normed data shown in Table 3-2 and plot it on a graph as shown in Figure 3-2. There are two units in this case which are found to be efficient: DMUs *D* and *G*. These units define the efficient frontier. The area of the graph beyond the frontier and containing the remainder of the DMUs is the **production possibility set**.

Table 3-2: Data for two input—single output model

	Input 1/Output	Input 2/Output
A	2.00	6.00
B	1.50	5.00
C	3.00	4.00
D	1.00	3.00
E	1.50	3.00
F	4.00	4.00
G	2.50	2.00

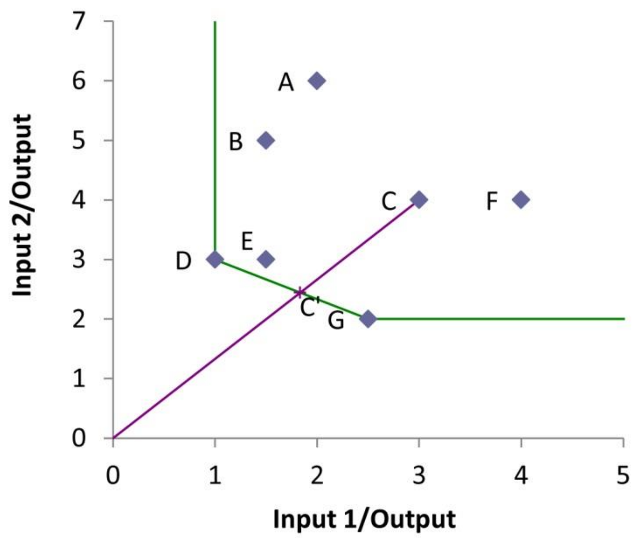


Figure 3-2: Model showing production possibility set and radial efficiency

By drawing a line from the origin through any inefficient DMU, e.g. *C*, we can identify a point on the frontier that corresponds to an equivalent efficient DMU, assuming that returns to scale are constant. The ratio OC'/OC tells us by how much input needs to be reduced in order to match the most efficient units. This is a measure of the DMU's **radial efficiency** or **technical efficiency**, so-called because the inputs can be simultaneously reduced in the same proportions, i.e. there is no need to change the "basic technological recipe" (Banker and Morey, 1986a: 513).

Once there are more than three factors involved – either one input and two outputs or two inputs and one output – it becomes increasingly difficult to visualise the frontier, the production possibility set and the projection of radial efficiency. In the two-dimensional chart above, we are looking at a piecewise linear boundary. In the *n*-dimensional space we need for more complex problems, the boundary consists of (*n*-1)-dimensional facets.

3.4.3 The CCR Model

The CCR model is so named for its originators, Charnes, Cooper and Rhodes. It is a general model for a set of *n* DMUs, each with *m* inputs and *s* outputs. All the data are assumed to be nonnegative, and it is further assumed that at least one input and one output for each DMU is nonzero (Cooper et al., 2006). The efficiency of each DMU is assessed relative to all the other DMUs by maximising the ratio of weighted outputs to weighted inputs (Charnes et al., 1978):

$$\begin{aligned} \max \quad & h_o = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \\ \text{subject to} \quad & \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1 \qquad j = 1, \ldots, n \\ & u_r, v_i \ge 0 \qquad r = 1,..,s;\ i = 1,..m \end{aligned} \tag{3.2}$$

For each DMU the weights u_r and v_i are selected to give it the best possible efficiency ratio. Charnes et al note that, "Under these observations and constraints no other set of common weights will give a more favourable rating relative to the reference set. Hence if a (relative) efficiency rating of 100% is not achieved under this set of weights then it will also not be attained from any other set" (1978: 431).

It is common to reformulate the fractional program given in (3.2) to obtain the equivalent linear program (3.3) which can be neatly expressed in matrix notation, where *v* and *u* are row vectors of input and output multipliers.

$$\begin{aligned} \max_{v,u} \quad & z_o = u y_o \\ \text{subject to} \quad & v x_o = 1 \\ & -vX + uY \le 0 \\ & v \ge 0, u \ge 0. \end{aligned} \tag{3.3}$$

This form of the problem is called the **multiplier form** and it is easy to see its relationship to the original fractional program. However, in general, a DEA problem will involve a large number n of DMUs and a much smaller number $m+s$ of inputs and outputs. Since (3.3) has $m+s$ variables and n + 1 constraints, it is computationally more efficient to solve its dual problem (3.4) which has n + 1 variables and only $m+s$ constraints.

$$\begin{aligned} &\min_{\theta,\lambda} && \theta \\ &\text{subject to} && \theta x_o - X\lambda \geq 0 \\ &&& Y\lambda \geq y_o \\ &&& \lambda \geq 0. \end{aligned} \tag{3.4}$$

The dual problem is called the **envelopment form**, and it is useful not only for its computational benefits; it also yields pertinent information. In particular, $\theta = z_o$ gives the efficiency of the DMU under consideration and $\lambda = (\lambda_1,....,\lambda_n)^T$ identifies the reference set, or peers, of this DMU. These peers are DMUs on the efficient frontier whose characteristics most closely match those of the DMU under evaluation. "It is as well to have the referents used for scoring the efficiency of each DMU as alike to it as possible, " note Charnes et al (1978: 437), since different units will be working under different constraints, e.g. with respect to the availability of inputs. Where $\lambda_k > 0$, DMU_k is a member of the reference set, and the convex combination $\sum_{k=1}^{n} \lambda_k DMU_k$ is the radial projection of the inefficient DMU onto the efficient frontier.

In the *Handbook on Data Envelopment Analysis* Thanassoulis et al (2004) describe the envelopment model as a "production framework", which emphasises the technology, and the multiplier model as a "value framework", which emphasises the free choice of weights. In the latter case, the values of the weights may be freely chosen to make the DMU look as good as it possibly can: if the process cannot discover weights that make the DMU appear efficient then it really must be inefficient.

The model we have just described is an input-oriented model. The corresponding output-oriented model is given in envelopment form by

$$\begin{aligned} &\max_{\eta,\mu} && \eta \\ &\text{subject to} && x_o - X\mu \geq 0 \\ &&& \eta y_o - Y\mu \leq 0 \\ &&& \mu \geq 0. \end{aligned} \tag{3.5}$$

Exploring this model more fully (see, for example, Cooper et al., 2006: 58-60) shows that what has been done is to exchange the numerator and denominator in our original formulation (3.2) It is therefore no surprise that the optimal solution of this model (η^*, μ^*) is very easily related to the optimal solution of (3.4) by $\eta^* = 1/\theta^*$ and $\mu^* = \lambda^* / \theta^*$.

The relationship seen here between input- and output-oriented efficiency depends on our original formulation of the problem and, in particular, on the assumption of constant returns to scale.

Returning to the input-oriented model in (3.4) let us take another look at the constraints. When the optimal $\theta^* < 1$ then we can reduce the inputs from x_o to $\theta^* x_o$ but by our first constraint the reduced value may still be greater than $X\lambda$, our convex combination of peer DMUs. Any such excess input, and similarly any shortfall in output, is referred to as **slack** and we write

$$\begin{aligned} s^- &= \theta^* x_o - X\lambda \\ s^+ &= Y\lambda - y_o \end{aligned} \tag{3.6}$$

In order to calculate the size of these slacks, we add an extra step to our linear program. Having minimised θ in (3.4) we then maximise the total slack as follows:

$$\begin{aligned} \max_{\lambda, s^-, s^+} \quad & \sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+ \\ \text{subject to} \quad & s^- = \theta^* x_o - X\lambda \\ & s^+ = Y\lambda - y_o \\ & \lambda, s^-, s^+ \geq 0 \end{aligned} \tag{3.7}$$

The concept of slacks is illustrated below in Figure 3-3. We have added a new DMU *H* to the dataset in Table 3-2. *H* is apparently on the frontier, but it is clearly not as efficient as *G* which is managing to use less of Input 1 to the same quantity of Input 2. Solving (3.7) for DMU *H* gives us the maximum slack, i.e. the largest quantity of Input 1 which can be removed within the constraints of the production possibility set. In this case, $H' = G$.

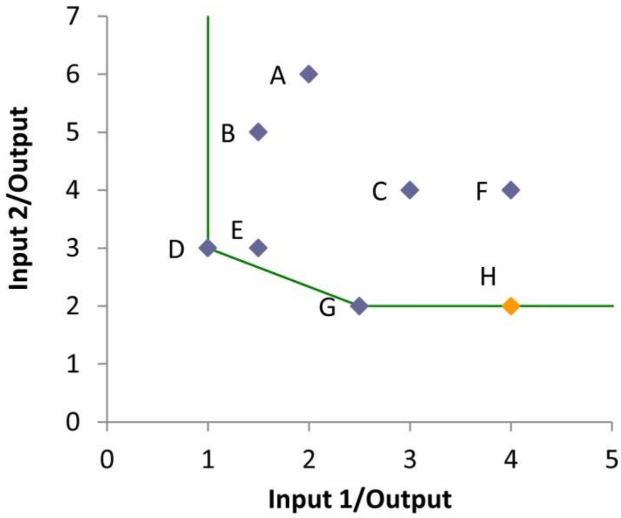


Figure 3-3: Model showing DMU with slack in Input 1

These two steps can be combined in a single linear program. The input- and output-oriented models are respectively

$$\begin{aligned} \min_{\theta,\lambda,s^-,s^+} \quad & \theta - \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right) \\ \text{subject to} \quad & s^- = \theta x_o - X\lambda \\ & s^+ = Y\lambda - y_o \\ & \lambda, s^-, s^+ \geq 0 \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} \max_{\eta,\mu,s^-,s^+} \quad & \eta + \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right) \\ \text{subject to} \quad & s^- = x_o - X\mu \\ & s^+ = Y\mu - \eta y_o \\ & \mu, s^-, s^+ \geq 0 \end{aligned} \tag{3.9}$$

where $\varepsilon > 0$ is a non-Archimidean element defined to be smaller than any real number (Cooper et al., 2004). Although this formulation provides a compact expression of the CCR model, in practice the computation is carried out in the two stages outlined above.

Taking into account the two elements of this model – radial efficiency and slacks – we arrive at a more explicit definition of efficiency.

3.4.4 CCR efficiency

A DMU is called CCR-efficient if there is an optimal solution to the CCR model such that (i) $\theta^* = 1$ and (ii) all slacks are zero. This is also referred to as **strong efficiency** or **Pareto-Koopmans efficiency** (see definition above).

3.4.5 Farrell efficiency

If a DMU satisfies condition (i) of CCR efficiency but not condition (ii) then it is said to have Farrell efficiency, also called **weak efficiency** (Cooper et al., 2006: 46).

3.5 Variant models

3.5.1 A chronology of the main DEA models

DEA has developed in a number of directions to address the requirements of different spheres of interest. Table 3-3 presents the main such developments in chronological order.

Table 3-3: Chronology of the main DEA models compiled from information in Cook and Seiford (2009)

Year	Authors	Model	Comments
1978	Charnes, Cooper & Rhodes	CCR	The first DEA model, gives radial efficiency under constant returns to scale.
1984	Banker, Charnes & Cooper	BCC	Gives radial efficiency under variable returns to scale: adds the constraint $\sum_k \lambda_k = 1$ to the CCR model.
1984	Deprins, Simar & Tulkens	Free disposal hull (FDH)	Rejects the possibility of convex combinations of DMUs on the efficient frontier, which is therefore stepped between observed DMUs: equivalent to the BCC model but with $\lambda_k \in \{0,1\}$.
1985	Charnes, Cooper, Golany, Seiford & Stutz	Additive or Pareto-Koopmans (PK)	Permits non-radial projection (efficiency change does not require inputs/outputs to be kept in proportion) but emphasis is on identifying slacks and the model does not provide a consistent measure of efficiency.

Year	Authors	Model	Comments
1996	Yu, Wei & Brockett	Generalized DEA (GDEA)	Incorporates three parameters which, when set, resolve the general model to one of a number of specific models, including the CCR and BCC models.
1999	Cooper, Park & Pastor	Range adjusted measure (RAM)	Another non-radial model similar to the additive model.
2001	Tone	Slacks-based measure (SBM)	Addresses shortcomings in additive model by providing a "legitimate PK efficiency score in the spirit of the CCR and BCC models" (2009: 5).

3.5.2 Variable returns to scale

The different configurations of returns to scale are best illustrated with the single input-single output model we first considered. In Figure 3-4 we see our original dataset from Table 3-1 and the line through *A* which gives us the efficient frontier under **constant returns to scale**. We also see a second frontier which again passes through *A* but this time is wrapped tightly around the data: it is defined by several DMUs all at the extremes of the dataset. This is the efficient frontier under **variable returns to scale,** i.e. when the achievable ratio of output to input varies with the overall size of the DMU.

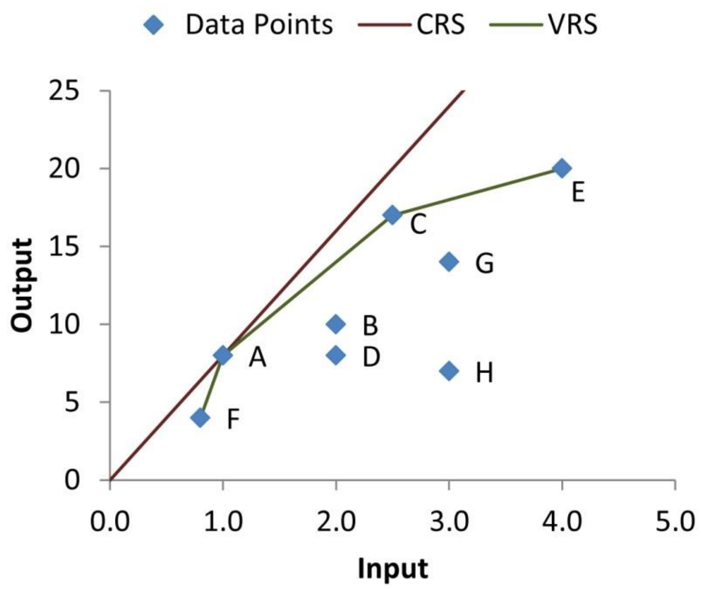


Figure 3-4: Constant Returns to Scale (CRS) and Variable Returns to Scale (VRS)

From *F* to *A* the VRS frontier is steeper than the CRS frontier, which reflects an increasing return to scale in this part of the dataset. From *A* to *E* the VRS frontier is shallower than the CRS frontier, reflecting a decrease in the returns to scale.

The CRS and VRS frontiers may in fact be combined to reflect assumptions about the dataset. In Figure 3-5a we combine the CRS frontier from the origin to *A* with the VRS frontier thereafter to obtain a frontier which captures **non-increasing returns to scale,** i.e. the returns to scale are either constant or decreasing. In Figure 3-5b we do the opposite to obtain the frontier under **non-decreasing returns to scale,** i.e. the returns to scale are either constant or increasing.

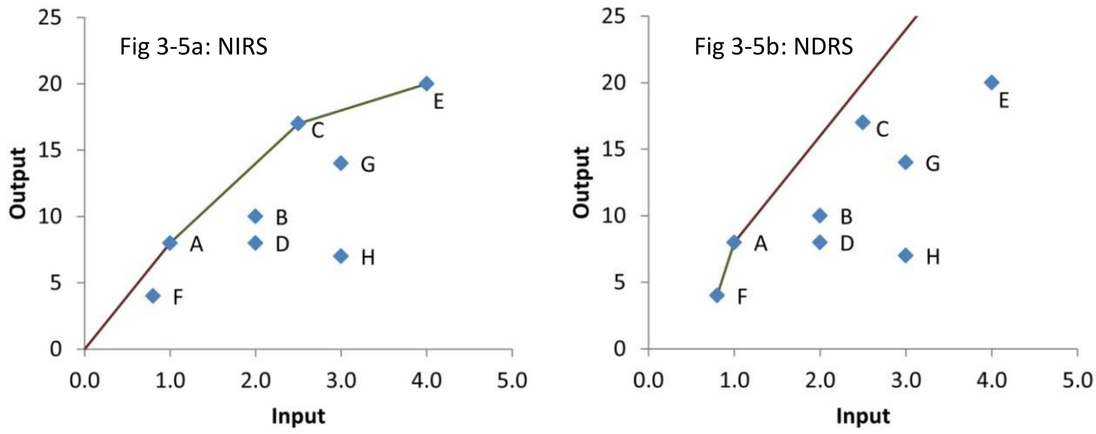


Figure 3-5: Non-Increasing Returns to Scale (NIRS) and Non-Decreasing Returns to Scale (NDRS)

Using the CCR model it is possible to identify for any given DMU whether it is operating under constant, increasing or decreasing returns to scale. The conditions are as follows (Banker et al., 2004: 49):

- Constant returns to scale prevail if $\sum_{k=1}^{n} \lambda_k = 1$ in any alternate optimum
- Decreasing returns to scale prevail if $\sum_{k=1}^{n} \lambda_k > 1$ for all alternate optima
- Increasing returns to scale prevail if $\sum_{k=1}^{n} \lambda_k < 1$ for all alternate optima

3.5.3 The BCC model

This model, developed by and named for Banker, Charnes and Cooper (1984), evaluates each DMU against the variable returns to scale frontier seen above in Figure 3-4. This is done by adding to the formulation (3.4) the additional constraint that $\sum \lambda = 1$.

For practical purposes the frontier is extended from the endpoints as shown below in Figure 3-6. A DMU with an efficient projection onto this extended part of the frontier has slacks which can be identified through the same two-stage process described for the CCR model.

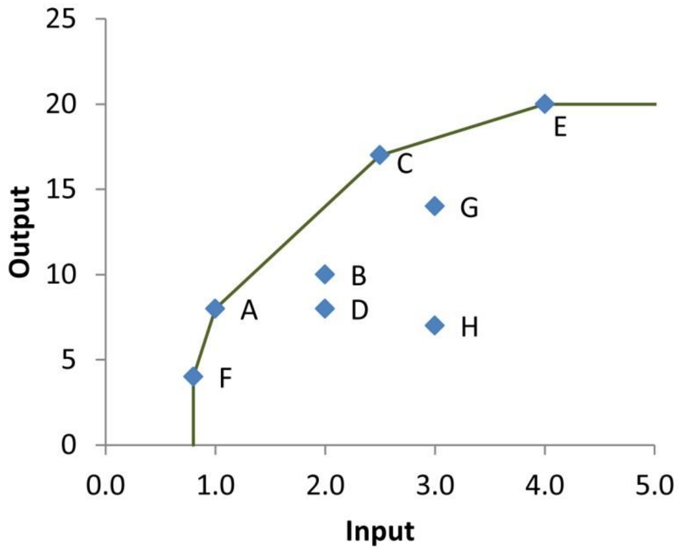


Figure 3-6: Extended VRS frontier of the BCC model

The full BCC model is described by

$$\begin{aligned}
\min_{\theta,\lambda,s^-,s^+} \quad & \theta - \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right) \\
\text{subject to} \quad & s^- = \theta x_o - X\lambda \\
& s^+ = Y\lambda - y_o \\
& \sum_k \lambda_k = 1 \\
& \lambda, s^-, s^+ \geq 0
\end{aligned} \tag{3.10}$$

When considering the choice between input and output orientation, note that constant returns to scale gives same efficiency in both cases but variable returns to scale may not.

Consider point *G* in Figure 3-6 above. If it is projected towards the frontier by moving parallel to the x-axis (i.e. reducing input) then it will be mapped to a point on the line *AC*. However, if it is projected by moving parallel to the y-axis (i.e. increasing output) then it will be mapped to a point on the line *CE*. These projections onto different linear segments will yield different efficiency scores. This is discussed in detail in Banker et al. (2004: 44).

3.5.4 Restricting multiplier weights

When working with a multiplier model, as in (3.3), it often makes sense to restrict the values taken by the weights *u* and *v* beyond stating that they must be non-negative. One of the criticisms of the CCR model is that units may be deemed to be efficient by considering only one or two performance factors while applying zero weights to all the others. While the mathematics permits this, in business terms such an evaluation is unrealistic. In practice there are a number of different ways of constraining the multiplier weights and each has advantages and disadvantages.

Absolute bounds

The most straightforward option would appear to be the imposition of absolute bounds, i.e. to specify a numeric range for each weight variable whose value you wish to constrain. However, the difficulty which immediately arises is knowing what range is appropriate (Thanassoulis et al., 2004): the various factors may be on very different scales and the weights will vary accordingly. At the very least it would be necessary to perform an unconstrained DEA first in order to establish the likely range of weights. Another alternative would be to do a regression analysis to investigate the relative contributions of each factor to the total performance (Dyson and Thanassoulis, 1988: 567); however, in a multiple input-multiple output model this poses the problems which DEA was developed to avoid. Furthermore, as in the unconstrained case, the selection of weights is easier to justify mathematically than managerially: the choice of bounds is made to suit the arithmetic rather than the business sense.

Assurance region

A more intuitively meaningful approach is embodied in the Assurance Region (AR) model (Cooper et al., 2006). In this case it is the ratio of the weights which is constrained, e.g.

$$LB_{21} \le \frac{v_2}{v_1} \le UB_{21}$$

where a lower bound and an upper bound are chosen to limit the ratio of v_2 to v_1. It is not necessary to specify a complete set of bounds for either the input or the output weights.

A variant of the AR model is the generalised Assurance Region Global (ARG) model (ibid.). In this model constraints are placed not directly on the weights but on the *weighted factors* so that, for example, a particular weighted input is restricted to a bounded proportion of the total weighted input, as proposed by Wong and Beasley (1990):

$$LB_1 \leq \frac{v_1 x_1}{v_1 x_1 + \ldots + v_n x_n} \leq UB_1$$

Again, it is not necessary to specify a complete set of bounds but, if all the inputs (or outputs) are constrained, then the sum of the lower bounds must be less than or equal to one, while the sum of the upper bounds must be greater than or equal to one.

The values of the lower and upper bounds could be derived from unconstrained DEA or by eliciting expert judgment. The ARG model may be particularly well-suited for elicitation as the interpretation of the limits is intuitive and readily explained.

The assurance region is a special case of the more general Cone-Ratio approach of Charnes et al (1990). This approach entails defining a restricted space *V* for the input weights by means of vectors. Another space *U* is defined for the output weights, and the constraints $v \in V$, $u \in U$ are added to the model in (3.3).This is a flexible approach which can be adapted for use with other DEA models, such as the additive model, but it is not as easily interpreted as the AR or ARG approach.

Trade-offs

Although this method is primarily motivated by consideration of the envelopment model, it can be implemented by restricting weights in the multiplier model and so we consider it here.

The method is presented in Podinovski (2004) where the following example of a trade-off is given. Suppose that a group of university departments all teach both undergraduates and masters students. They agree that

1. If one masters student is taken away then one undergraduate can be added without needing any additional resources.

2. If two undergraduates are taken away then one masters student can be added without needing any additional resources.

The asymmetry is deliberate and represents the judgment that the ratio between undergraduates and masters students is not precisely defined but lies somewhere between 1:1 and 2:1.

In a system of one input (staff) and two outputs (undergraduates and masters), these two trade-offs can be expressed as vectors

$$\mathbf{P_1} = (0), \mathbf{Q_1} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
$$\mathbf{P_2} = (0), \mathbf{Q_2} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

These vectors are then incorporated into the RHS of the envelopment constraints

$$\boldsymbol{\lambda}\mathbf{X} + \pi_1\mathbf{Q_1} + \pi_2\mathbf{Q_2} \leq \theta\mathbf{X_0}$$
$$\boldsymbol{\lambda}\mathbf{Y} + \pi_1\mathbf{P_1} + \pi_2\mathbf{P_2} \geq \mathbf{Y_0}$$

for some values of π_1 and π_2. The axioms which define the standard CRS technology are feasibility of observed data, convexity, free disposability and proportionality (Cooper et al., 2006). In addition to these, Podinovski introduces the following axiom for the feasibility of trade-offs (2004: 1314):

> "Let $(x, y) \in T_{CRS}$ [the CRS technology]. Then, for any trade-off *t* in the [above form] and any $\pi_t \geq 0$, the unit $(x + \pi_t P_t, y + \pi_t Q_t) \in T_{CRS}$, provided $x + \pi_t P_t \geq 0$ and $y + \pi_t Q_t \geq 0$."

In general, the radial efficiency model under CRS looks like this:

First, the input-oriented model, where the efficiency is given by θ.

$$\begin{aligned}
&\min_{\theta,\lambda,\pi} && \theta \\
&\text{subject to} && Y\lambda + \sum_{t=1}^{k} \pi_t Q_t \geq Y_0 \\
& && X\lambda + \sum_{t=1}^{k} \pi_t P_t \leq \theta X_0 \\
& && \lambda, \pi \geq 0 \\
& && \theta \text{ sign free}
\end{aligned} \tag{3.11}$$

Second, the output-oriented model, where the efficiency is given by $1/\theta$.

$$\begin{aligned}
&\max_{\theta,\lambda,\pi} && \theta \\
&\text{subject to} && Y\lambda + \sum_{t=1}^{k} \pi_t Q_t \geq \theta Y_0 \\
& && X\lambda + \sum_{t=1}^{k} \pi_t P_t \leq X_0 \\
& && \lambda, \pi \geq 0 \\
& && \theta \text{ sign free}
\end{aligned} \tag{3.12}$$

For VRS the usual constraint summing the λ_j to 1 can be added if the axiom of proportionality is dropped(ibid: 1316).

Translating the trade-offs into multiplier form gives a straightforward set of constraints, namely

$$u^T Q_t - v^T P_t \leq 0 \quad t = 1,2,\ldots k$$

which can be added either to the model in (3.3) or to the output-oriented model.

Like the ARG model, this is one which appears to make sense managerially so that expert judgment could be used to elicit the trade-offs. Podinovski argues (2004: 1316) that using trade-offs rather than weight restrictions keeps the PPS realistic: in other words, the trade-offs are about what is technically feasible rather than what is desirable.

Weight restrictions in practice

Models of the type discussed here have been developed to address one of the perceived shortcomings of DEA, that its mechanisms are not always suited to the real life organisations where it is applied. We noted as motivation the unsatisfactory situation arising when a DMU is deemed to be efficient, but the weights which contribute to this evaluation appear extreme, e.g. when a key input or output is assigned a zero weight in the DEA process. Introducing restrictions on the permitted weights is a method of incorporating prior knowledge or 'value judgements' into the DEA model. However, these judgements in their turn raise both theoretical and practical concerns.

Allen et al. (1997) identify several different motivations for incorporating value judgements, and note that some motivations (and the corresponding models) may be in direct opposition. For instance, a model which involves the preselection of DMUs for assessing efficiency will tend to support the prevailing managerial view of top performers, while a model which focuses on the relative value of inputs and outputs may produce an efficiency rating which surprises and challenges that view (pp 15-16).

Whichever approach is adopted, there are challenges for the modeller in identifying (a) where to impose restrictions and (b) what values those restrictions should take. These challenges were noted by Allen et al in 1997 and, although there have been some innovative developments informing weight selection (e.g. Cooper et al. (2007), Thanassoulis et al. (2012), Podinovski and Bouzdine-Chameeva (2013)) there is not yet a general method for addressing these challenges in any given modelling situation.

An important point to note about the trade-off approach proposed by Podinovski is that its basis in the envelopment model means that the focus is on 'technological thinking' rather than 'value thinking' (Førsund, 2012). This does not mean that the problems noted above do not apply, but it does mean that there is some basis for addressing them by considering what is technologically realistic. Podinovski emphasises (2005) that the trade-offs in his model are not equivalent to marginal rates of substitution as defined in production economics, as they are defined throughout the production space and not only on the efficient frontier. This is achieved by ensuring that they are "sufficiently undemanding" to be globally applicable (p 1410). Podinovski also notes (2004: 1320) that this requirement means that they are necessarily more relaxed than those used in Thanassoulis and Allen's approach (1998, 2012), which is based on judgements which are local to each DMU. Førsund (2012) accepts Podinovski's definition, but observes that such trade-offs may be regarded as bounds on the marginal rates (p 281).

It has been argued that, in general, the use of weight restrictions under variable returns to scale is inappropriate (Thanassoulis and Allen, 1998). Podinovski suggests that the rigorous formulation of the PPS in this model, which gives us $T_{VRS-TO} \subset T_{CRS-TO}$, overcomes this problem. He argues that if the trade-offs are regarded as realistic in the CRS space then they must also be realistic in the VRS space. However, he does limit his claim to the local area of the PPS where the DMUs can be found, since – like returns to scale – it doesn't make sense to extrapolate trade-offs to their outer limits.

3.5.5 Super-efficiency

Another approach to the problem of too many DMUs being classified as efficient is the concept of "super-efficiency" introduced (although not so named) by Andersen and Petersen (1993).

The essence of these methods is to find a way of ranking the efficient DMUs. The method proposed by Andersen and Petersen involves excluding each DMU from its own evaluation and comparing its frontier position with its frontier position in a conventional analysis. This gives a ratio greater than or equal to 1: the higher this ratio, the further the efficient DMU extends the frontier when it is included in the dataset.

As this research is not concerned with the actual measure of efficiency, further developments in this approach will not be considered here.

3.5.6 Measure-specific models

"Measure-specific" is the term used by Cook and Zhu (2005) to describe DEA models which selectively focus on minimising (or maximising) a subset of inputs (or outputs). The motivation for doing this may either be a preference for setting targets on some factors rather than others (Zhu, 1996) or a situation where some factors are outwith managerial control (Banker and Morey, 1986a).

The basic model given by Cook and Zhu allows the selection of a subset either of inputs or of outputs, depending on the orientation of the analysis.

Input-oriented model

Let $I = \{1, \ldots, m'\}$ be an index set for those inputs which are of particular interest.

$$\begin{aligned}
\min_{\theta,\lambda,s^-,s^+} \quad & \theta-\varepsilon\left(\sum_{i=1}^{m}s_i^- + \sum_{r=1}^{s}s_r^+\right) \\
\text{subject to} \quad & \sum_{j=1}^{n}\lambda_j x_{ij} + s_i^- = \theta x_{i0} \qquad i\in I \\
& \sum_{j=1}^{n}\lambda_j x_{ij} + s_i^- = x_{i0} \qquad i\notin I \\
& \sum_{j=1}^{n}\lambda_j y_{rj} - s_r^+ = y_{r0} \qquad r=1,\dots,s \\
& \lambda, s^-, s^+ \geq 0
\end{aligned} \tag{3.13}$$

Model (3.13) yields a set of targets for the evaluated DMU of the form

$$\begin{aligned}
\hat{x}_{i0} &= \theta^* x_{i0} - s_i^{-*} \quad & i\in I \\
\hat{x}_{i0} &= x_{i0} - s_i^{-*} \quad & i\notin I \\
\hat{y}_{r0} &= y_{r0} + s_r^{+*} \quad & r=1,\dots,s
\end{aligned}$$

The targeted inputs are reduced by the efficiency measure θ which lies between 0 and 1. Other inputs are maintained or reduced while outputs are maintained or increased.

Output-oriented model

Let $R=\{1,\dots,s'\}$ be an index set for those outputs which are of particular interest.

$$\begin{aligned}
\max_{\phi,\lambda,s^-,s^+} \quad & \phi+\varepsilon\left(\sum_{i=1}^{m}s_i^- + \sum_{r=1}^{s}s_r^+\right) \\
\text{subject to} \quad & \sum_{j=1}^{n}\lambda_j x_{ij} + s_i^- = x_{i0} \qquad i=1,\dots,m \\
& \sum_{j=1}^{n}\lambda_j y_{rj} - s_r^+ = \phi y_{r0} \qquad r\in R \\
& \sum_{j=1}^{n}\lambda_j y_{rj} - s_r^+ = y_{r0} \qquad r\notin R \\
& \lambda, s^-, s^+ \geq 0
\end{aligned} \tag{3.14}$$

In this case the model yields a set of targets for the evaluated DMU of the form

$$\begin{aligned}
\hat{x}_{i0} &= x_{i0} - s_i^{-*} \quad & i=1,\dots,m \\
\hat{y}_{r0} &= \phi y_{r0} + s_r^{+*} \quad & r\in R \\
\hat{y}_{r0} &= y_{r0} + s_r^{+*} \quad & r\notin R
\end{aligned}$$

The targeted outputs are increased by the factor ϕ which is greater than or equal to 1. Other outputs are maintained or increased while inputs are maintained or reduced.

Two limitations of this model are immediately apparent: the restriction, noted above, to improving either inputs or outputs at any one time, and the fact that a single multiplier is used to scale all the selected factors to the same degree.

Thanassoulis and Dyson (1992) address these limitations with a more general model which allows the user to specify a preference structure over both inputs and outputs at the same time. Furthermore, this model incorporates a user-specified weighting over the selected factors. They note that,

> "A general preference structure would allow a different degree of importance to be attached to the potential changes of individual input or output levels, it would allow changes to input as well as output levels, and it would not necessarily require such changes to be the same proportion of initial levels." (1992: 84)

3.5.7 Weights-based general preference structure

We define the following index sets:

$$I_0 \subseteq I = \{1, \dots, m\}$$
$$R_0 \subseteq R = \{1, \dots, s\}$$

I_0 is the subset of inputs we wish to improve, while R_0 is the subset of such outputs. For each $x_i, i \in I_0$ we specify a weight w_i^- and for each $y_r, r \in R_0$ we specify a weight w_r^+ "to reflect the relative degree of desirability of improvement of the corresponding input-output levels" (1992: 85)

$$\begin{aligned}
\max_{z,p,d^-,d^+} \quad & \sum_{r\in R_0} w_r^+ z_r - \sum_{i\in I_0} w_i^- p_i + \varepsilon\left(\sum_{r\in\overline{R}_0} d_r^+ + \sum_{i\in\overline{I}_0} d_i^-\right) \\
\text{subject to} \quad & z_r y_{rj_0} - \sum_{j=1}^{n} \beta_j y_{rj} = 0 \quad r\in R_0 \\
& p_i x_{ij_0} - \sum_{j=1}^{n} \beta_j x_{ij} = 0 \quad i\in I_0 \\
& \sum_{j=1}^{n} \beta_j y_{rj} - d_r^+ = y_{rj_0} \quad r\in\overline{R}_0 \\
& \sum_{j=1}^{n} \beta_j x_{ij} + d_i^- = x_{ij_0} \quad i\in\overline{I}_0 \\
& z_r \ge 1 \quad \forall r\in R_0 \\
& p_i \le 1 \quad \forall i\in I_0 \\
& \beta_j \ge 0 \quad \forall j \\
& p_i, z_r \text{ free} \quad \forall i\in I_0 \text{ and } r\in R_0 \\
& d_i^-, d_r^+ \ge 0 \quad \forall i\in\overline{I}_0 \text{ and } r\in\overline{R}_0
\end{aligned} \tag{3.15}$$

Given our weights, as described above, this process aims to maximise the difference between the increases to our preferred outputs $\sum_{r\in R_0} w_r^+ z_r$ and the decreases to our preferred inputs $\sum_{i\in I_0} w_i^- p_i$, while at the same time maximising the reductions to all other inputs $\sum_{i\in\overline{I}_0} d_i^-$ and the increases to all other outputs $\sum_{r\in\overline{R}_0} d_r^+$.

If the DMU is relatively efficient then we have $z_r^* = p_i^* = 1 \ \forall r\in R_0, i\in I_0$ and $d_r^{+*} = d_i^{-*} = 0$ $\forall r\in\overline{R}_0, i\in\overline{I}_0$. For a relatively inefficient DMU, we have a set of target inputs $\hat{x}_{ij_0}$ and a set of target outputs $\hat{y}_{rj_0}$ of the form

$$\begin{aligned}
\hat{x}_{ij_0} &= p_i^* x_{ij_0} && \forall i\in I_0 \\
\hat{x}_{ij_0} &= x_{ij_0} - d_i^{-*} && \forall i\in\overline{I}_0 \\
\hat{y}_{rj_0} &= z_r y_{rj_0} && \forall r\in R_0 \\
\hat{y}_{rj_0} &= y_{rj_0} + d_r^{+*} && \forall r\in\overline{R}_0
\end{aligned}$$

Note that the objective function does not yield a straightforward measure of efficiency comparable to the θ of the CCR model.

3.6 Requirements for successful analysis

A useful early guide to the practicalities of conducting a DEA study was provided by Golany and Roll (1989) and their treatment of the two main considerations – selection of DMUs and selection of input and output variables – remains very pertinent.

3.6.1 The population of DMUs

We have already noted that DEA is intended for the evaluation of *comparable* DMUs, and Golany and Roll specify in more detail what is meant by this. They define a homogeneous group of units as one where the DMUs perform the same tasks, have similar objectives, operate under the same market conditions and can be characterised by the same set of inputs and outputs (1989: 239). This is the ideal population of DMUs, but in practice it may be difficult to obtain. Haas and Murphy (2003) made an unsuccessful attempt to develop an adjustment mechanism which would compensate for heterogeneity across DMUs, but more recently another method has been proposed (Samoilenko and Osei-Bryson, 2008) which employs cluster analysis in tandem with DEA.

Another issue is the number of DMUs to be included in the study. If there are n inputs and m outputs in a DEA model then the efficient frontier has $n + m$ facets and this is the number of DMUs which can be found efficient if the weights are unrestricted. In general, then, the rule of thumb is that the population of DMUs should number at least $2(n + m)$ (Golany and Roll, 1989: 239). This requirement may, however, conflict with the requirement for homogeneity since the larger the population the more diverse the units are likely to be. In a study of US Air Force maintenance units, Charnes et al. (1985) increased the population of DMUs by breaking up five months' data for each unit into three overlapping three-month 'windows'. Each three-month segment was then treated in the analysis as an individual DMU, creating a population of 42 DMUs from an original set of 14.

3.6.2 The selection of variables

The choice of input and output variables for a successful DEA model is not as straightforward as one might imagine. There are several potential pitfalls, and an important subset of the DEA literature consists of papers which offer guidance to the would-be modeller on avoiding them.

A common stumbling block is the use of ratio data in variables. It is very common to find indicators of performance expressed as ratios or percentages and it is natural to think of incorporating them into a model of performance, such as DEA. However, as Emrouznejad and Amin (2009) point out, the presence of ratio data violates the convexity assumption on which DEA depends. Different authors propose different solutions to the problem. Emrouznejad and Amin (2009) offer an alternative non-linear model formulation, while Hollingsworth and Smith (2003) suggests using the BCC formulation in preference to CCR, pointing out that it does not function as a VRS model in the presence of ratio data. Dyson et al. (2001) note that using ratio data alone is not a problem, as long as all the variables are scaled in the same way, but warn against mixing ratio and volume measures. Their advice is either to use a proxy volume measure in place of the desired ratio variable or to scale the ratio, for example by using a key input measure as an indicator of size.

The term 'undesirable variable' is often used to designate another problem. If the production process produces an unwanted output, typically as a by-product of the desired output, then how should it be measured? The simplest method of handling this situation is to treat a negative output as an input and a negative input as an output (Dyson et al., 2001). In discussing the inclusion of environmental factors, Boussofiane et al. (1991) make a similar point: "The environmental factor which adds resource may be included as an input whereas one that requires resource to overcome a poor environment may be included as an output" (p 3). Other approaches, such as transforming the data through inversion or subtraction, are also feasible in principle but may lead to other difficulties (Dyson et al., 2001).

As well as choosing which variables to include, the DEA practitioner also needs to handle different types of variable in different ways. Two important considerations were first dealt with by Banker and Morey: identifying controllable versus non-controllable variables (1986a) and dealing with categorical variables (1986b).

One of the peculiarities of DEA is that it is possible for a unit to look efficient simply by having a different technical mix of inputs and outputs: this may be genuine, but poor performance on specific factors can also be hidden by a zero weighting which puts the DMU on a different part of the frontier. Weight-restriction techniques, as described in section 3.5.4 above, can be used to mitigate this effect.

3.7 Limitations and criticisms

3.7.1 Validation

Because DEA is a non-parametric approach there is often concern about model validation: how do you go about testing something when you have no template to test it against? Researchers have taken a number of different avenues of attack on this problem.

Rajiv Banker (1993) has developed a method which uses the efficiency estimates (which are calculated from the observed data but are not themselves observed) as stochastic variables and has shown that they can be treated as independent of the inputs. This approach permits the use of hypothesis testing and subsequently a number of testing strategies have been developed (Banker and Natarajan, 2004).

Simar and Wilson (1998, 2004) propose a bootstrapping procedure which repeatedly generates different sub-samples of the observed data. They suggest that this be used to test for significant differences between the efficiency scores of different DMUs.

Pastor, Ruiz and Sirvent (2002) have designed a test to assess the contribution of each potential variable to a DEA model. By evaluating each candidate variable in turn, the test allows the user to determine whether or not the inclusion of this variable in the model makes a significant difference to the calculated efficiency measures.

Some DEA users explicitly apply these more formal techniques to their models (e.g. Johnes, 2006), but many do not. Dyson and Shale recently reviewed some of the better known approaches to model validation in DEA (Dyson and Shale, 2010) along with some of their own contributions to the DEA literature and noted many sources of uncertainty which they had not addressed at the time. They remark that “In these and other applications, DEA has been used deterministically and any uncertainty in the situation has been handled only implicitly or by sensitivity analysis” (2010: 25).

3.7.2 Conservatism

Another consequence of a DEA model’s being based only on observed data is that it can appear too conservative in its evaluation of efficiency. The efficient projection of an inefficient DMU is no more efficient than the best current performance. This may be seen as a positive characteristic or a negative one, depending on the perspective of the modeller. For the economist Peter Bogetoft it is a limitation, which we will discuss in section 4.7.2 below.

3.7.3 Visualisation

The final issue we will consider here is the problem of visualisation. One of the attractions of the DEA model is the idea of visualising the efficient frontier as a surface wrapped around a population of DMUs. However, the multi-dimensional nature of the model means that in practice this is seldom achieved. If there are more than three factors involved (either two inputs and one output or vice versa) then the frontier cannot be realised on a two-dimensional graph. The first to tackle this issue was Farrell himself (Farrell, 1957), who proposed plotting contour lines of the frontier for the four-factor case. Since then other approaches have been suggested (Belton and Vickers, 1993, El-Mahgary and Lahdelma, 1995) and most recently Førsund et al (2009) have gone back to Farrell's original idea and proposed several ways of cutting through and plotting the frontier. Nonetheless, it remains a source of frustration that a simple visualisation is so difficult to achieve as this is an important means of communicating one's findings (Williams, 2008).

3.8 Recent developments

A recent paper by Liu et al. (2013) provides an interesting perspective on developments in DEA. Through an analysis of citations in the DEA literature they identify a 'main path' for the development of core ideas in DEA and four major areas of current theoretical work. These areas are:

3.8.1 Two stage contextual factor evaluation framework

The two stages referred to here are (1) a DEA study which provides efficiency scores for a set of DMUs and (2) a statistical process which correlates the DEA scores with selected contextual factors. While many DEA studies adopt this approach, Liu et al note that theoretical justification is limited. Work which aims to develop such a justification is seen to have its origins in the bootstrapping approaches developed by Simar and Wilson and mentioned in section 3.7.1 above.

3.8.2 Extending models

The group of papers identified under this heading are concerned with extending existing DEA models in order, for example, to increase their flexibility or their handling of ambiguous data. One example given is that of Cook and Zhu (2008), who extend the functionality of the assurance region model described above in section 3.5.4.. They a propose a methodology which can accommodate different sets of AR restrictions for different groups of DMUs, thus allowing a wider range of operating contexts to be incorporated into a single model.

3.8.3 Handling special types of data

The original DEA models are based on the assumption that all data are present, are numeric and are positive. We have already noted the remarks of Dyson and Shale on this subject (section 3.7.1), and commented on some of the practical requirements for successful DEA variables (section 3.6.2). The work discussed by Liu et al under this heading includes developments in handling interval data and other types of imprecise data as well as a model by Portela et al (2004) which permits the inclusion of negative data.

3.8.4 Examining the internal structure of DMUs

Under this heading Liu et al group a number of different concepts, the simplest of which is the two-stage DEA model. Unlike the two stages in section 3.8.1, of which only one involves DEA, the two-stage DEA model describes a situation where the outputs from one process become the inputs to a second process. A more complex situation may be modelled using network DEA (e.g. Chen et al. (2010)) or dynamic DEA (e.g. Chen (2009)).

3.9 Applications

A search of the *ISI Web of Knowledge* database for “Data Envelopment Analysis” retrieves over 3,500 references so it is clear that DEA has been widely used to date. Given its origins in the analysis of public sector efficiency, it is not surprising to find that it has been applied extensively in this area, notably to education (e.g. Portela and Thanassoulis, 2001, Waldo, 2007) and to health care (e.g. Athanassopoulos and Gounaris, 2001, Amado and Dyson, 2009). It has also been used to model the private sector, although less extensively. The exception to this general rule is retail banking, where there is a substantial body of work (e.g. Cook et al., 2004, Pastor et al., 2006).

There are a variety of motivations for the application of DEA. They include

- Separating out the efficiency of a programme from the efficiency of the people who manage it (Charnes et al., 1981). This idea has been extended to the idea of separating out the efficiency of an educational institution from the efficiency of the individual students (Portela and Thanassoulis, 2001, Johnes, 2003b).

- Identifying benchmark organisations and setting appropriate targets (Golany, 1988, Post and Spronk, 1999). This is linked with the development of selective weighting of inputs and outputs (Thanassoulis and Dyson, 1992) to allow, for example, some outputs to be prioritised over others depending on the circumstances of the DMU.
- Planning future provision, for example by identifying the most productive scale size for a particular service (Banker, 1984, Camanho and Dyson, 1999).
- Assessing the impact of policy change on efficiency (Herrero and Algarrada, 2009).

The examples given above are, in the main, examples from OR journals where the authors have an interest in extending the DEA theory and methods to accomplish a greater range of objectives. However, it is clear from the number of papers on DEA that it has also been widely picked up by specialists in other disciplines for application to their topic of interest. It is necessary to be alert to this distinction as there are a number of issues involved in setting up a DEA study (see section 3.6) and not all researchers appear to be aware of these, as we will see below.

3.9.1 Applications to Higher Education

The earliest studies

Three DEA studies of Higher Education were completed in the 1980s. Two of these involved Charnes and Cooper, originators of the DEA technique (Charnes et al., 1978). Ahn et al (1988) is a straightforward evaluation of doctoral-granting HEIs in the United States using the CCR model with some consideration given to the concept of most productive scale size (MPSS). A second study (Ahn et al., 1989) is focused on HEIs in Texas, and the results of the analysis are compared with a separate evaluation of the same institutions which had been carried out by the state's Select Committee on Higher Education (SCOHE). The SCOHE evaluation had been conducted with a view to identifying weaker institutions for closure or merger. Ahn et al again use the CCR model and extend it by using a window analysis to determine an efficiency measure which is representative of overall performance in a five-year period.

The third study, by Tomkins and Green (1988), is the first to apply DEA to UK Higher Education. They use the CCR model to evaluate departments of accounting in 20 UK universities, using data for 1982–85 from the British Accounting Research Register. They tested six models using different combinations of factors “in order to try to establish whether there was any consistency between different DEA results depending on the inputs and outputs used” (p 155). Overall they found some stability in their results and conclude that DEA is a useful tool for performance measurement in this context, although they note some qualms about the quality of the data available at this time (p 161).

These three studies are experimental in nature. While they succeed in applying the principles of DEA to a new domain, they do not attempt to extend its theoretical basis.

Economics studies

In the 1990s DEA was taken up by a number of different researchers interested in Higher Education. Within the realm of economics, this research is typically aiming to say something about the relative efficiency of different types of HEI or to compare various approaches to modelling efficiency. Studies are conducted using secondary data and are not particularly aiming to address issues of concern to HEIs themselves or to other stakeholders. In fact it is a feature of this literature as a whole that almost all the research is done without the active participation of universities or colleges. In the UK, at least, it is relatively easy to access aggregated administrative data about HEIs for research purposes through the Higher Education Statistics Agency (HESA) so the cooperation of HEIs is not a requirement for this kind of research.

Notable contributers to this literature are the economists Jill and Geraint Johnes of Lancaster University. Between 1993 and 2006, Johnes and Johnes produced several papers applying DEA to the HE context and comparing it with other techniques, such as multilevel modelling. As economists Johnes and Johnes are primarily concerned with the efficiency of the measured unit, whether that is a department of economics (Johnes and Johnes, 1993), a university (Johnes, 2006) or an individual student (Johnes, 2003a). They are less concerned with extending DEA theory, although they make use of a number of developments. In her working paper of 2003, Jill Johnes adopts the decomposition technique proposed by Portela and Thanassoulis (2001), and in her 2006 publication she uses the test developed by Pastor et al (2002) for nested DEA models.

The data used in these studies is mainly taken from datasets collected by HESA. HESA was a new organisation in the 1990s and the quality of the data still left something to be desired. Since many of the Johnes’ studies use the 1993 dataset, it is necessary to treat the conclusions with caution, particularly when students' personal information is used. In a working paper comparing DEA with multilevel modelling (Johnes, 2003b) several input variables such as marital status and type of residence are tested for their significance in explaining performance variation. However, it is important to bear in mind that this kind of information is not crucial to the adminstration processes in a university and for this reason is one of the weakest parts of the HESA dataset, although the quality has improved greatly in recent years.

The economics literature includes many other DEA studies of HE from around the world. Agasisti and colleagues have studied Italian, English and Spanish universities (Agasisti and Johnes, 2009, Agasisti and Pérez-Esparrells, 2010, Agasisti and Salerno, 2007), raising some interesting issues of the comparability of different systems of HE. For instance, in their comparison of English and Italian universities, Agasisti and Johnes are unable to treat undergraduate and postgraduate students as separate inputs. This practice is the norm in the UK-based studies reviewed here, but the curriculum structure in Italian universities makes the distinction impossible.

There have been several studies of Australian HE. The sector has undergone a number of shifts since the 1980s: in 1987 the binary divide between universities and colleges of advanced education was abolished, and since the 1990s Federal governement support for universities has declined while the number of student enrolments has increased. Madden et al (1997) investigated the effect of the 1987 restructuring, although their study is limited to Departments of Economics. More recent studies by Avkiran (2001) and by Abbott and Doucouliagos (2003) both use data from 1995 to evaluate the efficiency of universities, while Carrington et al (2005) examines performance over a five year period from 1996 to 2000 and concludes that "annual productivity growth for universities appears comparable or better than most sectors of the economy" (p 161). All these studies emply the CCR and BCC models, although interestingly Abbott and Doucouliagos choose an input-orientation. This is in contrast to the UK studies described above which use an output-oriented model on the grounds that an HEI typically has limited control over its inputs.

Other countries where the HE sector has been evaluated from an economics perspective using DEA include the USA (Haksever and Muragishi, 1998, Tauer et al., 2007), Canada (McMillan and Chan, 2006), the Netherlands (Salerno, 2006), Slovenia (Tajnikar and Debevec, 2008) and Turkey (Cokgezen, 2009). Access to data is not necessarily as straightforward elsewhere as it is in the UK. For instance, Johnes and Yu (2006) carried out a study of universities in China but the lack of available data meant that the population of institutions studied was only a small proportion of the sector: 115 out of over 1500.

Cross-country efficiency studies have been undertaken as well. The first of these was conducted by Joumady and Ris (2005), who based their research on the assumption that university is preparing students for the workplace. This study is unusual in that primary data was collected via a postal survey of graduates. The number of respondents is not given, but a total of 209 HEIs across eight European countries were evaluated for their efficiency in providing “competencies that match needs on the labor market” (p 195). In light of the debate currently taking place in the UK HE sector, it is interesting to note that Joumady and Ris found the UK universities to be very good at educating their students but less good at preparing them for work.

A comparative study by Kocher et al. (2006) evaluates research output in economics across 21 OECD countries. They measure output by counting articles published in ten high-ranking economics journals over nine years of an 18-year period, while the inputs are given by overall R&D expenditure, number of universities with economics departments and total population. Given the bias towards US-based publications which is built in to the ranking of journals, it is not surprising that the study finds the US to be the most efficient producer of economics research. One other attempt to use OECD data for a DEA study of HE has been made by Agasisti (2009), but in this case only EU countries are considered, the focus is on teaching ouputs rather than research and the candidate countries are evaluated in four separate groups based on the level of public sector support for HE. This approach seems more likely to yield a reliable comparison, but there is a significant flaw in the methodology. All the variables (input and output) are expressed as rates, such as “the percentage of the population that has attained tertiary education by [a certain age]”, which violates the assumption of convexity as we have seen above in section 3.6.2. Clearly, designing a robust cross-country study is a challenging task and more work remains to be done in this area.

There is one study in the economics literature which specifically considers universities from the persepctive of league tables. Breu and Raab (1994) used DEA to evaluate the top 25 universities from the *US News & World Report* rankings of 1992. The authors use the data published in *US News* and interpret each factor as either an input or an output, adding an extra input to account for tuition charges. Having run the DEA model they then give some consideration to the courses of action open to a university which is not on the frontier. Although these are described as “policy options” they are substantially restatements of the DEA results, such as the choice between a radial increase of outputs or a radial decrease of inputs. The authors also note a negative correlation between the league table and the results of their DEA, but are unsurprised “since [HEIs] with low [inputs] are measured as more efficient; but, from the *US News* perspective on quality, these changes imply schools of lower quality” (Breu and Raab, 1994: 42). It is worth noting that Johnes and Johnes (1993, 1995) also take the step of comparing their results with an external rating system, the Universities Funding Council’s research selectivity exercise of 1989, and find a good match between the most highly rated departments and efficient departments.

OR studies

There is no hard and fast division between the work of economists and the work of operational researchers in this area, and indeed some DEA studies by economists are published in OR journals and vice versa. However, making this distinction gives us a useful way to characterise two different approaches which we can observe in the literature. Although operational researchers are also applying DEA to the study of Higher Education, the resulting publications show that they are doing so with a different emphasis from that of the economists. OR-based work is less motivated by the desire to measure efficiency and more by the opportunity to test the boundaries of DEA and develop novel techniques and applications. (It should be noted that when operational researchers are developing new DEA techniques, they often use illlustrative examples based more or less loosely on data which describes their own department or institution: these are not the studies we are considering here. We confine ourselves to that work which is a genuine application of DEA to the study of Higher Education.)

Beasley’s 1990 paper is an early application of DEA to UK universities, in which he compares the performance of Physics and Chemistry departments at 62 universities using data from 1986-87. Beasley uses the CCR model for his study but develops it further by introducing restrictions on the values of input and output weights as presented in the theoretical paper by Wong and Beasley (1990) which has been discussed above. This process allows the author to incorporate some level of managerial judgment into the model based on “what is expected of university departments in a particular discipline” (Beasley, 1990, p 176).

The constraints employed in this model are of two kinds. The first kind specifies the relative importance of one variable compared with another so that, for instance, Beasley proposes that “the weight attached to a postgraduate doing research should be greater than (or equal to) the weight attached to a postgraduate on a taught course” (p 176). This gives rise to a constraint in the form

$$u_2 \geq \alpha u_1$$

for some multiplier α. This example links two output variables from one DMU. However, the form is also used to link the same output variable across several DMUs, so that if one department has a higher research rating than another, the weight attached to the output “research income” will be greater for the first department than for the second.

The other kind of constraint restricts the proportion of total output (or input) which one or more variables may supply. For example, Beasley suggests that teaching outputs should be constrained to a proportion of total output which is within a certain margin of the proportion of the total funding settlement which is intended to pay for teaching. This gives rise to a constraint in the form

$$\beta_1 \leq \frac{\sum_{p \in P} u_p y_p}{\sum_{r \in R} u_r y_r} \leq \beta_2$$

where *R* is the set of all outputs *r*, $P \subset R$ is the subset of outputs we wish to constrain and β_1, β_2 are the selected lower and upper bounds.

Wong and Beasley were the first to propose a method for restricting weight flexibility in DEA using virtual inputs and outputs constructed in this manner, and the application here appears to yield useful insights, such as the finding that small departments can operate as efficiently as larger ones under these constraints. However, the selection of parameters is somewhat arbitrary, and the constraint which imposes an ordering on DMUs is problematic. The intention is to introduce a measure of quality, but this would perhaps be better expressed by the introduction of an additional output variable, or by constructing a weighted variable before DEA is applied.

In a later piece of work, Beasley revisits the same variables for teaching and research, but this time develops a model to determine a separate measure of efficiency for each sphere of activity. This is done through a non-linear model which simultaneously determines the two efficiencies, *t* (for teaching efficiency) and *r* (for research efficiency), and the proportion of general expenditure devoted to each. The objective of the model is to maximise the expression

$$\lambda t + (1 - \lambda) r$$

where λ is the proportion of input resource given to teaching. This is a potentially rich approach, since it permits different groupings of variables according to one's interest. However, the development of non-linear models has remained a minority pursuit in DEA, where there is much that can be accomplished using linear programming models.

Another novel approach is taken by Sarrico et al. (1997) in their study of performance measurement and university selection. Their research is motivated by an awareness that different stakeholders will have different perspectives on what constitutes good performance and they consider in detail the viewpoint of a prospective student making an application to university. DEA is applied to data from the *Times Good University Guide 1996*, varying the subset of variables selected and employing different weight restrictions to suit certain hypothetical types of candidate whose interests are also assumed to differ. So, for instance, an overseas applicant with a strong academic record might prioritise one set of criteria, such as accommodation and teaching quality, while a mature UK applicant who has been out of education for many years might be more concerned about other factors, such as graduate employment prospects.

DEA is used in this study to create a personalised league table for each applicant based on preferences like these which are incorporated into the constraints. In order to achieve this, most of the *Times GUG* variables are treated as outputs even when the aspect they measure, such as library spending, would normally be considered an input to the HE process. This is typical of league table construction, as we saw in the previous chapter, and here the usual principles of DEA modelling are set aside in order to mimic the league table structure more closely. Interestingly, the authors note that some variables used in the *Times GUG*, such as the proportion of firsts awarded, were left out of the analysis altogether: potential applicants taking part in a case study expressed no interest in them.

In a second paper, Sarrico and Dyson (2000) explore the persepctive of the institution. Again a set of variables measuring different aspects of performance are grouped into subsets which represent different viewpoints. Only two such viewpoints are fully developed into DEA models; these are the "image" presented by an HEI to a prospective student and the "reality" which is delivered once they arrive. The implementation of this model consists in (1) evaluating each department against its peers in other HEIs and (2) comparing its performance from the two viewpoints with other departments in the same institution. Departments which have low scores from either viewpoint (or both) can be identified and marked for intervention. Another pair of viewpoints is suggested – that of "accountability" versus "autonomy" – but remains undeveloped due to the difficulties of obtaining suitable data at departmental level across comparable institutions.

It is clearly much easier for researchers to gain access to the now extensive datasets on student-related outputs, and to some extent research outputs, than it is to obtain detailed funding information, for example. Most OR research in this area, like the work in economics described above, is reliant on secondary data sources. The 1997 paper by Sarrico et al is a notable exception, since the research included a case study of school pupils in Years 12 and 13 (lower and upper sixth). The process of tailoring a DEA model to a candidate's particular concerns was piloted by asking pupils about the criteria they used in decision-making about university applications.

Another OR study published at this time was conducted by Colbert et al. (2000) and uses DEA to develop a ranking of MBA programmes in the USA. The authors consider 24 programmes which are highly ranked in the *Business Week Guide to the Best Business Schools* and use the magazine's data as their variables. The data includes surveys of student and recruiter satisfaction and the authors are able to develop three different models which focus in turn on student-related outputs, recruiter-related outputs and all outputs. The model used is the BCC model described in section 3.5.3 above, so this is not a study which offers any real theoretical development. The findings show that the ranking generated varies considerably depending on the selection of input and output variables, so although the authors conclude that "new rankings based on DEA will result in more complete, accurate representation of MBA programs" it is not clear how this accuracy is to be achieved or recognised.

Finally in this section we consider a pice of work by Bougnol and Dulá (2006) which explores whether DEA can be used as a classification and ranking tool for US universities. The authors start by looking at an annual report, *Top American Research Universities*, produced by The Center for Measuring University Performance (The Center) at the University of Florida. The Center's approach is to consider a number of attributes relating to research activity, rank the universities on each of these, and then count how many times a university appears in the top 25. The universities are thus classified into groups based on this count, rather than strictly ranked. Bougnol and Dulá adapt this approach to DEA by creating a vector of 1s and 0s for each university based on the same logic and using this vector as the set of output values. A nominal input variable is used. A DEA evaluation (using an additive model with variable returns to scale) serves to identify those universities which are efficient. The efficient universities are removed from the set and the evaluation is performed again. This step is repeated until all universities have been classified into tiers. The authors find that the results of their process produces results which closely match the original ranking and are satisfied that they have thereby demonstrated DEA's suitability for use as a ranking tool. However, it is not clear that there are any particular benefits from performing the analysis in this manner rather than as originally conceived by The Center. In setting up the vector, for instance, no consideration is given to whether a variable such as "Endowment Assets" is in fact an output or an input, so this distinctive feature of DEA is not being exploited to make any improvements in the ranking process. Nonetheless, this is a typical OR approach to DEA, which seeks to extend the boundaries of the theory through application to a particular problem.

League table studies

In the discussion above we have considered a number of studies which specifically address questions relating to league tables and ranking systems. We have looked in detail at the work of Sarrico et al. (1997) and Bougnol and Dulá (2006) and their imaginative approaches to the use of DEA in this context. We have also discussed the more conventional DEA modelling of Breu and Raab (1994) and Colbert et al. (2000).

There has been one other study which uses DEA to examine university league table rankings; it appears in the *Oxford Review of Education* so is aimed at a readership more interested in the educational isssues than the technical processes, which is perhaps why it includes very little reference to the DEA literature. Turner (2005) is critical of league table design and proposes DEA as a method for overcoming league table deficiencies. Unfortunately, his implementation of DEA leaves a lot to be desired. For his variables he takes the point scores awarded by the *Sunday Times University League Table* rather than going back to the underlying data so that any sense of scale is lost, and he struggles to incorporate the drop-out rate because the league table scoring uses negative numbers. This paper is a valiant attempt to take DEA to an HE audience, but it provides a poor model for them to follow.

3.10 Implications for this research

The purpose of this research is somewhat different from the majority of DEA studies of HE, since we have no interest in creating a new performance measure based on efficiency but are committed instead to the exploration of the space in which institutions operate. One of the most important features which DEA offers is the ability to construct a production possibility set (PPS) based on observed data in an environment where there are multiple inputs and multiple outputs. However, we know from Thanassoulis et al (2004) that DEA without additional constraints can provide an unrealistic evaluation of a DMU's activities, so we need to consider some of the options available for incorporating value judgments into the definition of the production possibility set.

The generalised Assurance Region Global (ARG) model (Cooper et al., 2006) was initially identified as a potential approach. This model is described in section 3.5.4 above: its key feature is that model constraints are placed on the weighted factors so that a particular weighted output, say, is restricted to a bounded proportion of the total weighted output. For example, undergraduate teaching might be considered to be an essential component of university activity which should not be allowed to carry a zero weight. The weighted teaching load could therefore be constrained to between, say, 30% and 90% of a university's output. The bounds themselves would need to be carefully elicited, but they have an intuitive meaning which is easier to grasp than some of the other weight-restriction approaches.

However, a preferred approach was found in the trade-off model (Podinovski, 2004), also described in section 3.5.4. As we note in that section, the trade-off model, like the ARG model, appears to make sense managerially. Podinovski argues (2004: 1316) that using trade-offs rather than weight restrictions keeps the production possibility set realistic: in other words, the trade-offs are about what is technically feasible rather than what is desirable. This offers the modeller the potential for using a variety of methods to identify trade-off values – such as statistical relationships between variables or the relative costs of different activities – in addition to expert judgment. This advantage arises because the method is developed from the perspective of the envelopment model, with its emphasis on a DMU's peers, rather than from the perspective of the multiplier model, which is focused on the DMU's own ideal weights. As our intention with regard to this model is to describe the space in which universities operate, rather than to generate measures of efficiency, this perspective is helpful and appropriate.

Having found a model which satisfactorily defines the boundaries of the PPS, the next issue concerns how the individual decision-making unit (DMU) manages its place within the PPS. The typical DEA model is focussed on evaluating the overall efficiency of a DMU and does this either by making a radial projection of the DMU onto the frontier of the PPS or by some other method that is intended to incorporate all the factors. Fortunately, developments in DEA focussed on target-setting have produced much more flexible models. These "measure-specific" models (Cook and Zhu, 2005) allow the manager of a DMU to choose a subset of factors to concentrate on rather than assuming that they all have equal priority. Examples of these types of model are discussed in section 3.5.6. The more flexible model is the weights-based general preference model (Thanassoulis and Dyson, 1992), which allows the user to specify differentiated preferences over a mixed selection of inputs and outputs.

The question therefore arises: can we combine two kinds of DEA model to create a new one which will permit selective target-setting in an environment which also accommodates value judgments? In particular, can this be achieved by combining the trade-off model (Podinovski, 2004) and the weights-based general preference model (Thanassoulis and Dyson, 1992)? This question is answered in the affirmative in Chapter 5, where we develop a new combined model.

3.11 Summary

In this chapter we have traced the origins and development of Data Envelopment Analysis and identified some of the key characteristics which make it suitable for the study of Higher Education. We have surveyed previous studies of the sector which use DEA, considering in detail the purpose of these studies and the methods which have been developed to achieve them. Finally, we have taken stock of the approaches which are relevant to this research and identified two DEA models of interest. We will return to these in Chapter 5, after consideration of the game theory literature.

4 Game Theory

The purpose of this chapter is to introduce some basic principles of game theory. Game theory is a mathematical approach to interactive decision-making and has been used to study a wide range of decision problems from politics to evolutionary biology. As set out in chapter 1, it is this property of interaction which leads us to draw on the principles of game theory for this research.

First the concepts of game theory are introduced and some examples of simple games are given. We then step back to consider some of the criticisms levelled at game theory and what these mean for management science research. Finally we investigate the game theory literature to see (1) how it has been applied to the study of higher education and (2) how it has been combined with DEA.

4.1 The origins of game theory

Although some aspects of the mathematics had already been noted in earlier times[3], the canonical date for the origin of contemporary game theory is 1944, the year when Von Neumann and Morgenstern published their book "Theory of Games and Economic Behavior". Von Neumann had been working on the subject for some time, having proved the minimax theorem as early as 1928 and provided the mathematical and theoretical skills for this project, but Morgenstern is credited with providing the focus on economics that brought the work together (Leonard, 1995).

After a few years, however, the connection between economics and game theory had not been cemented. Economists were developing their own models independently of game theorists (Rasmusen, 2001), while the game theorists turned their attention to modelling political situations including studies of voting, power, bargaining and conflict (Shubik, 1982). As this was the cold war era there was a great deal of interest in the process of international negotiation and the emphasis was very firmly on non-cooperative games, which Harsanyi later defined as games "where commitments have no binding force" (Harsanyi, 1966).

[3] Such as Cournot's model of a duopoly (1838) and Edgeworth's analysis of barter economies (1881) in which he showed that there were many equilibrium solutions.

Game Theory has now been adopted as a modelling tool across the social sciences. It has also contributed significantly to biology, through the study of evolutionary strategies of populations (Maynard Smith, 1982).

4.2 Concepts and definitions

A game is a situation in which interactive decisions are being made by two or more decision-makers. When we say that the decisions are interactive, we mean that each party's decision will affect the outcome of the other's decision and that both parties are aware of this interaction. This is the crucial factor which distinguishes a situation suitable for a game theoretic approach from other kinds of decision problem (Rasmusen, 2001).

4.2.1 The basic components of a game

Players are the individuals (e.g. people, companies, governments) who make decisions: each player's goal is to maximise her utility by her choice of actions. If there is an element of chance in the game, for example, an external event which may or may not happen, then this is represented by an additional player called **Nature**.

Actions are what the players do. An **action set** is the set of an individual player's actions. An **action combination** is an ordered set of *n* actions, one action for each of the *n* players in a game.

A **payoff** refers either to what players get or to what they *expect* to get out of the game – there are different usages of the term. It can be either

- the utility player *i* receives after all players have picked their strategies and the game is played out, or
- the expected utility of *i* as a function of the chosen strategies

4.2.2 Rational choice

It is assumed in game theory that a player acts **rationally** in order to maximise her utility. We will consider the issue of rationality in more detail in section 4.5 below, but for now we will note that no assumption is made about what the decision-maker's preferences *are*, only that they are consistent. In other words, if the player prefers *a* to *b* and also prefers *b* to *c* then it is assumed that she prefers *a* to *c*.

In order to make use of this assumption, it is also necessary to assume that each player knows that the other players in the game are rational: this allows each individual to make deductions about the likely behaviour of others. Reasoning further, we realise that each player needs to know that the others also know that all players are rational and so on. To avoid an ever-lengthening chain of assumptions, we invoke the assumption of **common knowledge of rationality** (Webb, 2007: 67).

4.2.3 Strategies

A key concept in game theory is the idea of strategies. A player's **strategy** is a rule which tells her which action to choose at each point in the game, depending on what information she has about the game at that point.

All the possible strategies for a player make her **strategy set** and it is given by the cross-product $S = A_1 \times A_2 \times ... \times A_n$ where A_i is the action set describing all the options available at decision node *i.* This set may also be called the set of **pure strategies**, a pure strategy being one in which there is no randomisation: it is selected directly from the set. A **reduced strategy set** is the set formed when all pure strategies that lead to indistinguishable outcomes are combined.

A **strategy combination** is an ordered set of *n* strategies, one for each of the *n* players. Crucially, it is the interaction of strategies which determine what happens when the game is played.

When there is only a single decision to be made, the action set and the strategy set are identical. There is also only one way of specifying randomising behaviour, which is to use a vector of probabilities $(p_1, p_2, ..., p_n)$ where $\sum p_i = 1$. This vector describes a **mixed strategy**, which is defined below.

When there is (potentially) more than one decision to be made, the action sets and strategy sets are no longer identical and randomising behaviour can be specified in two different ways (Webb, 2007).

1. A **mixed strategy** specifies the probability with which each of the pure strategies is used, in other words it is a vector of probabilities as above. A mixed strategy is effectively a linear combination of pure strategies. In this case, there is one

moment of randomisation: before the game begins a strategy is chosen (with a certain probability) and then it is followed to the conclusion of the game.

2. A **behavioural strategy** causes randomisation to take place several times as the game is played out. It is a *collection* of probability vectors $\{\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_n\}$ where each $\mathbf{p}_i$ determines the player's behaviour at node *i*.

Although theoretically distinct, in practice "a behavioural strategy and a mixed strategy are equivalent if they assign the same probabilities to each of the possible pure strategies that are available" (Webb, 2007: 29). Because any given behaviour can be represented in either form, Webb points out that we are able to use whichever is more appropriate to the given problem.

Equilibrium strategies are those strategies which players pick to maximise individual payoffs, so an **equilibrium** (written *s**) is a "strategy combination consisting of a best strategy for each of the *n* players in the game" (Rasmusen, 2001: 18). There are different ways to define "best" and these lead to different equilibrium concepts such as the **Nash equilibrium**, which is one of the few generally accepted. We will consider it in detail below.

If we are interested in finding equilibria and not concerned with the exact amount of the payoffs, then Webb (2007) points out that we can simplify a game by applying a generalised affine transformation (a linear transformation followed by a translation) so that the preference order is retained but not the size of the payoff.

4.2.4 Static and dynamic games

A **static game**, also known as a **simultaneous decision game**, is one in which a single decision is made by each player without knowing what decision has been made by the other players. To describe a static game you need to specify the players, a pure strategy set for each player and a complete set of payoffs for each player. Putting this information in a table gives you the **normal form** of the game, also called the **strategic form**.

Player 2's strategies

		S_{21}	S_{22}
Player 1's strategies	S_{11}	$p_1(S_{11}, S_{21})$, $p_2(S_{11}, S_{21})$	$p_1(S_{11}, S_{22})$, $p_2(S_{11}, S_{22})$
	S_{12}	$p_1(S_{12}, S_{21})$, $p_2(S_{12}, S_{21})$	$p_1(S_{12}, S_{22})$, $p_2(S_{12}, S_{22})$

Figure 4-1: A 2-player game in normal form. Each player *i* has two strategies S_{i1} and S_{i2} and four possible payoffs p_i which depend on the strategies selected by both players.

Alternatively, a game can be written as a directed graph, where each node represents a decision-making point and each arc a potential action. This is the **extensive form** of the game, and is more appropriate for a **dynamic game**, i.e. one "in which decisions are made at various times with at least some of the earlier choices being public knowledge when the later decisions are being made" (Webb, 2007: 89).

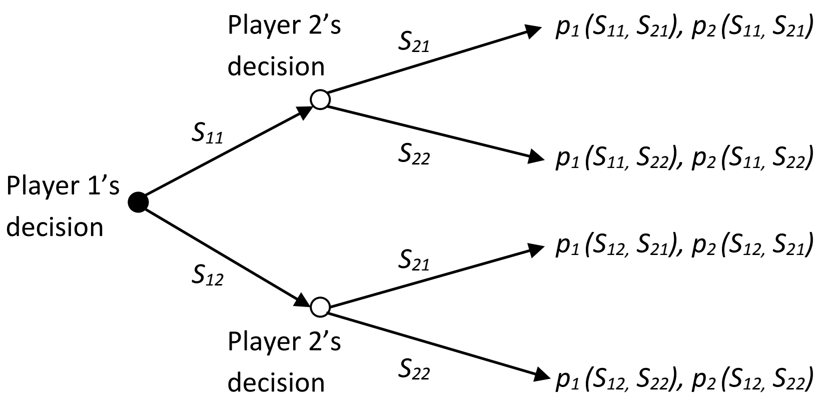


Figure 4-2: A 2-player game in extensive form. The strategies and payoffs are the same as in Figure 4-1, but in this example Player 1 makes her decision first.

4.2.5 Information

An **information set** refers to a player's knowledge of the values of the different variables at a particular point in the game: "at any particular point of the game [it] is the set of different nodes in the game tree that he knows might be the actual node, but between which he cannot distinguish by direct observation" (Rasmusen, 2001: 43).

Rasmusen defines the following types of information:

Perfect information	In a game of **perfect information** each information set is a singleton, i.e. each player knows exactly which decision node she is on at any given time. Otherwise the game is one of **imperfect information**.
Certain information	A game of **certainty** has no moves by Nature after any player moves. Otherwise the game is one of **uncertainty**.
Symmetric information	In a game of **symmetric information**, a player's information set at • any node where he chooses an action, or • an end node contains at least the same elements as the information sets of every other player. Otherwise the game is one of **asymmetric information**.
Complete information	In a game of **incomplete information,** Nature moves first and is unobserved by at least one of the players. Otherwise the game is one of **complete information**.

The distinction between complete and incomplete information was originally made by von Neumann and Morgenstern (1944) as a way of describing games where the players know all the important parameters and games where they do not. However, it was Harsanyi (Harsanyi, 1967-8) who noted that lack of information could occur in many different guises: one player might not know the other's payoffs, or strategies, or they might not know how much information the other player has, and so on. By expressing this uncertainty in different models of the game – one model for each possible situation – and then having an initial move by Nature which selects, with some probability, which game is to be played out, the scenario of incomplete information is transformed into one of complete but imperfect information. This means that, as the game is played out, each player has full information about what can happen, even though she still does not know with certainty what has happened to date. This insight allowed many games to be analysed which were previously intractable.

4.3 Milestones in game theory

Table 4-1 gives some of the key publications which have informed the development of game theory.

Table 4-1: Milestones in game theory, based on Walker (2005)

Name	Year	Title	Comments
von Neumann and Morgenstern	1944	Theory of Games and Economic Behavior	Introduced game theory and the minimax/maximin solution for zero-sum games
Nash	1950	Non-cooperative games	Distinguished between cooperative and non-cooperative games. Introduced the concept of a Pareto-optimal equilibrium for non-zero-sum games
Shapley	1953	A value for *n*-person games	Described a way of dividing the amount gained from a cooperative game between members of a coalition
Kuhn	1953	Extensive Games and the Problem of Information	Established the formulation for extensive form games
Harsanyi	1967-1968	Games with incomplete information played by "Bayesian" players	Showed how to convert a game with incomplete information into one with complete but imperfect information, therefore allowing them to be analysed.
Schmeidler	1969	The Nucleolus of a Characteristic Game	Introduced the nucleolus, showing that it always exists and is unique
Maynard Smith	1972	Game Theory and the Evolution of Fighting	Introduced the idea of an evolutionarily stable strategy
Harsanyi and Selten	1988	A General Theory of Equilibrium Selection in Games	Gave the first general theory for selecting one equilibrium point for any non-cooperative or cooperative game

A smaller development, but one which is potentially relevant to this research, is a piece of work by Inohara et al (1997). The authors consider the situation where a player is involved in more than one game at a time but only has limited resources to invest across his portfolio of games. For example, player *A* has the opportunity to invest in both game G_1

(for a payoff p_1) and also game G_2 (for a payoff p_2), but the total expected rewards $p_1 + p_2$ will depend on how much of his resources he decides to put into each game. If we think of player *A* as a university, and G_1 and G_2 as games which evaluate teaching and research respectively, then we can see how this model could be applied to the problem of competing performance measures in HE. Inohara et al propose a method for formulating this scenario as an integrated game so that it can be analysed using the tools of standard game theory.

4.4 Solving a game

A 2-player game such as the example in Figure 4-1 is said to have a solution if we can identify "a (not necessarily unique) pair of strategies that a rational pair of players might use" (Webb, 2007) (p 65). In this section we will look at some approaches to solving simple games.

4.4.1 Dominance

Consider the following 2 x 2 game (a game with two players, each with two strategies):

		Player 2	
		C	*D*
Player 1	*A*	(3, 1)	(4, 3)
	B	(1, 2)	(2, 4)

Figure 4-3: A 2 x 2 game with strictly dominated strategies

If Player 1 chooses strategy *A* then his payoff will be either 3 or 4, depending on the strategy chosen by Player 2. If Player 1 chooses strategy *B* then his payoff will be either 1 or 2. Strategy *A* clearly outperforms strategy *B* in either case, so we say that *B* is **strictly dominated** by *A*.

In general, strategy s_i is strictly dominated by s_i' for player *i* if $p_i\left(s_i', s_j\right) > p_i\left(s_i, s_j\right)\ \forall s_j \in S_j$ where S_j is the strategy set for player *j*. That is, irrespective of what player *j* does, player *i* will always get a better payoff from using strategy s_i' than from using strategy s_i.

In the example given in Figure 4-3 player 2 also has a dominant strategy: pursuing strategy *C* will give her a payoff of 1 or 2, while strategy *D* will lead to a payoff of 3 or 4. Based on our assumption that both players will rationally seek the highest payoff, we can eliminate strategies *B* and *C* from consideration and deduce that player 1 will adopt strategy *A* and player 2 strategy *D*. This is illustrated in Figure 4-4 below.

		Player 2	
		~~*C*~~	***D***
Player 1	***A***	~~(3, 1)~~	**(4, 3)**
	~~*B*~~	~~(1, 2)~~	~~(2, 4)~~

Figure 4-4: A 2 x 2 game with dominated strategies eliminated

If $p_i\left(s_i', s_j\right) \geq p_i\left(s_i, s_j\right)\ \forall s_j \in S_j$ then we say that s_i is **weakly dominated** by s_i' for player *i*. In this case, irrespective of what player *j* does, player *i* will get a payoff from strategy s_i' that is at least as good as his payoff from strategy s_i although it may not necessarily be better. Figure 4-5 gives an example of a 2 x 2 game with weakly dominated strategies.

		Player 2	
		C	*D*
Player 1	*A*	(3, 1)	(2, 3)
	B	(1, 2)	(2, 2)

Figure 4-5: A 2 x 2 game with weakly dominated strategies

In this simple game it is clear that strategies *A* and *D* are again the dominant strategies. In general, however, it is necessary to exercise caution when eliminating weakly dominated strategies as changing the sequence of elimination may change the solution to the game (Webb, 2007: 67).

Playing dominant strategies in a game does not necessarily deliver an optimal outcome for the players, as we can see in the classic game of Prisoner's Dilemma. This game is described in Poundstone (1995) as follows:

> "Two members of a criminal gang are arrested and imprisoned. Easch prisoner is in solitary confinement with no means of speaking to or exchanging messages with the other. The police admit they don't have enough evidence to convict the pair on the principal charge. They plan to sentence both to a year in prison on a lesser charge. Simultaneously, the police offer each prisoner a Faustian bargain. If he testifies against his partner, he will go free while the partner will get three years in

> prison on the main charge. Oh yes, there is a catch… If *both* prisoners testify against each other, both will be sentenced to two years in jail." (p 118)

The same information can be expressed in normal form as shown in Figure 4-6. For each player individually the dominant strategy is *Testify*, but if both players pursue this strategy then they will end up spending longer in prison than if they had both kept silent.

		Player 2	
		Keep silent	*Testify*
Player 1	*Keep silent*	(-1,-1)	(-3,0)
	Testify	(0,-3)	(-2,-2)

Figure 4-6: Prisoner's Dilemma

4.4.2 Nash Equilibrium

In many games, even those as simple as these 2 x 2 examples, a player will find that she has no a single strategy which dominates all others; rather, her own best choice of strategy depends on what choice the other player makes – and his best choice depends in turn on the option the first player chooses. We are in essence seeking a **best response** strategy for each player. The Nash equilibrium offers a defintion of such a strategy.

Nash (Nash, 1951) defines an equilibrium point for an *n*-person game as a vector of mixed strategies "such that each player's mixed strategy maximises his payoff if the strategies of the others are held fixed" p 287. Thus if σ_i is a mixed strategy for player *i*, the vector $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)$ is an equilibrium point if and only if for every *i*

$$p_i(\sigma^*) = \max_{\text{all } \sigma_i} p_i(\sigma^*; \sigma_i)$$

where $(\sigma^*; \sigma_i)$ denotes the vector σ^* with σ_i^* replaced by σ_i.

The Nash equilibrium can be demonstrated by considering the two-player game of "Boxed Pigs" (Rasmusen, 2001). The players in this game are Big Pig and Small Pig. They are in a box which has a panel at one end and a food dispenser at the other. Each pig has a choice of two strategies: they can press the panel and cause the dispenser to dispense food, or they can wait by the dispenser. Pressing the panel costs 2 energy units and releases 10 energy units of food.

- If Big Pig gets to the dispenser first he will eat 9 units of food leaving 1 unit for Small Pig.
- If Small Pig gets to the dispenser first he will eat 4 units and Big Pig 6 units.
- If they arrive at the same time, Small Pig eats 3 units and Big Pig 7 units.

The payoff table is shown Figure 4-7 below.

		Small Pig	
		Press	Wait
Big Pig	Press	(7 - 2 = **5**, 3 - 2 = **1**)	(6 - 2 = **4**, 4 - 0 = **4**)
	Wait	(9 - 0 = **9**, 1 - 2 = **-1**)	(**0**, **0**)

Figure 4-7: Payoffs in the game of Boxed Pigs

Considering each possibility in turn, we can see that

a) If Big Pig chooses to Press then Small Pig's best response is to Wait: it will give him a payoff of 4 units rather than 1.
b) If Big Pig chooses to Wait, then Small Pig's best response is also to Wait: it will give him 0 rather than -1 units.
c) If Small Pig chooses to Press then Big Pig's best response is to Wait: it will give him a payoff of 9 units rather than 5.
d) If Small Pig chooses to Wait then Big Pig's best response is to Press: it will give him 4 units rather than 0.

Putting (a) and (d) together, we find that if Big Pig chooses to Press then Small Pig's best response is to Wait, and that if Small Pig chooses to Wait then Big Pig's best response is to Press. Neither pig has incentive to deviate from the combination (Press, Wait), making it a Nash equilibrium.

A game may have more than one Nash equilibrium. For example, "coordination games" are a group of games which have multiple Nash equilibria in pure strategies. These are games in which players do better when they all adopt the same strategy than when they differ. In some examples, such as the classic "Stag Hunt" game, one equilibrium Pareto dominates the other (see Figure 4-8), while in other examples there is no single equilibrium which is Pareto efficient (see Figure 4-9).

		Player 2	
		Stag	*Hare*
Player 1	*Stag*	(10, 10)	(0, 5)
	Hare	(5, 0)	(5, 5)

Figure 4-8: Stag Hunt is a 2-player game in which the players can cooperate to hunt a stag or go their separate ways and each catch a hare. One Nash equilibrium (Stag, Stag) Pareto dominates the other (Hare, Hare).

		Player 2	
		Party	*Home*
Player 1	*Party*	(10, 5)	(0, 0)
	Home	(0, 0)	(5, 10)

Figure 4-9: Battle of the Sexes is a 2-player game in which the players have a choice of two activities, e.g. going out to a party or staying at home. Their preferences differ, but they would rather be together doing a less-preferred activity than alone. There are two Nash equilibria, but neither is Pareto efficient.

Nash demonstrates that every finite game has an equilibrium point in mixed strategies. However, a game does not necssarily have a solution in pure strategies (Rasmusen, 2001:70).

4.5 Criticisms of game theory

> **"Despite its intellectual power, Game Theory has been criticized as over-idealized, irrelevant, and even malign."**
>
> (Bennett et al., 1989: 290)

Some criticisms of game theory are justified, but many are the result of misunderstanding, such as this from Hurwitz (quoted by Wildavsky, 1992: 16-17):

> "It is a dubious tribute to U.S. social science," [Hurwitz] writes, "that its textual traditions have so easily transformed this inherently political parable [the Prisoner's Dilemma] into a technical problem."

This is a clear case of inverting the cart and the horse. The Prisoner's Dilemma has its origins in experiments at RAND conducted by Melvin Dresher and Merrill Flood. Rasmusen (1992: 84) explains how the story itself arose:

> "Starting with just the matrix, we are in the position of the RAND game theorists in the early 1950s who were perplexed by a certain 2 x 2 matrix that generated perverse results. Albert Tucker, on being asked to give a talk on game theory to the Stanford psychology department, decided to attach a story to the numbers. The result was the Prisoner's Dilemma and a deeper understanding than the mathematics alone could give."

In other words, it is not a real-world dilemma which has been abstracted, but a technical problem dressed in a parable to flesh it out and give it meaning. The parable succeeds in capturing our attention because we can readily visualise this situation, but Rapoport (1992: 74) sets us straight about the detail:

> "Prisoner's Dilemma is not about prisoners. For this reason, the psychology of prisoners, their social milieu, and their attitudes toward authority or any other circumstances, however relevant they may be to explaining the behavior of prisoners are entirely irrelevant to the theory that has suggested this game or derived its implications."

Rapoport addresses another criticism by making a clear distinction between descriptive (or predictive) theory and normative theory, firmly putting game theory in the latter category: when critics argue that game theory doesn't tell you what real people would do, Rapoport says, they "miss the point because game theory, being a normative theory, makes no predictions" (1992: 79). According to Sloman, "a normative statement is a statement of value: a statement about what ought or ought not to be [.... Such statements] cannot be proved or disproved by a simple appeal to the facts" (2006: 28). In terms of game theory we can read 'normative' as 'prescriptive': the model tells us how the rational player should play the game. Interestingly, though, Thomas (1986: 17) is reluctant to claim that game theory does this:

> "Let us emphasise again that game theory is not a prescriptive way of how to play a game. Rather it is a set of ideas and techniques for analysing these mathematical models of conflict of interest. It doesn't tell you how to play the game, but

> describes properties that certain ways of playing the game have, and which you might think desirable."

That is perhaps the management scientist's approach, in contrast to the economist's or the mathematician's: the management scientist aims to elucidate the problem and demonstrate to the decision-maker what might happen if she decided on a certain course of action. In any case, however, the descriptive work is carried out by the behavioural scientists, who have done rich work in the area of rationality. This has been in large part stimulated by the mismatch between game theoretical – as well as other theoretical – prescriptions and the observed behaviour of real people making decisions.

> "It is widely accepted that not every player behaves rationally in complex situations, so [the assumptions of game theory] are sometimes violated. For explaining consumer choices and other decisions, rationality may still be an adequate approximation … But game theory is different: the players' fates are intertwined. The presence of players who do not think strategically or optimize, even if there are very few such players, can change what rational players should do."

Camerer et al. (2004)

Camerer's stated aim is to use observational studies to develop more precise predictive models. These models need to account for the bounded rationality typically exhibited by human beings when they make decisions. He gives an interesting perspective on the concept of equilibrium, seeing it as "the limiting outcome of an unspecified learning or evolutionary process that unfolds over time" rather than as the immediate solution of a one-off game. In this view, players learn over time which strategies are more successful, and the more sophisticated players will look ahead, anticipating the learning of others (ibid: 153).

For now it will be sufficient to note the point made by Webb, that "rationality should not be equated with dispassionate reasoning (notwithstanding the view held by certain aliens from a popular science fiction series)" (2007: 12). In other words, a rational Captain Kirk is perfectly entitled to prefer *a* over *b* even if Mr Spock can conclusively demonstrate the objective superiority of *b* over *a*: the only requirement is that Kirk should prefer it consistently. This may mean that *a* and *b* need to be very tightly defined, but that is not a problem for game theory. Martin Osborne gives the following definition:

> "Allowing for the possibility that there are several equally attractive best actions, **the theory of rational choice** is
>
> > *the action chosen by a decision-maker is at least as good, according to her preferences, as every other available action.*"
>
> Osborne (2004: 6, author's emphasis)

So far we have seen two misperceptions of game theory; two things that game theory is *not*. It is not a science for predicting how people will behave, nor is it a method solely for the robotically rational. Another misperception comes about because of the name.

> "It is in a way an unfortunate choice of name, because it has the connotations of amusement, light-heartedness, and a recreational contest." (Thomas, 1986: 16)
>
> "To describe something as a 'game', in this context, is not to imply that it is trivial, or that there must be winners and losers. In this respect, the name is unfortunate." (Bennett et al., 1989: 287)

Game theory is certainly not limited to the trivial. Indeed the context in which it first flourished — analysis of the cold war and the arms race between the United States and the Soviet Union — could hardly be less trivial. Alternative names are sometimes employed and may well be more meaningful — interactive decision modelling is one such example — but "game theory" is well established for the time being and it is to be hoped that the plethora of authors who have written introductions to the subject will be heard.

> "Game theory aims to help us understand situations in which decision-makers interact." (Osborne, 2004: 1)
>
> "Game theory is concerned with the actions of decision makers who are conscious that their actions affect each other." (Rasmusen, 2001: 11)
>
> "A game is simply a model of interactive decisions, consisting of 'players', 'strategies', 'outcomes' and 'preferences'." (Bennett et al., 1989: 287)

> "An interactive decision problem involves two or more individuals making a decision in a situation where the payoff to each individual depends (at least in principle) on what every individual decides. Borrowing some terminology from recreational games … all such problems are termed 'games' and the individuals making the decisions are called 'players'." (Webb, 2007: 61)

Taken with the cautiously prescriptive approach described by Thomas, any one of these statements will serve as a good guide to the management scientist about what game theory is for.

4.6 Game Theory and Higher Education

There has been very little research on the application of game theory to higher education. In fact to date we have only uncovered two relevant studies in this category.

The first is Niklasson's evaluation of the HE regulatory framework in Sweden (Niklasson, 1996). The author uses the Prisoner's Dilemma to model the relationship between universities and government and applies four policy recommendations which the government should adopt if it is successfully to play a "tit-for-tat" strategy[4]. Comparing this framework with the government's actual strategy, as seen through policy documents, Niklasson finds that the government has not followed the recommended route for responsive regulation. However, there remains a question over whether it is the initial model which is faulty rather than the choice of strategy, and this is not addressed.

While Niklasson's study uses game theory at the level of the institution, the second paper (Langbein, 2008) uses game theory at the level of the individual lecturer and student. Langbein's hypothesis is that the use of student evaluations of teaching (SETs) in decisions about pay and promotion means "students, administrators and faculty are engaged in an individually rational but arguably socially destructive game" of grade inflation (2008: 419). She examines this situation using a principal-agent model and finds that higher grades do have a positive effect on SETs, and that the incentive to inflate grades is greater in private institutions which are more dependent on income from tuition fees. Because the game offers attractive payoffs to all participants, there seems to be little incentive to break the

[4] Tit-for-tat is a strategy which was devised by Anatole Rapoport and submitted to a game theoretic computer tournament organised by Robert Axelrod in 1980. It consists of the following moves: cooperate on the first move; on the next and subsequent moves, copy what the other player did on the previous move.

cycle even though the longer-term consequences may be the devaluation of grades in the job market.

4.7 Game Theory and DEA

There are several points of correspondence between game theory and DEA, most notably the competitive relationship between DMUs and the use of linear programming to find optimal values. These correspondences have been noted and explored by a number of different authors, although the literature which explicitly links the two areas is not extensive. A brief survey allows us to identify three main strands of thought.

4.7.1 Two-player zero-sum games

The first approach was initiated by Rajiv Banker (1980). He sets up a two-player zero sum game in which player I (the maximising player) is a DMU to be evaluated and player II (the minimising player) is a notional "evaluator". The payoff to player I is the CCR efficiency of the DMU. However, there are some problems with the resulting model as (i) it evaluates only radial efficiency and not slacks and (ii) it cannot handle multiple outputs except by combining them into a single virtual output. These weaknesses were subsequently addressed by Banker et al (1989). However, in order to achieve their revised model for the multiple-output case they had to drop the CCR model and use the BCC model of efficiency instead, a move which has been criticised (Rousseau and Semple, 1995).

Other models developed in this manner – using a two-player zero-sum game – include the work of Rousseau and Semple (1995, 1997) and of Hao et al (2000a, b). Rousseau and Semple return to the CCR model of efficiency but define the players of the game and the payoffs differently from Banker. In particular, their player I is still associated with a DMU but, breaking with game theoretic convention, is a minimising rather than a maximising player. The payoff (to player II) which player I seeks to minimise is the efficiency of an aggregated group of competitor DMUs under player I's selected weights. This payoff function does not include player I's own inputs and outputs which instead take the role of constraints on the weight selection. Further constraints can also be added by Semple's extension to this model (Semple, 1997).

Hao et al also propose a constrained game, but theirs is linked to the generalised DEA model of Yu et al (1996) rather than to a specific model as was done in both cases above. They develop a game in which player I is a maximising evaluator while player II is a minimising DMU. The payoff which they are seeking to influence is the efficiency of a "virtual DMU" against which the real DMU will be compared: the greater this efficiency, the lower the ratio will be.

All the work described thus far is theoretical: applications are few and far between. Rousseau and Semple illustrate their own model by applying it to the dataset of public school programs originally used by Charnes et al (1981) in the development of DEA. It was used again by Ray Chang in his unpublished doctoral thesis on managed healthcare in the USA which appears to be the basis of Brockett et al (2004). In both cases the essence of the application is that each DMU belongs to one of two subgroups, and the researcher is interested in the relative efficiency of these subgroups. The approach they take is to "play" each DMU against an aggregated competitor of DMUs from the *other* group. Brockett et al conclude that one kind of HMO (Independent Practice Associations) is more efficient than another (the more restrictive Group or Staff arrangements), based on a definition of efficiency which considers the number of outpatient visits and days in hospital obtained for the amount of money spent. This is clearly a limited perspective compared with, for example, the recent work of Amado and Dyson (2009) on primary diabetes care in England. Here the authors build a two-stage DEA model which incorporates effectiveness of the service as well as technical efficiency. Nonetheless, Brockett et al provides an interesting illustration of game theoretic DEA in practice: it is an illustration which is using the techniques but which is not drawing on a game structure underlying the model of the situation.

4.7.2 Agency theory

A second strand of work is that of the economist Peter Bogetoft. He uses a principal-agent model to inform his work with DEA. This model posits a principal who pays an agent to perform a specified task. It is assumed that the principal is interested in achieving the best results for the least cost, while the agent is motivated to get the most reward for the least effort. The principal's concern is therefore to identify the optimum incentive for the agent in order to accomplish the task satisfactorily. Bogetoft examines the use of DEA as a performance measure in determining the appropriate level of incentive (1994) and in directing the agent to implement the production plan which the principal prefers (1995). He finds that DEA does not provide a full range of incentives since it does not reward an agent for pushing the efficient frontier beyond the status quo, and suggests additional compensation for “over-efficient” performance. He warns against over-hasty implementation of such a system, however, since “high-powered incentive schemes developed from a misspecified model of the agents’ situation may ... end up being harmful” (Bogetoft, 1995: 77).

4.7.3 Cooperative game theory

A small amount of work has more recently been done which links DEA with cooperative game theory. Nakabayashi and Tone have devised a game called “the Egoist’s dilemma”, which pits several self-confident players against each other in bidding for a share of some resource. They need to reach a consensus, but each is convinced of their own superiority – expressed in DEA terms by the choice of weights they apply to the mix of inputs and outputs. The authors propose a scheme for solving such a game and show how the Shapely value can be applied. Although they suggest a number of possible fields of application, their own interest in developing the work further is clearly theoretical.

4.8 Implications for this research

We have described three main areas of interest – Higher Education, Data Envelopment Analysis and Game Theory – and considered the intersections between them. We can now place our own research problem in its context.

We have established that our primary aim is to move towards a better understanding of the tensions between performance measures in the UK Higher Education sector through a quantitative analysis of the league table space in which universities operate. We will do this by combining two kinds of DEA model to create a new one which will (1) permit selective target-setting in an environment which (2) also accommodates value judgments. In particular, this will be achieved by combining the trade-off model of (Podinovski, 2004) and the weights-based general preference model of (Thanassoulis and Dyson, 1992).

Our final question is therefore: can we use the tools of game theory to gain insight into the tensions we have identified and will quantify through DEA? We know from Rousseau and Semple (1995, 1997) and other examples of game theoretic DEA models that game theory and DEA can be used together to identify optimal strategies for a DMU which is under evaluation. However, with a few highly theoretical exceptions (e.g. Bogetoft (2000)), existing GTDEA models are generally focused on the evaluation of past performance rather than on planning future activity. Rather than pursue the theoretical work of extending our

DEA model to a GTDEA model in order to evaluate alternative strategies for selective target-setting, this research will take a more practical approach. In Chapter 7 we will undertake the evaluation of our DEA results using the basic principles of game theoretic reasoning: assessing the best strategies of other players and determining a best response to those strategies.

4.9 Summary

In this chapter we have presented the basic concepts of game theory and illustrated these with some simple examples. We have considered the relevance of game theory to management science research and reviewed a number of studies which are of particular interest in the context of higher education and the use of DEA. We have concluded that the most appropriate use of game theory in the present research is to draw on the principles of game theoretic reasoning in assessing alternative strategies for action.

5 A framework for investigating Higher Education performance

In the preceding three chapters the problem context (Chapter 2) and two management science modelling techniques, Data Envelopment Analysis (Chapter 3) and game theory (Chapter 4), were introduced. It was established that the aim of this research is to move towards a better understanding of the tensions between performance measures in the UK Higher Education sector and that DEA and game theory are appropriate tools for achieving this.

This chapter will describe the framework delevoped for this investigation and present a new DEA model which sits at the heart of this framework.

5.1 Overall framework

It has been established in previous chapters that the rewards which a university seeks from league table positioning are dependent not only on that institution's own decisions but also on the decisions of others and, therefore, that decision-making in this arena can usefully be studied through the lens of game theory. In this scenario an institution is considered as a player in a game and is expected to choose the strategy which will maximise their payoff. In order to use this approach, it is necessary to establish what strategies are available to the players and this is the purpose served by the DEA model developed below. However, for this research, the DEA model needs to be set within a larger model of the performance measurement process. To avoid confusion this larger model is hereafter referred to as the "framework" and the term "model" is reserved for the DEA component.

The framework sets out a process for using the model and consists of three stages:

1. Steps preceding the DEA model:

   - establishing the cohort of players, i.e. the institutions to be included in the DEA model
   - establishing the relevant variables so that the model is constructed appropriately
   - defining the trade-off space

2. Using the DEA model to generate potential strategies for the players

3. Steps following the DEA model:

   - for an individual player: analysing the impact of different strategies on league table performance
   - for a pair of players: analysing the impact each player's strategies has on the other's league table performance

There are a number of assumptions which are relevant to the implementation of this model and these will be considered in detail in Chapter 6. At this stage it is sufficient to note that the framework proposed above makes one significant assumption: that a strategy, once adopted, can be fully realised. In other words, if – in step 2 of the framework – we discover a potential strategy *s* for player *P*, then – in step 3 – we assume that *s* is fully achievable and evaluate its impact on this basis. This is clearly a simplification of the process of pursuing a given strategy and a more nuanced approach would be desirable. Some options for introducing more subtlety into this aspect of the model are considered in Chapter 8.

5.2 Background to the new DEA model

Having set out the overall framework, we now summarise the approaches which will be incorporated into the new DEA model.

We begin with a dataset suitable for DEA: a set of *n* DMUs, each with *m* inputs and *s* outputs. The input and output vectors for any DMU *j* are written $\left(x_j, y_j\right)$. The *m* x *n* input matrix *X* has columns x_j, and the *s* x *n* output matrix *Y* has columns y_j.

As described in Chapter 3, two DEA models have been selected and are here combined into a single, new model. The approaches chosen are

- the production trade-off model proposed by Podinovski (2004, 2007b), and
- the weights-based general preference structure of Thanassoulis and Dyson (1992)

5.2.1 The production trade-off model

The trade-off approach is concerned with the extent of the production possibility set. A trade-off *t* is represented as a pair of vectors $\left(P_t, Q_t\right)$ where the vector P_t modifies inputs and the vector Q_t modifies outputs according to some relationship which expresses a feasible exchange between them. For instance, consider a production technology with one input and two outputs, which permits a trade-off t_0 represented by the vectors $P_{t_0} = (0)$, $Q_{t_0} = (1, -1)$. These vectors express the judgment that one unit of output 2 can be removed and replaced with one unit of output 1 without any change in the quantity of input. This approach is described in detail in section 3.5.4.

The trade-off approach has been adopted because it offers a means of incorporating judgments about the relative costs of pursuing different types of production activity. These judgments are applied to the whole population of DMUs and need not be exact: very often an asymmetric form may be used which expresses the relative costs in terms of a range rather than a single precise value. This makes it both realistic and accessible, which is an important consideration as DEA can sometimes appear to be a "black box" method of modelling.

5.2.2 The weights-based preference structure

The weights-based preference structure operates at the level of the individual DMU rather than the population. It offers a way for managers to prioritise improvements to a subset of inputs and outputs which they choose themselves. This is done by specifying weights to be attached to the selected inputs and outputs "to reflect the relative degree of desirability of improvement of the corresponding input-output levels" (Thanassoulis and Dyson, 1992: 85). It is unlikely that the first attempt at specifying such weights will yield the preferred results, but the model can be used to test different sets of weights.

The model itself gives priority to maximising improvement in the chosen subset, i.e. to reducing selected inputs and increasing selected outputs. Each weighted factor is assigned a multiplier such that the weighted sum of the multipliers is maximised. The remaining inputs and outputs may then be adjusted to achieve full efficiency. In the basic model these factors are not permitted to deteriorate, although this is an option which can be incorporated if the DMU managers wish to sacrifice one output, say, to develop another.

By combining the attributes of these two models we will develop a model which incorporates trade-offs at the level of the population of DMUs as well as preference weighting for the individual DMU. In this model each DMU will be free to pursue a strategy of prioritising a subset of inputs and outputs within the constraints of the extended production technology provided by the trade-off approach.

5.3 New DEA model

In this section we first present the proposed model and then use a small example to show how the component parts work together. Then in two substantial subsections we will show how the model is computed and prove that it yields an efficient projection of any DMU in the set.

5.3.1 The proposed model

We have *n* DMUs, an *m* x *n* input matrix *X*, an *s* x *n* output matrix *Y*, and *T* trade-off vectors *P* and *Q*. Suppose we are evaluating DMU j_0. Managers have defined a subset of inputs $x_{ij_0} \in I_0$ which have been ascribed weights $w_i^- > 0$ and a subset of outputs $y_{rj_0} \in R_0$ with weights $w_r^+ > 0$.

We propose the following model:

$$
\begin{aligned}
\max \quad & \sum_{r \in R_0} w_r^+ z_r - \sum_{i \in I_0} w_i^- a_i + \varepsilon \left(\sum_{i \in \bar{I}_0} d_i + \sum_{r \in \bar{R}_0} e_r \right) & \\
\text{subject to} \quad & a_i x_{ij_0} - \left(\sum_{j=1}^{n} \lambda_j x_{ij} + \sum_{t=1}^{T} \pi_t P_t \right) = 0 & i \in I_0 \\
& z_r y_{rj_0} - \left(\sum_{j=1}^{n} \lambda_j y_{rj} + \sum_{t=1}^{T} \pi_t Q_t \right) = 0 & r \in R_0 \\
& \sum_{j=1}^{n} \lambda_j x_{ij} + \sum_{t=1}^{T} \pi_t P_t + c_i + d_i = x_{ij_0} & i \in \bar{I}_0 \\
& \sum_{j=1}^{n} \lambda_j y_{rj} + \sum_{t=1}^{T} \pi_t Q_t - e_r = y_{rj_0} & r \in \bar{R}_0 \\
& z_r \geq 1 & \\
& a_i \leq 1 & \\
& \lambda_j, \pi_t, c_i, d_i, e_r \geq 0 &
\end{aligned}
\qquad (5.1)
$$

The z_r are multipliers which will increase the selected outputs, while the a_i are multipliers to reduce the selected inputs. The first part of the objective function maximises the difference between the weighted combination of output multipliers and the weighted combination of input multipliers, thus ensuring that the output multipliers are as large as possible while the input multipliers are as small as possible within the given constraints. The second part of the objective function maximises the total adjustment of the remaining inputs and outputs with the slack variables d_i and e_r. Although it is written here in a single formulation, in fact the problem must be solved in three steps as outlined in section 5.3.3 below.

Solving the model will give us an optimal solution of the form $\lambda^*, \pi^*, a_i^*, z_r^*, c_i^*, d_i^*, e_r^*$ and allow us to find a projection of DMU j_0 which looks like this:

$$\begin{aligned}
x^*_{ij_0} &= a_i^* x_{ij_0} = \sum_{j=1}^{n} \lambda_j^* x_{ij} + \sum_{t=1}^{T} \pi_t^* P_t && i \in I_0 \\
y^*_{rj_0} &= z_r^* y_{rj_0} = \sum_{j=1}^{n} \lambda_j^* y_{rj} + \sum_{t=1}^{T} \pi_t^* Q_t && r \in R_0 \\
x^*_{ij_0} &= x_{ij_0} - d_i^* = \sum_{j=1}^{n} \lambda_j^* x_{ij} + \sum_{t=1}^{T} \pi_t^* P_t + c_i^* && i \in \overline{I}_0 \\
y^*_{rj_0} &= y_{rj_0} + e_r^* = \sum_{j=1}^{n} \lambda_j^* \hat{y}_{rj} + \sum_{t=1}^{T} \pi_t^* Q_t && r \in \overline{R}_0
\end{aligned} \tag{5.2}$$

Defining the projection $\left(x^*, y^*\right)$ in this way ensures that it belongs to the space $T_{CRS\text{-}TO}$ as described in section 3.5.4. A DMU is said to be efficient when $z_r^* = a_i^* = 1$ and $e_r^* = d_i^* = 0$, i.e. no changes are required to any of the inputs or outputs in order to maximise the objective function. For the purposes of this research the model is restricted to the CRS case.

5.3.2 Illustrative example

The following example illustrates how the framework and the model work together in practice.

Stage 1: Steps preceding the DEA model

Suppose the cohort of players which interests us consists of four institutions as shown in .

Table 5-1. We are considering their performance in an annual league table which includes three criteria: the number of staff, the number of graduates and the number of papers published in the preceding year. These three items will be used as the variables in our model: staff will be considered an input, graduates and publications will be considered as outputs. For the purpose of illustrating the model graphically, suppose that we have scaled the data so that each institution is credited with 100 staff. The original feasibility axiom states that the observed data belongs to the production technology so, for example, $\left(x_{U1}, y_{U1}\right) \in T_{CRS}$ where $x_{U1} = \left(100\right)$ and $y_{U1} = \left(1200, 25\right)$.

Table 5-1: Example DMUs

	Input	Output	
DMU	Staff (x_1)	Grads (y_1)	Papers (y_2)
U1	100	1200	25
U2	100	1100	50
U3	100	900	40
U4	100	600	45

Let us suppose further that managers have agreed the following trade-off judgments:

(1) that if one fewer paper was published in a year then, with no alteration to the number of staff, 10 additional graduates could be produced, and

(2) that if 20 fewer graduates were produced in a year then, with no alteration to the number of staff, one additional paper could be produced and published.

These judgments reflect the facts that (i) the relationship between teaching workload and the workload of producing a paper is better expressed as a range of values than as a single precise trade-off and (ii) in each exchange the more conservative judgment is preferred – in other words, we take the more pessimistic view of what can be obtained through the trade-off. They can be expressed as follows:

(1) t_1 is represented by the vectors $P_{t_1} = (0)$ and $Q_{t_1} = (10, -1)$

(2) t_2 is represented by the vectors $P_{t_2} = (0)$ and $Q_{t_2} = (-20, 1)$

Figure 5-1 shows the four DMUs U1, U2, U3 and U4 in red and the efficient frontier, incorporating the trade-offs t_1 and t_2, in blue. Note that the frontier is shown only in the area local to the observed data points and is not automatically extended to the axes (see section 3.5.4).

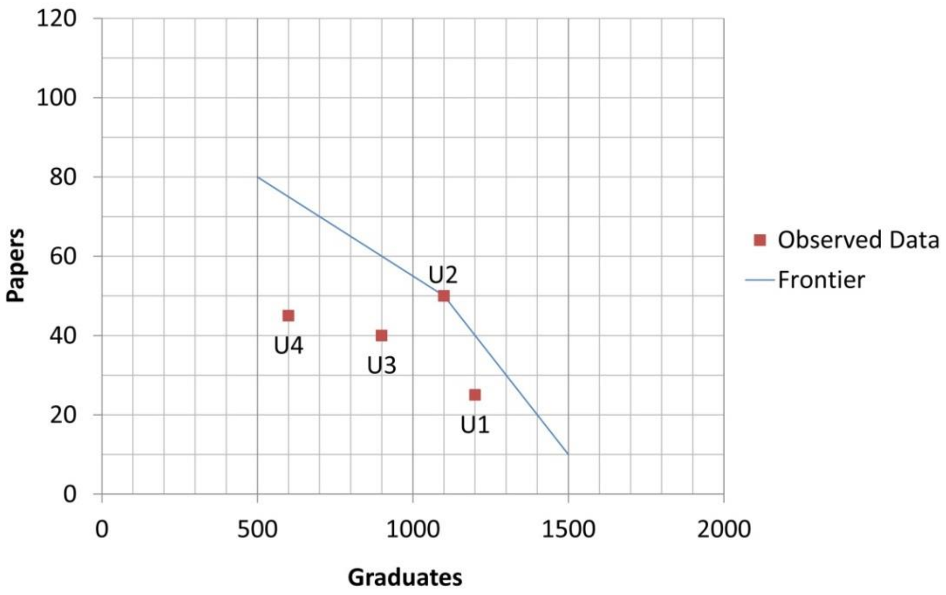


Figure 5-1: Observed data for U1 – U4 shown with an efficient frontier incorporating trade-offs t_1 and t_2

Stage 2: Using the DEA model to generate potential strategies for the players

We will concentrate on generating strategies for U4, which is currently performing poorly.

In this model the management team of U4 can choose whether to prioritise increasing graduates or papers. Let us suppose they elect to prioritise papers (y_2) and so choose a weight $w_2^+ > 0$. For this simple example, we will ignore the inputs and concentrate on this one improvement only. The first part of the objective function ($\max \sum_{r \in R_0} w_r^+ z_r - \sum_{i \in I_0} w_i^- a_i$) will therefore seek to find a value of z_2 which maximises $w_2^+ z_2$ while meeting all the constraints. Following the second equation in (5.2) we multiply the resultant value of z_2 by the original value of y_{2U4} and move U4 to A2 as shown in Figure 5-2. This equation also shows us that y'_{2U4} can be expressed in terms of the original set of observed data and the trade-off translations:

$$y'_{2U4} = z_2^* y_{2U4} = \sum_j \lambda_j^* y_{2j} + \sum_t \pi_t^* Q_t$$

In this case $\lambda^*_{U2} = 1$ while all the other λ^*_j are zero, and the active trade-off vector is $Q_{t_2} = (-20, 1)$ for which $\pi^*_{t_2} = 25$, so

$$\begin{aligned} y'_{2U4} &= \lambda^*_{U2} y_{2u2} + \pi^*_{t_2} Q_{t_2} \\ &= y_{2U2} + 25 Q_{t_2} \end{aligned}$$

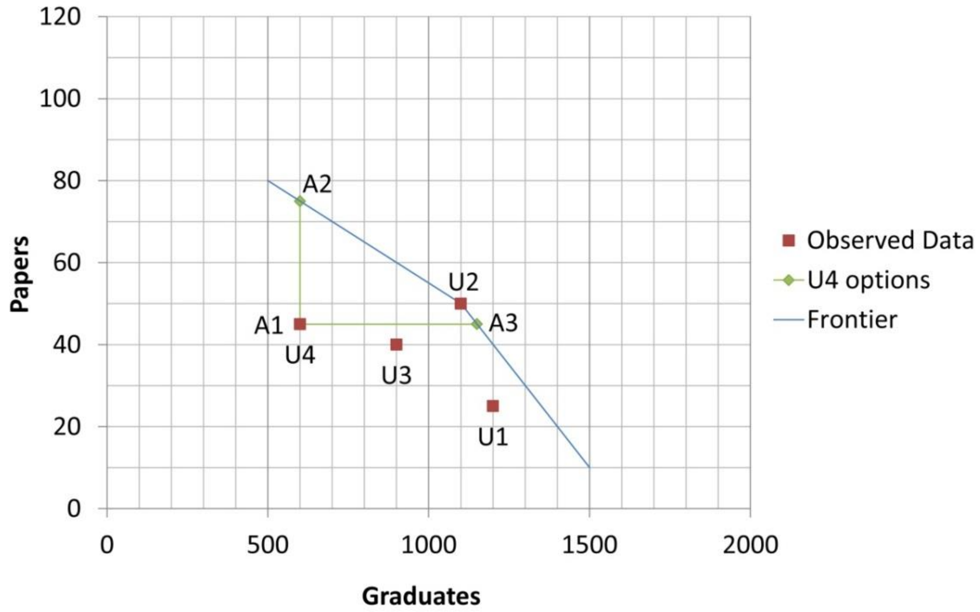


Figure 5-2: Generating options for U4 which prioritise different outputs

However, we do not wish to commit to a single strategy before evaluating its impact in stage 3, so we also consider the case where the management team decides to prioritise producing more graduates. This gives U4 a total of three options as shown in Figure 5-2 above and Table 5-2 below: to stay in the same place, to increase the number of papers produced or to increase the number of graduates produced. To keep this example small we will not add any further strategies, such as a mixed approach which aims to increase both graduates and papers.

Table 5-2: Strategies for U4

	Input	Output	
Strategy	Staff (x_1)	Grads (y_1)	Papers (y_2)
A1	100	600	45
A2	100	600	75
A3	100	1150	45

This demonstrates the way in which we are using the two aspects of the DEA model. The trade-off approach is used to determine the boundaries of the production possibility set, while the weighted preference structure is used to navigate through the PPS to achieve specific goals.

Stage 3: Steps following the DEA model

Returning to our imagined league table, the next step is to examine the potential impact of each of U4's options on their score and rank. Let us suppose that the three elements are equally weighted. For each element the institution with the best performance is awarded one point and the others are awarded a score between 0 and 1 based on the ratio of their performance to the maximum. Table 5-3 below shows the results.

Table 5-3: "League table" scores calculated relative to the best performance in each element. U4's score and rank are highlighted

	Strategy A1					Strategy A2					Strategy A3				
DMU	Staff	Grads	Papers	Score	Rank	Staff	Grads	Papers	Score	Rank	Staff	Grads	Papers	Score	Rank
U1	100	1200	25	2.50	3	100	1200	25	2.33	3	100	1200	25	2.50	4
U2	100	1100	50	2.92	1	100	1100	50	2.58	1	100	1100	50	2.92	1
U3	100	900	40	2.55	2	100	900	40	2.28	4	100	900	40	2.55	3
U4	100	600	45	2.40	4	100	600	75	2.50	2	100	1150	45	2.86	2
Best	100	1200	50			100	1200	75			100	1200	50		

In its current position, U4 ranks fourth in this league table. Adopting either A2 or A3 would take it to second place. However, we have so far assumed that every other institution is content to maintain its current performance. To examine the situation further we need to look at the possible strategies of other institutions. By applying the DEA model to U3 we can generate a similar set of three options as shown in the table and figure below.

Table 5-4: Strategies for U3

	Input	Output	
Strategy	Staff (x_1)	Grads (y_1)	Papers (y_2)
B1	100	900	40
B2	100	900	60
B3	100	1200	40

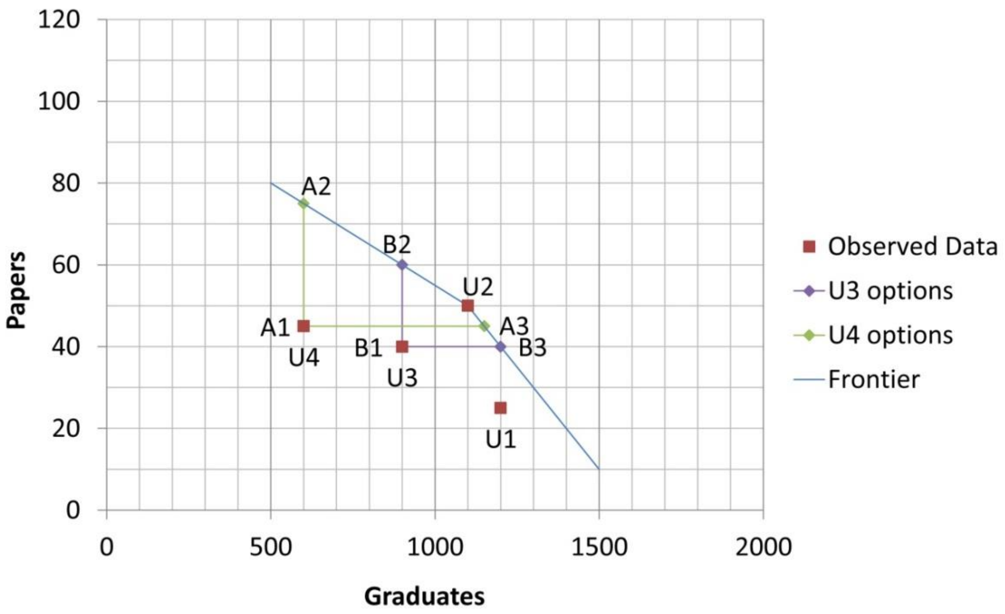


Figure 5-3: Three alternative options for each of U3 and U4

To keep this example as succinct as possible Table 5-5 omits the detail of the different strategies in order to focus on the scores and ranks which result. Note that where the payoff is expressed in terms of rank, a lower value is to be preferred to a higher one.

Table 5-5: Interaction of the alternative strategies for U3 and U4

U4 Strategies

U3 Strategies		A1		A2		A3	
	DMU	Score	Rank	Score	Rank	Score	Rank
B1	U1	2.50	3	2.33	3	2.50	4
	U2	2.92	1	2.58	1	2.92	1
	U3	2.55	2	2.28	4	2.55	3
	U4	2.40	4	2.50	2	2.86	2
B2	U1	2.42	3	2.33	4	2.42	4
	U2	2.75	1	2.58	1	2.75	1
	U3	2.75	1	2.55	2	2.75	1
	U4	2.25	4	2.50	3	2.71	3
B3	U1	2.50	3	2.33	4	2.50	4
	U2	2.92	1	2.58	1	2.92	1
	U3	2.80	2	2.53	2	2.80	3
	U4	2.40	4	2.50	3	2.86	2

Before U4 moves, we can see that U3 enjoys second place in the league tables, a position they will lose under either of U4's strategies. U3's best strategy is B2, which strongly dominates B1 and weakly dominates B3: it offers the opportunity to maintain or improve on their current position.

In response to this, U4 could adopt either A2 or A3 and still gain an improvement of one place from fourth to third. In fact, both (B2, A2) and (B2, A3) are equilibria. However, A3 weakly dominates A2 (it offers a better payoff against B3) and, if we look at the actual scores as well as the league table rank, we can see that against B2 it also "closes the gap" between U4 and the leading universities. This might also be an important factor in a management decision.

This is a very simple example, but it serves to illustrate the logic of the framework proposed here. We will now return to the DEA model in order to outline a way to obtain the preferred projection of a given DMU within the trade-off space and show that this projection is efficient.

5.3.3 Obtaining the projection

The usual approach to evaluating a DMU and obtaining an efficient projection is to undertake two computational steps. Podinovski has shown that in order to get a unique solution when using the trade-off model it is necessary to use three steps (Podinovski, 2007a). We adopt this three step process, but modify it to use the preference-weighted objective function in (5.1). This change is apparent mainly in step 1, but is continued through step 2 in the separation of the two groups of constraints: those which apply to the selected inputs and outputs and those which cover the remaining (unselected) factors.

Step 1

Maximise the weighted part of the objective function

$$\sum_{r\in R_0} w_r^+ z_r - \sum_{i\in I_0} w_i^- a_i$$

i.e. we want the input multipliers to be as small as possible and the output multipliers to be as large as possible within the constraints of the production technology. This means solving model (5.3) below. The constraints are defined so that the selected inputs and outputs are moved towards the frontier of $T_{CRS\text{-}TO}$ while the remaining inputs and outputs are not permitted to deteriorate.

$$
\begin{aligned}
\max_{z,a} \quad & \sum_{r\in R_0} w_r^+ z_r - \sum_{i\in I_0} w_i^- a_i \\
\text{subject to} \quad & a_i x_{ij_0} - \left(\sum_{j=1}^{n} \lambda_j x_{ij} + \sum_{t=1}^{T} \pi_t P_t \right) = 0 && i \in I_0 \\
& z_r y_{rj_0} - \left(\sum_{j=1}^{n} \lambda_j y_{rj} + \sum_{t=1}^{T} \pi_t Q_t \right) = 0 && r \in R_0 \\
& \sum_{j=1}^{n} \lambda_j x_{ij} + \sum_{t=1}^{T} \pi_t P_t \le x_{ij_0} && i \in \overline{I}_0 \\
& \sum_{j=1}^{n} \lambda_j y_{rj} + \sum_{t=1}^{T} \pi_t Q_t \ge y_{rj_0} && r \in \overline{R}_0 \\
& z_r \ge 1 \\
& a_i \le 1 \\
& \lambda_j, \pi_t \ge 0
\end{aligned}
\tag{5.3}
$$

Having obtained an optimal solution $z_r', a_i', \lambda_j', \pi_t'$ we then define a virtual DMU (x', y') with

$$
\begin{aligned}
x'_{ij_0} &= a_i' x_{ij_0} = \sum_{j=1}^{n} \lambda_j' x_{ij} + \sum_{t=1}^{T} \pi_t' P_t && i \in I_0 \\
y'_{ij_0} &= z_r' y_{rj_0} = \sum_{j=1}^{n} \lambda_j' y_{rj} + \sum_{t=1}^{T} \pi_t' Q_t && r \in R_0 \\
x'_{ij_0} &= x_{ij_0} && i \in \overline{I}_0 \\
y'_{rj_0} &= y_{rj_0} && r \in \overline{R}_0
\end{aligned}
$$

Step 2

Maximise the secondary part of the objective function

$$\sum_{i\in \overline{I}_0} d_i + \sum_{r\in \overline{R}_0} e_r$$

We now want to remove any slacks in the remaining inputs and outputs. By using the values we found in Step 1, we fix the prioritised inputs and outputs at their projected levels and look for any improvements we can make in the other factors.

$$\begin{aligned}
\max_{d,e} \quad & \sum_{i\in \bar{I}_0} d_i + \sum_{r\in \bar{R}_0} e_r \\
\text{subject to} \quad & \sum_{j=1}^{n} \lambda_j x_{ij} + \sum_{t=1}^{T} \pi_t P_t = x'_{ij_0} && i \in I_0 \\
& \sum_{j=1}^{n} \lambda_j y_{rj} + \sum_{t=1}^{T} \pi_t Q_t = y'_{rj_0} && r \in R_0 \\
& \sum_{j=1}^{n} \lambda_j x_{ij} + \sum_{t=1}^{T} \pi_t P_t + c_i + d_i = x'_{ij_0} && i \in \bar{I}_0 \qquad (5.4) \\
& \sum_{j=1}^{n} \lambda_j x_{ij} + \sum_{t=1}^{T} \pi_t P_t + c_i \geq 0 && i \in \bar{I}_0 \\
& \sum_{j=1}^{n} \lambda_j y_{rj} + \sum_{t=1}^{T} \pi_t Q_t - e_r = y'_{rj_0} && r \in \bar{R}_0 \\
& \lambda_j, \pi_t, c_i, d_i, e_r \geq 0
\end{aligned}$$

We obtain an optimal solution in the form $\lambda_j^*, \pi_t^*, c_i^*, d_i^*, e_r^*$ and then define a new virtual DMU $\left(x^*, y^*\right)$ with

$$\begin{aligned}
& x_{ij_0}^* = x'_{ij_0} && i \in I_0 \\
& y_{ij_0}^* = y'_{rj_0} && r \in R_0 \\
& x_{ij_0}^* = x'_{ij_0} - d_i^* = \sum_{j=1}^{n} \lambda_j^* x_{ij} + \sum_{t=1}^{T} \pi_t^* P_t + c_i^* && i \in \bar{I}_0 \\
& y_{ij_0}^* = y'_{rj_0} + e_r^* = \sum_{j=1}^{n} \lambda_j^* x_{ij} + \sum_{t=1}^{T} \pi_t^* Q_t && r \in \bar{R}_0
\end{aligned}$$

Step 3

Although we now have an optimal solution giving an efficient virtual DMU, this solution may not be unique. If it is not unique then the optimal λ_j may not satisfy the requirement for a CRS reference set that $\lambda_k = 0$ for inefficient DMU_k. Podinovski (2007a) shows that the solution which maximises the c_i terms satisfies this requirement.

$$
\begin{aligned}
\max_{c} \quad & \sum_{i\in\bar{I}_0} c_i \\
\text{subject to} \quad & \sum_{j=1}^{n}\lambda_j x_{ij} + \sum_{t=1}^{T}\pi_t P_t = x^*_{ij_0} && i \in I_0 \\
& \sum_{j=1}^{n}\lambda_j y_{rj} + \sum_{t=1}^{T}\pi_t Q_t = y^*_{rj_0} && r \in R_0 \\
& \sum_{j=1}^{n}\lambda_j x_{ij} + \sum_{t=1}^{T}\pi_t P_t + c_i = x^*_{ij_0} && i \in \bar{I}_0 \\
& \sum_{j=1}^{n}\lambda_j y_{rj} + \sum_{t=1}^{T}\pi_t Q_t = y^*_{rj_0} && r \in \bar{R}_0 \\
& \lambda_j, \pi_t, c_i \geq 0
\end{aligned} \tag{5.5}
$$

We solve model (5.5) to obtain an optimal $\lambda_j^\circ, \pi_j^\circ, c_i^\circ$, which gives us our complete optimal solution $z_r', a_i', d_i^*, e_r^*, \lambda_j^\circ, \pi_j^\circ, c_i^\circ$. We can now write $\left(x^*, y^*\right)$ as

$$
\begin{aligned}
x^*_{ij_0} &= a_i' x_{ij_0} = \sum_{j=1}^{n}\lambda_j^\circ x_{ij} + \sum_{t=1}^{T}\pi_t^\circ P_t && i \in I_0 \\
y^*_{ij_0} &= z_r' y_{rj_0} = \sum_{j=1}^{n}\lambda_j^\circ x_{ij} + \sum_{t=1}^{T}\pi_t^\circ Q_t && r \in R_0 \\
x^*_{ij_0} &= x_{ij_0} - d_i^* = \sum_{j=1}^{n}\lambda_j^\circ x_{ij} + \sum_{t=1}^{T}\pi_t^\circ P_t + c_i^\circ && i \in \bar{I}_0 \\
y^*_{ij_0} &= y_{rj_0} + e_r^* = \sum_{j=1}^{n}\lambda_j^\circ x_{ij} + \sum_{t=1}^{T}\pi_t^\circ Q_t && r \in \bar{R}_0
\end{aligned} \tag{5.6}
$$

5.3.4 Showing that the projection is efficient

We noted above that Podinovski has proved that when $\lambda_k^\circ > 0$ in the optimal solution to step 3 then DMU *k* is efficient. First we need to demonstrate that this is true for our composite model.

Theorem

If $\lambda_k^\circ > 0$ in the solution to (5.5), then DMU *k* is Pareto-efficient in $T_{CRS\text{-}TO}$.

Proof

Assume that *k* is inefficient. Then (x_k, y_k) is dominated by some $(\hat{x}, \hat{y})$ in $T_{CRS\text{-}TO}$. This means that (x_k, y_k) can be expressed as $x_k = \hat{x} + \delta$, $y_k = \hat{y} - \varepsilon$ where δ and ε are non-negative vectors and $(\delta, \varepsilon) \neq 0$.

Since $(\hat{x}, \hat{y})$ is in $T_{CRS\text{-}TO}$ it can be written in the standard form with some $\hat{\lambda}, \hat{\pi}, \hat{d}, \hat{e}$ so that it looks like this:

$$\hat{x} = X\hat{\lambda} + \sum_t \hat{\pi}_t P_t + \hat{d}$$
$$\hat{y} = Y\hat{\lambda} + \sum_t \hat{\pi}_t Q_t - \hat{e}$$

Then (x_k, y_k) looks like this:

$$x_k = X\hat{\lambda} + \sum_t \hat{\pi}_t P_t + \hat{d} + \delta$$
$$y_k = Y\hat{\lambda} + \sum_t \hat{\pi}_t Q_t - \hat{e} - \varepsilon$$

with each x_{ik} and y_{rk} written:

$$x_{ik} = \sum_{j=1}^{n} \hat{\lambda}_j x_{ij} + \sum_{t=1}^{T} \hat{\pi}_t P_t + \hat{d}_i + \delta_i$$
$$y_{rk} = \sum_{j=1}^{n} \hat{\lambda}_j y_{rj} + \sum_{t=1}^{T} \hat{\pi}_t Q_t - \hat{e}_r - \varepsilon_r$$

We take these expressions defining DMU *k* and we use them in place of x_k and y_k in the expressions (5.6) for our optimised DMU (x^*, y^*).

For all the selected factors $i \in I_0, r \in R_0$ we have

$$\begin{aligned}
x_{ij_0}^{*} &= \sum_{j=1}^{n} \lambda_j^{\circ} x_{ij} + \sum_{t=1}^{T} \pi_t^{\circ} P_t \\
&= \sum_{j \neq k} \lambda_j^{\circ} x_{ij} + \lambda_k^{\circ} \left(\sum_{j=1}^{n} \hat{\lambda}_j x_{ij} + \sum_{t=1}^{T} \hat{\pi}_t P_t + \hat{d}_i + \delta_i \right) + \sum_{t=1}^{T} \pi_t^{\circ} P_t \\
&= \sum_{j \neq k} \lambda_j^{\circ} x_{ij} + \sum_{j=1}^{n} \lambda_k^{\circ} \hat{\lambda}_j x_{ij} + \sum_{t=1}^{T} \lambda_k^{\circ} \hat{\pi}_t P_t + \sum_{t=1}^{T} \pi_t^{\circ} P_t + \lambda_k^{\circ} \left(\hat{d}_i + \delta_i \right) \\
&= \sum_{j=1}^{n} \lambda_j'' x_{ij} + \sum_{t=1}^{T} \pi_t'' P_t + c_i''
\end{aligned}$$

$$\begin{aligned}
y_{rj_0}^{*} &= \sum_{j=1}^{n} \lambda_j^{\circ} y_{rj} + \sum_{t=1}^{T} \pi_t^{\circ} Q_t \\
&= \sum_{j \neq k} \lambda_j^{\circ} y_{rj} + \lambda_k^{\circ} \left(\sum_{j=1}^{n} \hat{\lambda}_j y_{rj} + \sum_{t=1}^{T} \hat{\pi}_t Q_t - \hat{e}_r - \varepsilon_r \right) + \sum_{t=1}^{T} \pi_t^{\circ} Q_t \\
&= \sum_{j \neq k} \lambda_j^{\circ} y_{rj} + \sum_{j=1}^{n} \lambda_k^{\circ} \hat{\lambda}_j y_{rj} + \sum_{t=1}^{T} \lambda_k^{\circ} \hat{\pi}_t Q_t + \sum_{t=1}^{T} \pi_t^{\circ} Q_t - \lambda_k^{\circ} \left(\hat{e}_r + \varepsilon_r \right) \\
&= \sum_{j=1}^{n} \lambda_j'' y_{rj} + \sum_{t=1}^{T} \pi_t'' Q_t - e_r''
\end{aligned}$$

where $e_r'' = \lambda_k^{\circ} \left(\hat{e}_r + \varepsilon_r \right)$, $c_i'' = \lambda_k^{\circ} \left(\hat{d}_i + \delta_i \right)$, $\pi'' = \lambda_k^{\circ} \hat{\pi}_t + \pi_t^{\circ}$, $\lambda_j'' = \lambda_j^{\circ} + \lambda_k^{\circ} \hat{\lambda}_j$ for $j \neq k$ and $\lambda_k'' = \lambda_k^{\circ} \hat{\lambda}_k$.

For the unselected factors $i \in \overline{I_0}, r \in \overline{R}_0$ we have

$$\begin{aligned}
x_{ij_0} &= \sum_{j=1}^{n} \lambda_j^{\circ} x_{ij} + \sum_{t=1}^{T} \pi_t^{\circ} P_t + c_i^{\circ} + d_i^{*} \\
&= \sum_{j \neq k} \lambda_j^{\circ} x_{ij} + \lambda_k^{\circ} \left(\sum_{j=1}^{n} \hat{\lambda}_j x_{ij} + \sum_{t=1}^{T} \hat{\pi}_t P_t + \hat{d}_i + \delta_i \right) + \sum_{t=1}^{T} \pi_t^{\circ} P_t + c_i^{\circ} + d_i^{*} \\
&= \sum_{j \neq k} \lambda_j^{\circ} x_{ij} + \sum_{j=1}^{n} \lambda_k^{\circ} \hat{\lambda}_j x_{ij} + \sum_{t=1}^{T} \pi_t^{\circ} P_t + \sum_{t=1}^{T} \lambda_k^{\circ} \hat{\pi}_t P_t + \lambda_k^{\circ} \left(\hat{d}_i + \delta_i \right) + c_i^{\circ} + d_i^{*} \\
&= \sum_{j=1}^{n} \lambda_j'' x_{ij} + \sum_{t=1}^{T} \pi_t'' P_t + c_i'' + d_i''
\end{aligned}$$

$$
\begin{aligned}
y_{ij_0} &= \sum_{j=1}^{n} \lambda_j^{\circ} x_{ij} + \sum_{t=1}^{T} \pi_t^{\circ} Q_t - e_r^* \\
&= \sum_{j \neq k} \lambda_j^{\circ} y_{rj} + \lambda_k^{\circ} \left(\sum_{j=1}^{n} \hat{\lambda}_j y_{rj} + \sum_{t=1}^{T} \hat{\pi}_t Q_t - \hat{e}_r - \varepsilon_r \right) + \sum_{t=1}^{T} \pi_t^{\circ} Q_t - e_r^* \\
&= \sum_{j \neq k} \lambda_j^{\circ} y_{rj} + \sum_{j=1}^{n} \lambda_k^{\circ} \hat{\lambda}_j y_{rj} + \sum_{t=1}^{T} \lambda_k^{\circ} \hat{\pi}_t Q_t + \sum_{t=1}^{T} \pi_t^{\circ} Q_t - \lambda_k^{\circ} \left(\hat{e}_r + \varepsilon_r \right) - e_r^* \\
&= \sum_{j=1}^{n} \lambda_j'' y_{rj} + \sum_{t=1}^{T} \pi_t'' Q_t - e_r''
\end{aligned}
$$

where $e_r'' = \lambda_k^{\circ} \left(\hat{e}_r + \varepsilon_r \right) + e_r^*$, $c_i'' = \lambda_k^{\circ} \left(\hat{d}_i + \delta_i \right) + c_i^{\circ}$, $d_i'' = d_i^*$, $\pi'' = \lambda_k^{\circ} \hat{\pi}_t + \pi_t^{\circ}$, $\lambda_j'' = \lambda_j^{\circ} + \lambda_k^{\circ} \hat{\lambda}_j$ for $j \neq k$ and $\lambda_k'' = \lambda_k^{\circ} \hat{\lambda}_k$.

Since $\lambda_k^{\circ} > 0$ and $\left(\delta, \varepsilon \right) \neq 0$, in the case of the selected factors we must have either $e_r'' > 0$ or $c_i'' > 0$ and in the other case we must have either $e_r'' > e_r^*$ or $c_i'' > c_i^{\circ}$.

However, $c_i'' > c_i^{\circ}$ contradicts the optimality of $\lambda_j^{\circ}, \pi_j^{\circ}, c_i^{\circ}$ as a solution of (5.5) while $e_r'' > e_r^*$ contradicts the optimality of $\lambda_j^{\circ}, \pi_j^{\circ}, c_i^{\circ}, e_r^*, d_i^*$ as a solution of (5.4), and if we have $e_r'' > 0$ or $c_i'' > 0$ then we have contradicted the optimality of our original z_r', a_i' in the solution to (5.3).

So we cannot have $\left(\delta, \varepsilon \right) \neq 0$ when $\lambda_k^{\circ} > 0$, therefore we must have $\delta = \varepsilon = 0$ and DMU *k* is efficient.

Now we turn to the virtual DMU $\left(x^*, y^* \right)$ and show that it is efficient.

Theorem

The virtual DMU $\left(x^*, y^*\right)$ defined in (5.6) is Pareto-efficient in the $T_{CRS\text{-}TO}$ production possibility set.

Proof

Suppose we have solved model (5.1) for DMU j_0 and obtained an optimal solution $\left(z'_r, a'_i, \lambda'_j, \pi'_t, d'_i, e'_r, c'_i\right)$ where $z'_r > 1$, $a'_i < 1$, $e'_r > 0$ or $d'_i > 0$ for some *i* or *r*. We know, therefore, that j_0 is an inefficient DMU. We use the solution to create a virtual DMU $\left(\hat{x}, \hat{y}\right)$ with

$$
\begin{aligned}
\hat{x}_{ij_0} &= a'_i x_{ij_0} && i \in I_0 \\
\hat{y}_{ij_0} &= z'_r y_{rj_0} && r \in R_0 \\
\hat{x}_{ij_0} &= x_{ij_0} - d'_i && i \in \bar{I}_0 \\
\hat{y}_{rj_0} &= y_{rj_0} + e'_r && r \in \bar{R}_0
\end{aligned}
\tag{5.7}
$$

We now substitute this DMU for j_0 in the set of *n* DMUs, and evaluate it using the following model:

$$
\begin{aligned}
\max \quad & \sum_{r \in R_0} w_r^+ v_r - \sum_{i \in I_0} w_i^- b_i + \varepsilon \left(\sum_{i \in \bar{I}_0} g_i + \sum_{r \in \bar{R}_0} h_r \right) \\
\text{subject to} \quad & b_i \hat{x}_{ij_0} - \left(\sum_{j \neq j_0} \mu_j x_{ij} + \mu_{j_0} \hat{x}_{ij_0} + \sum_{t=1}^{T} \rho_t P_t \right) = 0 && i \in I_0 \\
& v_r \hat{y}_{rj_0} - \left(\sum_{j \neq j_0} \mu_j y_{rj} + \mu_{j_0} \hat{y}_{rj_0} + \sum_{t=1}^{T} \rho_t Q_t \right) = 0 && r \in R_0 \\
& \sum_{j \neq j_0} \mu_j x_{ij} + \mu_{j_0} \hat{x}_{ij_0} + \sum_{t=1}^{T} \rho_t P_t + f_i + g_i = \hat{x}_{ij_0} && i \in \bar{I}_0 \\
& \sum_{j \neq j_0} \mu_j y_{rj} + \mu_{j_0} \hat{y}_{rj_0} + \sum_{t=1}^{T} \rho_t Q_t - h_r = \hat{y}_{rj_0} && r \in \bar{R}_0 \\
& v_r \geq 1 \\
& b_i \leq 1 \\
& \mu_j, \rho_t, f_i, g_i, h_r \geq 0
\end{aligned}
\tag{5.8}
$$

Let the optimal solution be $\left(v_r^*, b_i^*, \mu_j^*, \rho_t^*, f_i^*, g_i^*, h_r^*\right)$ such that $v_r^* > 1$, $b_i^* < 1$, $h_r^* > 0$ or $g_i^* > 0$ for some *i* or *r*. In other words, we suppose that the optimal solution for the model shows our virtual DMU to be inefficient. Since we have ensured (through step 3 of the model's implementation and the proof above) that $\mu_k = 0$ for inefficient DMU *k* we know that $\mu_{j_0}^*$ must be equal to zero.

If we write out the constraints from model (5.8) as equations incorporating our optimal values we have

$$b_i^* \hat{x}_{ij_0} - \left(\sum_{j \neq j_0} \mu_j^* x_{ij} + \mu_{j_0}^* \hat{x}_{ij_0} + \sum_{t=1}^{T} \rho_t^* P_t \right) = 0 \qquad i \in I_0$$

$$v_r^* \hat{y}_{rj_0} - \left(\sum_{j \neq j_0} \mu_j^* y_{rj} + \mu_{j_0}^* \hat{y}_{rj_0} + \sum_{t=1}^{T} \rho_t^* Q_t \right) = 0 \qquad r \in R_0$$

$$\sum_{j \neq j_0} \mu_j^* x_{ij} + \mu_{j_0}^* \hat{x}_{ij_0} + \sum_{t=1}^{T} \rho_t^* P_t + f_i^* + g_i^* = \hat{x}_{ij_0} \qquad i \in \overline{I}_0$$

$$\sum_{j \neq j_0} \mu_j^* y_{rj} + \mu_{j_0}^* \hat{y}_{rj_0} + \sum_{t=1}^{T} \rho_t^* Q_t - h_r^* = \hat{y}_{rj_0} \qquad r \in \overline{R}_0$$

We can then replace $(\hat{x}, \hat{y})$ with the values from (5.7) and, since $\mu_{j_0}^* = 0$, we obtain

$$b_i^* a_i' x_{ij_0} - \left(\sum_{j \neq j_0} \mu_j^* x_{ij} + \sum_{t=1}^{T} \rho_t^* P_t \right) = 0 \qquad i \in I_0$$

$$v_r^* z_r' y_{rj_0} - \left(\sum_{j \neq j_0} \mu_j^* y_{rj} + \sum_{t=1}^{T} \rho_t^* Q_t \right) = 0 \qquad r \in R_0$$

$$\sum_{j \neq j_0} \mu_j^* x_{ij} + \sum_{t=1}^{T} \rho_t^* P_t + f_i^* + g_i^* + d_i' = x_{ij_0} \qquad i \in \overline{I}_0$$

$$\sum_{j \neq j_0} \mu_j^* y_{rj} + \sum_{t=1}^{T} \rho_t^* Q_t - h_r^* - e_r' = y_{rj_0} \qquad r \in \overline{R}_0$$

This is clearly a feasible solution to our original model (5.1) with the following values

$$\begin{aligned} \tilde{a}_i &= b_i^* a_i' \\ \tilde{z}_r &= v_r^* z_r' \\ \tilde{\lambda}_j &= \mu_j^* \\ \tilde{\pi}_t &= \rho_t^* \\ \tilde{d}_i &= g_i^* + d_i' \\ \tilde{e}_r &= h_r^* + e_r' \\ \tilde{c}_i &= f_i^* \end{aligned}$$

Since we assumed that $v_r^* > 1$, $b_i^* < 1$, $h_r^* > 0$ or $g_i^* > 0$ we must have $\tilde{z}_r > z'_r$, $\tilde{a}_i < a'_i$, $\tilde{d}_i > d'_i$ or $\tilde{e}_r > e'_r$. However, this would mean that the objective function is greater for our new solution than for our original solution, which was optimal. So we have a contradiction and we must have $v_r^* = 1$, $b_i^* = 1$, $h_r^* = 0$ and $g_i^* = 0$. Thus the projected DMU is efficient.

5.4 Implementation

A key aspect of implementation was to code the DEA model using suitable software. The optimisation package XPRESS-MP and its programmimg language Mosel were found to be appropriate for this purpose. A program was developed to implement stage 2 of the framework; Excel spreadsheets were used for reading and writing the data and for final analysis and presentation of the results.

There is a significant technical challenge in checking the correctness of any implementation of a DEA model. For standard models it is possible to test the results against those given by other programs such as DEA-solver and DEAFrontier, although even here there is room for discrepancy. The results of a DEA analysis are not necessarily uniquely defined and Cooper et al (2007) have demonstrated the range of values obtained from different DEA programs for the same data. To test the implementation of the new model (see Appendix A), simple numerical examples were devised which could be checked using Excel's Solver add-in before more realistic and challenging datasets were used.

The testing process revealed a problem with the initial model. In step 2 the total sum of output slack $\sum_{r \in \bar{R}_0} e_r$ was found to be unbounded, due to the formulation of the final constraint in (5.4) where

$$\sum_{j=1}^{n} \lambda_j y_{rj} + \sum_{t=1}^{T} \pi_t Q_t - e_r = y'_{rj_0} \qquad r \in \bar{R}_0$$

In order to overcome the problem of an unbounded variable, a cap was introduced. The value of the cap is arbitrary but needs to be sufficiently large to allow for any reasonable values of e_r. The additional constraint introduced was therefore

$$\sum_{r \in \bar{R}_0} e_r \le \sum_{r \in \bar{R}_0} y_r$$

i.e. the total output slack cannot excced the total output.

One further modification was made, namely to insert a loop so that the program could accept multiple alternative sets of weights and run the model for each set and each DMU without interruption.

5.5 Summary

In this chapter we have presented in full the framework for our theoretical model.

We have examined the existing models and described how they will work together. We have used a detailed illustration to work through the proposed framework step by step. A new DEA model has been created and proofs provided to show that it provides an efficient projection of the DMU under evaluation. We have also described the implementation of this model in XPRESS-MP.

Having completed this foundational work, we are now ready to operationalise the model in Chapter 6.

6 Operationalising the Model

Having developed a theoretical framework in Chapter 5, we now turn to the task of operationalising the proposed model. First we compare the technical literature with the problem context and identify a number of issues to be addressed. Then we consider in detail the data specification of previous models as well as the limitations of the available dataset before we develop an appropriate specification for the proposed new model.

6.1 The problem context and the requirements of DEA

In our review of the literature we have considered both the context of the problem and the ideal specification of a DEA model. Setting these side by side, we can immediately see some key differences.

Table 6-1 Comparison of the problem context with the requirements of DEA

The problem context	Requirements of DEA
League tables in the UK include institutions from all four constituent nations of the United Kingdom in spite of differences in the way their degree programmes are structured	DEA assumes that the decision making units being analysed are comparable because they are homogeneous, i.e. “the factors … characterizing the performance of all units in the group, are identical, except for differences in intensity or magnitude” (Golany and Roll, 1989)
The data used in league tables relates to inputs, processes and outputs and credit is given for larger scores in all three categories	DEA uses a production model in which credit is given for larger scores when outputs are measured but for smaller scores when inputs are measured
League tables may use several variables which are highly correlated	In DEA we aim to exclude redundant variables

The first issue noted in Table 6-1 concerns the selection of DMUs, while the second and third concern the selection of variables. We will consider these two aspects of operationalising the model in turn.

6.2 Selection of variables

In the first place, our selection of variables is constrained by the system we have chosen to study. Although we have already established that the measures which appear in league tables are not ideal for characterising the performance of institutions (see section 152.5 above), we do need to use these variables in order to see what happens when we set targets to increase or reduce them.

Before turning to a detailed consideration of this problem, we will review the data employed in the DEA studies we surveyed in Chapter 3.

6.2.1 DEA studies

Our purpose in revisiting these studies is to examine the data used to specify previous models. Many authors rightly give this topic a great deal of attention, as measuring the inputs and outputs of a higher education institution is by no means straightforward even when one accepts this modelling approach as valid. Agasisti and Pérez-Esparrells (2010) provide a useful model of the HE production process, characterising universities as "organizations using financial and human resources as inputs to produce human capital and research products as outputs" (p 89).

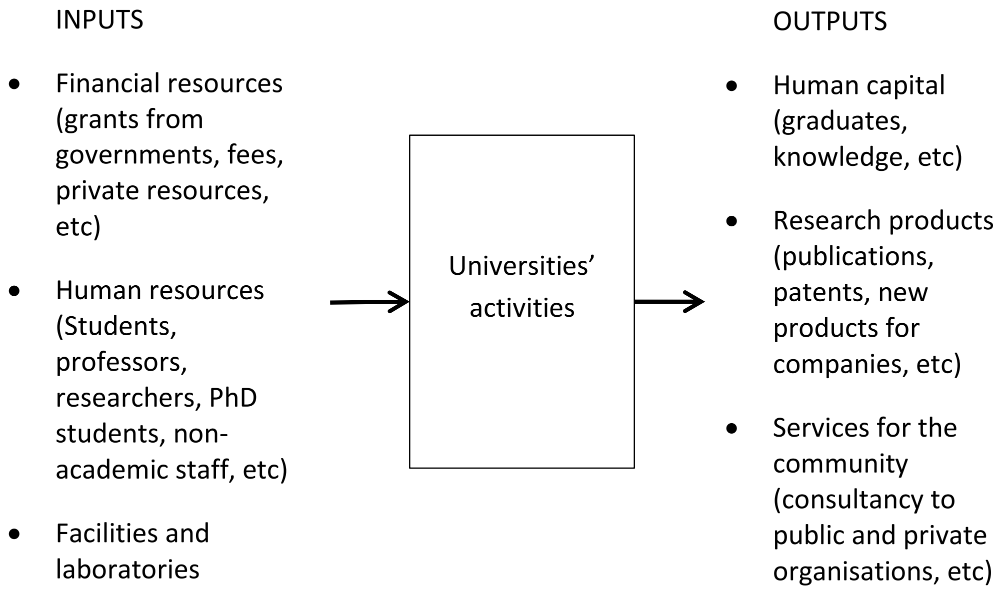


Figure 6-1: Simplified model of the production process (Agasisti and Pérez-Esparrells, 2010, p 90)

Output Variables

In this model, the outputs of teaching activities are broadly covered by the category "Human capital". In practice, typical indicators used as proxies for teaching outputs include

Number of graduates

This is usually subdivided by level of qualification, such as first degrees and higher degrees (see e.g. Madden et al., 1997, Johnes, 2006, Agasisti and Johnes, 2009) and sometimes an effort is made to adjust a simple count to incorporate a measure of quality (see e.g. Johnes, 2003a, b, Flegg et al., 2004, Turner, 2005). It is worth noting that in league tables teaching outputs are typically measured not only by the number of students taught but by defining certain threshholds of achievement, such as a minimum class of degree or type of employment, and calculating the proportion of students who pass this threshhold.

Number of students

This is typically given in terms of Full-Time Equivalents (FTE) rather than as a head count and may be subdivided by level (see e.g. Beasley, 1990, 1995, Avkiran, 2001) and also by subject (see e.g. Agasisti and Salerno, 2007, Thanassoulis et al., 2011). Sometimes the number of students is expressed as a rate, such as the average enrolment per class (Arcelus and Coleman, 1997), but there is a problem with this approach as we will discuss below.

Employment indicators

These are used by Haksever and Muragishi (1998) and Colbert et al. (2000), in their evaluations of MBA programmes. Average starting salary is used in both of these studies, and Haksever and Muragishi also use the percentage of students who have a job by graduation. Employment rates are often components of league tables, and both Sarrico et al. (1997) and Turner (2005) include these indicators in their studies.

Student Satisfaction

This is a relatively new measure in UK league tables and has not yet been picked up in any DEA studies of the sector. However, it has been used by Colbert et al. (2000) in their evaluation of MBA programmes.

It is interesting that some researchers take the number of *students* to be an output while others take the number of *graduates*. Those who opt for graduates as an output also include some measure of the student FTE but as an input. Those who opt for students as an ouput argue that even if a student does not complete their course, they still benefit from the teaching they have received. The use of student numbers as an output might be justified if one were specifically modelling the workload of academic staff, but, as Johnes rightly notes (2003a, p 8), students are also making a contribution to their education and some part of the output from HE can be attributed to them. For this reason, we agree with those who categorise students as an input and graduates as an output.

Some DEA studies use a more specialised production model than the one proposed by Agasisti and Pérez-Esparrells. For example, we have already noted that Joumady and Ris (2005) place preparation of graduates for the workplace at the centre of their research into university efficiency. They define two sets of competencies: generic competencies, such as reflective thinking and problem solving, and vocational competencies, which include subject-specific knowledge and practice. For each item their survey of graduates asks the respondent to identify their own level of competency at the time they graduated and the level of competency they need in their current job. The authors then develop three models, each of which takes as its outputs a set of variables derived from these data.

Turning to measures of research output – "Research products" in Figure 6-1 – we find the following indicators in use:

Publications

These are used both by studies whose interest is solely in research efficiency, such as Johnes and Johnes (1993) and Kocher et al. (2006), and by those where both teaching and research are considered (Madden et al., 1997, Sellers-Rubio et al., 2010). Publications are always divided into categories, although the choice of categories varies greatly. Johnes & Johnes distinguish primarily between types of journal based on their intended audience (academic, professional or popular) while Sellers-Rubio et al distinguish between national and international publications. Three of the four studies identified here consider as a category articles published in "core" journals of Economics, although each uses a different protocol for identifying such journals and the total number varies from 10 journals (Kocher et al., 2006) to 93 (Madden et al., 1997).

Research grants

Research income was used as an output variable in some of the earliest DEA studies of Higher Education (Ahn et al., 1989, Tomkins and Green, 1988) where it served "as a proxy for research output" (Ahn et al., 1989:172) and it has continued to be used in this way in many subsequent studies (e.g. Agasisti and Pérez-Esparrells, 2010, Flegg et al., 2004, McMillan and Chan, 2006). The rationale for this is clearly stated by Flegg et al.

(2004) who explains that "since universities sell their services to government and industry, the income received can be used to estimate the value of the output produced" (p 234). However, several authors express a desire for a better alternative and only "reluctantly resort" (McMillan and Chan, 2006:9) to using this as an output variable, acknowledging the objections set out by Johnes and Johnes (1993) who "strongly favour the view that grants are an input not an output" since they purchase "research assistance and other facilities which are an input into the production process" (p 338). Authors who follow Johnes and Johnes (1993) in this preference include Beasley (1990) and Athanassopoulos and Shale (1997).

Research quality ratings

Quality ratings are an alternative output measure favoured by those who classify research grants as inputs, including Beasley and Athanassopoulos. An advantage of the quality rating as a variable is that difficult decisions, such as which publications to include and how to classify them, have already been taken and are not therefore the responsibility of the researcher. This works well for studies based in the UK since the results of the Research Assessment Exercise are readily available and offer a basis for comparing institutions. In studies of Australian universities the "Research Quantum" can be seen performing a similar role e.g. Abbott and Doucouliagos (2003), Carrington et al. (2005). Although it is a research grant, it is based on a composite index of performance which includes measures such as higher degree completions and publications (Carrington et al., 2005:150). However, a policy change in 2002 means that this variable is no longer available (Department of Education Employment and Workplace Relations, 2011) and there are many countries where no such scheme exists, leaving the researcher to fall back on the previous two options of publications and grants.

None of the DEA studies examined include any variables relating to the third area of output in Figure 6-1, "Services for the community". This is an aspect of university life around which little data has so far been gathered, although recent work by Ursula Kelly et al (Kelly et al., 2008) has explored some possibilities in this area, e.g. events and performances open to the public, external usage of library and sports facilities and engagement with public policy.

Input Variables

Our discussion of output variables has already introduced some of the input variables used in existing studies. In general they can be summarised under three headings which overlap but do not quite match the headings proposed by Agasisti and Pérez-Esparrells (2010).

Staff

The most commonly used indicator of human resource input is the number of full-time equivalent academic staff. However, the qualifiers used in the preceding statement – "full-time equivalent" and "academic" – hint at the complexity of this group of variables. In Italy, for example, FTE is not recorded and so in some comparative studies a headcount of staff has been used (Agasisti and Johnes, 2009, Agasisti and Pérez-Esparrells, 2010). Sometimes non-academic staff are included as a separate variable (e.g. Abbott and Doucouliagos, 2003, Avkiran, 2001) but they may be restricted to academic-related staff (e.g. Flegg et al., 2004). The total number of academic staff may be subdivided into teaching and research (e.g. Johnes and Johnes, 1993, Johnes and Johnes, 1995, Madden et al., 1997) or limited to full-time staff only (e.g. Tajnikar and Debevec, 2008, Tomkins and Green, 1988). There is clearly a close relationship between human and financial resources, and some studies have used the salary bill as an alternative to counting the number of staff (e.g. Tomkins and Green, 1988, Ahn et al., 1989, Agasisti and Salerno, 2007).

Students

We have already reviewed the arguments over whether students are inputs to the HE system or outputs from it. In the many studies which treat students as inputs, that treatment is as diverse as for staff. A headcount may be used (e.g. Agasisti and Johnes, 2009) although FTE is more typical (e.g. Athanassopoulos and Shale, 1997, Flegg et al., 2004). It is quite likely to be subdivided into undergraduate and postgraduate (Athanassopoulos and Shale, 1997, Johnes, 2006),and postgraduate may be further subdivided into taught and research postgraduate (Flegg et al., 2004). Johnes (2006) applies a weighting to the undergraduate FTE based on A-level performance and in other studies (Johnes, 2003b, a), where the DMU is the individual student rather than the institution, she uses a range of qualifying variables such as mode of study, marital status and gender.

Finance

We have already noted that expenditure on staff salaries is sometimes used instead of a headcount/FTE of staff. Tomkins and Green (1988) experiment with several different models and obtain fairly consistent results when salaries are substituted for a headcount of full-time staff. However staff are accounted for, it is also common to see a financial input variable defined which covers everything *except* staff. For example, in addition to two variables for staff (academic and non-academic) Abbott and Doucouliagos (2003) define a third variable to include "expenditure on all other inputs other than labor inputs" (p 93). Arcelus and Coleman (1997), Flegg et al. (2004), Agasisti and Salerno (2007) and others also adopt this approach. Johnes (2006) subdivides non-staff expenditure further into spending on information services (libraries, computing) and on administration and central services as well as capital depreciation. Alternatively, some studies use total operating cost as their only input variable, e.g. Thanassoulis et al. (2011) and Carrington et al. (2005), in which case it includes staff salaries as well as other types of expenditure.

One problem in the handling of input data is exemplified by Agasisti's attempt to evaluate the overall efficiency of the HE sector in a number of European countries (2009). The model is established at national rather than institutional level so both inputs and outputs are population-based statistics taken from the OECD's *Education at a Glance* publications. Unfortunately, all of the selected data is expressed either as a ratio (e.g. staff-student ratio) or as a percentage (such as the percentage of the population who are tertiary graduates at the typical age of graduation) which means that it unsuitable for use in DEA (Hollingsworth and Smith, 2003) and so the resulting model is unsound.

This flaw is found in other studies as well. Both Breu and Raab (1994) and Turner (2005) are using their DEA studies to investigate league tables, and they make the mistake of depending on the ratios and percentages used in these tables for their inputs and outputs. Other studies also employ some ratio data, such as Arcelus and Coleman (1997) and Haksever and Muragishi (1998) as described above. In general, later studies are more likely to conform to the DEA requirement for non-ratio data.

6.2.2 Variables for modelling league tables

Turning to the variables which are used in the leagues tables themselves, we can see that the range of functions covered is similar, although, at least in UK-based tables, there is much more emphasis on teaching outputs than on research outputs. For this analysis we will focus on four main published league tables in their latest incarnation, as shown in Table 6-2.

Table 6-2: Major UK university league tables

League Table	Date of Publication	Date of Source Data
The Guardian University Guide 2013	May 2012	HESA 2010-11, 2009-10 NSS 2011
The Times Good University Guide 2013	May 2012	HESA 2010-11 NSS 2011 RAE 2008
The Sunday Times University League Table 2012	September 2011	HESA 2008-9, 2009-10 NSS 2011 RAE 2008
The Complete University Guide 2013	April 2012	HESA 2009-10, 2010-11 NSS 2011 RAE 2008

The first three of these are all published by major national newspapers. The fourth, the Complete University Guide (CUG), has been published independently since 2007 when its organisers parted company with the Times newspaper, but has been variously "in association with" the Daily Telegraph (2007 and 2011) and The Independent (2008, 2009 and 2010) (Mayfield University Consultants, 2012a).

Table 6-3 summarises the indicators used in these tables using the categories employed in 6.2.1 above. We can immediately see strong similarities between the tables, particularly between the Times and the CUG which have a common origin. Some key points of difference are

- The Sunday Times has the only compilers who collect their own data. Their table uses a survey of academics which asks them "to rate departments in their subject field on a five-point scale for the quality of their undergraduate provision". The indicator is compiled by calculating a mean score for each institution and coverting it to a 100-point scale.

- The Guardian explicitly omits any measures of research performance (Centre for Higher Education Research and Information et al., 2008b:37).
- The Guardian uses a measure of "Value Added" instead of counting the number of First and Upper Second class degrees awarded.
- The Sunday Times does not include any input measures relating to resources, while the other tables count staff and spending.

Table 6-3: Indicators used in current league tables

		Source	Sunday Times	Times	Guardian	CUG
Teaching Outputs	Rates of Completion/Dropout	HESA	X	X		X
	First & Upper Second class degrees awarded	HESA	X	X		X
	"Value Added"	HESA			X	
	Rates of Employment/ Unemployment	HESA	X	X	X	X
	Satisfaction scores from the National Student Survey	NSS	X	X	X	X
	Survey of academics on quality of undergraduate provision	Survey	X			
Research outputs	Results of the 2008 Research Assessment Exercise	RAE	X	X		X
Inputs	Spend per student on academic services & facilities	HESA		X	X	X
	Student : Staff ratio	HESA		X	X	X
	Students' qualifications on entry	HESA	X	X	X	X

Examining the content of these indicators we find a mix of different data types.

Ratios and Percentages

Many of the measures in Table 6-3 involve defining a category, such as “graduates awarded a First or a 2:i” or “staff with teaching responsibilities”, counting the number of individuals in the category and then expressing it as a percentage (typically of the total number of students) or as a ratio (typically per individual student).

These measures are derived from the HESA records for each institution which are a comprehensive source for this type of information, although defining the required categories and assigning individuals to them is not necessarily straightforward. For example, the staff record will state whether a staff member is employed to do “teaching”, “research” or “teaching and research” but in the latter case will not specify the amount of time allotted to each area of work. Thus the league table compiler has to decide how to count these individuals: should each “teaching and research” staff member count as one teacher or as half of one teacher?

Scores

Measures of student satisfaction and research quality are derived from scores obtained from the National Student Survey (NSS) and the Research Assessment Exercise (RAE).

The NSS consists of 23 statements on eight themes. Participants respond to each statement with a response on a five point scale from “strongly agree” to “strongly disagree”. Although the formal presentation of the results is based on the percentage who “agree” or “strongly agree” with each statement (HEFCE, 2012d), in the league table context these responses are typically ascribed a numerical value, e.g. 5 for “strongly agree” and an overall score calculated. An average may be calculated over all or most of the questions (Mayfield University Consultants, 2012b) or for specific themes (IntelligentMetrix, 2012).

The formal results of the RAE are expressed as “quality profiles”: the percentage of submitted work which is found to be at each of four levels, from 1* (recognised nationally) to 4* (world-leading) (HEFCE et al., 2008). The approach taken by league table compilers to this scoring system varies. The Times gives a weight of 3 to the percentage of papers rated 4* and a weight of 1 to the percentage rated 3* to calculate an overall score; 2* and 1* are unweighted (Times Newspapers Ltd, 2012). The Complete University Guide weights all four categories, giving 4* a weight of 4, 3* a weight of 3 and so on (Mayfield University Consultants, 2012b) and the Sunday Times does the same (Times Newspapers Ltd, 2011).

The compilers then adjust the scores depending on the number of staff submitted for assessment by the institution concerned.

Performance Indicators

The measure of completion (or of dropout, in the case of the Sunday Times) is much more complex than the other measures used. This indicator is taken directly from the Performance Indicators published annually by HESA on behalf of the funding councils (HESA, 2012a).

The Times and the Complete University Guide use the “Projected Outcomes” indicator, which is based on a Markov chain analysis of linked student records. Future outcomes for the current cohort of students are estimated based on current progression patterns at their institution (HESA, 2012c). The Sunday Times uses the “Non-continuation” indicator, which is a somewhat simpler measure as it based only on historical data. The calculation looks at every student who enters an institution in a given year and then assesses where they are in the following year (HESA, 2012b).

The Guardian’s “Value Added” indicator is not a formal performance indicator, but takes advantage of HESA’s system of linking records in order to connect a student’s qualifications on entry to their final degree result.

Correlated Variables

The CHERI report (Centre for Higher Education Research and Information et al., 2008b) examined the relationship between variables in some detail and noted that several are highly correlated. The most striking example, also noted by Smith and Naylor (2001), is that degree outcomes are very highly correlated with students’ qualifications on entry, i.e. those institutions which admit students with a stronger academic record also award more first class and upper second class degrees.

In DEA we generally wish to avoid highly correlated factors, in order to increase discrimination between DMUs, and in section 3.7.1 we considered some of the technical tests for selecting appropriate variables. Dyson et al. (2001) warn that in some cases the omission of a highly correlated variable may lead to “significant changes in efficiencies” (p 249) since in general DEA models are not translation invariant.

6.2.3 Conclusions on variable selection

We noted in section 6.1 that inputs and outputs are treated differently in DEA models and in league tables. Nonetheless, our survey of the data used shows that very similar raw material is used in the variables of each system, bearing in mind the league table emphasis on teaching. Another issue that has come to light in this survey is that some league table variables are very complicated indicators. In a DEA model we aim to keep the variables as simple as possible and to avoid ratios, percentages or other compound constructions which may lead us to erroneous results. Finally, we wish to avoid highly correlated variables, but without jeopardising the validity of the model. Remembering that this model is not in fact intended to measure efficiency, we will take a pragmatic approach to variable selection and, where there is a high degree of correlation, choose the simplest variable available to represent that component of the problem.

6.2.4 Final choice of variables

Reviewing the contents of Table 6-3 we can identify suitable raw material for variables in a DEA model which reflects the league table environment.

Teaching Outputs

Rates of Completion/Dropout (HESA)

The published performance indicators are too complex to be used directly in a DEA model. It would be very ambitious to attempt to gather the data in a more suitable format, so for the time being we set this variable aside. It is worth noting that the CHERI report found this variable to be highly correlated with the Degrees Awarded variable (Centre for Higher Education Research and Information et al., 2008b:71).

Conclusion: **Exclude**.

First & Upper Second class degrees awarded (HESA)

A basic measure of teaching output which can readily be obtained. The CHERI report found that this was very highly correlated with a number of other variables, so including this measure allows us to exclude others which are more controversial.

Conclusion: **Include**.

“Value Added” (HESA)

This is another complex measure and it would be very ambitious to attempt to reproduce the Guardian’s calculations.

Conclusion: **Exclude**.

Rates of Employment/Unemployment (HESA)

Another common measure of teaching output which can be expressed as a count rather than a percentage. However, as data is not available for every student, some extrapolation may be required.

Conclusion: **Include**.

Satisfaction scores from the National Student Survey (NSS)

A measure of teaching output which can readily be obtained. It is more appropriate to follow the lead of the formal results reporting and use a count of the satisfied students than to employ an average score.

Conclusion: **Include**.

Survey of academics on quality of undergraduate provision (Survey)

As a survey of university reputation, this variable could be argued to be as much an outcome of the league table process as an input to it. It was also found to be very highly correlated with other variables, particularly Entry Standards, Degrees Awarded and RAE Results (Centre for Higher Education Research and Information et al., 2008b:64).

Conclusion: **Exclude**.

Research outputs

Results of the 2008 Research Assessment Exercise (RAE)

It is important to include some measure of research output and the results of the RAE are widely used for this purpose. However, it is not a straightforward indicator and it would be appropriate to experiment with different formats for this variable.

Conclusion: **Include**.

Inputs

Spend per student on academic services & facilities (HESA)

We saw in section 6.2.1 that it is typical of DEA modelling to include some measure of expenditure on items other than staff salaries. Although this would need to be expressed as a total sum rather than "per student", the data is very accessible and a good candidate for inclusion here. One aspect of an institution's resources which is not accounted for here, however, is accumulated wealth, e.g. in the form of endowments or facilities, which might be a significant factor in giving older universities an advantage over newer ones. If it is feasible, then this would be a useful alternative or additional variable.

Conclusion: **Include**.

Student : Staff ratio (HESA)

Although the ratio format is inappropriate, both parts of this indicator are basic measures in establishing the operating parameters of an institution. The data for both staff and student variables are readily available, but, as discussed above, defining the scope is not necessarily straightforward. In particular, decisions need to be made on whether a headcount or FTE is appropriate and some experimentation with the data will be required to inform these decisions. However, given that FTE is used by league table compilers to arrive at this ratio, this would be an appropriate starting point.

Conclusion: **Include**.

Students' qualifications on entry (HESA)

This is a controversial indicator as it is available only for those students who have the traditional qualifications of A-levels or Highers/Advanced Highers. It has also been found to be very highly correlated with Degrees Awarded (Centre for Higher Education Research and Information et al., 2008b:64,71).

Conclusion: **Exclude**.

We have thus established a basic set of variables for a DEA model of the league table environment which includes three inputs – Staff, Students and Expenditure – and four outputs – Degrees Awarded, Employment, Student Satisfaction and RAE scores. This model will be implemented and fine-tuned in Chapter 7.

6.3 Selection of DMUs

To determine a selection policy for DMUs we first need to decide at what level to analyse the data: university or subject. Since we are concerned about homogeneity, an institution-level model may pose problems. Some subjects, such as medicine and experimental sciences, are clearly more expensive than others, and an institution's overall profile will reflect its subject mix. However, modelling at subject level is by no means as straightforward as it might initially appear, because of the way in which this data is reported and collected. It is worth considering this aspect in some detail, as it is one of the major sources of tension between league table compilers and the HE sector.

6.3.1 Subject classification systems

Subject classification in HESA data is based on two systems of coding. First, each student's qualification aim (or qualification achieved, in the case of graduates) is recorded using the Joint Academic Coding System (JACS). This system of codes is used throughout the life cycle of a student, from the point where they make their application through UCAS until they graduate. The "degrees awarded" measure described in section 6.2.2 above is based on this code and the National Student Survey results are also grouped by JACS. The version currently in use is JACS 2.0, which was introduced in 2007 (HESA, 2006); a review has recently been completed and JACS 3.0 will be implemented from 2012/13 (HESA, 2011c). Each JACS code is a string in the form *LNNN*, where *L* is a letter from A to Z and *N* is a digit from 0 to 9. The codes are structured so that each subject can be described at several levels of detail. For example, Table 6-4 shows some of the codes found within the letter H, which denotes Engineering.

When this data is compared across universities we can identify students who are studying similar subjects, but it tells us nothing about who is teaching them. For example, a student with a qualification aim of V110 (Ancient History) might be studying in a Department of History or a Department of Classics.

The second system consists of "cost centres" defined by HESA and was originally intended to capture financial information: "the definition of a cost centre was driven by the notion of grouping together of activities with a similar cost" (HESA, 2011b). The current list of codes was introduced in 2004 and consists of 35 academic subject centres and a further six administrative cost centres (HESA, 2004).

Table 6-4: Examples of nested JACS codes within the Engineering category (letter H)

Code	Subject	Description
H200	Civil Engineering	The study of the principles of engineering as they apply to the designing and construction of public works, eg buildings, bridges, pipelines etc. Involves the study and application of specialist mathematics.
H210	Structural Engineering	The study of the principles of engineering as they apply to the design and construction of physical shapes and forms. Involves the study and application of specialist mathematics.
H300	Mechanical Engineering	The study of the principles of engineering as they apply to the design, development manufacture and operation of machinery.
H330	Automotive Engineering	The study of mechanical self propulsion in vehicles.
H331	Road Vehicle Engineering	The study of mechanical self propulsion in road vehicles.

Table 6-5 below illustrates the codes corresponding to engineering subjects. Where numbers are "missing" from the sequence, this indicates that a reorganisation of the code list has taken place.

Table 6-5: Examples of cost centre codes for engineering and related subjects

Engineering & technology
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 25 IT & systems sciences, computer software engineering

Every student record and every staff record which is returned to HESA is associated with one or more cost centres based on the proportions of teaching and/or research activities in which they have participated during the year. For students this division across cost centres may take place at the level of the individual module studied and thus be quite different from their stated qualification aim. And we are still none the wiser about where our student of Ancient History is taught, unless they are in a Department of Archaeology (cost centre 37), since History and Classics are both subsumed under the extremely broad category of "Humanities & language based studies" (cost centre 31).

In English institutions the cost centre codes are highly significant as HEFCE uses them in its allocation of funds for teaching. However, in Scottish institutions cost centre codes are simply an artefact of the reporting system: the SFC uses a separate mechanism for the allocation of funds and its audit of this mechanism is based on the JACS codes described above.

When data is retrieved from HESA's databases, students and staff can be grouped according to cost centre codes. League table compilers make use of this feature in order to calculate the student:staff ratios described in section 6.2.2 above. It is important to note that this use was not intended when the system was designed, and that this "by-product" has been the cause of much anxiety in the sector. This system has also been the subject of a review and a revised set of cost centres will be implemented in 2012/13 (HESA, 2011b). The new coding structure is also intended to provide a better match to a third set of subject codes: the Units of Assessment used in the RAE/REF.

Neither of these systems necessarily reflects the university's departmental structures and it is up to the individual institution to determine the mapping between them. HEFCE regularly asks universities in England to submit information showing how their departments are mapped onto cost centres; the results show that there is enormous variation within the system. HEFCE (2009-10) reports on the number and percentage of institutions "mapping subjects in the sector norm cost centre" (HEFCE, 2012a). The figures range from 100% in dentistry to 22% in "other technologies" (which includes subjects such as ergonomics, transport logistics and music recording).

A further level of complexity is introduced when league table compilers attempt to combine JACS coded data with data based on cost centres. They typically map *both* sets of codes onto a subject classification of their own devising. By the end of this process it can be very unclear to a university which of their departments is actually being represented in any given part of a league table. The Guardian was heavily criticised for their handling of subject data, which has led to an annual "consultation" with institutions. In this process, each institution is asked to review the default mappings proposed by the newspaper and "to make sensible adjustments … to reflect their particular course or departmental structures" (Guardian, 2011).

6.3.2 Subject classification in practice

JACS and Cost Centres

Over the 146 subjects reported by HEFCE (2009-10), the mean percentage of institutions who map to the sector norm is 67% and the median 70%. In order to test the feasibility of modelling at subject level, data was requested from HESA for a subject that sits in the middle of the distribution: chemistry. The JACS code for chemistry is F100 and it contains 19 sub-groups, such as organic chemistry and inorganic chemistry as shown in Table 6-6. In the HEFCE analysis this entire group of codes is referred to as F1 and 68% of institutions reporting this subject mapped it to the sector norm cost centre, 11. No other subjects have cost centre 11 as the sector norm.

Three datasets were obtained:

1. Staff FTE for cost centre 11 "Chemistry" in 2009-10
2. Student FTE for cost centre 11 "Chemistry" in 2009-10
3. Student FPE for JACS code F1 "Chemistry" in 2009-10.

FPE stands for Full-Person Equivalent. Each student, whether full- or part-time, is counted as one FPE. This unit is then divided across the component parts of the student's qualification aim. If the student's degree is a single subject course in chemistry then 1 FPE is added to the total for code F1, if it is a joint course in chemistry and one other subject then 0.5 FPE is added to code F1.

Combining the datasets, we find that 85 UK institutions reported data in at least one of these three categories. 61 reported staff FTE in cost centre 11; 63 reported student FTE in cost centre 11; and 78 reported student FPE in JACS code F1.

Table 6-6: JACS codes for Chemistry

Code	Subject	Description
F100	Chemistry	The study of individual atoms and molecules and the way they react together naturally and synthetically.
F110	Applied chemistry	Topics in chemistry of commercial or social importance.
F111	Industrial chemistry	The study of chemical processes of industrial significance.
F112	Colour chemistry	The chemical science of dyes and pigments.
F120	Inorganic chemistry	The study of inorganic elements, compounds and reaction mechanisms.
F130	Structural chemistry	Determination and analysis of chemical structures.
F131	Crystallography	The study and application of techniques for determining crystal structure.
F140	Environmental chemistry	Concerned with environmental issues related to the chemical sciences.
F141	Marine chemistry	Topics in the chemical sciences concerned with understanding the marine environment.
F150	Medicinal chemistry	Aspects of Chemistry, such as drug design, of importance to medical science.
F151	Pharmaceutical chemistry	The study of drug function.
F160	Organic chemistry	The study of organic compounds and their reaction mechanisms.
F161	Organometallic chemistry	The study of reactions between organic compounds and metals.
F162	Polymer chemistry	The study of the properties of macromolecular compounds and their synthesis.
F163	Bio-organic chemistry	The study of natural organic compounds.
F164	Petrochemical chemistry	The chemical science of petroleum and petroleum compounds.
F165	Biomolecular chemistry	The chemical science of biological materials at the molecular level.
F170	Physical chemistry	The study of atomic and molecular structure, chemical bonding, energetics and dynamics.
F180	Analytical chemistry	The study of chemical and instrumental analysis.
F190	Chemistry not elsewhere classified	Miscellaneous grouping for related subjects which do not fit into the other Chemistry categories. To be used sparingly.

We can investigate the relationship between student FTE and student FPE by considering the ratio of FTE to FPE. We would not expect this to be precisely equal to one, since the dataset is likely to include part-time students, students studying some chemistry modules as part of another degree and students of chemistry taking modules in other subjects. In fact, the values were found to range from less than one to more than 400. A description and analysis of the findings is presented in Table 6-7.

Table 6-7: Analysis of student FTE : FPE ratio in institutions reporting on chemistry

	No. of HEIs	Description of the Data	Analysis
a	5	0.5 < FTE : FPE ratio < 0.8	This group includes institutions which primarily serve part-time students. The others in the group are post-1992 institutions which may also have part-time students or students doing industrial placements.
b	9	0.8 <= FTE : FPE ratio < 1.0	This is a feasible range for institutions which have some part-time students or chemistry students undertaking some study in other departments.
c	21	1.0 <= FTE : FPE ratio < 1.2	This is a feasible range for institutions where the chemistry department is teaching a number of students for whom chemistry is not a priniciple subject.
d	9	1.2 <= FTE : FPE ratio < 2	The number of non-chemistry students taught by the chemistry department increases. At the top end of this group the FTE is nearly 5 times the FPE. It seems likely that the institutions are assigning several additional JACS codes to cost centre 11, e.g. pharmacology or forensic science.
e	12	2 <= FTE : FPE ratio	Some institutions have very high ratios. They have very few students enrolled under subject F1 and appear to be using cost centre 11 either to report subjects other than chemistry, or to report activity in chemistry for students enrolled on general science courses.
f	10	FTE = 0, FPE > 10	In these institutions the chemistry cost centre is not used at all, although students are registered with a principal subject of chemistry. Teaching is likely to be captured in another cost centre, possibly pharmacology (08) or chemical engineering (17).
g	12	FTE = 0, FPE < 10	As above but FPE numbers are very small, typically 1 or 2. Many of these institutions were formerly teacher training colleges so they may be reporting mainly under cost centre 34 (Education).
h	2	FTE < 10, FPE = 0	Two institutions report a very small FTE of students, matched by an even smaller FTE of staff, but no FPE. These students are likely to be enrolled on a non-chemistry degree course mainly taught in another cost centre.

	No. of HEIs	Description of the Data	Analysis
i	2	FTE > 10, FPE = 0	These institutions report a substantial student FTE (over 150 in both cases) along with an appropriate staff FTE, but no FPE. As in the high ratio group (e), these institutions may be using the chemistry cost centre to record students studying other subjects, e.g. forensic science.
j	3	Student FTE > 0, Staff FTE = 0	There are two institutions which report a high student FTE but no staff FTE in cost centre 11. Given the way in which FTE data is used in league tables, this is now unusual. However, over the three year period 2008-09 to 2010-11, both institutions show considerable change in the numbers reported which suggests that restructuring has been taking place.

In order to use both sets of data – cost centre FTE for staff and student numbers and JACS coded FPE data for degrees awarded – without unreasonably distorting the model, we would need to limit ourselves to those institutions where the ratio is close to one, i.e. groups b, c and d in Table 6-7. This would restrict us to a population of 39 DMUs out of our original pool of 85. 39 is an adequate number for the technical requirements of a DEA model, but we would need to bear in mind that it is far from giving the whole picture of activity in this subject area.

RAE Units of Assessment

We now turn to the subject classification used in the Research Assessment Exercise. The last such exercise was conducted in 2007 and the results published in December 2008. Those results have formed the backbone of research funding formulae since that date with only minor updates made to account for e.g. variation in the number of PhD students and research assistants. In this period there has also been much debate about the shape of future assessments (HEFCE, 2012b), but the principle of grouping research activity into Units of Assessment remains unchanged even though the number of these units is planned to be reduced from 67 to 36.

Those wishing to examine the data from the 2008 RAE, including those who were involved in conducting the assessment (HEFCE et al., 2005:9), have needed to match the categories used here to those used in HESA datasets and other reports. No ‘standard’ or ‘recommended’ mapping has been published, even where the data has been used for statutory reporting (e.g. HESA, 2012d), leaving those who want to use the data to devise their own methods (e.g. Targeting Innovation, 2009).

Our example subject of chemistry, from section 6.3.2 above, is easily extracted from the RAE dataset as it is represented by a single Unit of Assessment, “Chemistry” (number 18). Comparing Table 6-7 with data on RAE submissions we find that 32 of the 33 institutions which made a submission in Unit of Assessment 18 are in groups b, c and d, i.e. those we have already identified as most suitable for use in the DEA model. Of the nine institutions in group b, eight made an RAE submission in chemistry. Of the 21 in group c, 19 made a submission, and of the 9 in group d, 5 made a submission. This distribution offers some additional support for our earlier analysis.

6.3.3 Conclusions on DMU selection

We have found that DMU selection is far from straightforward in the higher education context. Institutions are diverse and have a high degree of autonomy so that any homogeneity is at best superficial. Probing the data very quickly reveals a range of different activities and behaviours even in a relatively clearcut subject area. However, we have taken sensible steps to constrain our population of DMUs to those which are most comparable and can be confident that we do have the basis for a suitable DEA dataset, even if it does not give a comprehensive picture of the sector.

6.4 Other issues

The final component of the model is the set of trade-off vectors. As discussed in Chapters 3 and 5 above, the need for trade-offs arises from the need to constrain the DEA model so that certain inputs and outputs are not 'ignored' by being assigned zero weights. The use of trade-offs reflects managerial judgment about the relationships between different factors in the model, where one factor might feasibly be exchanged for another. For instance, suppose we have a system where there are two possible outputs and output A costs three times as much to produce as output B. Given a fixed level of input, then, one unit of output A could be exchanged for three units of output B: we can express this relationship using trade-off vectors. In general, how are such trade-offs to be specified?

Some of the motivations for introducing trade-offs (or weight restrictions) were discussed above in section 3.5.4. In practical terms these motivations may arise in two different ways: either from observation of the results of an unrestricted model and the conclusion that they are in some way unsatisfactory or from prior assumptions about the problem space and the relationship between variables. In the first case, Podinovski offers some guidance (Podinovski, 2007b) on interpreting the optimal weights from a standard CRS model and using them to inform the development of trade-off vectors. The second case is perhaps more challenging. As with the other aspects of constructing a model in DEA, there are many individual small judgements to be made, each potentially presenting "traps for the unwary" (Dyson et al., 2001:257).

It is important to note that the purpose in using trade-offs is to make the production space more realistic, not to incorporate judgements about the relative value of different activities (Podinovski, 2007b). Because they are applied to the whole dataset they must necessarily be fairly conservative. In some contexts, such as modelling departments within an organisation, it would be appropriate to discuss and agree what trade-offs can realistically be applied. In the present context this is not feasible, so only the most cautious statements can be made.

The simplest approach is to consider the variables pair-wise and ask whether there is any production relationship between them. This might be a relationship between two outputs, as in the illustration in section 5.3.2 where research output is related to teaching output. However, one can also consider relationships between two inputs or between an input and an output. For example, we might say that adding one member of teaching staff (an input variable) to a department should increase the number of graduates (an output variable) by at least 10.

By working methodically through the pairs of variables in this way, we may identify relationships which in fact involve more than two variables. For example, if we replaced one member of teaching staff (an input variable) with a researcher (another input variable) we would expect there to be an impact on the outputs produced by the department.

It is not necessary to specify every possible production trade-off and care should be taken, if multiple trade-offs are specified, that we do not inadvertently introduce inconsistencies which may lead to problems (Podinovski and Bouzdine-Chameeva, 2013).

In the previous sections of this chapter we have identified a number of factors necessary for a model of UK Higher Education. In the following sections we briefly review some possible sources of information about trade-off relationships between them.

6.4.1 Cost-based trade-offs

One option is to evaluate the relative costs of different activities. This could be approached from the top down, by considering the rates at which teaching and research activities are funded (e.g. HEFCE (2009), SFC (2009)), or from the bottom up, by looking at how monies are actually spent. A key strength of this approach is that it is highly relevant to management decisions about the use of financial resources to improve performance. However, a weakness is that there is a considerable difference between the top-down and bottom-up approaches, as we can see in the analysis of 2010-11 data published by HEFCE (2012c). "Full Economic Cost recovery" can be seen to be approximately 100% for publicly funded teaching activity. However, research funding falls short of paying for all research activity, with only 78% of costs covered by income. This shortfall is recouped chiefly from teaching activity which is not publicly funded, i.e. mainly international students. The gap between these two views of higher education costs means that establishing realistic trade-off relationships on this basis would be very challenging.

6.4.2 Time-based trade-offs

A related option is to analyse the time costs of different activities rather than the financial costs. A potentially fruitful source of data for this purpose is the academic workload model. This is a model used in universities, usually at school or department level, to distribute teaching, research and administrative tasks across the academic staff. Such a model typically takes the form of a spreadsheet which converts inputs, such as the number of preparation and contact hours involved in teaching a module to a certain number of students, into standard units by means of "agreed coefficients" (Barrett and Barrett, 2009:15). Validation formulae in the spreadsheet will flag up where staff are significantly under- or over-loaded. Workload models are based on a standard annual workload of 1650 hours, which is the figure used by Research Councils UK (2012) and HEFCE (KPMG, 2012) when funding allocations are made. This option has the same key strength as the one above, namely its relevance to management decision-making. However, the division of this workload into hours designated for different activities varies across universities, across subjects and across individual members of staff. This again poses a challenge for the establishment of realistic trade-offs.

6.4.3 Goal-based trade-offs

It is also worth considering whether a trade-off relationship needs to be strictly realistic to be of value in this context. The approaches summarised above all provide information about the trade-offs involved in current practice. However, as we have already noted (in section 3.7.2 above), one of the characteristics of DEA is that it tends to be conservative, since the space it constructs is bounded by current performance. Indeed, the concern that this might constrain effort − since a typical evaluation on DEA principles does not reward "over-efficiency" − has led to a number of suggested modifications of the basic DEA model (e.g. Andersen and Petersen (1993), Bogetoft (1995)). One could argue that it would be better to allow the model to reflect different ways of managing universities through an appropriate choice of trade-offs. For example, one approach might be to start from a possible goal, such as a certain level of performance in the forthcoming REF, and establish what trade-offs an institution would consider making in pursuit of that goal.

The strength of this option is that the model would not be constrained by conservative measures based on current practice. A key weakness, however, is that it would be difficult to formulate relevant, plausible and achievable trade-offs. There is a risk of going too far and positing over-aspirational relationships between inputs and outputs. It is also important to remember that the trade-offs are applied to the whole population of DMUs. While they need not be exact, and can be expressed in an asymmetric form to permit a range of values, the goals of a particular institution may not extrapolate well to the whole dataset. Podinovski is very clear that if a trade-off violates production realities, then the model is jeopardised (2007:1269).

Any of the approaches outlined above has potential. There are strengths and weaknesses to each one, but the significance of these will vary depending on the context of the problem. For the current model, we have chosen to work with highly simplified time-based trade-offs and will develop an example of this more fully in Chapter 7 below. This approach uses known management information in a realistic manner, but also has the potential to be extended to model alternative ways of working which have not yet been implemented.

6.5 Summary

In this chapter we have worked through the task of operationalising the proposed model.

We have considered the existing literature which applies DEA to the higher education context and examined the types of variables which are typically used. We have also looked at the specifications of some of the major UK league tables. This has led to us a detailed evaluation of the advantages and disadvantages of different potential variables and thence to the establishment of a basic set of variables suitable for our purpose.

An equally important step is the evaluation of potential DMUs, and we have conducted this in the light of a detailed analysis of subject classification and data structures. Finally, we have considered some of the options for defining trade-offs in this context.

Having determined the outline structure of our dataset, in Chapter 7 we will populate it with appropriate data and execute the model.

7 Findings and Discussion

In this chapter we take the data specification developed in chapter 6 and apply it to an illustrative example. Firstly, the specification is considered in more detail as the dataset is refined. Then the framework outlined in chapter 5 is considered step by step. Additional parameters are defined for the DEA model, the dataset is then processed using the DEA model and, finally, the results are viewed through the lens of game theoretic reasoning.

7.1 Establishing a suitable dataset

In section 6.2 we identified a set of variables appropriate for modelling the university league table environment using DEA. We identified three inputs – staff, students and expenditure – and four outputs – degrees awarded, employment destinations, student satisfaction and RAE scores.

We explored the issues connected with analysis at subject level versus university level, and discovered that even with a relatively well-defined subject, namely chemistry, it would be necessary to restrict our dataset to a subset of institutions where the reporting of the subject provision falls within certain parameters (section 6.3.2).

We also noted the difficulties of comparing institutions across national boundaries withint the UK, but concluded that it was important to attempt to do so given the league table context.

7.1.1 Data Sources

In order to populate these variables it is necessary to obtain data from three sources: HESA, the Research Assessment Exercise (RAE) and the National Student Survey (NSS). At the point when this dataset was compiled, the latest date for which all the necessary HESA returns had been collected and published was the academic year 2009/10. This year was therefore selected for study.

The NSS dataset is available for downloading from the website of the Higher Education Funding Council for England (http://hefce.ac.uk) in a series of Excel spreadsheets. These spreadsheets contain a breakdown of the responses to each survey question by institution and subject. The survey for 2009/10 was conducted in early 2010 and is published as the “2010 National Student Survey” (HEFCE, 2010).

The last RAE was conducted at the end of 2007 and the results published in 2008. This data has been used in league tables every year since that date and will continue to be used until the results of the first Research Excellence Framework (REF) are published at the end of 2014 (REF, 2012). The RAE dataset is also available for download (http://rae.ac.uk). The submission files for all institutions and Units of Assessment, as well as the results file, are provided in csv format as the number of individual records exceeds the limit for Excel. In order to access the data in a more user-friendly format, the csv files were imported into a relational database in Filemaker Pro. The data could then be summarised as required and the relevant information exported to Excel.

In order to obtain the HESA data required, a request was made to the HESA Information Provision Service. As discussed in chapter 6, the HESA datasets are complex and apparently simple variables such as "number of staff" are often challenging to define in practice. The complexity also has cost implications. Our initial request was discussed in detail with an information analyst at HESA and gradually refined to give the final specification set out in Appendix B. In summary, this specification covers three items:

1 Staff FTE by institution, broken down by academic employment function

2 Student FTE and FPE by institution, with the FPE data further broken down by level of study, degree awarded and destination after graduation

3 Financial summary by institution

Within each item the data relating to chemistry (cost centre 11, JACS code F1) was separately identified. This data was supplied in the form of Excel pivot tables.

Once these datasets were in hand, the variables could be developed in more detail.

7.1.2 Inputs

Staff

In order to enable trade-offs to be made, it was decided to separate research FTE from teaching FTE. Non-academic FTE was ignored for the purposes of this example. The relevant data was thus taken from three categories: teaching only, research only and teaching and research. Data in the last category was split so that half of the FTE was assigned to teaching and half to research. This approach is clearly a simplification for two reasons: firstly, it does not account for any time spent by an academic on other tasks, such as administration or outreach, and secondly, the proportion of time spent by academic staff on teaching or research varies greatly by institution. We will consider this second point in more detail below, but at this stage simply note that to attempt a more accurate breakdown of academic employment function is beyond the scope of this research. We therefore proceeded with two staff variables:

Research FTE = 1 x Research only FTE + 0.5 x Teaching and Research FTE

Teaching FTE = 1 x Teaching only FTE + 0.5 x Teaching and Research FTE

Students

In the first place, the choice of data for this input variable is made more complex by the alternatives of FTE and FPE. Using FTE would potentially provide a better match for the corresponding staff input, while using FPE would match more closely to the student-related output variables on degrees awarded and employment destinations as these are only available by FPE. Noting that we have already constrained our set of DMUs based on the ratio between FTE and FPE (see section 6.3.2), we will proceed using FPE in this illustration.

A second complication arises when we consider the league table context, which is focused on the undergraduate experience. The output variables are mainly concerned with the end results of this experience, so one option would be to restrict the input variable to the students who are most directly concerned in those outputs, i.e. those in their final year of study for a first degree. However, this would not take into account any students who had dropped out of their course or proceeded at a different rate. Nor is it a straightforward matter to 'turn the clock back' and select instead the first year students from a set number of years ago. We have already noted that the typical duration of degree courses in England is shorter than in Scotland, but in science subjects there is even more scope for variability with students having the option of an integrated masters course, such as an MSci or an MChem, which generally lasts one year longer than a BSc. Using the total number of current first degree students is a simpler alternative, but, given the variation in course lengths, tends to hide the size of the cohort.

With the aim of keeping the model as simple as possible, the variable adopted here is therefore the number of students enrolled in the first year of a first degree programme in the current year, which is taken from the HESA standard registration population.

Expenditure

Within each cost centre, this is broken down into staff costs, other operating costs and depreciation. Since staff have already been included above, we will proceed with one variable encompassing the other operating costs (expressed in thousands of pounds).

7.1.3 Outputs

RAE Scores

In order to represent RAE performance we consider the 'research outputs' produced at each institution. The RAE database allows us to count the total number of research outputs included in each submission, but does not tell us which individual outputs were awarded four stars, three stars and so on. This assessment is given as a percentage of the institution's total submission, rounded to the nearest 5%. For this illustration we will apply these percentages to the number of research outputs in order to estimate how many outputs were rated at each level. For example, an institution with the profile shown in line one of Table 7-1, would be estimated to have produced the number of outputs shown in line two of the table.

Table 7-1: Example of a research profile used to estimate the number of outputs produced at each level

	4*	3*	2*	1*	Unclassified
Research profile (expressed as a percentage)	20	50	25	5	0
Estimated number of outputs (out of a total of 152 outputs)	30	76	38	8	0

As we saw in section 6.2.2, in a typical league table context research outputs rated 3* and 4* are given much greater weight than others. These highly rated outputs are therefore the most desirable. To capture this in our illustration we have a number of different alternatives.

One option is to construct a weighted indicator similar to those used in some league tables. There are two main disadvantages to this approach. One is the selection of the weights, which vary from league table to league table (and also from league table to funding council): any weights chosen are necessarily arbitrary. Given that we make such a selection, a further disadvantage of this indicator is that there is not a unique interpretation of the score. If the value of a 4* research output is considered to be, say, 4 times the value of a 1* output, then a score of 20 could be interpreted as five 4* outputs, 20 1* outputs or some intermediate combination of these and of 2* and 3* outputs.

Another option is to disregard the outputs rated 2* and lower. This is perhaps the most straightforward and easily interpreted alternative: we focus our attention on 'desirable' 3* and 4* outputs and seek to maximise this total. However, in doing so we would disregard a potentially significant proportion of the research workload which is not captured elsewhere in the model.

A third option is to divide the research outputs into two separate variables: the 3* and 4* outputs grouped together as one variable and the 'undesirable' outputs, rated 2* and below, grouped into a second variable. Through the weighted preference structure we can ensure that highly rated outputs are given more weight than the low rated outputs and therefore prioritised in the process of setting targets.

However, there is a potential drawback to this option. If we have two institutions which are identical in their performance, except that one produces more low rated research outputs than the other, then the impetus of the DEA process is to identify slacks in the one with the smaller quantity of research output. The target resulting from such an analysis would be that the insititution should increase its low rated outputs. In the league table context, though, this would not be beneficial as it would decrease the *proportion* of that institution's research which is rated 3* and 4*.

Given that, in this context, outputs rated 2* and below can be seen as 'undesirable'we have a further set of options, described by Dyson et al (2001), which include inverting the variable, subtracting it from a large constant or moving the variable to the 'opposite side' of the model. The last of these might be considered the simplest as it does not alter the scale of the data. We can treat the highly rated outputs as an output variable: they will be subject to maximisation under the model, i.e. the target will be to produce more of these outputs. At the same time, we can treat the low rated outputs as an input variable: they will be subject to minimisation under the model, i.e. the target will be to produce fewer of these outputs.

Again, there is a potential drawback to this option, which relates to interpretation of the model. If we consider scaling up from a small institution to a large one, the logic is that a greater quantity of inputs will lead to a greater quantity of outputs. Does increasing the number of low-rated research outputs therefore lead to a larger number of high-rated research outputs? It is very unlikely. However, it is important to remember that DEA has been adopted in this study because the units under consideration have multiple inputs and multiple outputs and it cannot be assumed that there is a direct causal link between any individual input-output pair. If an institution were suddenly to be doubled in size, with twice as many staff and students as before, we would expect there to be a greater quantity of research output. However, with no other information about this transformation, we would expect the proportions of high-rated versus low-rated outputs to remain the same, i.e. that both the input and the output variables would be increased to scale.

It is clear that this is a complex situation with no ideal solution. For the purposes of this illustration, we will adopt the fourth option. It is appropriate to the league table context to consider research outputs rated 3* and 4* to be desirable outputs and outputs rated 1*, 2* or unclassified to be undesirable outputs. We wish to minimise the latter while maximising the former. We also need to scale these variables to account for the fact that research assessment is conducted approximately every five years, while all the other activities we are incorporating into the model are evaluated annually. For the sake of simplicity we simply divide the total by five to arrive at an approximate annual rate of production of research outputs.

Degrees awarded

The issues to consider in populating this variable have some similarity with those discussed above. We have noted (in section 6.2.1 above) that, in the league table context, only those graduates awarded first class or upper second class honours degrees are counted: any award below that level is ignored. This clearly falls short of a complete assessment of the teaching output of a university.

However, unlike the RAE case, our model does already incorporate some measure of the total load by including student FPE as an input. We will therefore stay close to the league table model by adopting the total number of first class and uppper second class honours degrees awarded in a given year as an output variable. This figure is readily obtained from the student FPE dataset supplied by HESA. It should be noted that the number is taken from the HESA qualifications obtained population (HESA, 2012e) and therefore includes any qualifications awarded to students who were dormant, writing up or on sabbatical in the year in question. In the case of undergraduates, the number of students graduating in this way is typically vanishingly small.

For other purposes, this exclusion of part of the graduating population would not be acceptable. Refinement of the relevant variables would be a useful topic for further work.

Employment destinations

Employment destination information is collected by HESA by means of a survey conducted approximately six months after students have graduated. The data gathered is very detailed, but has typically been analysed by dividing destinations into “graduate” and “non-graduate” based on a classification by (Elias and Purcell, 2004). In common with league table compilers, we will consider “graduate” employment and further study to be positive employment destinations and count the total number of respondents in these categories.

Student Satisfaction

The NSS dataset contains three levels of information. Level three consists of individual subjects, which are aggregated to give level two and aggregated further to give level one. Level three of the NSS dataset provides the information we need at subject level. For each institution we can obtain the number of students who responded with “Agree” or “Strongly Agree” to the statement “Overall, I am satisfied with the quality of the course”. This is question 22 on the survey and the one used in all the major league tables. These results are only available if the number of respondents, and the size of the department overall, reach certain threshholds. In the case of our example, there are two institutions where data is missing. For the purposes of this illustration we will exclude these institutions from our analysis, although a future development could consider substituting this data from level two.

7.1.4 The final dataset

Our final dataset, then, differs somewhat from our more general plan in that we have increased the number of inputs from three – staff, students and expenditure – to five – teaching staff, research staff, students, expenditure and low-rated research outputs. We still have four outputs, namely degrees awarded, employment destinations, student satisfaction and highly-rated research outputs. These are summarised in Table 7-2 below and given brief variable names which we will use from now on.

From our original pool of 85 DMUs, we first excluded those where the student FTE:FPE ratio was beyond certain limits (see Table 6-7) and have since excluded those where data from the NSS was unobtainable. This gives us a group of 37 DMUs, less than half the total number who are reporting activity in chemistry. However, we can be confident that by focusing on this smaller group we are considering institutions whose activity is reasonably comparable, with one caveat.

Table 7-2: Variables included in the final dataset

	Variable Name	Variable Description
Input variables	TFTE	Number of teaching staff (Full time equivalent)
	RFTE	Number of research staff (Full time equivalent)
	SFPE	Number of students (Full person equivalent)
	Spend	Non-staff spend (in £000)
	RAELow	RAE outputs rated 1*, 2* or unclassified
Output variables	GoodHons	First and Upper Second class degrees awarded
	EmpDest	Graduates employed in positive destinations
	Satis	Students expressing overall satisfaction with their course
	RAEHigh	RAE outputs rated 3* or 4*

There are some DMUs in the set who did not make a submission to the RAE in the chemistry Unit of Assessment. Non-submission is a legitimate choice for an institution which is focused primarily on teaching in this area. However, we cannot accept a zero value into our DEA model, so we need to make a substitution and in this case have chosen 0.01 as a suitably small number.

7.2 The test process

We have considered in detail the construction of the core dataset required for the DEA model. We now turn to the framework set out in Chapter 5 and complete the remaining necessary steps.

7.2.1 Steps preceding the DEA model

Besides the core dataset, the DEA model requires several other pieces of data to be identified. Of paramount importance are the trade-off vectors and the user-selected weights which prioritise certain variables.

Trade-off vectors

In establishing the trade-off vectors we have taken into consideration the issues discussed in section 6.4. It has been assumed that one TFTE and one RFTE have the same cost but produce different outputs, so that substituting one for the other would affect the relevant variables of student satisfaction and research outputs. Research on the topic of class sizes has not yielded clear-cut results (e.g. Machado and Vera-Hernandez, 2008, Bandiera et al., 2009), but it seems reasonable in this context to suppose that more teaching FTE would not necessarily improve degree results but that the associated smaller class sizes would improve student satisfaction. On the other hand, more research FTE would be expected to produce more highly rated papers in the RAE. Following Podinovski (2005, 2007b), a range of values was adopted to express this relationship in two trade-off vectors, T_1 and T_2.

$$T_1 : P_1 = (-0.5, 0.5, 0, 0), Q_1 = (0, 0, -4, 0.2)$$

i.e. moving 0.5 FTE from teaching to research would have no effect on degrees awarded or on employment, but would reduce the number of satisfied students by 4 and increase the total of highly rated RAE papers by 0.2 in one year.

$$T_2 : P_2 = (0.5, -0.5, 0, 0), Q_2 = (0, 0, 2, -0.6)$$

i.e. moving 0.5 FTE from research to teaching would have no effect on degrees awarded or on employment, but would reduce the total of highly rated RAE papers by 0.6 and increase the number of satisfied students by 2 in one year.

Note the conservative framing of these relationships, so that in each case the transfer of input resource is expected to lead to the 'worst case' in the range: the greatest loss of one output and the least gain in the other.

User-selected weights

In order to keep the example as simple as possible, it was decided to evaluate the model only twice for each DMU. In one evaluation the DMU would prioritise student satisfaction and in the other they would prioritise high-rated research outputs. For those DMUs found not to be on the frontier of the production space, this process would identify two possible 'strategies' for them to take forward into the next stage of the test: a league table 'game' in which they will play their strategies against one another.

With only one variable to be prioritised at a time, the weight assigned to it can be completely arbitrary. 1 was therefore chosen as the weight for the prioritised variable and zero for all the other variables.

7.2.2 Using the DEA model

Having developed the theory in detail in chapter 5, we now note only a few practical points arising in the test process.

The dataset for the test was set up in an Excel spreadsheet using named ranges to facilitate reading into and writing from the XPRESS program. This entailed presenting separate ranges for the input and output variables, the P and Q vectors, and the proposed weights for each evaluation round. For this test constant returns to scale were assumed.

Each DMU was labelled with an anonymised code, numbered U1 to U37. Each separate round was labelled with an R: RRAE when the RAE outputs were prioritised, RSatis when student satisfaction was prioritised. Altogether 74 evaluation rounds were conducted, being two for each DMU.

The output from the model took the form of projected input and output variables as well as a reference set of peer DMUs for each DMU in the dataset.

7.2.3 Steps following the DEA model

The purpose of restricting our priorities to two variables was to allow the DEA output to furnish alternative strategies for the DMUs when they play against each other in league table games. We continue to keep the example as simple as possible by proposing two single-variable league tables.

Student Satisfaction League Table

For this league table we find the percentage of students at each DMU who are satisfied with their course by using the output variable Satis as numerator and the input variable SFPE as denominator. A relative score is then calculated for each DMU by

$$Satisfaction\ Score_{DMU} = \frac{Satisfaction_{DMU}}{Satisfaction_{\max}}$$

RAE Performance League Table

For this league table we find the proportion of RAE output from each DMU which is rated 3* or 4*. This is done by using the output variable RAEHigh as numerator and the sum of RAEHigh and RAELow as denominator. A relative score is then calculated for each DMU by

$$RAE\ Score_{DMU} = \frac{RAE_{DMU}}{RAE_{\max}}$$

Choice of players

The output from the DEA model shows that 18 of the 37 DMUs are on the frontier of the production space and therefore have no scope to 'move' in this game. The players need to be selected from the remaining 19 DMUs which are not on the frontier. To use all 19 is not feasible since, even when the strategies are as tightly restricted as proposed, this would require us to evaluate 2^{19} (524,288) games. A smaller example offers a better opportunity to look in detail at the outcomes and to develop some insights into the characteristics of different players and their strategies.

Seven players yields $2^7 = 128$ games, which is sufficient to generate a variety of results within a mangeable scope. The players were chosen to be as varied as possible based on their 'rankings' for research and student satisfaction using the initial dataset.

Table 7-3: Selected players and their starting ranks

DMU	Student Satisfaction	Research	
U34	5	9	'high-high'
U22	4	27	'high-low'
U5	15	15	'middle-middle'
U37	13	25	'middle-low'
U23	35	6	'low-high'
U26	30	17	'low-middle'
U33	37	31	'low-low'

7.3 The starting point

The following figures illustrate the positions of the active DMUs in the two league tables as well as indicating the overall spread of performance.

7.3.1 Student Satisfaction League Table

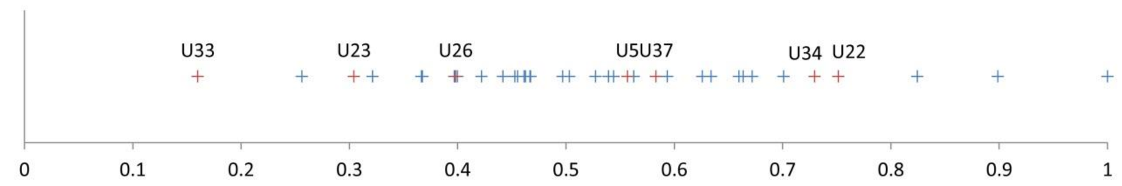


Figure 7-1: Relative Student Satisfaction scores at start. Active DMUs are shown as red crosses, inactive DMUs as blue crosses.

U16 has the highest proportion of satisfied students, namely 84%. U16 scores 1.0 for this performance and ranks at the top of the table. The lowest proportion is 13%, which gives DMU U33 a score of 0.13/0.84 = 0.16.

7.3.2 RAE Performance League Table

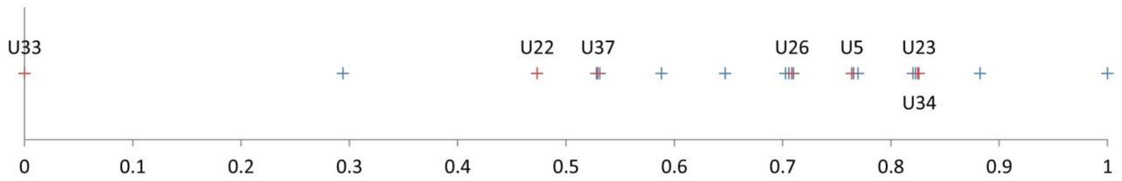


Figure 7-2: Relative RAE scores at start. Active DMUs are shown as red crosses, inactive DMUs as blue crosses.

U25 has the highest proprtion, 85%, of its RAE submission classified as 3* or 4*, giving it a score of 1.0 and top place in the table. Having no RAE submission at all yields the lowest score, which applies to six DMUs (U9, U10, U21, U30, U32, U33 and U35). The lowest non-zero score is 0.25/0.85 = 0.29, which applies to U16, U19 and U28.

U33 is the only one of our active DMUs which enters the process having no RAE submission. As noted above, we need to assign a token small number to its RAE variables in order to use the DEA model. However, the DEA model is still limited in what it can achieve. If asked to target the RAEHigh variable then it will calculate a feasible increase in this output, but it will not match this increase with an increase in the RAELow variable, because RAElow is an input. Therefore, even if the proposed increase to RAEHigh were very modest, the DMU's proportion of highly ranked RAE output would suddenly become very large.

There are a number of different scenarios which might lead a DMU to submit research to the RAE for the first time, but it is unlikely that any of them would result in work which was all classified as 3* or 4* with none at 2* or below, so we need to substitute a more feasible proportion. In the starting dataset the proportion of high ranking RAE output varies from 25% to 85%, with a median of 60%. We will therefore begin by assuming that U33 achieves 60%.

7.4 Results

In the seven sections below we review the results for each active DMU in turn. The same format is used for each.

Profile at start

First, a profile is presented. The initial values for each variable are given and, to put these into context, the main input variables – which offer the best guide to the scale of the DMU – are shown against a boxplot of the entire dataset. The starting position of the DMU in the two league tables is also described.

Note that the data which was obtained from HESA is presented in accordance with the HESA Services Standard Rounding Methodology. This methodology requires that values are rounded to the nearest 5. It applies to SFPE, EmpDest and GoodHons as financial data and FTE figures do not have to be rounded. All calculations were, of course, based on the true value and not the rounded figure which is used for presentation only. Full details of the rounding methodology are given in Appendix B (see p 184).

Performance measures and relative scores

The next section consists of a table which presents the key output variables and the performance measures derived from them, both at the start of the process and after the implementation of each proposed strategy.

Outcome frequencies

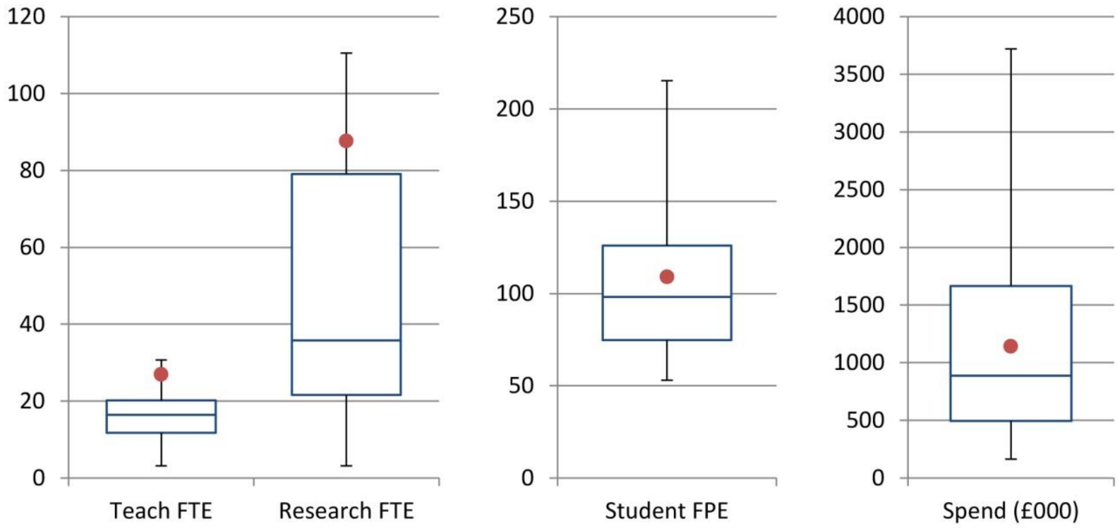
The outcome in terms of league table position depends not only on the DMU's own strategy but also on the strategies of other DMUs. A frequency table summarises the different outcomes for the 128 games. These are also presented graphically on a scatterplot. The x-axis represents the rank position in the Student Satisfaction league table (lowest ranked on the left, highest on the right) and the y-axis the RAE performance league table (lowest ranked at the bottom, highest at the top). The top right corner of the graph represents the top of both tables.

Finally, the outcomes are described in some detail with attention paid to any unusual findings. However, further discussion of the results and their implications is reserved for section 7.6 below.

7.4.2 Results for U34

Profile at start

TFTE	RFTE	SFPE	Spend	RAELow	GoodHons	EmpDest	Satis	RAEHigh
26.98	87.68	110	1143	9.8	45	45	67	23



U34 is one of the largest DMUs in the dataset. Teaching FTE is nearly 90% of the maximum value and Research FTE 80% of the maximum, although non-staff spending and student FPE are both very close to average.

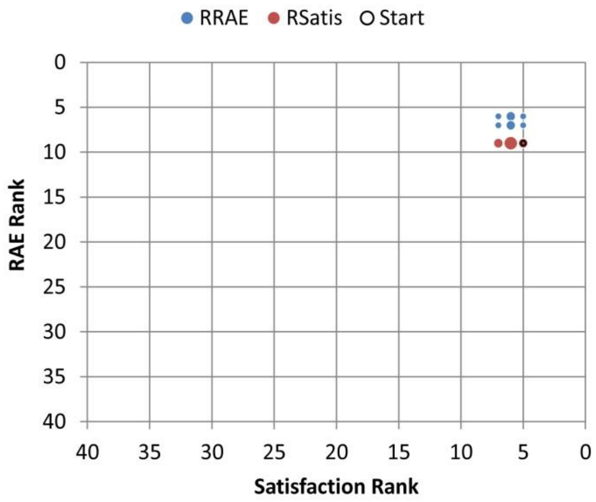
At the starting point U34 is the highest ranking of our active DMUs. Its position is 5th in the student satisfaction table and 9th in the RAE table. This is based on 61.5% of students expressing satisfaction with their course (giving a relative score of 0.729) and on 70.1% of its RAE submission being classified as 3* or 4* (giving a relative score of 0.825).

Performance measures and relative scores

	Inputs		Outputs		Performance measures		Best possible scores (relative to max performance)	
	Student FPE	RAE Low	Satisfied Students	RAE High	% Satisfied	% RAE High	Student satisfaction	RAE
Starting postition	110	9.8	67	23	61.5	70.1	0.729	0.825
Focus on RAE	110	9.8	67	23.1	61.5	70.2	0.729	0.826
Focus on Satisfaction	110	9.8	67.4	23	61.8	70.1	0.734	0.825

Outcome frequencies

Strategy	Satis Rank	RAE Rank	Frequency
Focus on RAE	5	6	8
	5	7	8
	6	6	16
	6	7	16
	7	6	8
	7	7	8
Focus on Satisfaction	5	9	16
	6	9	32
	7	9	16

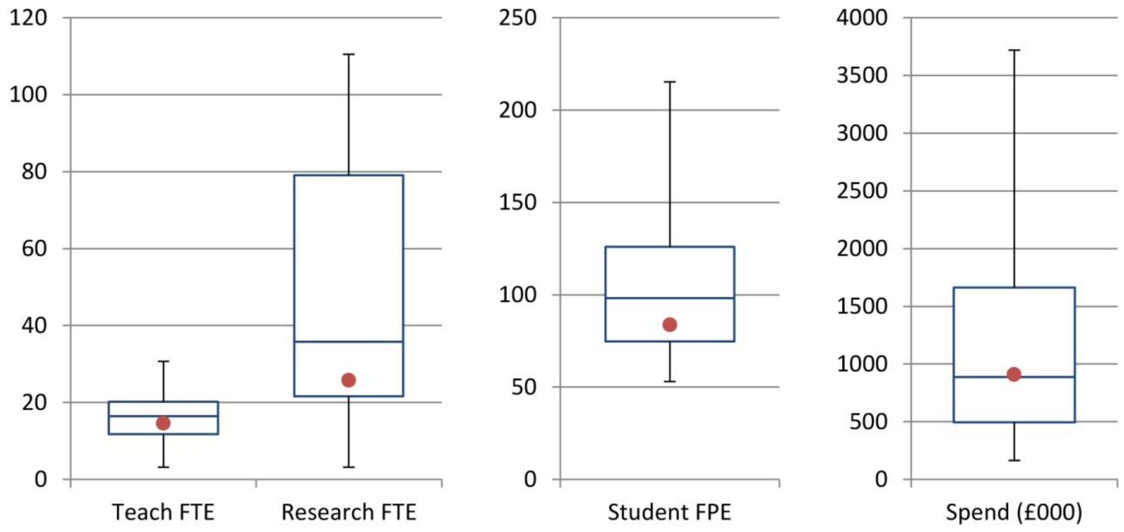


The DEA model suggests that U34 is very close to the frontier, having only a very small scope for improvement in either its RAE performance or student satisfaction without sacrificing some other output. If the suggested improvement in RAE performance were achieved, U34 could increase its score from 0.825 to 0.826 – a very slight change but enough to increase its ranking in the RAE table from 9 to 6 or 7, depending on the strategies of other DMUs. However, although the potential improvement in student satisfaction is somewhat greater – in the best case, its score could increase from 0.729 to 0.734 – this is not enough to improve its rank and the DMU risks losing ground in this table, whichever strategy it adopts. Indeed, the frequency table shows that in either case it will only be able to retain 5th place in 16 of the 64 games.

7.4.3 Results for U22

Profile at start

TFTE	RFTE	SFPE	Spend	RAELow	GoodHons	EmpDest	Satis	RAEHigh
14.61	25.81	85	911	10.4	30	30	53	7



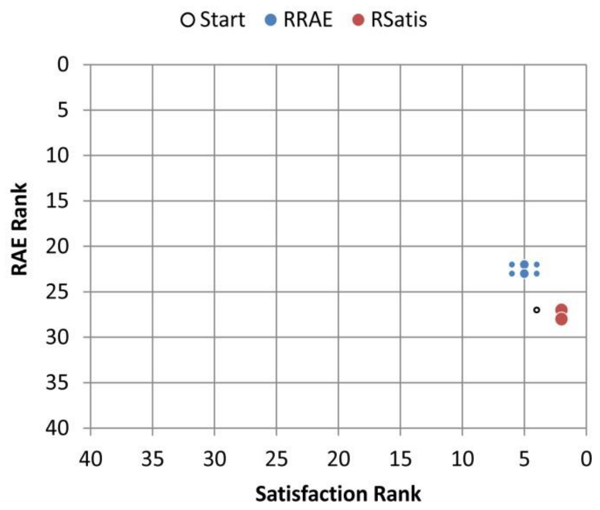
U22 is a smaller than average institution on all the main input measures except for spend, where it is almost exactly at the median. At the starting point U22 ranks 4th for student satisfaction but only 27th for RAE performance. This is based on 63.3% of students expressing satisfaction with their course (giving a relative score of 0.751) and on 40.2% of its RAE submission being classified as 3* or 4* (giving a relative score of 0.473).

Performance measures and relative scores

	Inputs		Outputs		Performance measures		Best possible scores (relative to max performance)	
	Student FPE	RAE Low	Satisfied Students	RAE High	% Satisfied	% RAE High	Student satisfaction	RAE
Starting postition	85	10.4	53	7	63.3	40.2	0.751	0.473
Focus on RAE	85	10.4	53	11.6	63.3	52.8	0.751	0.621
Focus on Satisfaction	85	10.4	64.8	7	77.4	40.2	0.919	0.473

Outcome frequencies

Strategy	Satis Rank	RAE Rank	Frequency
Focus on RAE	4	22	8
	4	23	8
	5	22	16
	5	23	16
	6	22	8
	6	23	8
Focus on Satisfaction	2	27	32
	2	28	32



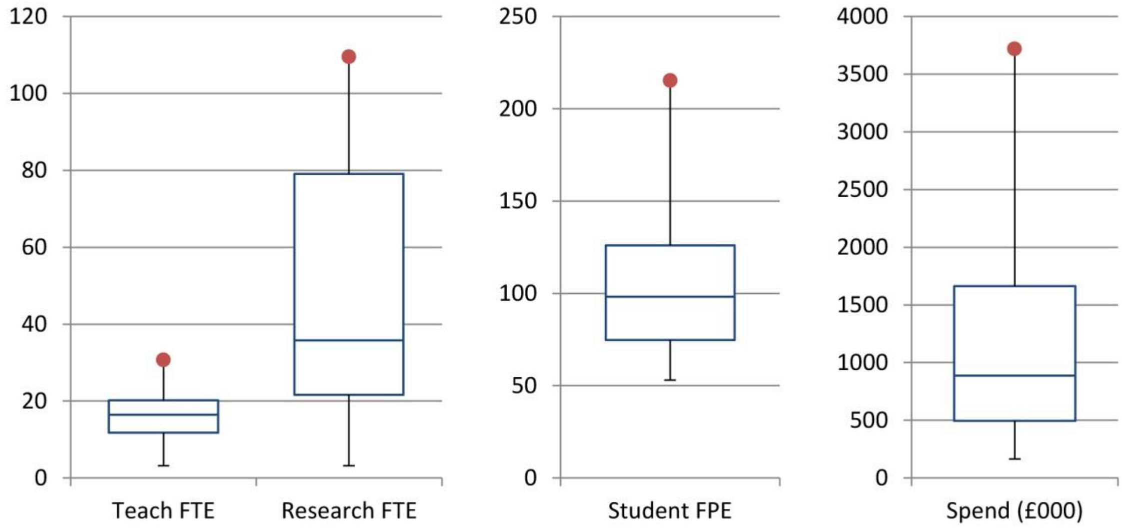
The output from the DEA model suggests that a significant improvement in RAE performance would be required to reach the frontier of the production space: from 7 papers per year at the 3* and 4* level to nearly 12. This would increase the proportion of high value RAE output to 52.8% giving a relative score of 0.621. However, the rank associated with this performance is no higher than 22 and in half of the games this drops to 23. At the same time, U22 risks slipping to 5th or 6th place in student satisfaction. Its initial rank is maintained in only 16 out of 64 games.

There is also scope for a large increase in the percentage of satisfied students from 63.3% to 77.4%. This would give U22 a relative score of 0.919 in the games with U37 and second place in the ranking. In half of these games the RAE ranking of 27 is maintained and in half it is reduced to 28.

7.4.4 Results for U5

Profile at start

TFTE	RFTE	SFPE	Spend	RAELow	GoodHons	EmpDest	Satis	RAEHigh
30.7	109.6	215	3719	17.6	120	125	101	32.6



U5 is the largest institution in the dataset. It has the most staff and students and the highest spend. The staff FTE is highly skewed towards research, but when the total is broken down it is not quite the highest on this variable: U2 has 110.4 compared with 109.6 here. U5 does, however, have the highest teaching FTE.

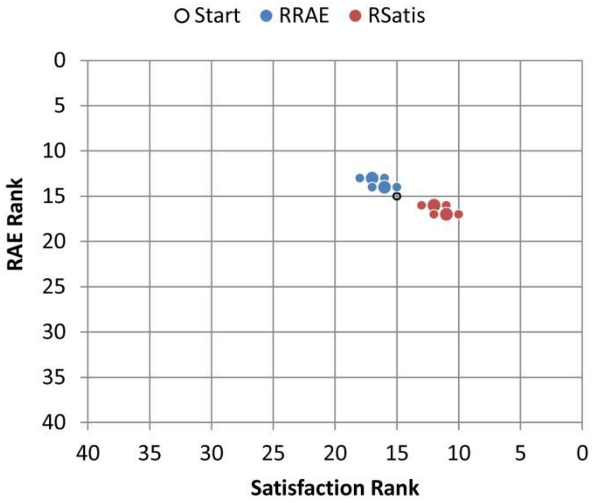
It starts in the middle of our league tables, ranked 15th in both. This is based on 46.9% of students expressing satisfaction with their course (giving a relative score of 0.557) and on 64.9% of its RAE submission being classified as 3* or 4* (giving a relative score of 0.764).

Performance measures and relative scores

	Inputs		Outputs		Performance measures		Best possible scores (relative to max performance)	
	Student FPE	RAE Low	Satisfied Students	RAE High	% Satisfied	% RAE High	Student satisfaction	RAE
Starting postition	215	17.6	101	32.6	46.9	64.9	0.557	0.764
Focus on RAE	215	16.9	101	36.1	46.9	68.1	0.557	0.801
Focus on Satisfaction	215	17.6	118.5	32.6	55.0	64.9	0.653	0.764

Outcome frequencies

Strategy	Satis Rank	RAE Rank	Frequency
Focus on RAE	15	14	8
	16	13	8
	16	14	16
	17	13	16
	17	14	8
	18	13	8
Focus on Satisfaction	10	17	8
	11	16	8
	11	17	16
	12	16	16
	12	17	8
	13	16	8



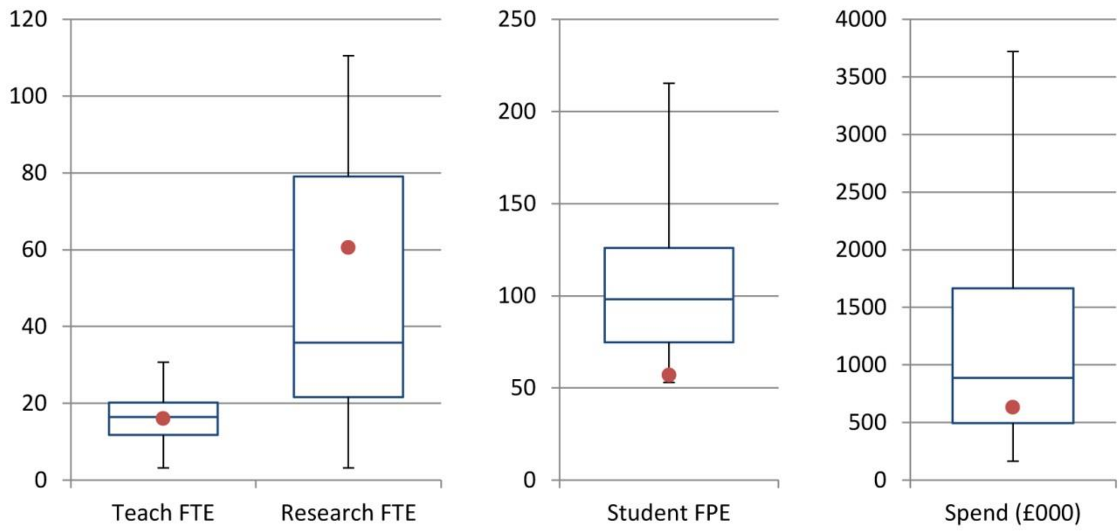
The DEA model suggests an increase from 32.6 to 36.1 in high ranking RAE output and it also suggests a small reduction in the RAE output rated below 3*. If this were achieved then the total RAE output would be increased from 50.2 papers per year to 53, of which 68.1% would be classified as 3* or 4*. This improvement would raise the relative score to 0.801, apparently close to the maximum of 0.850 but only enough to rise at most two places in this table. There are 32 out of the 64 games in which a rank of 13 is achieved and 32 in which 14 is achieved. Meanwhile, there are only 8 games in which U5 is able to maintain its ranking for student satisfaction. In the majority of games it slips to 16 or 17, and in 8 out of the 64 games it drops to 18.

The proposed increase in satisfied students is quite substantial – from 101 to 118.5 – but this would still mean only 55% of students were satisfied with their course. This improvement would increase U5's relative score from 0.557 to 0.653 and achieve a rise of at least two places in the ranking. There are 8 out of 64 games in which a rank of 13 is achieved, 24 in which 12 is achieved, 24 in which 11 is achieved and 8 in which a rank of 10 is achieved. At the same time, though, it is dropping to 16th or 17th place in the RAE table.

7.4.5 Results for U37

Profile at start

TFTE	RFTE	SFPE	Spend	RAELow	GoodHons	EmpDest	Satis	RAEHigh
16.0	60.6	55	632	15.4	25	15	28	12.6



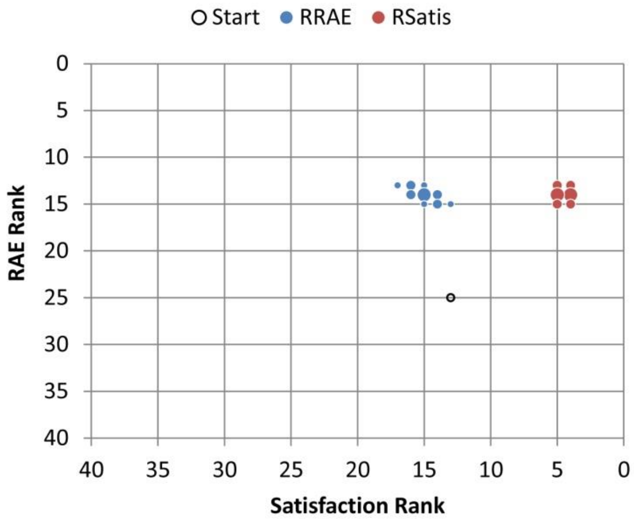
U37 is a medium-sized institution as far as staff numbers are concerned, but has one of the smallest student cohorts in the datatset. Perhaps unsurprisingly in this context, Research FTE outweighs Teaching FTE by almost four to one. However, its initial performance in our RAE league table is poor. With 45% of its submission classified as 3* or 4* U37 achieves a relative score of 0.529 and ranks 25th. The rate of student satisfaction is similar at 49.1%, but the relative score of 0.583 is sufficient for a rank of 13th in the table.

Performance measures and relative scores

	Inputs		Outputs		Performance measures		Best possible scores (relative to max performance)	
	Student FPE	RAE Low	Satisfied Students	RAE High	% Satisfied	% RAE High	Student satisfaction	RAE
Starting postition	55	15.4	28	12.6	49.1	45.0	0.583	0.529
Focus on RAE	55	7.8	28	15.7	49.1	66.9	0.583	0.787
Focus on Satisfaction	55	6.4	38.4	12.6	67.4	66.4	0.800	0.781

Outcome frequencies

Strategy	Satis Rank	RAE Rank	Frequency
Focus on RAE	13	15	4
	14	14	8
	14	15	8
	15	13	4
	15	14	16
	15	15	4
	16	13	8
	16	14	8
	17	13	4
Focus on Satisfaction	4	13	8
	4	14	16
	4	15	8
	5	13	8
	5	14	16
	5	15	8



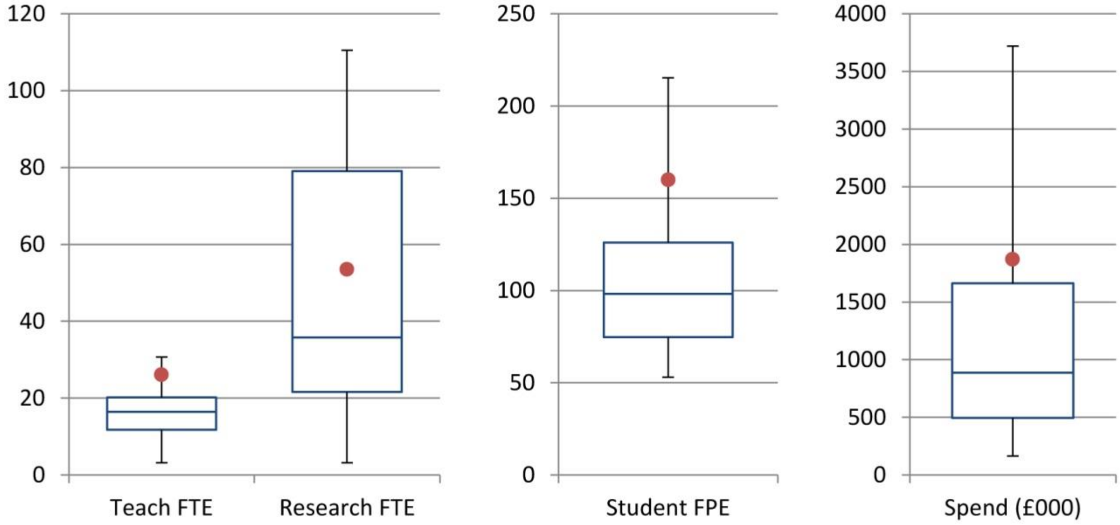
When the priority is RAE performance, the DEA model proposes an increase in high ranking RAE output from 12.6 papers per year to 15.7 and, more dramatically, a decrease in low ranking RAE output from 15.4 to 7.8 papers per year. U37's total submissible output would therefore be reduced from 28 papers to 23.5. The effect of this would be to increase the proportion of high ranking RAE output to 66.9% of the total, giving a relative score of 0.787 and a substantial improvement in rank to at least 15th. The plot shows that there is a wide range of possible outcome pairs. A rank of 14 in the RAE table being the most likely (32 out of 64 games). However, U37 is likely to slip one or two places in the student satisfaction table and drops as low as 17th in 4 out of the 64 games.

The model proposes a considerable increase in student satisfaction, from 28 to 38.4 students. This would yield a satisfaction rate of 67.4%, a relative score of 0.800 and a rank of 4 or 5 in the table. Even when the focus is on students, however, the model also finds a significant degree of slack in the RAE performance and proposes an even more substantial reduction in low ranking RAE output than in the research-focused scenario. Decreasing this by 9 papers per year, from 15.4 to 6.4, would reduce the total submissible output to 19 papers but increase the proportion classified 3* or 4* to 66.4% − almost as much as was seen in the previous strategy. The result is that the student satisfaction strategy leads to a much better overall outcome in the league table rankings. While achieving 4th or 5th in the student satisfaction table, U37 is also climbing to 13, 14 or 15 in the RAE table.

7.4.6 Results for U23

Profile at start

TFTE	RFTE	SFPE	Spend	RAELow	GoodHons	EmpDest	Satis	RAEHigh
26.1	53.5	160	1871	7.1	35	35	41	16.7



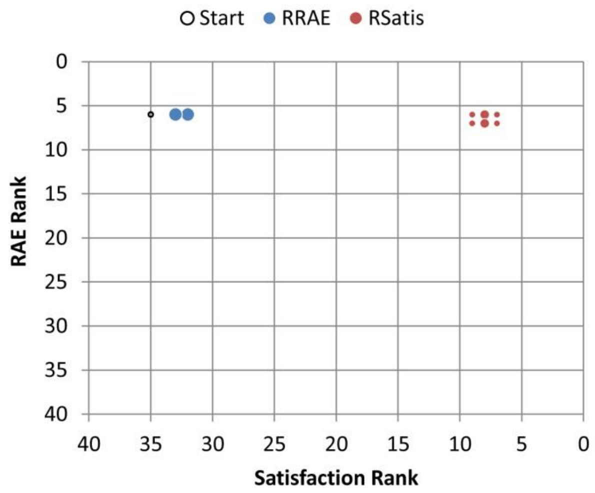
U23 has a larger than average student population and a correspondingly high Teaching FTE, while its Resarch FTE is closer to the dataset average. Its RAE performance is very strong, with 70.2% of its output being classified as 3* or 4*. This gives a relative score of 0.826 and a rank of 6 in our RAE table. However, student satisfaction is very low at 25.6%. Its relative score of 0.304 puts U23 at 35th place in this table.

Performance measures and relative scores

	Inputs		Outputs		Performance measures		Best possible scores (relative to max performance)	
	Student FPE	RAE Low	Satisfied Students	RAE High	% Satisfied	% RAE High	Student satisfaction	RAE
Starting postition	160	7.1	41	16.7	25.6	70.2	0.304	0.826
Focus on RAE	131.6	7.1	41	20.8	31.2	74.6	0.370	0.877
Focus on Satisfaction	160	7.1	94	16.7	58.7	70.2	0.697	0.826

Outcome frequencies

Strategy	Satis Rank	RAE Rank	Frequency
Focus on RAE	32	6	32
	33	6	32
Focus on Satisfaction	7	6	8
	7	7	8
	8	6	16
	8	7	16
	9	6	8
	9	7	8



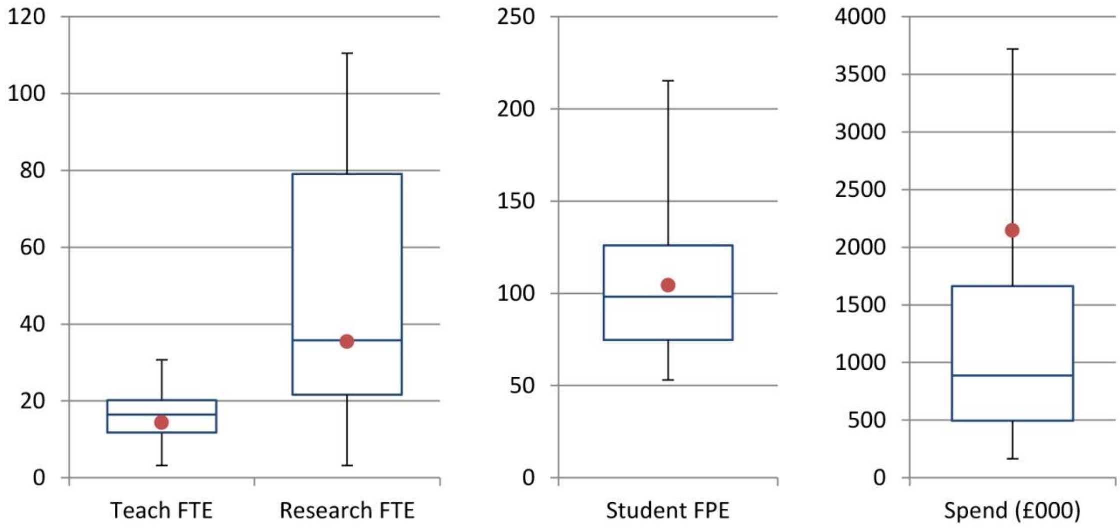
The DEA model finds some scope for increasing U23's 3* and 4* RAE output, from 16.7 papers per year to 20.8. This would be 74.6% of the DMU's total RAE output and yield a slightly higher relative score of 0.877. However, it would make no difference at all to the rank achieved. While the focus is on RAE output, the model nonetheless finds that there is slack in other variables. It is proposed that the number of students should be reduced from 160 to 131.6, which would have the effect of increasing the proportion satisfied with their course to 31.2% and improve U23's rank on student satisfaction by two or three places.

When the DEA model is focused on student satisfaction, then a much bigger change in this area is proposed. It suggests that 94 out of the population of 160 should be satisfied with their course, which is 58.7%. Achieving this rate would yield a relative score of 0.697 and a rank of at least 9 in the table. There is a risk that U23 might lose out in the RAE table, but its current rank is maintained in half the games and it only drops one place (to 7th) in the other half.

7.4.7 Results for U26

Profile at start

TFTE	RFTE	SFPE	Spend	RAELow	GoodHons	EmpDest	Satis	RAEHigh
14.4	35.4	105	2145	7	35	35	35	10.6



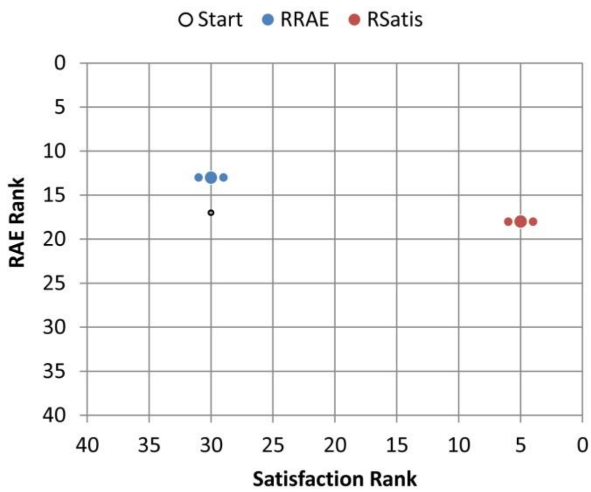
U26 is very close to being the 'average institution' in this dataset. Its teaching and research FTE and student FPE are all close to the median. However, its non-staff spend is relatively high: in the upper quartile and more than twice the median value. U26 starts with a satisfaction rate of 33.5% which gives a relative score of 0.398 and a rank of 30. Its research performance is slightly better, with 60.2% of its RAE output rated 3* or 4*. This yields a relative score of 0.709 and 17th place in the table.

Performance measures and relative scores

	Inputs		Outputs		Performance measures		Best possible scores (relative to max performance)	
	Student FPE	RAE Low	Satisfied Students	RAE High	% Satisfied	% RAE High	Student satisfaction	RAE
Starting postition	105	7	35	10.6	33.5	60.2	0.398	0.709
Focus on RAE	103.9	7	35	15.6	33.7	69.0	0.398	0.812
Focus on Satisfaction	105	7	69.8	10.6	66.9	60.2	0.794	0.709

Outcome frequencies

Strategy	Satis Rank	RAE Rank	Frequency
Focus on RAE	30	13	32
	31	13	16
	29	13	16
Focus on Satisfaction	4	18	16
	5	18	32
	6	18	16



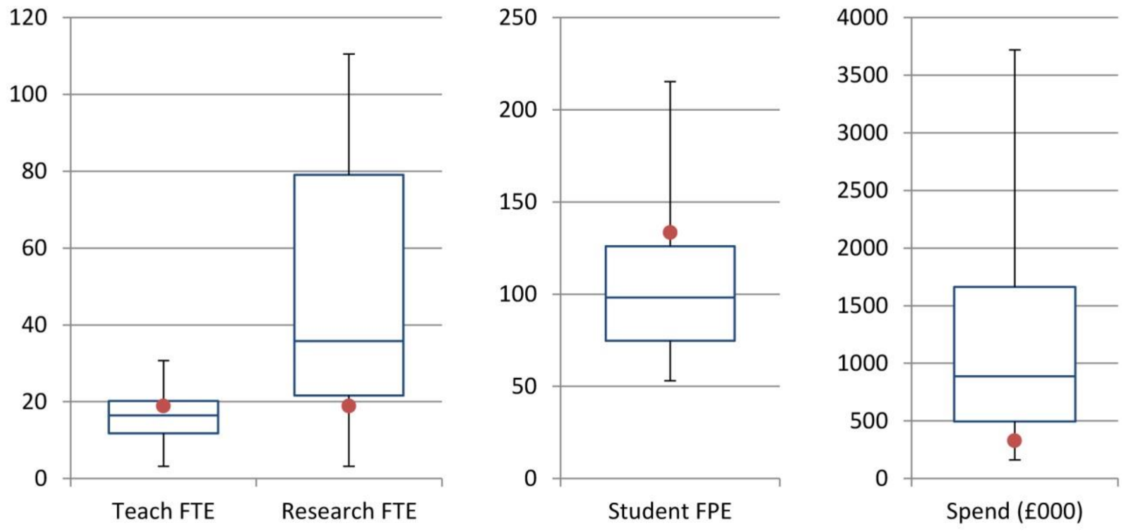
When the DEA model is focused on improving RAE performance it suggests a considerable increase of nearly 50% in highly rated output, from 10.6 to 15.6 papers per year. This would increase the percentage from 60.2% to 69% but has only a modest impact on rank, taking U26 from 17th to 13th in the table. A very small adjustment is also made to the student numbers, but this is not guaranteed to improve or even maintain the starting rank for student satisfaction. In 16 out of the 64 games U26 drops a place from 30th to 31st.

The proposed increase in student satisfaction, when this is the focus, is even more substantial: the model suggests that it should be doubled from 35 students (33.5%) to 69.8 students (66.9%). This would have a significant impact on the position of U26 in the table. The minimum rank achieved for this improvement would be 6 (in 16 out of 64 games) and could be as high as 4 (also in 16 games). With no attention paid to the RAE performance, however, the model shows that U26 would drop one place to 18th.

7.4.8 Results for U33

Profile at start

TFTE	RFTE	SFPE	Spend	RAELow	GoodHons	EmpDest	Satis	RAEHigh
18.9	18.9	135	328	None	30	30	18	None



U33's profile clearly shows that it is a teaching-focused institution. Thenumbers of teaching staff and students at the top of the upper middle quartile, while non-staff spend and research FTE are low. However, although the number of research staff falls in the lower quartile, it is a long way above the minimum and the lack of RAE submission need not be interpreted as a total absence of research activity. U33 starts at the bottom of both tables. It ranks 31st in the RAE table along with all the other six institutions which made no submission in RAE 2008 and 37th out of 37 in the student satisfaction table. This position is based on a satisfaction rate of 13.5% and a relative score of 0.160.

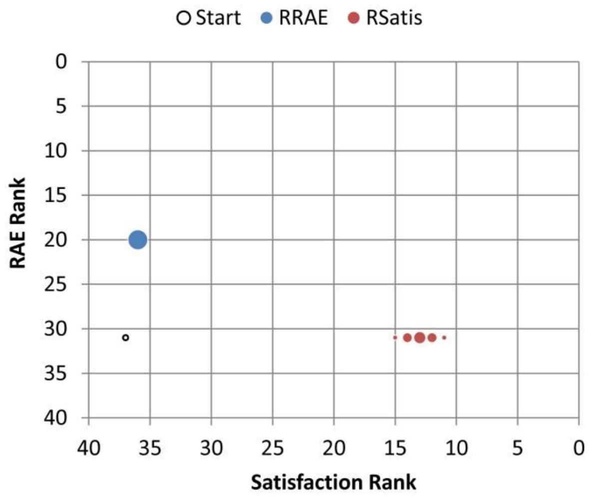
Performance measures and relative scores

	Inputs		Outputs		Performance measures		Best possible scores (relative to max performance)	
	Student FPE	RAE Low	Satisfied Students	RAE High	% Satisfied	% RAE High	Student satisfaction	RAE
Starting postition	135	0.01	18	0.01	13.5	0	0.160	0
Focus on RAE	76.0	0.01	18	1.1	23.7	60*	0.281	0.706
Focus on Satisfaction	76.0	0.01	40.3	0.01	53.0	0	0.630	0

*60% is adopted here as it is the median value – see section 7.3.2 above

Outcome frequencies

Strategy	Satis Rank	RAE Rank	Frequency
Focus on RAE	36	20	64
Focus on Satisfaction	11	31	4
	12	31	16
	13	31	24
	14	31	16
	15	31	4



We have already noted that the DEA model cannot fully account for a 'new entrant' into the RAE because the low ranking output is treated as an input. We have therefore assumed that U33 will achieve the median performance of 60%. This would yield a relative score of 0.706 and a rank of 20. At the same time, the model proposes a dramatic cut in student numbers in order to improve the rate of satisfaction. However, in the context of focusing on research, this is only enough to achieve an improvement of one position from 37 to 36.

When the emphasis is on student satisfaction, the model still proposes the same cut in student numbers, from 135 to 76 FPE. However, this strategy also includes doubling the number of satisfied students from 18 to 40.3, thus achieving a satisfaction rate of 53%. This would have a substantial impact on U33's rank in this table, taking them to at least 15 (in 4 out of 64 games) and possibly as high as 11 (also in 4 games) with the most likely outcome being a rank of 13 (24 games).

7.5 The league table game

Our development of this scenario has been conducted with game theory in mind. We have given each DMU a choice of two strategies and established the ranks they will achieve if they play each of these strategies. To examine the results further through the lens of game theory we need to interpret the rank outcomes as payoffs. For each DMU *i* the outcome can be expressed as an ordered pair $\boldsymbol{\omega}_i = \left(\omega_{S_i}, \omega_{R_i}\right)$ where ω_{S_i} is the rank achieved in the student satisfaction table and ω_{R_i} is the rank achieved in the RAE table.

The next step is to define a payoff function over these outcomes in order to represent the decision-maker's preferences for certain outcomes over others. However, it is immediately obvious that such a function may not be straightforward. A given rank ω_S^1 is likely to be preferred to another rank ω_S^2 if $\omega_S^1 < \omega_S^2$, but it is not necessarily the case that ω_S^1 will be preferred to ω_R^1 simply because $\omega_S^1 < \omega_R^1$ because a decision-maker may attach greater value to a good performance in one area than to an equally good performance in another.

There may also be significant threshhold values, e.g. a place in the top ten in the RAE table, in which case $\omega_R^1 \leq 10$ would be greatly preferred to $\omega_R^1 = 11$. This is illustrated with a hypothetical utility curve in Figure 7-3 below. A rank of 1 is the preferred outcome, wth a utility of 1. Then there is a sharp drop to ranks 2 – 10, followed by another sharp drop to ranks 11 – 20 and so on.

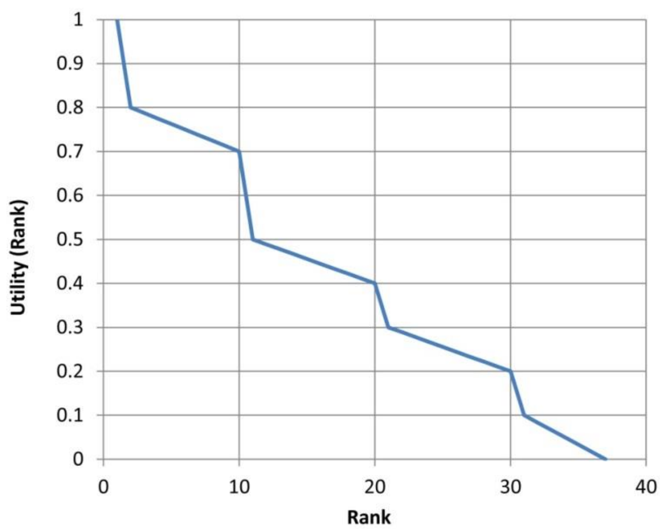


Figure 7-3: Possible payoff function for ranks in a league table, where there is a step change in utility at key threshhold values

The approach described so far is based on the idea that a given rank has a fixed value to a DMU, irrespective of the DMU's starting position. However, it is a well known finding of Kahneman and Tversky that "losses loom larger than gains" (1979:279), and thus the utility function in Figure 7-3 may be too naïve to capture the relative benefits of alternative possible outcomes to specific DMUs. Figure 7-4 shows a simple piecewise linear utility function which is based on change in rank rather than the absolute value of the rank achieved.

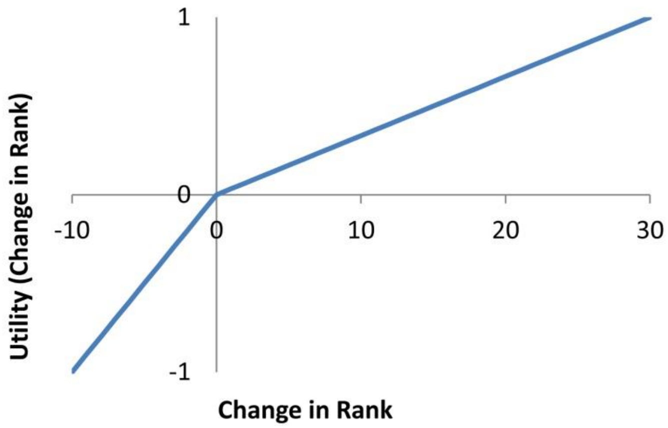


Figure 7-4: Possible payoff function for ranks in a league table, where a loss in rank is felt more acutely than a gain in rank

We will consider the results of section 7.4 using four different approaches to utility:

1. A utility function based on the absolute rank achieved, where both league tables have equal value
2. A utility function based on the absolute rank achieved, where performance in one league table is valued more than the other
3. A utility function based on change in rank, where both league tables have equal value
4. A utility function based on change in rank, where performance in one league table is valued more than the other

7.5.1 Absolute rank, Equal value

We will first assume that the ranks in both tables carry equal weight and will use the function illustrated in Figure 7-3 to incorporate step changes in preferences at threshhold values. We will further assume that all the DMUs have the same payoff function. These assumptions allow us to turn each ordered pair into a single value for each DMU *i*, namely $u_i = u\left(\omega_{S_i}\right) + u\left(\omega_{R_i}\right)$.

Examining the values of u_i we find that five of our DMUs achieve their highest utility by playing the strategy RSatis, while two achieve their highest utility by playing RRAE.

Table 7-4: Maximum utilities and associated strategies for approach 1

DMU	Maximum utility	Associated with	
		Rank outcomes	Strategy
U5	1.13	(10, 17)	RSatis
U22	1.06	(4, 22)	RRAE
U23	1.49	(7, 6)	RSatis
U26	1.20	(4, 18)	RSatis
U33	0.60	(11, 31)	RSatis
U34	1.51	(5, 6)	RRAE
U37	1.25	(4, 13)	RSatis

We can see immediately from Table 7-4, that maximum utility for three DMUs (U22, U26 and U37) depends on them all achieving fourth place in the student satisfaction table. If they play their preferred strategies then, predictably, only one DMU achieves fourth place, namely U37. U26 comes fifth and U22 drops to sixth. Since U22 is choosing to play RRAE, it is not surprising that in this scenario it loses ground on student satisfaction. Its utility is still maximised by this strategy under the assumptions we have adopted, although the benefit is very slight. Setting aside the other four DMUs for the moment, playing RRAE will allow U22 to achieve (6, 22) which carries a utility of 1.04 while playing RSatis achieves (2, 27) which has a utility of 1.03. If losses loom larger than gains for the decision-makers at U22, then they may prefer to secure their status in the student satisfaction table rather than make modest gains in the RAE table.

In general, we find that all the DMUs' best responses to the other DMUs' strategies is to play the strategy shown in Table 7-4. However, except in the case of U37, the reward will not be as great as anticipated if they only consider their own actions. If we take U5 as an example, we see that its greatest utility is associated with the rank outcomes (10, 17). This is the only outcome which includes a top ten place for U5 and it is achieved by playing RSatis. If we suppose that all the other DMUs play their own best strategies, however, then we find that the best outcomes U5 can achieve with RSatis are (13, 16). Since the alternative strategy RRAE would yield (18, 13), which has a slightly lower utility, RSatis is still the best response strategy for U5 under these circumstances.

Table 7-5: Rank outcomes and utilities achieved when best strategy is played

DMU	Strategy	Associated with	
		Rank outcomes	Utility
U5	RSatis	(13, 16)	0.92
U22	RRAE	(6, 22)	1.04
U23	RSatis	(9, 7)	1.45
U26	RSatis	(5, 18)	1.18
U33	RSatis	(15, 31)	0.56
U34	RRAE	(7, 6)	1.49
U37	RSatis	(4, 13)	1.25

7.5.2 Absolute rank, Different value

For the second evaluation we will suppose that performance in the RAE league table is in general valued more highly than performance in the Student Satisfaction league table. We will assume that the payoff function from Figure 7-3 still applies to ranks in the RAE table and introduce a second function to represent the value of ranks in the Satisfaction table. We assume that first place in both tables is equally highly prized, but that below first place there is a steep drop in the utility of places in the Satisfaction table. We continue to assume that all the DMUs have the same payoff function.

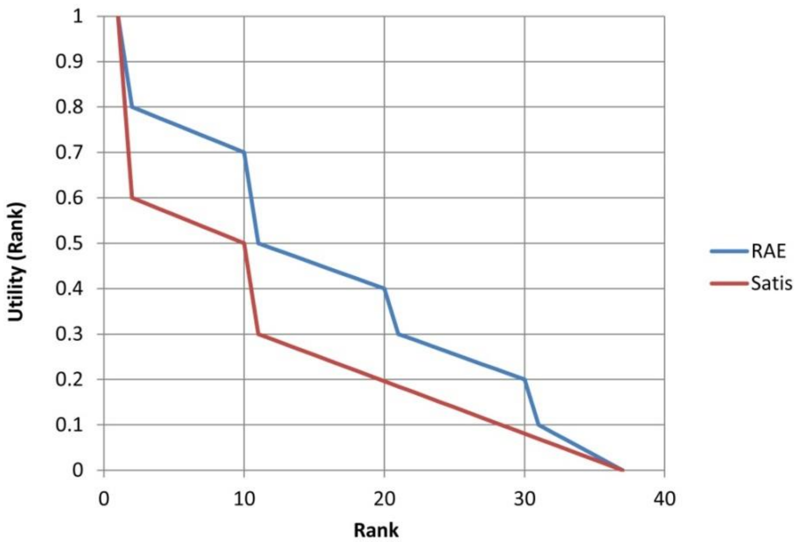


Figure 7-5: Alternative payoff functions for ranks in a league table, where RAE performance is valued more highly than Student Satisfaction

Under these assumptions, we find that the best outcomes, and hence the best strategies, for most of our DMUs are unchanged. Only U33, the lowest performing of all the DMUs, now finds that its best strategy is to pursue research performance rather than student satisfaction.

Table 7-6: Maximum utilities and associated strategies for approach 2

DMU	Maximum utility	Associated with	
		Rank outcomes	Strategy
U5	0.93	(10, 17)	RSatis
U22	0.94	(4, 22)	RRAE
U23	1.33	(7, 6)	RSatis
U26	1.07	(4, 18)	RSatis
U33	0.41	(36, 20)	RRAE
U34	1.38	(5, 6)	RRAE
U37	1.13	(4, 13)	RSatis

As before, we find that the DMUs' best responses to the other DMUs' strategies is to play the strategy shown in Table 7-6, even though the payoff will be less than their best possible outcome.

7.5.3 Change in rank, Equal value

We now change our approach to utility by considering the gains and losses of a DMU rather than the absolute rank achieved. The function shown in Figure 7-4 above is piecewise linear, with a much steeper slope for negative x (when rank position deteriorates) than for positive x (when rank position improves). We assume that the same function holds for both league tables and is shared by all DMUs.

Under these assumptions we again find that for most of the DMUs their best outcomes and associated strategy are unchanged. Interestingly, from having the lowest maximum utility, U33 now has one of the highest. Its low starting position means that it is much less at risk of losing rank in either table than those DMUs which are placed higher up. U33 is now best served by the RSatis strategy, as in our first example.

Table 7-7: Maximum utilities and associated strategies for approach 3

DMU	Maximum utility	Associated with	
		Rank outcomes	Strategy
U5	0.03	(11, 16)	RSatis
		(15, 14)	RRAE
U22	0.17	(4, 22)	RRAE
U23	0.93	(7, 6)	RSatis
U26	0.77	(4, 18)	RSatis
U33	0.87	(11, 31)	RSatis
U34	0.10	(5, 6)	RRAE
U37	0.70	(4, 13)	RSatis

U5 is now in an interesting position as it appears to have two alternative best outcomes. From its starting position in the middle of both tables (15, 15) it might choose to aim for a substantial gain in student satisfaction at the expense of a slight drop in RAE performance (11, 16), or to achieve a more modest improvement in RAE performance while maintaining its standing in student satisfaction (15, 14). However, if we assume that all the other DMUs play their best strategies first, we find that, as before, U5 will actually be faced with an outcome of either (13, 16) or (18, 13). Under the current payoff function, (13, 16) is rated more highly and RSatis is therefore the better response strategy.

We also noted in 7.5.1 that, after all the other DMUs have played their strategies, U22 cannot achieve their best outcome of (4, 22) but is left with a very close decision between an outcome of (2, 27) or (6, 22). Under the current payoff function, we find that the losses being weighted more heavily than the gains has tipped the balance towards the student satisfaction strategy, although the two outcomes are still extremely close. Playing RSatis to achieve (2, 27) has a utility of 0.067, while playing RRAE to achieve (6, 22) has a utility of -0.03.

7.5.4 Change in rank, Different value

Suppose that losses are weighted more heavily than gains and that performance in the RAE league table is also valued more highly than performance in the Student Satisfaction league table. We adjust the functions as shown in Figure 7-6 below. Losses in both tables are assumed to be equally unpleasant, but gains in RAE ranking are more highly valued than gains in student satisfaction. We continue to assume that the payoff functions are shared by all DMUs.

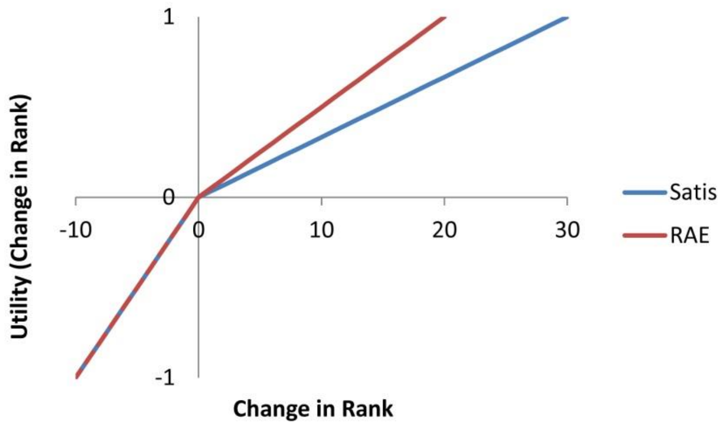


Figure 7-6: Payoff function where losses are weighted more heavily than gains and RAE performance is valued more highly than Student Satisfaction

Although this alteration to the function changes some of the maximum utility values (see Table 7-8), it makes little difference to the associated strategies and rank outcomes. Only U5 now finds that its best strategy has moved from one focused on student satisfaction to one focused on RAE performance.

Table 7-8: Maximum utilities and associated strategies for approach 4

DMU	Maximum utility	Associated with	
		Rank outcomes	Strategy
U5	0.05	(15, 14)	RRAE
U22	0.25	(4, 22)	RRAE
U23	0.93	(7, 6)	RSatis
U26	0.77	(4, 18)	RSatis
U33	0.87	(11, 31)	RSatis
U34	0.15	(5, 6)	RRAE
U37	0.90	(4, 13)	RSatis

Playing out the game, we find that most of the DMUs do maximise their utility by following these strategies. However U22 and U5 are again faced with borderline decisions. In practice they will each maximise utility by following the opposite strategy to the one designated ‘best’ in Table 7-8, but the difference in payoff is still slight (see Table 7-9).

Table 7-9: Rank outcomes and payoffs associated with each strategy for U5 and U22

DMU	Strategy	Associated with	
		Rank outcomes	Utility
U5	RRAE	(18, 13)	-0.2
	RSatis	(13, 16)	-0.03
U22	RRAE	(6, 22)	0.05
	RSatis	(2, 27)	0.07

7.6 Discussion

In 7.4 and 7.5 above we have described in some detail the different outcomes which pertain to the alternative strategies of our set of active DMUs. In this section we discuss a number of issues arising from these observations.

7.6.1 Insights

In general we can see that most strategies, if successfully implemented, have a positive effect on league table rank, even if there is some uncertainty about the precise rank which will be achieved. This is to be expected since the majority of the DMUs in our test dataset (30 out of 37) are static and not competing to improve performance. We can see the reverse of this effect when our active DMUs focus on one variable to the exclusion of the other: in the majority of cases they lose ground in the table where they have not improved because others are actively focusing in this area.

However, it is noticeable in some cases, such as the RAE performance strategy of U5, that successfully implementing a strategy and achieving a high relative score is not necessarily sufficient to achieve a significant improvement in rank. We can gain further insight into this situation by re-considering the linear plot of all the DMUs' performance in each area. Looking back at Figure 7-2, we can see that there are large numbers of DMUs with scores between 0.7 and 0.85. If we update this figure to reflect the situation should all the DMUs choose to focus on RAE performance, then we can see that five out of the seven of them have moved into, or improved within, this very busy section of the graph.

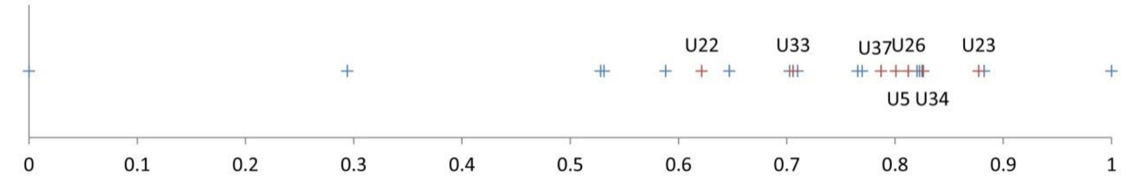


Figure 7-7: Relative RAE scores when all play RAE strategy

In this crowded field it is difficult for a DMU to distinguish itself. Among our active DMUs, including U5, none of those which start outside the top ten are able to break into it solely by achieving the level of improvement suggested by the DEA model. Those who start in the top ten – U34 and U23 – have to focus on their RAE performance simply to maintain their position, or each risks losing ground to the other.

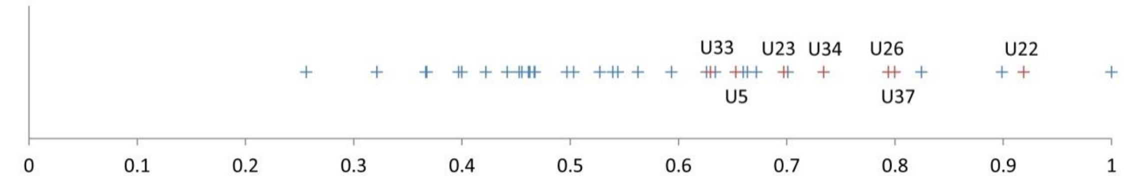


Figure 7-8: Relative Student Satisfaction scores when all play Satisfaction strategy

On the other hand, Figure 7-1 shows that the majority of starting scores in the Student Satisfaction table are ranged between 0.4 and 0.7. The improvements proposed by the DEA model take every DMU to a score of at least 0.6 (see Figure 7-8). If they can achieve this level of improvement then they will overtake the majority of inactive DMUs in this example. Furthermore, those DMUs, such as U26 and U37, that are able to move into the relatively open space above 0.7 can potentially see dramatic improvements in their rank.

The practical implications of these two scenarios are seen particularly clearly in the example of U23. This DMU starts with a high rank in the RAE table and a low rank in the satisfaction table, so one might expect that there would be greater scope for improvement in student satisfaction. The output from the DEA model shows that this is indeed the case, but the difference is magnified when it is viewed in terms of league table rank. No improvement at all in RAE rank is achieved even when it is the main focus of U23's attention and everyone else is looking elsewhere. However, concentrating on student satisfaction could potentially take U23 into the top ten of this table. This strategy risks a slight fall in RAE rank but the example shows U23 maintaining a place in the top ten here as well. The game theory models show that even when other DMUs are playing their best strategies, U23 is still able to hold onto two top ten places. If top ten ranking were a priority for the university, then a focus on student satisfaction would seem to be a sensible strategy.

It is important at this point to consider the value of insights such as these. It is not the intention of this research to direct the decision-makers of U23 or any other HEI to adopt gaming strategies in order to improve league table performance. Indeed, the complexity of this analysis highlights some of the dangers inherent in pursuing rank status whether for its own sake or for the perceived benefits associated with it. However, a deeper understanding of the structures of league tables and the interaction of institutions' ongoing performance is potentially useful both to those who are measured and to those who do the measuring.

7.6.2 Limitations

A key limitation of this work arises from the difficulty in establishing a workable dataset. We considered the needs of the model in some detail in chapter 6 and populated it as precisely as we could within the constraints of time and budget. However, the process of transforming the raw data into something appropriate for use has necessitated a number of assumptions and approximations. The effect of these assumptions on the model could perhaps best be tested by substituting alternatives, but this is also an expensive and time-consuming business. We have some suggestions for further work in chapter 8 which might offer improvements.

Given the dataset we elected to work with, the next most significant limitation is the size of the active cohort in our analysis. In general, our game theoretic reasoning showed that the best response for a DMU in each game was to pursue its own best strategy irrespective of what the other DMUs were doing; although we found that U5 and U23 were both marginal cases whose interests were finely balanced between the two alternatives. In our game scenario we have only seven active DMUs out of a population of 37: the impact on the system of six 'others' has turned out to be less than the impact of the DMU's own choices.

Furthermore, what seems to be sensible from a league table game-playing perspective may be less appealing from another point of view. We found that U37's RAE performance was improved by reducing the overall size of its submission rather than increasing it. But decreasing the number of research outputs by nine papers per year, as the DEA model proposes, would, over five years, be roughly equivalent to omitting ten full-time members of staff from the RAE submission. As well as the personal impact of such an omission on individual staff members, a reduction on this scale would be likely to challenge U37's sense of identity as a research-active department. The application of the DEA model to this DMU raises questions about its size, scale and performance that are not readily answered by looking only at the league table outcomes.

The payoff functions we used in sections 7.5.1 – 7.5.4 were abitrarily chosen and very simply applied, with the assumption that all DMUs would attribute the same value to the possible outcomes. Although we incorporated a number of step changes into the absolute value model in sections 7.5.1 and 7.5.2, our model of gains and losses in 7.5.3 and 7.5.4 was not sophisticated enough to take this into account. It is likely that a step up (or down) of a single rank position at a threshhold, e.g. from 11 to 10 or vice versa, would carry more significance than such a step in general. A more subtle payoff function would vary for each DMU depending on its starting point.

7.7 Summary

In this chapter we have taken the framework developed in chapter 5 and the data specification from chapter 6 and applied them to a specific case using data which describes chemistry teaching and research in UK universities. The results have been considered in detail and various game theoretic approaches used to analyse the outcome. We have presented insights arising from this research and identified some of its limitations. Further work to overcome these limitations is proposed in chapter 8.

8 Conclusion

In this chapter we reflect on the objectives for this research and how they have been met through the work described in this thesis. Suggestions are made for further work which would address some of the limitations identified in chapter 7.

8.1 Research objectives

The aim of this research has been to use quantitative modelling of the league table environment to improve our understanding of the tensions in the university sector which result from the competing demands of different performance measures. Game theory was identified as a suitable framework for analysing the decision process, since the decisions of each individual institution clearly interact to affect the outcomes for all institutions. Data Envelopment Analysis was identified as an appropriate tool to furnish the strategies available to a university in the league table context.

In the introduction to this thesis we identified two key research objectives, namely

> **1.2.1 Develop a quantitative framework for evaluating university decision making in relation to league table rankings**
>
> **1.2.2 Develop a DEA model suitable for determining targets in the context of the university sector**

In the following section we will review these objectives and identify the contribution to knowledge made by this research in respect of each one.

8.2 Contribution to knowledge

8.2.1 Develop a quantitative framework for evaluating university decision making in relation to league table rankings

The development of a quantitative framework is a new approach to this area. We established in Chapter 2 that the research to date has been largely qualitative, with quantitative research focused on the structure of league tables and on individual measures rather than on their interaction. We identified two sorts of tension created by the influence of league tables. Firstly, there is the tension between two different goals which might be sought by one institution and, secondly, the tension between two different institutions seeking the same goal.

The framework proposed in Chapter 5 and implemented in Chapters 6 and 7 enables us to address both these concerns by furnishing us with the resources necessary for a game theoretic analysis of the situation. In section 7.5 our illustration of 'the league table game' clearly shows how this analysis can be applied to institutions with different profiles. It is also shown to operate under a range of alternative assumptions about the value ascribed to teaching and research-related activities. This flexibility is important in the higher education context, which is constantly evolving in the face of new political and economic challenges.

We have noted that the insights obtained from this modelling process are of benefit both to the institutions whose performance is under scrutiny and to those who design the monitoring frameworks. In the present case these include universities, league table compilers and other interested bodies such as the funding councils. However, this framework has wider applicability. In any context where 'league tables' of performance are used there is pressure on the individual units to appear as successful as possible. This pressure can cause tension within and between organisations and give rise to gaming behaviours which may be unhelpful in the longer term – and, indeed, ineffective in the short term. In particular, Camerer and Lovallo (1999) have identified the phenomenon of 'competition neglect' which describes the tendency of decision-makers to overlook the impact of others' actions on their own goals, such as a 'top ten position'. The application of this framework can serve to illuminate these competitive situations by identifying both realistic and unrealistic expectations of performance. It is particularly well-suited to public sector contexts such as education and health, where there is an obligation to share performance data and where external performance measures are widespread, but it could also be employed in other environments, e.g. within a single organisation that encompasses independent decision-making units.

8.2.2 Develop a DEA model suitable for determining targets in the context of the university sector

DEA was identified as a suitable modelling tool for the higher education environment because of its ability to handle multiple inputs and outputs. In order to populate the framework established under our first objective, we needed to develop a DEA model which would permit selective target-setting within a constrained production possibility set. Two models (Podinovski, 2004, Thanassoulis and Dyson, 1992) were identified which each provided one part of the desired model but no existing model offered all the required features. A new and original model was therefore developed which combines both attributes: a PPS constrained by trade-offs and the ability to specify preferences over a mixed selection of inputs and outputs. This model was designed to support the framework described above, but is transferrable to many other contexts.

8.3 Further work

In section 7.6.2 we identified some limitations of this present work. Many of these could be addressed by further research.

8.3.1 Extensions of the present model

Firstly, some steps could be taken which follow directly from the work in Chapters 5, 6 and 7. For instance, the model structure used in this chapter has been constructed around a single game in which each DMU has the choice of two strategies. We have not permitted any mixed strategies in our analysis but this option could be incorporated, for instance by more fully exploiting the capacity of the DEA model to mix changes to inputs and outputs. Alternatively, the game could be reconfigured as two separate games – one focused on the RAE league table and one on the Student Satisfaction league table – and the model analysed using a hypergame structure as proposed by Inohara et al. (1997).

The single example developed in Chapter 7 would be enhanced by comparison with other examples focused on other subjects. Some subjects, such as Medicine, are much more clearly defined in the available datasets while others, such as Information Services, are less so. We noted (in section 6.3) that modelling at the level of the whole institution poses problems, but a careful selection of comparable institutions would make this a useful experiment.

8.3.2 Theoretical developments

Secondly, more work could be done to develop the theoretical model. For example, we could modify the DEA model to allow some inputs to increase and/or some outputs to decrease in order to vary the production mix for each DMU (Thanassoulis and Dyson, 1992). We could also draw on the game theory literature more deeply to develop more sophisticated payoff functions and on the literature of behavioural game theory to incorporate the implications of bounded rationality. Since league tables are an annual phenomenon, it would seem appropriate to consider ‘learning models’ (Camerer et al., 2004) in which players adapt their strategies as a result of prior experience. An interesting feature of more advanced models of this kind is that they accommodate players who have varying levels of sophistication in their approach to forecasting what other players are likely to do (p 153).

In our outline of the analytical framework in Chapter 5 we anticipated a more complex league table structure than we were ultimately able to develop in our worked example in Chapter 7, when we needed to address the challenge of working with a large and complex dataset. Further developments to extend this part of the analysis would be particularly beneficial. This might be achieved through a standalone optimisation process which 'plays out' the league table game for *n* players, but it would potentially be fruitful to revisit the idea of Game Theoretic DEA (see section 4.7) with this application in mind.

In section 5.1 we noted that the proposed framework incorporates the assumption that a strategy, once adopted, can be fully realised. In other words, if we discover a potential strategy *s* for player *P*, then we assume that *s* is fully achievable and evaluate its impact on this basis. A more realistic assumption is that a given target would not be met immediately, but that there would be movement towards this target over several time periods. Therefore, rather than use the deterministic outcomes specified in section 7.4, it might be preferable to specify a probability function over a range of possible values and use this to derive an 'expected rank' for each DMU given a particular strategy. Another option would be to use the structure of Bayesian games (Harsanyi, 1967-8) to model uncertainty about which players have achieved their goals, for example. Finally, Dynamic DEA (e.g. Sengupta (1999), de Mateo et al. (2006)) is a tool which supports an incremental approach to achieving specified goals and it would be interesting to explore its potential in this context.

8.3.3 Empirical developments

Thirdly, more empirical work could be done to strengthen the insights into the university context. A survey of universities would help to establish a better subject mapping between HESA data and departmental structures. This is a large piece of work, however, and as neither the HESA structures nor the universities' own structures are unchanging, it may prove difficult to put such a survey to practical use.

The trade-offs we incorporated into our example in Chapter 7 were highly simplified. The literature on workload allocation in universities is thin and further empirical work would be needed before a more sophisticated trade-off space could be defined.

It would also be interesting to elicit empirical data on the utility of rank outcomes to different university managers. However, as it is not the intention of this research to recommend ‘gaming’ as an institutional response to league tables and ranking systems, a cautious approach to this topic is advised.

8.4 Conclusion

This thesis presents an original and insightful piece of research into the league table environment in which UK universities are operating.

Through the development of a quantitative framework for evaluating university decision making in relation to league table rankings we have provided a new approach to examining the tensions caused by competing and conflicting performance measures. The framework is supported by a new and original DEA model which combines a realistic production possibility set constrained by trade-offs with a weights-based preference structure. The DEA model has many potential applications beyond the present context.

Overall, while limitations exist, this research makes a valuable contribution to knowledge through the development of a new DEA tool and through increasing understanding of the tensions inherent in the current regime of league tables and ranking systems.

9 References

Abbott, M. & Doucouliagos, C. (2003) The efficiency of Australian universities: a data envelopment analysis. *Economics of Education Review,* 22(1)**,** 89-97.

Ackoff, R. L. (1977) Optimization + objectivity = opt out. *European Journal of Operational Research,* 1(1)**,** 1-7.

Ackoff, R. L. (1979) The future of operational research is past. *Journal of the Operational Research Society,* 30**,** 273–280.

Agasisti, T. (2009) Performances and spending efficiency in higher education: a European comparison through non-parametric approaches. *Education Economics*.

Agasisti, T. & Johnes, G. (2009) Beyond frontiers: comparing the efficiency of higher education decision-making units across more than one country. *Education Economics,* 17(1)**,** 59 - 79.

Agasisti, T. & Pérez-Esparrells, C. (2010) Comparing efficiency in a cross-country perspective: the case of Italian and Spanish state universities. *Higher Education,* 59**,** 85-103.

Agasisti, T. & Salerno, C. (2007) Assessing the Cost Efficiency of Italian Universities. *Education Economics,* 15(4)**,** 455 - 471.

Ahn, T., Arnold, V., Charnes, A. & Cooper, W. (1989) DEA and ratio efficiency analyses for public institutions of higher learning in Texas. *Research in Governmental and Nonprofit Accounting,* 5**,** 165-85.

Ahn, T., Charnes, A. & Cooper, W. (1988) Some statistical and DEA evaluations of relative efficiencies of public and private institutions of higher learning. *Socio-Economic Planning Sciences,* 22(6)**,** 259-269.

Allen, R., Athanassopoulos, A., Dyson, R. G. & Thanassoulis, E. (1997) Weight restrictions and value judgments in Data Envelopment Analysis: Evolution, development and future directions. *Annals of Operations Research,* 73**,** 13-34.

Amado, C. A. F. & Dyson, R. G. (2009) Exploring the use of DEA for formative evaluation in primary diabetes care: An application to compare English practices. *Journal of the Operational Research Society,* 60(11)**,** 1469-1482.

Andersen, P. & Petersen, N. C. (1993) A procedure for ranking efficient units in data envelopment analysis. *Management Science,* 39(10)**,** 1261-1264.

Arcelus, F. J. & Coleman, D. F. (1997) An efficiency review of university departments. *International Journal of Systems Science,* 28(7)**,** 721-9.

Athanassopoulos, A. & Gounaris, C. (2001) Assessing the technical and allocative efficiency of hospital operations in Greece and its resource allocation implications. *European Journal of Operational Research,* 133(2)**,** 416-431.

Athanassopoulos, A. D. & Shale, E. (1997) Assessing the Comparative Efficiency of Higher Education in the UK by Means of Data Envelopment Analysis. *Education Economics,* 5(2)**,** 117-134.

Avkiran, N. K. (2001) Investigating technical and scale efficiencies of Australian universities through data envelopment analysis. *Socio-Economic Planning Sciences,* 35(1)**,** 57-80.

Bandiera, O., Larcinese, V. & Rasul, I. (2009) Heterogeneous Class Size Effects: New Evidence from a Panel of University Students. *IZA Discussion Papers.* Bonn: IZA.

Banker, R. D. (1980) A game theoretic approach to measuring efficiency. *European Journal of Operational Research,* 5**,** 262-268.

Banker, R. D. (1984) Estimating most productive scale size using Data Envelopment Analysis. *European Journal of Operational Research,* 17(1)**,** 35-44.

Banker, R. D. (1993) Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation. *Management Science,* 39(10)**,** 1265-1273.

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

Banker, R. D., Charnes, A., Cooper, W. W. & Clarke, R. (1989) Constrained Game Formulations and Interpretations for Data Envelopment Analysis. *European Journal of Operational Research,* 40(3)**,** 299.

Banker, R. D., Cooper, W., Seiford, L. M. & Zhu, J. (2004) Returns to Scale in DEA. In Cooper, W., Seiford, L. M. & Zhu, J. (Eds.) *Handbook on Data Envelopment Analysis.* Dordrecht: Kluwer Academic Publishers.

Banker, R. D. & Morey, R. C. (1986a) Efficiency Analysis for Exogenously Fixed Inputs and Outputs. *Operations Research,* 34(4)**,** 513-521.

Banker, R. D. & Morey, R. C. (1986b) The use of categorical variables in Data Envelopment Analysis. *Management Science,* 32(12)**,** 1613-1627.

Banker, R. D. & Natarajan, R. (2004) Statistical tests based on DEA efficiency scores. In Cooper, W., Seiford, L. M. & Zhu, J. (Eds.) *Handbook on Data Envelopment Analysis.* Dordrecht: Kluwer Academic Publishers.

Barrett, L. & Barrett, P. (2009) The Management of Academic Workloads: Improving Practice in the Sector. Final Report. *Research and Development Series.* London: Leadership Foundation for Higher Education.

Beasley, J. M. (1990) Comparing university departments. *Omega,* 18**,** 171-83.

Beasley, J. M. (1995) Determining the teaching and research efficiencies. *Journal of the Operational Research Society,* 46(4)**,** 441-52.

Belfield, C. R. & Fielding, A. (2001) Measuring the relationship between resources and outcomes in higher education in the UK. *Economics of Education Review,* 20**,** 589-602.

Belton, V. & Vickers, S. P. (1993) Demystifying DEA-A Visual Interactive Approach Based on Multiple Criteria Analysis. *The Journal of the Operational Research Society,* 44(9)**,** 883-896.

Bennett, P. G., Cropper, S. & Huxham, C. S. (1989) Modelling interactive decisions: the hypergame focus. In Rosenhead, J. (Ed.) *Rational Analysis for a Problematic World.* Chichester: John Wiley & Sons.

Bogetoft, P. (1994) Incentive Efficient Production Frontiers: An Agency Perspective on DEA. *Management Science,* 40(8)**,** 959-968.

Bogetoft, P. (1995) Incentives and productivity measurements. *International Journal of Production Economics,* 39(1-2)**,** 67-77.

Bogetoft, P. (2000) DEA and activity planning under asymmetric information. *Journal of Productivity Analysis,* 13(1)**,** 7-48.

Bougnol, M.-L. & Dulá, J. H. (2006) Validating DEA as a ranking tool: An application of DEA to assess performance in higher education. *Annals of Operations Research,* 145(1)**,** 339.

Boussofiane, A., Dyson, R. G. & Thanassoulis, E. (1991) Applied data envelopment analysis. *European Journal of Operational Research,* 52(1)**,** 1-15.

Bratti, M., McKnight, A., Naylor, R. & Smith, J. (2004) Higher education outcomes, graduate employment and university performance indicators. *Journal of the Royal Statistical Society, Series A,* 167(3)**,** 475-496.

Breu, T. M. & Raab, R. L. (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-Economic Planning Sciences,* 28(1)**,** 33-45.

Broad, M., Goddard, A. & Von Alberti, L. (2007) Performance, Strategy and Accounting in Local Government and Higher Education in the UK. *Public Money & Management,* 27(2)**,** 119-126.

Broadbent, J. & Loughlin, R. (2010) *Performance Management of Higher Education: A Conceptual and Empirical Analysis,* Edinburgh: University of Edinburgh Business School Seminar Series

Brockett, P. L., Chang, R. E., Rousseau, J. J., Semple, J., H. & Yang, C. (2004) A comparison of HMO efficiencies as a function of provider autonomy. *Journal of Risk and Insurance,* 71(1)**,** 1.

Browne, J. (2010) *Securing a Sustainable Future for Higher Education: An Independent Review of Higher Education Funding and Student Finance,* www.independent.gov.uk/browne-report [Accessed 12/10/10].

Burrell, G. & Morgan, G. (1979) *Sociological paradigms and organisational analysis : elements of the sociology of corporate life,* Aldershot: Ashgate.

Camanho, A. S. & Dyson, R. G. (1999) Efficiency, size, benchmarks and targets for bank branches: an application of data envelopment analysis. *Journal of the Operational Research Society,* 50(9)**,** 903-915.

Camerer, C. F., Ho, T.-H. & Chong, J. K. (2004) Behavioural Game Theory: Thinking, Learning and Teaching. *Advances in understanding strategic behaviour: game theory, experiments, and bounded rationality: essays in honor of Werner Güth.* New York: Palgrave Macmillan.

Camerer, C. F. & Lovallo, D. (1999) Overconfidence and Excess Entry: An Experimental Approach. *American Economic Review,* 89(1)**,** 306-318.

Carrington, R., Coelli, T. & Rao, D. S. P. (2005) The performance of Australian universities: Conceptual issues and preliminary results. *Economic Papers - Economic Society of Australia,* 24(2)**,** 145-163.

Cave, M., Hanney, S., Kogan, M. & Trevett, G. (1988) *The Use of Performance Indicators in Higher Educatioon: A Critical Analysis of Developing Practice,* London: Jessica Kingsley Publishers.

Centre for Higher Education Research and Information, Open University & Hobsons Research (2008a) Counting what is measured or measuring what counts? Bristol: HEFCE.

Centre for Higher Education Research and Information, Open University & Hobsons Research (2008b) Counting what is measured or measuring what counts? Appendix C: Detailed findings of the analyses of the five league tables. Bristol: HEFCE.

Charnes, A., Clark, C. T., Cooper, W. & Golany, B. (1985) A developmental study of Data Envelopment Analysis in measuring the efficiency of maintenance units in the US Air Forces. *Annals of Operations Research,* 2**,** 95-112.

Charnes, A. & Cooper, W. (1962) Programming with linear fractional functionals. *Naval Research Logistics Quarterly,* 9(3,4).

Charnes, A. & Cooper, W. (1973) An explicit general solution in linear fractional programming. *Naval Research Logistics Quarterly,* 20(3).

Charnes, A., Cooper, W., Lewin, A. & Seiford, L. M. (Eds.) (1994) *Data envelopment analysis: theory, methodology, and application,* Dordrecht: Kluwer Academic Publishers.

Charnes, A., Cooper, W. W., Huang, Z. M. & Sun, D. B. (1990) Polyhedral Cone-Ratio DEA Models with an illustrative application to large commercial banks. *Journal of Econometrics,* 46(1-2)**,** 73-91.

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

Charnes, A., Cooper, W. W. & Rhodes, E. (1981) Evaluating program and managerial efficiency: An application of DEA to program follow-through. *Management Science,* 27(6)**,** 668-697.

Chen, C.-M. (2009) A network-DEA model with new efficiency measures to incorporate the dynamic effect in production networks. *European Journal of Operational Research,* 194(3)**,** 687-699.

Chen, Y., Du, J., Sherman, H. D. & Zhu, J. (2010) DEA model with shared resources and efficiency decomposition. *European Journal of Operational Research,* 207(1)**,** 339-349.

Churchman, C. W. (1979) Paradise regained: a hope for the future of systems design education. In Bayraktar, B., Muller-Merbach, H., Roberts, J. & Simpson, M. (Eds.) *Education in Systems Science.* London: Taylor and Francis.

Coates, H. (2007) Excellent measures precede measures of excellence. *Journal of Higher Education Policy and Management,* 29(1)**,** 87-94.

Coelli, T., Rao, D. S. P. & Battese, G. E. (1998) *An Introduction to Efficiency and Productivity Analysis,* Norwell, MA: Kluwer Academic.

Cokgezen, M. (2009) Technical efficiencies of faculties of economics in Turkey. *Education Economics,* 17(1)**,** 81 - 94.

Colbert, A., Levary, R. R. & Shaner, M. C. (2000) Determining the relative efficiency of MBA programs using DEA. *European Journal of Operational Research,* 125(3)**,** 656-669.

Cook, W. & Zhu, J. (2005) *Modeling Performance Measurement: Applications and Implementation Issues in DEA,* New York: Springer.

Cook, W. & Zhu, J. (2008) CAR-DEA: context-dependent assurance regions in DEA. *Operations Research,* 56(1)**,** 69-78.

Cook, W. D., Seiford, L. A. & Zhu, J. (2004) Models for performance benchmarking: measuring the effect of e-business activities on banking performance. *Omega-International Journal of Management Science,* 32(4)**,** 313-322.

Cook, W. D. & Seiford, L. M. (2009) Data envelopment analysis (DEA) - Thirty years on. *European Journal of Operational Research,* 192(1)**,** 1-17.

Cooper, W., Ruiz, J. L. & Sirvent, I. (2007) Choosing weights from alternative optimal solutions of dual multiplier models in DEA. *European Journal of Operational Research,* 180**,** 443-458.

Cooper, W., Seiford, L. M. & Zhu, J. (2004) Data Envelopment Analysis: History, Models and Interpretations. In Cooper, W., Seiford, L. M. & Zhu, J. (Eds.) *Handbook on Data Envelopment Analysis.* Dordrecht: Kluwer Academic Publishers.

Cooper, W. W., Seiford, L. M. & Tone, K. (2006) *Introduction to data envelopment analysis and its uses : with DEA-solver software and references,* New York: Springer.

Cyrenne, P. & Grant, H. (2009) University decision making and prestige: An empirical study. *Economics of Education Review,* 28(2)**,** 237-248.

Danermark, B., Ekstrom, M., Jakobsen, L. & Karlsson, J. C. (2002) *Explaining Society: Critical realism in the social sciences,* London: Routledge.

de Mateo, F., Coelli, T. & O'Donnell, C. J. (2006) Optimal paths and costs of adjustment in dynamic DEA models: with application to Chilean department stores. *Annals of Operations Research,* 145**,** 211-227.

Department of Education Employment and Workplace Relations (2011) *Higher Education Base Funding Review: Final Report,* http://www.deewr.gov.au/HigherEducation/Policy/BaseReview/Documents/HigherEd_FundingReviewReport.pdf.

Dill, D. & Soo, M. (2005) Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems. *Higher Education,* 49(4)**,** 495-533.

Dixon, S. (2006) The University of Leeds and the British higher education system, 1963-2004. *Northern History,* 43(2)**,** 303-325.

Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S. & Shale, E. (2001) Pitfalls and protocols in DEA. *European Journal of Operational Research,* 132**,** 245-259.

Dyson, R. G. & Shale, E. (2009) *DEA and uncertainty,* Warwick University, 8 - 10 September: OR51

Dyson, R. G. & Shale, E. (2010) Data envelopment analysis, operational research and uncertainty. *Journal of the Operational Research Society,* 61(1)**,** 25-34.

Dyson, R. G. & Thanassoulis, E. (1988) Reducing Weight Flexibility in Data Envelopment Analysis. *Journal of the Operational Research Society,* 39(6)**,** 563-576.

El-Mahgary, S. & Lahdelma, R. (1995) Data envelopment analysis: Visualizing the results. *European Journal of Operational Research,* 83(3), 700-710.

Elias, P. & Purcell, K. (2004) SOC (HE): A classification of occupations for studying the graduate labour market. *Researching Graduate Careers Seven Years On.* Warwick Institute for Employment Research.

Emrouznejad, A. & Amin, G. R. (2009) DEA models for ratio data: Convexity consideration. *Applied Mathematical Modelling,* 33, 486-498.

Farrell, M. J. (1957) The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society, Series A,* 120(III), 253-290.

Filippakou, O., Salter, B. & Tapper, T. (2012) Higher Education as a System: The English Experience. *Higher Education Quarterly,* 66(1), 106-122.

Flegg, A. T., Allen, D. O., Field, K. & Thurlow, T. W. (2004) Measuring the efficiency of British universities: a multi-period data envelopment analysis. *Education Economics,* 12(3), 231 - 249.

Førsund, F. R. (2012) Weight restrictions in DEA: misplaced emphasis? *Journal of Productivity Analysis,* 40(3), 271-283.

Førsund, F. R., Kittelsen, S. A. C. & Krivonozhko, V. E. (2009) Farrell revisited - Visualizing properties of DEA production frontiers. *Journal of the Operational Research Society,* 60(11), 1535-1545.

Fried, H. O., Lovell, C. A. K. & Schmidt, S. (2008) Efficiency and Productivity. In Fried, H. O., Lovell, C. A. K. & Schmidt, S. (Eds.) *The Measurement of Productive Efficiency and Productivity Growth.* Oxford: Oxford University Press.

Golany, B. (1988) An Interactive MOLP Procedure for the Extension of DEA to Effectiveness Analysis. *Journal of the Operational Research Society,* 39(8), 725-734.

Golany, B. & Roll, Y. (1989) An application procedure for DEA. *Omega,* 17(3), 237-250.

Goldstein, H. & Spiegelhalter, D. J. (1996) League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance. *Journal of the Royal Statistical Society. Series A (Statistics in Society),* 159(3)**,** 385-443.

Guardian (2011) *University Guide 2013: Submit your details,* http://www.guardian.co.uk/education/2011/nov/08/university-guide-mapping [Accessed 20/3/12].

Haas, D. A. & Murphy, F. H. (2003) Compensating for non-homogeneity in decision-making units in data envelopment analysis. *European Journal of Operational Research,* 144**,** 530-544.

Haksever, C. & Muragishi, Y. (1998) Measuring value in MBA programmes. *Education Economics,* 6(1)**,** 11-25.

Hao, G., Wei, Q. L. & Yan, H. (2000a) A game theoretical model of DEA efficiency. *The Journal of the Operational Research Society,* 51(11)**,** 1319.

Hao, G., Wei, Q. L. & Yan, H. (2000b) The generalized DEA model and the convex cone constrained game. *European Journal of Operational Research,* 126(3)**,** 515-525.

Harsanyi, J. C. (1966) A General Theory of Rationality in Game Situations. *Econometrica,* 34(3)**,** 613-634.

Harsanyi, J. C. (1967-8) Games with incomplete information played by "Bayesian" players, I - III. *Management Science,* 14(3, 5, 7).

Hazelkorn, E. (2007) The Impact of League Tables and Ranking Systems on Higher Education Decision Making. *Higher Education Management and Policy,* 19**,** 81-105.

Hazelkorn, E. (2009) Rankings and the Battle for World-Class Excellence: Institutional Strategies and Policy Choices. *Higher Education Management and Policy,* 21**,** 47-68.

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

HEFCE (2007) Review of Performance Indicators. HEFCE.

HEFCE (2009) *Recurrent grants for 2009-10: Final allocations*: HEFCE

HEFCE (2009-10) *Percentage of institutions mapping subjects in the sector norm cost centre,* http://www.hefce.ac.uk/media/hefce/content/data/fundingandmonitoringdata/fundingandmonitoringdataexercise/PercentageSN.xls [Accessed 31/7/12].

HEFCE (2010) *2010 Teaching Quality Information data,* http://www.hefce.ac.uk/whatwedo/lt/publicinfo/nationalstudentsurvey/nationalstudentsurveydata/2010/ [Accessed 29/4/2012].

HEFCE (2012a) *HESA funding and monitoring data exercise,* http://www.hefce.ac.uk/data/famd/thefundingandmonitoringdataexercise/hesadata/ [Accessed 31/7/12].

HEFCE (2012b) *REF2014: Research Excellence Framework,* http://www.ref.ac.uk/ [Accessed 22/10/2012].

HEFCE (2012c) TRAC Income and Costs by Activity 2010-11. Bristol: HEFCE.

HEFCE (2012d) *Unistats,* http://unistats.direct.gov.uk [Accessed 2/7/2012].

HEFCE, SFC, HEFCW & DEL (2005) RAE2008: Guidance on submissions. Bristol: HEFCE.

HEFCE, SFC, HEFCW & DEL (2008) *Introduction to the RAE 2008 Quality Profiles,* http://www.rae.ac.uk/results/intro.aspx [Accessed 2/7/12].

Herrero, I. & Algarrada, I. (2009) Is the new ECTS system better than the traditional one? An application to the ECTS pilot-project at the University Pablo de Olavide. *European Journal of Operational Research,* In Press, Accepted Manuscript.

HESA (2004) *Cost Centres and Groupings,* http://www.hesa.ac.uk/content/view/102/136/1/1/ [Accessed 20/3/12].

HESA (2006) *JACS2 - Full listing,* http://www.hesa.ac.uk/index.php?option=com_studrec&task=show_file&Itemid=233&mnl=07051&href=jacs2.html [Accessed 20/3/12].

HESA (2011a) *Headline Statistics 2009/10,* http://www.hesa.ac.uk/ [Accessed 23/01/12].

HESA (2011b) *HESA Cost Centres consultation,* http://www.hesa.ac.uk/index.php/content/view/1953/128/ [Accessed 20/3/12].

HESA (2011c) *Joint Academic Coding System (JACS) Version 3.0,* http://www.hesa.ac.uk/content/view/1776/649/ [Accessed 20/03/12].

HESA (2012a) *Performance Indicators in Higher Education in the UK,* http://www.hesa.ac.uk/content/view/2072/141/ [Accessed 2/7/12].

HESA (2012b) *PIs: Non-continuation rates - technical details,* http://www.hesa.ac.uk/index.php?option=com_content&task=view&id=2053&Itemid=141 [Accessed 2/7/12].

HESA (2012c) *PIs: Projected outcomes - technical notes and detailed information,* http://www.hesa.ac.uk/index.php?option=com_content&task=view&id=2066&Itemid=141 [Accessed 2/7/12].

HESA (2012d) *PIs: Research indicators - technical notes and detailed information,* http://www.hesa.ac.uk/index.php?option=com_content&task=view&id=2070&Itemid=141 [Accessed 22/10/12].

HESA (2012e) Students, qualifiers and leavers in Higher Education: Defintions. (Enquiry 32822 item 2).

Hiely-Rayner, M. (2011) Guardian University Guide 2012: methodology. *The Guardian.* London: Guardian News and Media Ltd.

Hollingsworth, B. & Smith, P. (2003) Use of ratios in data envelopment analysis. *Applied Economics Letters,* 10(11)**,** 733-735.

Inohara, T., Takahashi, S. & Nakano, B. (1997) Integration of games and hypergames generated from a class of games. *Journal of the Operational Research Society,* 48(4)**,** 423-432.

Institute for Higher Education Policy (2009) Impact of College Rankings on Institutional Decision Making: Four Country Case Studies. Washington, DC: Institute for Higher Education Policy.

IntelligentMetrix (2012) *The Guardian University League Table 2013 - Methodology,* http://image.guardian.co.uk/sys-files/Guardian/documents/2012/05/21/TablesMethodology.doc [Accessed 26/6/12].

Jackson, M. C. (2006) John Mingers is a critical realist imperialist. *Journal of the Operational Research Society,* 57**,** 1370-1371.

James, W. (1907) *Pragmatism: A New Name for Some Old Ways of Thinking,* London: Penguin Books.

Jardine, N. (1978) 'Realistic' Realism and the Progress of Science. In Hookway, C. & Pettit, P. (Eds.) *Action and Interpretation: Studies in the Philosophy of the Social Sciences.* Cambridge: Cambridge University Press.

Johnes, G. & Johnes, J. (1993) Measuring the research performance of UK economics departments: an application of data envelopment analysis. *Oxford Economic Papers,* 45**,** 332-47.

Johnes, J. (2003a) Measuring Teaching Efficiency in Higher Education: an Application of Data Envelopment Analysis to Graduates from UK Universities 1993. *Lancaster University Management School Working Paper 2003/007.*

Johnes, J. (2003b) Measuring the Efficiency of Universities: A Comparison of Multilevel Modelling and Data Envelopment Analysis. *Lancaster University Management School Working Paper 2003/082.*

Johnes, J. (2006) Data envelopment analysis and its application to the measurement of efficiency in higher education. *Economics of Education Review,* 25(3)**,** 273-288.

Johnes, J. & Johnes, G. (1995) Research funding and performance in U.K. University Departments of Economics: A frontier analysis. *Economics of Education Review,* 14(3)**,** 301-314.

Johnes, J. & Taylor, J. (1990) *Performance Indicators in Higher Education: UK Universities,* Milton Keynes: Open University Press and The Society for Research into Higher Education.

Johnes, J. & Yu, L. (2006) Measuring the research performance of Chinese higher education institutions using data envelopment analysis. *Lancaster University Management School Working Paper 20036/025.*

Johnson, P. & Duberley, J. (2000) *Understanding management research : an introduction to epistemology,* London: Sage.

Joumady, O. & Ris, C. (2005) Performance in European higher education: A non-parametric production frontier approach. *Education Economics,* 13(2)**,** 189 - 205.

Kahneman, D. & Tversky, A. (1979) Prospect Theory: An Analysis of Decision under Risk. *Econometrica,* 47(2)**,** 263 - 292.

Keating, M. (2005) Higher Education in Scotland and England after Devolution. *Regional and Federal Studies,* 15(4)**,** 423-435.

Kelly, U., McNicoll, I. & Brooks, R. (2008) Towards the estimation of the economic value of the outputs of Scottish HEIs: Next Steps Project. University of Strathclyde/Scottish Funding Council.

King, R. (2009) *Governing Universities Globally: Organizations, Regulation and Rankings,* Cheltenham: Edward Elgar Publishing.

Kocher, M. G., Luptacik, M. & Sutter, M. (2006) Measuring productivity of research in economics: A cross-country study using DEA. *Socio-Economic Planning Sciences,* 40**,** 314-332.

Koopmans, T. C. (1951) An Analysis of Production as an Efficient Combination of Activities. In Koopmans, T. C. (Ed.) *Activity Analysis of Production and Allocation.* New York: Wiley.

KPMG (2012) Review of time allocation methods. Bristol: HEFCE.

Langbein, L. (2008) Management by results: Student evaluation of faculty teaching and the mis-measurement of performance. *Economics of Education Review,* 27(4)**,** 417-428.

Lee, C. & Buckthorpe, S. (2008) Robust Performance Indicators for Non-completion in Higher Education. *Quality in Higher Education,* 14(1)**,** 67-77.

Leonard, R. J. (1995) From Parlor Games to Social Science: von Neumann, Morgenstern and the Creation of Game Theory 1928-1944. *Journal of Economic Literature,* 33(2)**,** 730-761.

Liu, J. S., Lu, L. Y. Y., Lu, W.-M. & Lin, B. J. Y. (2013) Data envelopment analysis 1978–2010: A citation-based literature survey. *Omega.*

Machado, M. P. & Vera-Hernandez, M. (2008) Does class-size affect the academic performance of first year college students? *Work in Progress,* Universidad Carlos III de Madrid.

Madden, G., Savage, S. & Kemp, S. (1997) Measuring Public Sector Efficiency: A Study of Economics Departments at Australian Universities. *Education Economics,* 5(2)**,** 153 - 168.

Marginson, S. (2007) *Global university rankings: where to from here?,* National University of Singapore: Asia-Pacific Association for International Education

Mayfield University Consultants (2012a) *About The Complete University Guide,* http://www.thecompleteuniversityguide.co.uk/about-us/ [Accessed 26/6/12].

Mayfield University Consultants (2012b) *How the League Table Works,* http://www.thecompleteuniversityguide.co.uk/league-tables/key/ [Accessed 26/6/12].

Maynard Smith, J. (1982) *Evolution and the Theory of Games,* Cambridge: Cambridge University Press.

McMillan, M. L. & Chan, W. H. (2006) University Efficiency: A Comparison and Consolidation of Results from Stochastic and Non-stochastic Methods. *Education Economics,* 14(1)**,** 1 - 30.

Michael, S. O. (2005) The cost of excellence: The financial implications of institutional rankings. *The International Journal of Educational Management,* 19(4/5)**,** 365.

Mingers, J. (1992) Technical, Practical and Critical OR: Past, Present and Future? In Alvesson, M. & Wilmott, H. (Eds.) *Critical Management Studies.* London: SAGE Publications.

Mingers, J. (2000) The contribution of critical realism as an underpinning philosophy for OR/MS and systems. *Journal of the Operational Research Society,* 51**,** 1256-1270.

Mingers, J. (2006) A critique of statistical modelling in management science from a critical realist perspective: its role within multimethodology. *Journal of the Operational Research Society,* 57**,** 202-219.

Nash, J. (1951) Non-Cooperative Games. *Annals of Mathematics,* 54(2)**,** 286-295.

Newman, M. (2007) LSE to refocus on teaching. *Times Higher Education.* London: TSL Education Ltd.

Niklasson, L. (1996) Game-like regulation of universities: Will the new regulatory framework for higher education in Sweden work? *Higher Education,* 32(3)**,** 267-282.

Osborne, M. J. (2004) *An Introduction to Game Theory,* Oxford: Oxford University Press.

Pastor, J. T., Lovell, C. A. K. & Tulkens, H. (2006) Evaluating the financial performance of bank branches. *Annals of Operations Research,* 145**,** 321-337.

Pastor, J. T., Ruiz, J. L. & Sirvent, I. (2002) A statistical test for nested radial DEA models. *Operations Research,* 50(4)**,** 728-735.

Podinovski, V. (2005) The explicit role of weight bounds in models of data envelopment analysis. *Journal of the Operational Research Society,* 56(12)**,** 1408-1418.

Podinovski, V. V. (2004) Production trade-offs and weight restrictions in data envelopment analysis. *Journal of the Operational Research Society,* 55**,** 1311-1322.

Podinovski, V. V. (2007a) Computation of efficient targets in DEA models with production trade-offs and weight restrictions. *European Journal of Operational Research,* 181(2)**,** 586-591.

Podinovski, V. V. (2007b) Improving data envelopment analysis by the use of production trade-offs. *Journal of the Operational Research Society,* 58(10)**,** 1261-1270.

Podinovski, V. V. & Bouzdine-Chameeva, T. (2013) Weight Restrictions and Free Production in Data Envelopment Analysis. *Operations Research,* 61(2)**,** 426-437.

Portela, M., Thanassoulis, E. & Simpson, G. (2004) Negative data in DEA: a directional distance approach applied to bank branches. *Journal of the Operational Research Society,* 55(10)**,** 1111–1121.

Portela, M. C. & Thanassoulis, E. (2001) Decomposing school and school-type efficiency. *European Journal of Operational Research,* 132(2)**,** 357-373.

Post, T. & Spronk, J. (1999) Performance benchmarking using interactive data envelopment analysis. *European Journal of Operational Research,* 115(3)**,** 472-487.

Poundstone, W. (1995) *Prisoner's Dilemma: John von Neumann, Game Theory and the Puzzle of the Bomb,* Oxford: Oxford University Press.

QS Intelligence Unit (2011) *Methodology - Faculty Student,* http://www.iu.qs.com/university-rankings/rankings-indicators/methodology-faculty-student/ [Accessed 22/02/2012].

Rapoport, A. (1992) Game Theory Defined: What it is and is Not. *Rationality and Society,* 4(1)**,** 74-82.

Rasmusen, E. (1992) Heterogeneous Players and Specialized Models. *Rationality and Society,* 4(1)**,** 83-94.

Rasmusen, E. (2001) *Games and Information: an introduction to game theory,* Oxford: Blackwell.

REF (2012) *Timetable,* http://www.ref.ac.uk/timetable/ [Accessed 8/1/13].

Research Councils UK (2012) *Frequently Asked Questions on Full Economic Costs (FEC)*: RCUK

Rousseau, J. J. & Semple, J. H. (1995) Two-person ratio efficiency games. *Management Science,* 41(3), 435.

Rousseau, J. J. & Semple, J. H. (1997) Dominant Competitive Factors for evaluating program efficiency in grouped data. *Annals of Operations Research,* 73, 253-276.

Russell Group (2012) *About Us,* http://www.russellgroup.ac.uk/about-russell-group/ [Accessed 25/01/2012].

Salerno, C. (2006) Using Data Envelopment Analysis to Improve Estimates of Higher Education Institutionâ€™s Per-student Education Costs. *Education Economics,* 14(3), 281 - 295.

Samoilenko, S. & Osei-Bryson, K.-M. (2008) Increasing the discriminatory power of DEA in the presence of the sample heterogeneity with cluster analysis and decision trees. *Expert Systems with Applications,* 34, 1568-1581.

Sarrico, C. S. & Dyson, R. G. (2000) Using data envelopment analysis for planning in UK universities - an institutional perspective. *Journal of the Operational Research Society,* 51, 789-800.

Sarrico, C. S., Hogan, S. M., Dyson, R. G. & Athanassopoulos, A. D. (1997) Data envelopment analysis and university selection. *Journal of the Operational Research Society,* 49(12), 1163-1177.

Sayer, R. A. (1992) *Method in social science : a realist approach,* London ; New York: Routledge.

Sellers-Rubio, R., Mas-Ruiz, F. J. & Casado-Diaz, A. B. (2010) University Efficiency: Complementariness versus Trade-off between Teaching, Research and Administrative Activities. *Higher Education Quarterly,* 64(4), 373-391.

Semple, J. (1997) Constrained games for evaluating organizational performance. *European Journal of Operational Research,* 96(1), 103.

Sengupta, J. K. (1999) A dynamic efficiency model using data envelopment analysis. *International Journal of Production Economics,* 62, 209-218.

SFC (2009) *General Fund in support of teaching and research for higher education institutions for academic year 2009-10*: SFC

Shubik, M. (1982) *Game Theory in the Social Sciences: Concepts and Solutions,* Cambridge, MA: The MIT Press.

Simar, L. & Wilson, P. W. (1998) Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models. *Management Science,* 44(1)**,** 49-61.

Simar, L. & Wilson, P. W. (2004) Performance of the bootstrap for DEA estimators and iterating the principle. In Cooper, W. W., Seiford, L. M. & Zhu, J. (Eds.) *Handbook on data envelopment analysis.* Boston: Kluwer.

Slack, N., Chambers, S. & Johnston, R. (2010) *Operations Management,* Harlow, England: Financial Times Prentice Hall.

Sloman, J. (2006) *Economics,* Harlow: Pearson Education Limited.

Smith, J., McKnight, A. & Naylor, R. (2000) Graduate employability: policy and performance in higher education in the UK. *Economic Journal,* 110**,** F382-411.

Smith, J. & Naylor, R. (2001) Determinants of degree performance in UK universities: a statistical analysis of the 1993 student cohort. *Oxford Bulletin of Economics and Statistics,* 63**,** 29-60.

Tajnikar, M. & Debevec, J. (2008) Funding system of full-time higher education and technical efficiency: case of the University of Ljubljana. *Education Economics,* 16(3)**,** 289 - 303.

Targeting Innovation (2009) Number of Postgraduate Research Students in Scotland, Appendix 5: RAE2008 Analysis. Glasgow: The Scottish Further & Higher Education Funding Council.

Tauer, L. W., Fried, H. O. & Fry, W. E. (2007) Measuring Efficiencies of Academic Departments within a College. *Education Economics,* 15(4)**,** 473 - 489.

Thanassoulis, E. & Allen, R. (1998) Simulating Weights Restrictions in Data Envelopment Analysis by Means of Unobserved DMUs. *Management Science,* 44(4)**,** 586-594.

Thanassoulis, E. & Dyson, R. G. (1992) Estimating preferred target input-output levels using data envelopment analysis. *European Journal of Operational Research,* 56(1)**,** 80-97.

Thanassoulis, E., Kortelainen, M. & Allen, R. (2012) Improving envelopment in Data Envelopment Analysis under variable returns to scale. *European Journal of Operational Research,* 218**,** 175-185.

Thanassoulis, E., Kortelainen, M., Johnes, G. & Johnes, J. (2011) Costs and Efficiency of Higher Education Institutions in England: A DEA Analysis. *Journal of the Operational Research Society,* 62**,** 1282-1297.

Thanassoulis, E., Portela, M. C. & Allen, R. (2004) Incorporating value judgments in DEA. In Cooper, W., Seiford, L. & Zhu, J. (Eds.) *Handbook on Data Envelopment Analysis.* Boston: Kluwer Academic Press.

Thomas, L. C. (1986) *Games, Theory and Applications,* Chichester: Ellis Horwood/John Wiley.

Tight, M. (2000) Do League Tables Contribute to the Development of a Quality Culture? Football and Higher Education Compared. *Higher Education Quarterly,* 54(1)**,** 22-42.

Times Newspapers Ltd (2011) *Sunday Times University Guide: Methodology,* http://features.thesundaytimes.co.uk/stug/2011/live/methodology [Accessed 26/6/12].

Times Newspapers Ltd (2012) *The Times Good University Guide: About the Tables,* http://extras.thetimes.co.uk/gooduniversityguide/about [Accessed 26/6/12].

Tofallis, C. (2012) A different approach to university rankings. *Higher Education,* 63(1)**,** 1-18.

Tomkins, C. & Green, R. (1988) An experiment in the use of DEA for evaluating the efficiency of UK university departments of accounting. *Financial Accountability and Management,* 4(2)**,** 147-64.

Trigg, R. (1980) *Reality at risk: a defence of realism in philosophy and the sciences*: Harvester press.

Turner, D. (2005) Benchmarking in universities: league tables revisited. *Oxford Review of Education,* 31(3)**,** 353-371.

Universities UK (2011a) Higher Education in Facts and Figures: Summer 2011. London: Universities UK.

Universities UK (2011b) *An overview of the higher education sector,* http://www.universitiesuk.ac.uk/UKHESector/Pages/OverviewSector.aspx [Accessed 23/01/12].

Usher, A. & Savino, M. (2006) A World of Difference: A Global Survey of University League Tables. *Canadian Education Report Series.* Toronto, ON: Educational Policy Institute.

von Neumann, J. & Morgenstern, O. (1944) *Theory of Games and Economic Behavior,* Princeton, NJ: Princeton University Press.

Waldo, S. (2007) Efficiency in Swedish Public Education: Competition and Voter Monitoring. *Education Economics,* 15(2)**,** 231-251.

Walker, P. (2005) *A Chronology of Game Theory,* http://www.econ.canterbury.ac.nz/personal_pages/paul_walker/gt/hist.htm [Accessed 31 October 2009].

Webb, J. N. (2007) *Game Theory: Decisions, Interaction and Evolution,* London: Springer-Verlag.

West, P. (2009) A Faustian Bargain Institutional Responses to National and International Rankings. *Higher Education Management and Policy,* 21**,** 1-10.

Wildavsky, A. (1992) Indispensable Framework or Just another Ideology?: Prisoner's Dilemma as an Antihierarchical Game. *Rationality and Society,* 4(1)**,** 8-23.

Williams, T. (2008) *Management Science in Practice,* Chichester: John Wiley & Sons.

Wong, Y.-H. B. & Beasley, J. M. (1990) Restricting Weight Flexibility in Data Envelopment Analysis. *Journal of the Operational Research Society,* 41(9)**,** 829-835.

Yorke, M. & Longden, B. (2005) Significant Figures - Performance Indicators and 'League Tables'. London: Standing Conference of Principals.

Yu, G., Wei, Q. L. & Brockett, P. L. (1996) A generalized Data Envelopment Analysis model: a unification and extension of existing methods for efficiency analysis of decision making units. *Annals of Operations Research,* 66(1)**,** 47-89.

Zemsky, R. (2008) The Rain Man Cometh-Again. *The Academy of Management Perspectives,* 22(1)**,** 5.

Zhu, J. (1996) Data Envelopment Analysis with Preference Structure. *The Journal of the Operational Research Society,* 47(1)**,** 136-150.

Appendix A

Implementation of the new DEA model in XPRESS-MP (see Chapter 5)

```
model "Preference model with trade-offs"
uses "mmodbc","mmxprs";

! Three-step computation procedure for a preference-based DEA model with trade-offs
! Final version October 2011

parameters
    ! These values can be overwritten at runtime
    DATAFILE = "operational data"           ! Excel file with names defined for InputData, OutputData etc
    VRS = FALSE                             ! If VRS is true then BCC model is used instead of CCR
end-parameters

! The code for these procedures is given after the main program

forward procedure step_1(rnd:string)
forward procedure step_2(rnd:string)
forward procedure step_3(rnd:string)

! These declarations set up the principal structures for the dataset

declarations
    FILENAME: string
    ! The dataset to be evaluated
    DMU: set of string                          ! Set of DMUs
    OUTPUT: set of string                       ! The names of the output measures
    INPUT: set of string                        ! The names of the input measures
    TRADEOFF: set of string                     ! Identifiers for the trade-off vectors
    EVALDMU: set of string                      ! DMUs to be evaluated
    ROUND: set of string
    INPUTVAL: array(DMU, INPUT) of real             ! Values of the input measures
    OUTPUTVAL: array(DMU, OUTPUT) of real           ! Values of the output measures
    INPUTPREF: array(ROUND, DMU, INPUT) of real     ! Preference weightings for selected inputs
    OUTPUTPREF: array(ROUND, DMU, OUTPUT) of real   ! Preference weightings for selected outputs
    PVAL: array(TRADEOFF, INPUT) of real            ! Trade-off vectors for inputs
    QVAL: array(TRADEOFF, OUTPUT) of real           ! Trade-off vectors for outputs
    ! Results of the evaluation
    NEWINPUTVAL: array(ROUND, EVALDMU, INPUT) of real       ! Input values modified according to preferences
```

```
    NEWOUTPUTVAL: array(ROUND, EVALDMU, OUTPUT) of real     ! Output values modified according to preferences
    REFSET: array(ROUND, EVALDMU, DMU) of real              ! Set of efficient peers for each DMU
end-declarations

! The dataset is read in from the Excel file

FILENAME := "mmodbc.excel:" + DATAFILE + ".xls"

initializations from FILENAME
    INPUTVAL as "InputData"
    OUTPUTVAL as "OutputData"
    EVALDMU as "EvaluationDMUs"
    INPUTPREF as "InputPreferences"
    OUTPUTPREF as "OutputPreferences"
    PVAL as "PValues"
    QVAL as "QValues"
end-initializations

finalize(DMU); finalize(OUTPUT); finalize(INPUT); finalize (TRADEOFF); finalize(ROUND)

! These declarations set up structures for interim results of the process

declarations
    RESULTS = DMU + OUTPUT + INPUT + TRADEOFF
    STEP1RESULTVAL: array(ROUND, EVALDMU, RESULTS) of real
    STEP2RESULTVAL: array(ROUND, EVALDMU, RESULTS) of real
    STEP3RESULTVAL: array(ROUND, EVALDMU, RESULTS) of real
end-declarations

! The three-step process

forall (r in ROUND) do

    step_1(r)

    step_2(r)
```

```
    step_3(r)

end-do

! The results are exported to the Excel file
initializations to FILENAME
    STEP1RESULTVAL as "grow;Step1Results"
    STEP2RESULTVAL as "grow;Step2Results"
    STEP3RESULTVAL as "grow;Step3Results"
    NEWINPUTVAL as "grow;NewInputData"
    NEWOUTPUTVAL as "grow;NewOutputData"
    REFSET as "grow;ReferenceSet"
end-initializations

procedure step_1(rnd:string)
! In step 1 we maximise the changes to the factors selected in
! INPUTPREF and OUTPUTPREF. Selected outputs are increased as much as
! possible, while selected inputs are decreased as much as possible.
! Other inputs and outputs are not changed.

    ! These declarations set up the variables which will be used in the optimisation
    declarations

        ! These variables for the input and output multipliers will only be used for those
        ! factors which have a preference weighting
        PreferredINPUT: set of string
        PreferredOUTPUT: set of string
        InputDecrease: dynamic array(INPUT) of mpvar
        OutputIncrease: dynamic array(OUTPUT) of mpvar

        ! These variables will define the projection of the DMU which is under evaluation
        Lambda: array(DMU) of mpvar             ! reference set coefficients
        Pi: array(TRADEOFF) of mpvar            ! trade-off coefficients

        ! Composite variables used to define the LHS of the constraints
        cOutput: array(OUTPUT) of mpvar
        cInput: array(INPUT) of mpvar
```

```
    ! The constraint variables
    LimOut: array(OUTPUT) of linctr
    LimIn: array(INPUT) of linctr
    LimDec: dynamic array(INPUT) of linctr
    LimInc: dynamic array(OUTPUT) of linctr

    ! The objective function
    WeightedTotal: linctr

end-declarations

writeln("Starting Step 1 ", rnd)
writeln

! Use BCC model if variable returns to scale is required
if VRS then
    sum(d in DMU) Lambda(d) = 1
end-if

! We multiply the lambdas by inputs and outputs and add the trade-off terms.
! These form the LHS of the constraints and are the same for all DMUs
! in the dataset.
forall (o in OUTPUT) cOutput(o) = sum(d in DMU) (OUTPUTVAL(d,o) * Lambda(d)) + sum(t in TRADEOFF) (QVAL(t,o) * Pi(t))
forall (i in INPUT) cInput(i) = sum(d in DMU) (INPUTVAL(d,i) * Lambda(d)) + sum(t in TRADEOFF) (PVAL(t,i) * Pi(t))

    ! We solve the problem for each DMU in turn
    forall (d in EVALDMU) do

        PreferredINPUT := {}
        PreferredOUTPUT := {}
        ! Constraints are set with the appropriate RHS for current DMU

        forall (o in OUTPUT) do
            if OUTPUTPREF(rnd,d,o) <= 0 then
                ! If the output has no preference weighting then the constraint is simply:
                LimOut(o) := cOutput(o) >= OUTPUTVAL(d,o)
```

```
        else
            ! Otherwise we add it to the set of preferred outputs and a multiplier is included on the RHS
            ! of the constraint
            PreferredOUTPUT += {o}
            create(OutputIncrease(o))
            LimOut(o) := cOutput(o) = OUTPUTVAL(d,o) * OutputIncrease(o)

            ! Any output multiplier must be at least 1 (outputs are not allowed to decrease)
            create(LimInc(o))
            LimInc(o) := OutputIncrease(o) >= 1
        end-if
    end-do

    forall (i in INPUT) do
        if INPUTPREF(rnd,d,i) <= 0 then
            ! If the input has no preference weighting then the constraint is simply:
            LimIn(i) := cInput(i) <= INPUTVAL(d,i)
        else
            ! Otherwise we add it to the set of preferred inputs and a multiplier is included on the RHS
            ! of the constraint
            PreferredINPUT += {i}
            create(InputDecrease(i))
            LimIn(i) := cInput(i) = INPUTVAL(d,i) * InputDecrease(i)

            ! No input multiplier can exceed 1 (inputs are not allowed to increase)
            create(LimDec(i))
            LimDec(i) := InputDecrease(i) <= 1
        end-if
    end-do

    ! The objective function is the weighted sum of output multipliers
    ! less the weighted sum of input multipliers
    WeightedTotal := sum(o in PreferredOUTPUT) (OutputIncrease(o) * OUTPUTPREF(rnd,d,o)) - sum(i in PreferredINPUT)
(InputDecrease(i) * INPUTPREF(rnd,d,i))

    maximize(WeightedTotal)
```

```
! Output is written to the interim results file and shown in the output log
writeln("DMU ", d, " ", rnd)
writeln
!forall (p in PreferredOUTPUT) writeln(p)

writeln("Reference set: ")
forall (m in DMU) writeln(m, " = ", getsol(Lambda(m)))
forall (m in DMU) STEP1RESULTVAL(rnd,d,m):= getsol(Lambda(m))
writeln

writeln("Trade-offs: ")
forall (t in TRADEOFF) writeln(t, " = ", getsol(Pi(t)))
forall (t in TRADEOFF) STEP1RESULTVAL(rnd,d,t):= getsol(Pi(t))
writeln

writeln("Factors: ")
forall (i in PreferredINPUT) writeln(i, " = ", getsol(InputDecrease(i)))
forall (i in PreferredINPUT) STEP1RESULTVAL(rnd,d,i):= getsol(InputDecrease(i))
forall (o in PreferredOUTPUT) writeln(o, " = ", getsol(OutputIncrease(o)))
forall (o in PreferredOUTPUT) STEP1RESULTVAL(rnd,d,o):= getsol(OutputIncrease(o))
writeln

! For the prioritised inputs the multipliers are now applied.
! Other inputs are left unchanged.
forall (i in INPUT) do
    if INPUTPREF(rnd,d,i) <= 0 then
        NEWINPUTVAL(rnd,d,i):= INPUTVAL(d,i)
     else
        NEWINPUTVAL(rnd,d,i):= INPUTVAL(d,i) * STEP1RESULTVAL(rnd,d,i)
     end-if
end-do

! For the prioritised outputs the multipliers are now applied.
! Other outputs are left unchanged.
forall (o in OUTPUT) do
    if OUTPUTPREF(rnd,d,o) <= 0 then
        NEWOUTPUTVAL(rnd,d,o):= OUTPUTVAL(d,o)
```

```
                else
                   NEWOUTPUTVAL(rnd,d,o):= OUTPUTVAL(d,o) * STEP1RESULTVAL(rnd,d,o)
                end-if
           end-do

       end-do

   writeln("Step 1 finished ", rnd)
   writeln

end-procedure

procedure step_2(rnd: string)
! In step 2 we maximise the slacks remaining in the unselected factors. The preferred
! inputs and outputs are held at their new levels from step 1 while the remaining outputs
! are increased and inputs decreased in order to move the projection to the efficient
! frontier.

   ! These declarations set up the variables which will be used in the optimisation
   declarations

       ! These variables for the input and output slacks will only be used for those
       ! factors which were not selected in step 1
       OutputSlack: dynamic array(OUTPUT) of mpvar
       InputSlack: dynamic array(INPUT) of mpvar
       MaximumOutputSlack: dynamic array(OUTPUT) of real

       UnpreferredINPUT: set of string
       UnpreferredOUTPUT: set of string

       ! An additional buffer variable is required for the inputs to ensure that
       ! they are not reduced below zero
       InputBuffer: dynamic array(INPUT) of mpvar

       ! These variables will define the projection of the DMU which is under evaluation
       Lambda: array(DMU) of mpvar             ! reference set coefficients
       Pi: array(TRADEOFF) of mpvar            ! trade-off coefficients
```

```
    ! Composite variables used to define the LHS of the constraints
    cOutput: array(OUTPUT) of mpvar
    cInput: array(INPUT) of mpvar
    bInput: dynamic array(INPUT) of mpvar

    ! The constraint variables
    LimOut: array(OUTPUT) of linctr
    LimIn: array(INPUT) of linctr
    LimBuffer: dynamic array(INPUT) of linctr

    ! The objective function
    TotalSlack: linctr

end-declarations

writeln("Starting Step 2 ", rnd)
writeln

! Use BCC model if variable returns to scale is required
if VRS then
    sum(d in DMU) Lambda(d) = 1
end-if

    ! Solve the problem for each DMU in turn
    forall (d in EVALDMU) do

        UnpreferredOUTPUT := {}
        UnpreferredINPUT := {}

        forall (o in OUTPUT) do
            if OUTPUTPREF(rnd,d,o) <= 0 then
                ! For the unpreferred outputs a slack term is introduced to the LHS of the constraint
                UnpreferredOUTPUT += {o}
                create(OutputSlack(o))
                cOutput(o) = sum(m in DMU) (OUTPUTVAL(m,o) * Lambda(m)) + sum(t in TRADEOFF) (QVAL(t,o) * Pi(t)) - OutputSlack(o)
                ! An additional variable captures the current maximum output level
```

```
            create(MaximumOutputSlack(o))
            MaximumOutputSlack(o) := max(m in DMU) OUTPUTVAL(m,o)
            else
            ! For the preferred outputs the LHS of the constraints is the same as in step 1
            cOutput(o) = sum(m in DMU) (OUTPUTVAL(m,o) * Lambda(m)) + sum(t in TRADEOFF) (QVAL(t,o) * Pi(t))
        end-if
        ! For all outputs the constraint is now an equality
        LimOut(o) := cOutput(o) = NEWOUTPUTVAL(rnd,d,o)
    end-do

    ! The total output slack is capped at the total maximum output to prevent an unbounded problem
    sum (o in UnpreferredOUTPUT) OutputSlack(o) <= sum(o in UnpreferredOUTPUT) MaximumOutputSlack(o)

    forall (i in INPUT) do
        if INPUTPREF(rnd,d,i) <= 0 then
            ! For the unpreferred inputs a slack term and a buffer term are introduced on the LHS of the constraint
            UnpreferredINPUT += {i}
            create(InputSlack(i))
            create(InputBuffer(i))
            cInput(i) = sum(m in DMU) (INPUTVAL(m,i) * Lambda(m)) + sum(t in TRADEOFF) (PVAL(t,i) * Pi(t)) + InputBuffer(i) +
InputSlack(i)

            ! The additional buffer constraint is defined for each non-preferred input
            create(bInput(i))
            bInput(i) = sum(m in DMU) (INPUTVAL(m,i) * Lambda(m)) + sum(t in TRADEOFF) (PVAL(t,i) * Pi(t)) + InputBuffer(i)
            create(LimBuffer(i))
            LimBuffer(i) := bInput(i) >= 0
        else
            ! For the preferred inputs the LHS of the constraints is the same as in step 1...
            cInput(i) = sum(m in DMU) (INPUTVAL(m,i) * Lambda(m)) + sum(t in TRADEOFF) (PVAL(t,i) * Pi(t))
        end-if
        ! For all inputs the constraint is an equality
        LimIn(i) := cInput(i) = NEWINPUTVAL(rnd,d,i)
    end-do

    ! The objective function is the total sum of input and output slacks
    TotalSlack := sum(i in UnpreferredINPUT)(InputSlack(i)) + sum(o in UnpreferredOUTPUT) (OutputSlack(o))
```

```
maximize(TotalSlack)

! Output is written to the interim results file and shown in the output log
writeln("DMU ", d, " ", rnd)
writeln

writeln("Reference set: ")
forall (m in DMU) writeln(m, " = ", getsol(Lambda(m)))
forall (m in DMU) STEP2RESULTVAL(rnd,d,m):= getsol(Lambda(m))
writeln

writeln("Trade-offs: ")
forall (t in TRADEOFF) writeln(t, " = ", getsol(Pi(t)))
forall (t in TRADEOFF) STEP2RESULTVAL(rnd,d,t):= getsol(Pi(t))
writeln

writeln("Slacks: ")

forall (i in INPUT) do
    if INPUTPREF(rnd,d,i) <= 0 then
        writeln(i, " = ", getsol(InputSlack(i)))
        writeln("buffer for ",i, " = ", getsol(InputBuffer(i)))
        STEP2RESULTVAL(rnd,d,i) := getsol(InputSlack(i))
    else
        STEP2RESULTVAL(rnd,d,i):= 0
    end-if
    ! For the non-preferred inputs the slacks are now applied...
    NEWINPUTVAL(rnd,d,i) := NEWINPUTVAL(rnd,d,i) - STEP2RESULTVAL(rnd,d,i)
    ! ...and the buffer value is also reported
end-do

forall (o in OUTPUT) do
    if OUTPUTPREF(rnd,d,o) <= 0 then
        writeln(o, " = ", getsol(OutputSlack(o)))
        STEP2RESULTVAL(rnd,d,o) := getsol(OutputSlack(o))
    else
```

```
                    STEP2RESULTVAL(rnd,d,o):= 0
                end-if
                ! For the non-preferred outputs the slacks are now applied
                NEWOUTPUTVAL(rnd,d,o) := NEWOUTPUTVAL(rnd,d,o) + STEP2RESULTVAL(rnd,d,o)
            end-do

            writeln

            writeln("Projected DMU: ")
            forall (i in INPUT) writeln(i, " = ", NEWINPUTVAL(rnd,d,i))
            forall (o in OUTPUT) writeln(o, " = ", NEWOUTPUTVAL(rnd,d,o))
            writeln

            ! Empty the buffer constraint before evaluating the next DMU
            forall(i in INPUT) LimBuffer(i):= 0

        end-do

    writeln("Step 2 finished ", rnd)
    writeln

end-procedure

procedure step_3(rnd: string)
! In step 3 we hold all the inputs and outputs at their new levels and maximse the total of
! the input buffer variables. This gives us optimal values for lambda and pi such that any
! inefficient DMUs have a lambda coefficient of zero and are not included in the reference
! set for the DMU under evaluation.

    ! These declarations set up the variables which will be used in the optimisation
    declarations

        ! The input buffer variables
        BufferedINPUT: set of string
        InputBuffer: dynamic array(INPUT) of mpvar

        ! These variables will define the projection of the DMU which is under evaluation
```

```
    Lambda: array(DMU) of mpvar              ! reference set coefficients
    Pi: array(TRADEOFF) of mpvar             ! trade-off coefficients

    ! Composite variables used to define the LHS of the constraints
    cOutput: array(OUTPUT) of mpvar
    cInput: array(INPUT) of mpvar

    ! The constraint variables
    LimOut: array(OUTPUT) of linctr          ! output constraints for DMU under evaluation
    LimIn: array(INPUT) of linctr            ! input constraints for DMU under evaluation

    ! The objective function
    TotalBuffer: linctr

end-declarations

writeln("Starting Step 3 ", rnd)
writeln

! Use BCC model if variable returns to scale is required
if VRS then
    sum(d in DMU) Lambda(d) = 1
end-if


    ! Solve the problem for each DMU in turn
    forall (d in EVALDMU) do

        BufferedINPUT := {}

        ! The LHS of the output constraints is a straightforward combination of output values
        ! and trade-offs
        forall (o in OUTPUT) cOutput(o) = sum(m in DMU) (OUTPUTVAL(m,o) * Lambda(m)) + sum(t in TRADEOFF) (QVAL(t,o) * Pi(t))

        forall (i in INPUT) do
            ! For the unpreferred inputs the LHS of the constraint includes the buffer term
            if INPUTPREF(rnd,d,i) <= 0 then
```

```
        BufferedINPUT += {i}
        create(InputBuffer(i))
        cInput(i) = sum(m in DMU) (INPUTVAL(m,i) * Lambda(m)) + sum(t in TRADEOFF) (PVAL(t,i) * Pi(t)) + InputBuffer(i)
    else
        ! The LHS for the preferred inputs is straightforward
        cInput(i) = sum(m in DMU) (INPUTVAL(m,i) * Lambda(m)) + sum(t in TRADEOFF) (PVAL(t,i) * Pi(t))
    end-if
end-do

! All the constraints are equalities, with the RHS set to the new projected values
! for the DMU under evaluation
forall (o in OUTPUT) LimOut(o) := cOutput(o) = NEWOUTPUTVAL(rnd,d,o)
forall (i in INPUT) LimIn(i) := cInput(i) = NEWINPUTVAL(rnd,d,i)

! The objective function is the total sum of the input buffer variables
TotalBuffer := sum(i in BufferedINPUT)(InputBuffer(i))

maximize(TotalBuffer)

! Output is written to the interim results file and shown in the output log
writeln("DMU ", d, " ", rnd)
writeln

writeln("Reference set: ")
forall (m in DMU) writeln(m, " = ", getsol(Lambda(m)))
forall (m in DMU) STEP3RESULTVAL(rnd,d,m):= getsol(Lambda(m))
forall (m in DMU) REFSET(rnd,d,m):= STEP3RESULTVAL(rnd,d,m)
writeln

writeln("Trade-offs: ")
forall (t in TRADEOFF) writeln(t, " = ", getsol(Pi(t)))
forall (t in TRADEOFF) STEP3RESULTVAL(rnd,d,t):= getsol(Pi(t))
writeln

writeln("Buffer: ")
forall (i in BufferedINPUT) writeln(i, " = ", getsol(InputBuffer(i)))
forall (i in BufferedINPUT) STEP3RESULTVAL(rnd,d,i):= getsol(InputBuffer(i))
```

```
            forall (i in BufferedINPUT) writeln(i, " = ", getsol(InputBuffer(i)))
            writeln

        end-do

    writeln("Step 3 finished ", rnd)
    writeln

end-procedure

! That's it
end-model
```

Appendix B

The specification of the dataset supplied by HESA Information Provision (see Chapters 6 and 7).

Important Note: All intellectual property rights in the Data supplied by HESA Services and in any database containing the Data compiled by HESA services or HESA are vested and shall remain vested in HESA Services and/or HESA. HESA does not accept responsibility for any inferences or conclusions derived from the data by third parties.

Staff data definitions 2008/09 - 2010/11

Definitions

Enquiry 32822 item 1

Coverage

The HESA Staff record provides data in respect of the characteristics of members of all academic and non-academic staff employed under a contract of employment at a reporting higher education institution (HEI) in the UK. Staff employed under consultancy contracts, or on the basis of payment of fees for services without a contract of employment are not included in the record.

The reporting period for the HESA Staff record is 1 August to 31 July.

The record is collected in three sections; staff person, staff contract and staff grade table. The person table contains one record for every person employed by an institution during the HESA reporting period and contains attributes of the individual such as birth date, gender and ethnicity. Each person's employment with an institution will be governed by a legally-binding contract and each contract that exists is recorded on the contract table. If a person has a single contract with the institution there will be one record on the person table and one record on the contract table. If a person has three contracts with an institution there will be one record on the person table and three records on the contract table.

The range of data required about an individual and the contract(s) that they hold will depend on the nature of those contracts and also the classification of the activity for which the contract exists.

Atypical staff are those members of staff whose contracts involve working arrangements that are not permanent, involve complex employment relationships and/or involve work away from the supervision of the normal work provider. For atypical staff only a minimum data set is required.

Staff (excluding atypical) are those members of staff where one or more of the contracts held during the reporting period cannot be defined as atypical, and includes open-ended/permanent and fixed-term contracts. For these staff there is a requirement to return a wider range of data (which may include salary information and start and end dates of employment and contracts).

Academic staff are defined as academic professionals who are responsible for planning, directing and undertaking academic teaching and research within higher education institutions (HEIs). They also include vice-chancellors, medical practitioners, dentists, veterinarians and other health care professionals who undertake lecturing or research activities.

Non-academic staff are defined as those that do not have an academic employment function. They include managers, non-academic professionals, student welfare workers, secretaries, caretakers and cleaners.

The **HESA staff contract session population** is an indicator of those contracts that were active during the reporting period. Atypical staff contracts are counted in this population.

The HESA staff contract session population is only used in analyses of staff cost centre activity, or when summing full-time equivalents (FTE) from the contract table, during the reporting period.

London Metropolitan University and Liverpool Hope University have requested that their individual level data is not released at this time.

Rounding strategy

Due to the provisions of the Data Protection Act 1998 and the Human Rights Act 1998, HESA implements a strategy in published and released tabulations designed to prevent the disclosure of personal information about any individual. This strategy involves rounding all numbers to the nearest multiple of 5. A summary of this strategy is as follows:

- 0, 1, 2 are rounded to 0
- All other numbers are rounded to the nearest multiple of 5.

So for example 3 is represented as 5, 22 is represented as 20, 3286 is represented as 3285 while 0, 20, 55, 3510 remain unchanged.

This rounding strategy is also applied to total figures, the consequence of which is that the sum of numbers in each row or column rarely matches the total shown precisely. Note that staff data calculated by full person equivalents (FPE) will also be rounded in accordance with this strategy.

Average values, proportions and FTE values prepared by HESA are not usually affected by the above strategy, and are calculated on precise raw numbers. However, percentages calculated on populations which contain 52 or fewer individuals will be suppressed and represented as '..' as will averages based on populations of 7 or fewer.

Full-time equivalent

Staff **full-time equivalent (FTE)** is defined by the contract(s) of employment and is proportioned to each activity's cost centre. FTE indicates the proportion of a full-time year being undertaken over the course of the reporting period 1 August to 31 July. The FTE is therefore counted using a population of staff who were active during the reporting period, not just on a given snapshot date, and uses the HESA staff contract session population.

Terms of employment

Terms of employment describe the type of contract(s) a member of staff has with the higher education institution (HEI) at the date the data is returned to HESA, or date of leaving if earlier.

Open-ended/permanent staff are those who are employed on a contract of employment that states the member of staff as permanent or on an open-ended contract. This includes term-time only staff who are employed on an open-ended contract.

Fixed-term contract staff are those employed for a fixed period of time or have an end date on their contract of employment. This includes staff on rolling fixed-term contracts.

Atypical staff are those whose working arrangements are not permanent, involve complex employment relationships and/or involve work away from the supervision of the normal work provider. These may be characterised by a high degree of flexibility for both the work provider and the working person, and may involve a triangular relationship that includes an agent. Source: Department of Trade and Industry (DTI) Discussion Document on Employment Status, July 2003, paragraph 23.

In addition to this definition from the DTI, some HE specific guidance has been devised by HESA in consultation with HEIs. Atypical contracts meet one or more of the following conditions:

- are for less than four consecutive weeks - meaning that no statement of terms and conditions needs to be issued,
- are for one-off/short-term tasks - for example answering phones during clearing, staging an exhibition, organising a conference. There is no mutual obligation between the work provider and working person beyond the given period of work or project. In some cases individuals will be paid a fixed fee for the piece of work unrelated to hours/time spent,
- involve work away from the supervision of the normal work provider - but not as part of teaching company schemes or for teaching and research supervision associated with the provision of distance learning education,

- involve a high degree of flexibility often in a contract to work as-and-when required - for example conference catering, student ambassadors, student demonstrators.

Terms of employment are grouped as 'Non-atypical' and 'Atypical'

Academic employment function

The academic employment function of a member of staff relates to the academic contract of employment and not the actual work undertaken.

Teaching only staff are those whose contracts of employment state that they are employed only to undertake teaching.

Teaching and research staff are those whose contracts of employment state that they are employed to undertake both teaching and research.

Research only staff are those whose contracts of employment state that the primary academic employment function is research only, even though the contract may include a limited number of hours teaching.

Neither teaching nor research staff are those whose contracted academic employment function is neither teaching nor research, e.g. Vice-Chancellor.

Cost centre groups

In certain analyses cost centres have been assigned into cost centre groups, which reflect both academic similarities and comparable resource requirements.

Medicine, dentistry & health
01 Clinical medicine
02 Clinical dentistry
04 Anatomy & physiology
05 Nursing & paramedical studies
06 Health & community studies
07 Psychology & behavioural sciences
08 Pharmacy & pharmacology.

Agriculture, forestry & veterinary science

03 Veterinary science

13 Agriculture & forestry.

Biological, mathematical & physical sciences

10 Biosciences

11 Chemistry

12 Physics

14 Earth, marine & environmental sciences

24 Mathematics.

Engineering & technology

16 General engineering

17 Chemical engineering

18 Mineral, metallurgy & materials engineering

19 Civil engineering

20 Electrical, electronic & computer engineering

21 Mechanical, aero & production engineering

25 Information technology & systems sciences & computer software engineering.

Architecture & planning

23 Architecture, built environment & planning.

Administrative, business & social studies

26 Catering & hospitality management

27 Business & management studies

28 Geography

29 Social studies

30 Media studies.

Humanities & language based studies & archaeology

31 Humanities & language based studies

35 Modern languages

37 Archaeology.

Design, creative & performing arts

33 Design & creative arts.

Education

34 Education

38 Sports science & leisure studies

41 Continuing education.

In certain analyses cost centres 01 to 41 may be grouped together as **academic cost centres**.

Academic services

51 Total academic services.

Administration & central services

54 Central administration & services

55 Staff & student facilities.

Premises

56 Premises.

Residences & catering

57 Residences & catering.

Cost centres are grouped as ’11 Chemistry’ and ‘Other’

Institution identifiers

INSTID - Institution identifier (INSTID) is the unique identifier allocated to institutions by HESA.

SOC - Occupational coding for higher education staff

The Standard Occupational Classification (SOC) provides a national standard for categorising occupational information. SOC forms the basis of occupational classification in a variety of national surveys that collect statistical information such as the Labour Force

Survey and New Earnings Survey. The utilisation of SOC for classifying occupations within the HE sector therefore both allows for the heterogeneity of occupations that exist and enables comparisons to be made with other sectors of the economy and from a variety of data sources.

However, some difficulties emerge in the direct application of SOC for occupational coding within the HE sector. At the most aggregate level, SOC distinguishes nine broad categories termed Major Groups. The titles associated with these Major Groups, which by necessity have to be general in their nature to encompass all occupations, do not provide an intuitive method of classifying the occupations within HE. Additionally, the coding manuals of the Standard Occupational Classification contain information on many occupations and job titles that are not relevant to the HE sector.

The classification of occupations within higher education has therefore necessitated the development of a variant of the national standard that is relevant for the HE sector. This enables the classification of job titles found within the HE sector to fall into one of 13 broad occupational activities. In certain analyses these 13 activities may also be assigned to one of four activity groups.

Activity group/Activity

Academic staff
2A Academic professionals

Non-academic staff:

Managerial, professional and technical staff
1 Managers
2B Non-academic professionals
3A Laboratory, engineering, building, IT and medical technicians (including nurses)
3B Student welfare workers, careers advisers, vocational training instructors, personnel and planning officers
3C Artistic, media, public relations, marketing and sports occupations

Clerical staff

4A Library assistants, clerks and general administrative assistants

4B Secretaries, typists, receptionists and telephonists

Manual staff

5 Chefs, gardeners, electrical and construction trades, mechanical fitters and printers

6 Caretakers, residential wardens, sports and leisure attendants, nursery nurses and care occupations

7 Retail and customer service occupations

8 Drivers, maintenance supervisors and plant operatives

9 Cleaners, catering assistants, security officers, porters and maintenance workers

In certain analyses the 13 activities may also be grouped as **academic** and **non-academic**:

Academic staff are defined as academic professionals who are responsible for planning, directing and undertaking academic teaching and research within HEIs. They also include vice-chancellors, medical practitioners, dentists, veterinarians and other health care professionals who undertake lecturing or research activities. All academic staff fall into group 2A of the SOC classification, regardless of their discipline (e.g. science, engineering, social sciences, humanities, languages).

Non-academic staff are defined as members of staff who fall into one of the remaining 12 occupational activities such as managers, non-academic professionals, student welfare workers, secretaries, caretakers and cleaners.

Data is restricted to 2A Academic professionals.

2010/11

MERGERS

0176 The University of Wales, Lampeter merged with Trinity University College (0092). This has led to a name change (see below).

NAME CHANGES

0079 The University of Teesside changed to Teesside University.

0080 Thames Valley University changed to The University of West London.
0089 University of Wales Institute, Cardiff changed to Cardiff Metropolitan University.
0101 The Royal Scottish Academy of Music and Drama changed to Royal Conservatoire of Scotland.
0196 UHI Millennium Institute changed to University of the Highlands and Islands.
0176 Due to University of Wales, Lampeter merging with 0092 Trinity University College, University of Wales, Lampeter changed to The University of Wales Trinity Saint David.

2009/10

NAME CHANGE

0030 Ravensbourne College of Design and Communication changed to Ravensbourne

2008/09

MERGERS

0015 Dartington College of Arts merged with 0017 University College Falmouth in 2007/08, but continued to make separate returns for that collection year. A single return has been made in 2008/09.

NAME CHANGES

0092 Trinity College, Carmarthen changed to Trinity University College
0107 Napier University changed to Edinburgh Napier University
0197 The Arts Institute at Bournemouth changed to The Arts University College at Bournemouth
0041 Trinity Laban changed to Trinity Laban Conservatoire of Music and Dance
0040 Leeds Trinity and All Saints changed to Leeds Trinity University College

Students, qualifiers and leavers in Higher Education Institutions 2008/09 - 2009/10

Definitions

Enquiry 32822 item 2

Coverage – student

In general, the HESA Student record is collected in respect of all students registered at a reporting higher education institution (HE institution) who follow courses that lead to the award of a qualification(s) or institutional credit, excluding those registered as studying wholly overseas. The data specification of the record uses the term 'instance' to describe a student's engagement with the institution, which, because a student can have more than one instance of engagement, will exceed the number of students. Unless stated otherwise, student data is based on an instance of engagement. Postdoctoral students are not included in the HESA Student record. Courses involving collaborative or franchising arrangements are administration specific:

In England and Northern Ireland all students included on the Higher Education Students Early Statistics Survey (HESES) return to the Higher Education Funding Council (HEFCE), whether fundable or not, are returned to HESA. This includes all students funded through franchised, associate and regional college arrangements. Students funded through a HEFCE recognised funding consortium or students registered at another institution, although included in the HESES return of the lead institution, are not included within the HESA return of that institution. These students are included in the HESA (or the Data Service) return of the registering institution.

In Wales students included on the HESES return to the Higher Education Funding Council for Wales (HEFCW), whether fundable or not, are returned to HESA regardless of where the student is registered. This includes all students funded through franchise arrangements where the provision is franchised out from the institution. Students who are franchised in to the institution are excluded. The term franchise, also referred to as outreach, in HE in

Wales refers to a HE course taught at an institution (the franchisee) which is not directly in receipt of funding from HEFCW for that course, and for which quality assurance is provided by another Welsh institution (the franchisor). Students taught at institutions in Wales may be registered at the franchisee or franchisor institution. However, students registered at institutions outside Wales, with a Welsh institution providing quality assurance, are not included within the definition of franchised students.

In Scotland students taking articulated or franchised courses at further education (FE) colleges, or other courses at other HE institutions or FE colleges, for the years of such courses for which the institution does not provide any of the teaching input, does not receive any funding or does not receive any tuition fee payment (e.g. from the Student Awards Agency for Scotland) are excluded from the HE institution's return to HESA. In the case of those years of a course for which two or more HE institutions are involved in providing the teaching input and/or receiving funding or tuition fees, only one of the HE institutions includes the students in its returns to HESA. It is up to the institutions concerned to agree between themselves who should be responsible for making the returns to HESA, and for which years of the course (or for which students on a particular year of the course), as seems most appropriate given their administrative arrangements.

If it is known at the beginning of the course that a student will spend a block of eight weeks or more in the UK as part of their programme then they are included on the Student record throughout, and not included in the Aggregate offshore record. For the reporting years in which their location of study is identified as being abroad, the student instance, whilst being collected in the year's Student return, is however excluded from the standard HESA student populations and hence from the standard publication figures.

The reporting period for the HESA Student record is 1 August to 31 July.

Higher education (HE) students for the purpose of HESA's data collection are those students on courses for which the level of instruction is above that of level 3 of the Qualifications and Curriculum Authority (QCA) National Qualifications Framework (NQF) (e.g. courses at the level of Certificate of HE and above).

Further education (FE) students are those students on programmes of study for which the level of instruction is equal to or below that of level 3 of the NQF.

The **HESA standard registration population** has been derived from the HESA Student record, from all registered higher education and further education student instances active at a reporting institution in the reporting period 1 August to 31 July, following courses that lead to the award of a qualification or institutional credit, and ensures that similar activity is counted in a similar way irrespective of when it occurs.

The population splits the student experience into years of study. The first year is deemed to start on the commencement date of the student instance, with second and subsequent years starting on, or near, the anniversary of that date. Student instances are counted once for each year of study. However students who leave within two weeks of their instance start date, or anniversary of their start date, and are on a course of more than two weeks duration, are not included in the standard registration population.

Also excluded from this population are:

1. dormant students (those who have ceased studying but have not formally de-registered)
2. incoming visiting and exchange students
3. postdoctoral student instances
4. instances where the whole of the programme of study is outside of the UK
5. instances where the student has spent, or will spend, more than 8 weeks in the UK but the study programme is primarily outside the UK
6. Training and Development Agency for Schools (TDA) Student Associates Scheme (SAS) and Subject Knowledge Enhancement (SKE) student instances
7. students on sabbatical, and
8. writing-up students.

The HESA standard registration population forms the basis for most counts of first year and continuing student instances.

The **HESA qualifications obtained population** is a count of student instances associated with the award of an HE qualification (excluding HE institutional credits) during the HESA reporting period 1 August to 31 July, which were returned to HESA by 31 October. This includes qualifications awarded from dormant, writing-up and sabbatical status.

Incoming visiting and exchange students are excluded from this population.

London Metropolitan University and Liverpool Hope University have requested that their individual level data is not released at this time.

Coverage – leavers

The **HESA Destinations of Leavers from Higher Education (DLHE) target population** contains all United Kingdom (UK) and European Union (EU) domiciled students reported to HESA for the reporting period 1 August to 31 July as obtaining relevant qualifications and whose study was full-time or part-time (including sandwich students and those writing-up theses). Awards from dormant status are not included in the target population. Relevant qualifications exclude professional qualifications. Officially, the Crown Dependencies of Guernsey, Jersey and the Isle of Man are not part of the UK or the EU. However, they are grouped with and assumed to be part of the UK in the HESA DLHE record.

The data specifications of the Student and DLHE records use the term 'instance' to describe a student's engagement with the institution, which, because a student can have more than one instance of engagement, will exceed the number of students. Unless stated otherwise, DLHE data is based on an instance of engagement.

Relevant qualifications for inclusion in the DLHE record are taken from the qualification awarded to the student instance during the reporting year, usually at the end of an instance. The qualification awarded may be different to the student's qualification aim, and each student instance may have a maximum of two qualifications awarded. Where two relevant qualifications are awarded, the highest award is selected as the relevant qualification for DLHE.

HESA classifies courses according to a framework which aligns with the framework for HE qualifications in England, Wales and Northern Ireland (FHEQ), the Scottish Credit and Qualifications Framework (SCQF) (of which the framework for qualifications of HE institutions in Scotland is a constituent part) and the International Standard Classification of Education (ISCED) and Bologna frameworks. Details are available on the Course.COURSAIM field notes in the HESA Student record coding manual. It includes level M for taught masters degrees, and level H for honours degrees.

Relevant qualifications include: doctorate and masters degrees; other postgraduate qualifications obtained primarily through supervised research at level L; qualifications leading towards obtaining eligibility to register to practice with a health or social care or veterinary statutory regulatory body (at level M, H, I and J); integrated undergraduate/postgraduate taught masters degrees on the enhanced/extended pattern; postgraduate bachelors degrees (at level M and level H); Postgraduate Certificates in Education/Professional Graduate Diplomas in Education and Professional Graduate Certificates in Education; other taught qualifications at level M; qualifications leading towards registration with the Architects Registration Board (Parts 2 and 1) (at level M and level H); Diplomas at level M and H (but excluding those specifically for Teaching in the Lifelong Learning Sector); first degrees with honours/ordinary first degrees (including those leading to qualified teacher status (QTS)/registration with a General Teaching Council (GTC), but excluding those from the intercalated pattern); first degrees with honours on the enhanced/extended pattern at level H; first degrees with honours and diploma; Certificates at level H, graduate diploma/certificates at level H and level I; other qualifications at level H; foundation degrees (including those which on completion meet entry requirement for pre-registration health or social care qualification); Diplomas of Higher Education (DipHE); Higher National Diplomas (HND); Certificates of Education (CertHE); Higher National Certificates (HNC).

The population for the DLHE return does not necessarily represent the full cohort graduating during the reporting period; examples of those excluded are professional qualifications (e.g. associate membership or membership of a body such as the Institute of Bankers).

Eligible DLHE population includes those instances identified in the HESA Student record, that met criteria within the DLHE target population based on location of study, domicile, mode of study, end date of instance and qualification awarded.

Known destination includes leavers within the eligible DLHE population who replied to the DLHE questionnaire providing destination information.

Percentage with known destination is the total of known destination expressed as a percentage of the eligible DLHE population.

Explicit refusal includes leavers within the eligible DLHE population who replied to the DLHE questionnaire explicitly refusing to provide information.

Response includes leavers who replied to the DLHE questionnaire (i.e. known destination *plus* explicit refusals).

Rounding strategy

Due to the provisions of the Data Protection Act 1998 and the Human Rights Act 1998, HESA implements a strategy in published and released tabulations designed to prevent the disclosure of personal information about any individual. This strategy involves rounding all numbers to the nearest multiple of 5. A summary of this strategy is as follows:

- 0, 1, 2 are rounded to 0
- All other numbers are rounded to the nearest multiple of 5

So for example 3 is represented as 5, 22 is represented as 20, 3286 is represented as 3285 while 0, 20, 55, 3510 remain unchanged.

This rounding strategy is also applied to total figures, the consequence of which is that the sum of numbers in each row or column rarely matches the total shown precisely. Note that subject level data calculated by apportionment will also be rounded in accordance with this strategy.

Average values, proportions and FTE values prepared by HESA are not usually affected by the above strategy, and are calculated on precise raw numbers. However, percentages calculated on populations which contain 52 or fewer individuals will be suppressed and represented as '..' as will averages based on populations of 7 or fewer.

Mode of study

(Applicable to HESA populations except the qualifications obtained population)

Full-time includes students recorded as studying full-time, normally required to attend an institution for periods amounting to at least 24 weeks within the year of study, plus those enrolled on a sandwich course (thick or thin), irrespective of whether or not they are in attendance at the institution or engaged in industrial training, and those on a study-related year out of their institution. During that time students are normally expected to undertake periods of study, tuition or work experience which amount to an average of at least 21 hours per week for a minimum of 24 weeks study/placement.

In certain analysis sandwich mode of study is shown separately, defined as follows:

Sandwich includes students enrolled on a sandwich course (thick or thin), irrespective of whether they are in attendance at the institution or engaged in industrial training. During that time students are normally expected to undertake periods of study, tuition or work experience which amount to an average of at least 21 hours per week for a minimum of 24 weeks study/placement.

Part-time includes students recorded as studying part-time, or studying full-time on courses lasting less than 24 weeks, on block release, or studying during the evenings only.

Where analysis includes FE level students, part-time includes those recorded as studying part-time, or studying full-time on courses lasting less than 24 weeks, on block release, or studying during the evenings only, plus those students on FE continuous delivery.

Writing-up and sabbatical includes students who are normally expected to submit a thesis to the institution for examination, have completed the work of their course and are not making significant demands on institutional resources, plus those on sabbatical.

Writing-up students and students on sabbatical are excluded from the HESA standard registration population.

(Applicable to HESA qualifications obtained and leavers populations)

Full-time includes students whose study was recorded as full-time (as described as above), and also includes awards from dormant and writing-up status where the student's mode of study was previously full-time.

Part-time students are those whose study was recorded as part-time (as described above), and also includes awards from dormant and writing-up status where the student's mode of study was previously part-time, and awards given to those on sabbatical.

Level of study - student

Level of study is taken from the course aim of the student.

HESA classifies courses according to a framework which aligns with the framework for HE qualifications in England, Wales and Northern Ireland (FHEQ), the Scottish Credit and Qualifications Framework (SCQF) (of which the framework for qualifications of HE institutions in Scotland is a constituent part) and the International Standard Classification of Education (ISCED) and Bologna frameworks. Details are available at www.hesa.ac.uk/C10051/a/COURSEAIM. It includes level M for taught masters degrees, and level H for honours degrees.

Postgraduate (research) includes doctorate (incorporating New Route PhD), masters degrees and postgraduate diplomas or certificates (not Postgraduate Certificate in Education (PGCE) at level M) studied primarily through research.

Postgraduate (taught) includes doctorate, and masters degrees, postgraduate bachelors degrees at level M and postgraduate diplomas or certificates not studied primarily through research, including Postgraduate Certificate in Education (PGCE) at level M (unless shown separately), Masters in Teaching and Learning, level M Diploma in Teaching in the Lifelong Learning Sector, and professional qualifications.

First degree includes first degrees (including eligibility to register to practice with a health or social care or veterinary statutory regulatory body), first degrees with Qualified Teacher Status (QTS)/registration with a General Teaching Council (GTC), postgraduate bachelors degree at level H, enhanced first degrees (including those leading towards obtaining eligibility to register to practice with a health or social care or veterinary statutory regulatory body), first degrees obtained concurrently with a diploma and intercalated first degrees.

Other undergraduate includes qualification aims equivalent to and below first degree level, including, but not limited to, Professional Graduate Certificate in Education (PGCE) at level H (unless shown separately), foundation degrees (unless shown separately), diplomas in higher education (including those with eligibility to register to practice with a health or social care or veterinary statutory regulatory body), Higher National Diploma (HND), Higher

National Certificate (HNC), Diploma of Higher Education (DipHE), Certificate of Higher Education (CertHE), foundation courses at higher education level, National Vocational Qualification (NVQ)/Scottish Vocational Qualification (SVQ) at NQF levels 4 and 5, post-degree diplomas and certificates at undergraduate level (including those in Teaching in the Lifelong Learning Sector), professional qualifications at undergraduate level, other undergraduate diplomas and certificates including post-registration health and social care courses, other formal higher education qualifications of less than degree standard, institutional undergraduate credit and non-formal undergraduate qualifications.

Level of study – qualifier (Qualification obtained)

Qualification obtained is taken from the qualification awarded to the student during the reporting year, usually at the end of an instance. The qualification awarded may be different to the student's qualification aim, and the student may be awarded more than one qualification during the reporting period.

Qualification obtained is based on the HESA Qualification obtained population and therefore also includes qualifications awarded from dormant, writing-up and sabbatical status.

The groupings are as **Level of study**, except in certain analysis where the following groupings may be used:

Doctorate includes doctorate degrees obtained/not obtained primarily through research and New Route PhD.

Other higher degree includes masters degrees obtained/not obtained primarily through research, Masters in Teaching and Learning, pre-registration masters degrees leading towards obtaining eligibility to register to practice with a health or social care or veterinary statutory regulatory body and postgraduate bachelors degrees at level M.

Other postgraduate qualifications includes supervised research at level D, E and L for institutional credits, National Vocational Qualifications (NVQ) at level M and E, other postgraduate qualifications obtained primarily through research, fellowships, diplomas and certificates at level M, Scottish Vocational Qualification (SVQ) 5, professional taught qualifications at level M other than a masters degrees, Level M Diplomas in Teaching in the Lifelong Learning Sector, and other taught qualifications at level M.

HND/DipHE includes Diplomas of Higher Education (DipHE) (including those leading towards obtaining eligibility to register to practice with a health or social care or veterinary statutory regulatory body) and Higher National Diplomas (HND).

Level of study – leavers (Level of qualification obtained)

Postgraduate qualifications obtained includes doctorate degrees obtained/not obtained primarily through research and New Route PhD; masters degrees obtained/not obtained primarily through research, Masters in Teaching and Learning, pre-registration masters degrees leading towards obtaining eligibility to register to practice with a health or social care or veterinary statutory regulatory body and postgraduate bachelors degrees; postgraduate diplomas, certificates and professional qualifications, Postgraduate Certificates in Education or Professional Graduate Diplomas in Education (unless shown separately); other taught qualifications at level M including those leading towards registration with the Architects Registration Board (Part 2 qualification); Diplomas at level M (but excluding those specifically for Teaching in the Lifelong Learning Sector).

Where **Postgraduate Certificate in Education** is shown separately, this is taken to mean both Postgraduate Certificate in Education and Professional Graduate Diploma in Education.

In analyses where postgraduate qualification obtained is disaggregated into **Doctorate degree**, **Other higher degree** and **Other postgraduate** the following groupings are used:

Doctorate degree qualifications obtained includes doctorate degrees studied primarily through advanced supervised research and those not studied primarily through advanced supervised research, plus New Route PhD.

Other higher degree qualifications obtained includes masters degrees obtained/not obtained primarily through research, Masters in Teaching and Learning, Masters of Business Administration (MBA), pre-registration masters degrees leading towards obtaining eligibility to register to practice with a health or social care or veterinary statutory regulatory body, plus postgraduate bachelors degrees at level M.

Other postgraduate degree qualifications obtained includes other postgraduate qualifications obtained primarily through advanced supervised research; diplomas at level M; other taught qualifications at level M including those leading towards obtaining eligibility to register to practice with a health or social care or veterinary statutory regulatory body, and those leading towards registration with the Architects Registration Board (Part 2 qualification); Diplomas at level M (but excluding those specifically for Teaching in the Lifelong Learning Sector); plus Postgraduate Certificates in Education or Professional Graduate Diplomas in Education.

First degree qualifications obtained includes integrated undergraduate/postgraduate taught masters degrees on the enhanced/extended pattern, including those leading towards obtaining eligibility to register to practice with a health or social care or veterinary statutory regulatory body, and first degrees with honours on the enhanced/extended pattern at level H; first degrees with honours/ordinary first degrees (including those leading to qualified teacher status (QTS)/registration with a General Teaching Council (GTC), but excluding those from the intercalated pattern); first degrees with honours leading towards registration with the Architects Registration Board (Part 1 qualification); pre-registration first degrees with honours/ordinary first degrees leading towards obtaining eligibility to register to practice with a health or social care or veterinary statutory regulatory body; first degrees with honours and diploma; postgraduate bachelors degrees at level H.

Other undergraduate qualifications obtained includes graduate diplomas/certificates at level H; Professional Graduate Certificates in Education (unless shown separately); other qualifications at level H including those leading towards registration with the Architects Registration Board (Part 2 qualification); Certificates at level H, graduate diplomas/certificates at level I; foundation degrees (including those which on completion meet the entry requirement for pre-registration health or social care qualification); Diplomas of Higher Education (DipHE) (including those leading towards obtaining eligibility to register to practice with a health or social care or veterinary statutory regulatory body); Higher National Diplomas (HND); Certificates of Higher Education (CertHE); Higher National

Certificates (HNC); Diplomas at level H (but excluding those specifically for Teaching in the Lifelong Learning Sector).

Class of first degree

The classification of a first degree indicates the qualification class obtained. Certain qualifications obtained at first degree level are not subject to classification of award, notably medical and general degrees. These, together with ordinary degrees and aegrotat qualifications have been included within **Unclassified**. Third class honours, fourth class honours and the pass have been aggregated as **Third/pass**. Lower second and undivided second class honours have been aggregated as **Lower second**.

Domicile

Domicile data is supplied to HESA in the form of postcodes (UK, Guernsey, Jersey and the Isle of Man domiciled students) or country codes. Postcodes are mapped to counties, unitary authorities and UK nations using the Office National Statistics Postcode Directory (ONSPD). Countries are mapped to geographical regions, informed by the National Statistics Country Classification 2006 grouping of countries (www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/national-statistics-country-classification/index.html). Where no data is supplied about the student's domicile, fee eligibility is used to assign to either UK region unknown or Non-European-Union unknown.

United Kingdom domiciled students are those whose normal residence prior to commencing their programme of study was in the UK, and for the purpose of HESA analysis includes Guernsey, Jersey and the Isle of Man. (Officially, the Crown Dependencies of Guernsey, Jersey and the Isle of Man are not part of the UK or the EU.)

Other European Union domiciled students are those whose normal residence prior to commencing their programme of study was in countries which were European Union (EU) members (excluding the UK) at 1 December of the reporting period. This includes Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Gibraltar, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain and Sweden.

Where European Union countries are shown separately, individual country figures exclude those domiciled in the Åland Islands, the Canary Islands, and the French overseas departments of French Guiana, Guadeloupe, Martinique and Réunion. These figures are included in **European Union not otherwise specified**.

Other EEA countries includes the European Economic Area countries of Iceland, Liechtenstein and Norway.

Other Europe includes Albania, Andorra, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Croatia, Cyprus (Non-European-Union), Faroe Islands, Georgia, Kosovo, Macedonia, Moldova, Monaco, Montenegro, Russia, San Marino, Serbia, Svalbard and Jan Mayen, Switzerland, Turkey, Ukraine, Vatican City and Europe not otherwise specified.

Non-European-Union students are those whose normal residence prior to commencing their programme of study was outside the EU. Where Non-EU countries are shown separately, individual country figures exclude the country's overseas territories. These individual country figures are listed within the geographic region in which they lie.

Institution identifiers

INSTID - Institution identifier (INSTID) is the unique identifier allocated to institutions by HESA.

Subject of study and JACS codes

Specification of JACS

All JACS subject codes consist of a letter followed by three digits, the first of them non-zero (except the generic codes described below). The initial letter identifies the subject group, for example F for physical sciences. The initial letter and immediately following digit identify the principal subject, for example F5 astronomy. F500 is a valid JACS code used where there is no need for a higher level of precision, but subjects can be identified more precisely using a second non-zero digit, for example F520 space and planetary sciences, and with even more precision, F521 space science and F522 planetary science. Often it is necessary to consider together all the codes, or all the student numbers, falling within a principal subject, and this is done by referring to it using just the first two characters, so F5 refers to all of astronomy and to total numbers in it, by no means all of which will have code F500. Similarly, F52 refers to the whole of space and planetary sciences.

In 2007/08 a review of a selection of subject areas resulted in the implementation of a revision of the JACS subject codes, JACS2. The full listing of JACS2 can be found at www.hesa.ac.uk/jacs2.

Apportionment at Principal subject level

Although subject areas provide a broad-brush framework for presenting information, a more detailed breakdown to JACS Principal subjects is used in some tables. Again, a process of apportionment is necessary, and the procedure is consistent with that used for Subject areas, as follows:

For split courses not involving an initial teacher training (ITT) component, the apportionment algorithm is as reported by the institution.

ITT students at undergraduate level who also have a specialism subject recorded (typically, secondary ITT students) are apportioned 50% to the 'X1 Training teachers' Principal subject and the remaining 50% is further apportioned according to the algorithm for non-ITT students. Where no subject other than education is recorded, or where the student is on a PGCE course, apportionment is 100% to the X1 Training teachers Principal subject.

Principal subject is grouped as (F1) Chemistry, and 'Other'.

Year of course/ programme

This field indicates the year number of the course that the student is currently studying. This could be different from the year of student if the student has changed course or re-taken a year.

Activity

Activity describes the employment category of the leaver based on the values in the Matrix of employment circumstance and further study as follows:

Full-time paid work only (including self-employed)	A
Part-time paid work only	B
Voluntary/unpaid work only	C
Work and further study	D
Further study only	E
Assumed to be unemployed	F
Not available for employment	G
Other	O
Explicit refusal	X

Work and further study includes those who reported that they were in full-time paid work only (including self-employed), part-time paid work only, voluntary/unpaid work only plus work and further study.

Further study only includes those who gave their employment circumstances as temporarily sick or unable to work/looking after the home or family, not employed but not looking for employment, further study or training, or something else, and who were also either in full-time or part-time study, training or research. It also includes those who were due to start a job within the next month or unemployed and looking for employment, further study or training, and who were also in full-time study, training or research.

Assumed to be unemployed includes those students who gave their employment circumstances as unemployed and looking for employment, further study or training, and who were also either in part-time study, training or research or not studying, plus those who were due to start a job within the next month and who were also either in part-time study, training or research or not studying.

In certain analyses the following groupings of activity may be displayed:

Work only includes those who reported that they were in full-time paid work only (including self-employed), part-time paid work only or voluntary/unpaid work only.

Of those working (including work and further study) includes those who reported that they were in full-time paid work only (including self-employed), part-time paid work only, voluntary/unpaid work only or those who reported that they were in work and further study

Of those studying (including work and further study)/Further study (including work and study) includes those who reported that they were in either further study only or work and further study.

Not available for employment and other includes those who reported that they were either not available for employment or other.

Employment overseas includes those who reported that they were in full-time paid work only (including self-employed), part-time paid work only, voluntary/unpaid work only plus work and further study.

Explicit refusal indicates that a leaver did not answer the question relating to the employment circumstances or study.

Graduate employment marker

This is defined in Elias & Purcell's report 'SOC (HE) A Classification of occupations for studying the graduate labour market'. Categorisations are as follows:

- Graduate
    - which can be further classified as traditional graduate occupations, modern graduate occupations, new graduate occupations, or niche graduate occupations
- Non-graduate

These figures are extracted from the HESA Destination of Leavers from HE (DLHE) Record, using fields 5 Employment circumstances (EMPCIR) and 11 Standard Occupational Classification (SOCDLHE)

This split of SOC 2000 codes was derived from Elias & Purcell's report 'SOC (HE) A Classification of occupations for studying the graduate labour market' (Institute for Employment Research, Warwick). This split of SOC 2000 codes produces four categories of graduate level employment ('Traditional graduate occupations', 'Modern graduate occupations', 'New graduate occupations', 'Niche graduate occupations') which are grouped as follows:

Defined using four character SOC 2000 code groups:

Graduate:

'2216','2212','2314','2113','2311','2214','2322','2432','2112','2444','2411','2329','2213','2321','1182','2211','2111','2215','2431','1137','3551','2451','2313','2423','2312','2315','3215','3223','2125','2452','1114','2419','1212','1113','3431','3412','2316','3432','3411','2126','2319','1134','2121','3229','2131','3564','1181','1112','2132','2442','3416','1184','1136','1111','3221','3214','3222','3552','3433','2434','3568','2124','2443','3543','2422','3415','2433','2129','1171','1135','3539','1222','1123','3422','2122','3512','2421','3561','1131','3414','3232','3121','3567','2127','3111','3449','1235','1132','2317','1141','2441','3535','5414','1172','3212','3565','4114','1231','3562','2128','3421','3534','2123','3231','3566','1133','3520','4137','3123','3537','3563','3413','3115','3132','3532','1225','4111','1151','1183','3319','3541','3114','3119','1121','1185','1142','1122','3442','3434','3536','3544','1173','1152','3542','1239','3533','3531','3218','5245','3211','3312','1221','1226','1162','1163','3113','1224','1211'

This category can be further split into the following:

Traditional graduate occupations:

'2216', '2212', '2314', '2113', '2311', '2214', '2322', '2432', '2112', '2444', '2411', '2329', '2213', '2321', '1182', '2211', '2111', '2215', '2431', '1137', '3551', '2451', '2313', '2423','2312'

Modern graduate occupations:

'2315', '3215', '3223', '2125', '2452', '1114', '2419', '1212', '1113', '3431', '3412', '2316', '3432', '3411', '2126', '2319', '1134', '2121', '3229', '2131', '3564', '1181', '1112', '2132', '2442', '3416', '1184', '1136','1111'

New graduate occupations:

'3221', '3214', '3222', '3552', '3433', '2434', '3568', '2124', '2443', '3543', '2422', '3415', '2433', '2129', '1171', '1135', '3539', '1222', '1123', '3422', '2122', '3512', '2421', '3561', '1131', '3414', '3232', '3121', '3567', '2127', '3111', '3449', '1235','1132'

Niche graduate occupations:

'2317', '1141', '2441', '3535', '5414', '1172', '3212', '3565', '4114', '1231', '3562', '2128', '3421', '3534', '2123', '3231', '3566', '1133', '3520', '4137', '3123', '3537', '3563', '3413', '3115', '3132', '3532', '1225', '4111', '1151', '1183', '3319', '3541', '3114', '3119', '1121', '1185', '1142', '1122', '3442', '3434', '3536', '3544', '1173', '1152', '3542', '1239', '3533', '3531', '3218', '5245', '3211', '3312', '1221', '1226', '1162', '1163', '3113', '1224','1211'

Non-graduate:

'4215','8138','4135','3131','3511','6219','3122','3216','4136','4131','3443','6214','6212','4112','4122','1174','5112','7211','5499','4213','4113','4150','5496','9259','7129','4214','6124','9249','5411','6211','7212','5242','5244','9226','1219','4132','7113','3441','4123','9224','6215','4142','1161','9225','4134','5494','1234','6123','6114','9112','4212','9221','4217','1223','4141','4216','9229','3112','4211','3313','5419','7121','7125','7112','4133','5249','4121','6291','5433','7111','5111','5421','8218','7123','6139','9241','3213','6213','8133','9219','9251','5491','8143','5119','6131','6115','9133','6111','6112','9111','5492','6231','7124','6122','5113','8139','9119','8119','6232','8124','8114','3314','9121','5495','9211','5319','5311','9223','5432','3513','8215','9141','8149','5224','9139','6222','9132','9222','5241','3311','7122','5434','5412','8214','5422','9149','9234','8219','8111','5413','9233','8131','8136','8212','5423','1232','6121','5315','5223','5216','5316','9134','3217','8216','8141','8123','8129','8137','5243','5231','5323','8125','9244','5431','5312','8222','6221','6113','5314','5213','5321','9129','8213','5322','8132','8211','5221','1233','3514','5211','5212','5214','5215','5222','5232','5233','5234','5313','5424','5493','6292','8112','8113','8115','8116','8117','8118','8121','8122','8126','8134','8135','8142','8217','8221','8223',’8229','9131','9231','9232','9235','9239','9242','9243','9245'

2009/10

NAME CHANGE

0030 Ravensbourne College of Design and Communication changed to Ravensbourne

2008/09

MERGERS

0015 Dartington College of Arts merged with 0017 University College Falmouth in 2007/08, but continued to make separate returns for that collection year. A single return has been made in 2008/09.

NAME CHANGES

0092 Trinity College, Carmarthen changed to Trinity University College
0107 Napier University changed to Edinburgh Napier University
0197 The Arts Institute at Bournemouth changed to The Arts University College at Bournemouth
0041 Trinity Laban changed to Trinity Laban Conservatoire of Music and Dance
0040 Leeds Trinity and All Saints changed to Leeds Trinity University College

Students in Higher Education Institutions 2008/09 - 2010/11

Definitions

Enquiry 32822 Item 2A

Coverage

In general, the HESA Student record is collected in respect of all students registered at a reporting higher education institution (HE institution) who follow courses that lead to the award of a qualification(s) or institutional credit, excluding those registered as studying wholly overseas. The data specification of the record uses the term 'instance' to describe a student's engagement with the institution, which, because a student can have more than one instance of engagement, will exceed the number of students. Unless stated otherwise, student data is based on an instance of engagement. Postdoctoral students are not included in the HESA Student record. Courses involving collaborative or franchising arrangements are administration specific:

In England and Northern Ireland all students included on the Higher Education Students Early Statistics Survey (HESES) return to the Higher Education Funding Council (HEFCE), whether fundable or not, are returned to HESA. This includes all students funded through franchised, associate and regional college arrangements. Students funded through a HEFCE recognised funding consortium or students registered at another institution, although included in the HESES return of the lead institution, are not included within the HESA return of that institution. These students are included in the HESA (or the Data Service) return of the registering institution.

In Wales students included on the HESES return to the Higher Education Funding Council for Wales (HEFCW), whether fundable or not, are returned to HESA regardless of where the student is registered. This includes all students funded through franchise arrangements where the provision is franchised out from the institution. Students who are franchised in to the institution are excluded. The term franchise, also referred to as outreach, in HE in Wales refers to a HE course taught at an institution (the franchisee) which is not directly in receipt of funding from HEFCW for that course, and for which quality assurance is provided by another Welsh institution (the franchisor). Students taught at institutions in Wales may be registered at the franchisee or franchisor institution. However, students registered at institutions outside Wales, with a Welsh institution providing quality assurance, are not included within the definition of franchised students.

In Scotland students taking articulated or franchised courses at further education (FE) colleges, or other courses at other HE institutions or FE colleges, for the years of such courses for which the institution does not provide any of the teaching input, does not receive any funding or does not receive any tuition fee payment (e.g. from the Student Awards Agency for Scotland) are excluded from the HE institution's return to HESA. In the case of those years of a course for which two or more HE institutions are involved in providing the teaching input and/or receiving funding or tuition fees, only one of the HE institutions includes the students in its returns to HESA. It is up to the institutions concerned to agree between themselves who should be responsible for making the returns to HESA, and for which years of the course (or for which students on a particular year of the course), as seems most appropriate given their administrative arrangements.

If it is known at the beginning of the course that a student will spend a block of eight weeks or more in the UK as part of their programme then they are included on the Student record throughout, and not included in the Aggregate offshore record. For the reporting years in which their location of study is identified as being abroad, the student instance, whilst being collected in the year's Student return, is however excluded from the standard HESA student populations and hence from the standard publication figures.

The reporting period for the HESA Student record is 1 August to 31 July.

Higher education (HE) students for the purpose of HESA's data collection are those students on courses for which the level of instruction is above that of level 3 of the Qualifications and Curriculum Authority (QCA) National Qualifications Framework (NQF) (e.g. courses at the level of Certificate of HE and above).

The **HESA session population** has been derived from the HESA Student record. It includes all registered higher education and further education student instances active at a reporting institution at any point in the reporting period 1 August to 31 July, following courses that lead to the award of a qualification or institutional credit, except:

1. dormant student instances (those who have ceased studying but have not formally de-registered)
2. incoming visiting and exchange student instances
3. postdoctoral student instances
4. instances where the whole of the programme of study is outside of the UK
5. instances where the student has spent, or will spend, more than 8 weeks in the UK but the study programme is primarily outside the UK
6. Training and Development Agency for Schools (TDA) Student Associates Scheme (SAS) and Subject Knowledge Enhancement (SKE) student instances, and
7. students on sabbatical.

Incoming visiting and exchange students are excluded from the session population in order to avoid an element of double-counting with both outgoing and incoming students being included.

The HESA session population forms the basis for counts of full-time equivalent (FTE) student instances.

London Metropolitan University and Liverpool Hope University have requested that their individual level data is not released at this time.

Rounding strategy

Due to the provisions of the Data Protection Act 1998 and the Human Rights Act 1998, HESA implements a strategy in published and released tabulations designed to prevent the disclosure of personal information about any individual. This strategy involves rounding all numbers to the nearest multiple of 5. A summary of this strategy is as follows:

- 0, 1, 2 are rounded to 0
- All other numbers are rounded to the nearest multiple of 5

So for example 3 is represented as 5, 22 is represented as 20, 3286 is represented as 3285 while 0, 20, 55, 3510 remain unchanged.

This rounding strategy is also applied to total figures, the consequence of which is that the sum of numbers in each row or column rarely matches the total shown precisely. Note that subject level data calculated by apportionment will also be rounded in accordance with this strategy.

Average values, proportions and FTE values prepared by HESA are not usually affected by the above strategy, and are calculated on precise raw numbers. However, percentages calculated on populations which contain 52 or fewer individuals will be suppressed and represented as '..' as will averages based on populations of 7 or fewer.

Full-time equivalent

Full-time equivalent (FTE) data represents the institution's assessment of the full-time equivalence of the student instance during the reporting period 1 August to 31 July.

FTE data is based on the HESA session population, and includes writing-up students.

Institution identifiers

INSTID - Institution identifier (INSTID) is the unique identifier allocated to institutions by HESA.

Cost centre groups

In certain analyses cost centres have been assigned into cost centre groups, which reflect both academic similarities and comparable resource requirements.

Medicine, dentistry & health

01 Clinical medicine

02 Clinical dentistry

04 Anatomy & physiology

05 Nursing & paramedical studies

06 Health & community studies

07 Psychology & behavioural sciences

08 Pharmacy & pharmacology.

Agriculture, forestry & veterinary science

03 Veterinary science

13 Agriculture & forestry.

Biological, mathematical & physical sciences

10 Biosciences

11 Chemistry

12 Physics

14 Earth, marine & environmental sciences

24 Mathematics.

Engineering & technology

16 General engineering

17 Chemical engineering

18 Mineral, metallurgy & materials engineering

19 Civil engineering

20 Electrical, electronic & computer engineering

21 Mechanical, aero & production engineering

25 Information technology & systems sciences & computer software engineering.

Architecture & planning

23 Architecture, built environment & planning.

Administrative, business & social studies

26 Catering & hospitality management

27 Business & management studies

28 Geography

29 Social studies

30 Media studies.

Humanities & language based studies & archaeology

31 Humanities & language based studies

35 Modern languages

37 Archaeology.

Design, creative & performing arts

33 Design & creative arts.

Education

34 Education

38 Sports science & leisure studies

41 Continuing education.

In certain analyses cost centres 01 to 41 may be grouped together as **academic cost centres**.

Academic services

51 Total academic services.

Administration & central services

54 Central administration & services

55 Staff & student facilities.

Premises

56 Premises.

Residences & catering

57 Residences & catering.

Cost centre is grouped as 11 Chemistry and 'Other'

2010/11

MERGERS

0176 The University of Wales, Lampeter merged with Trinity University College (0092). This has led to a name change (see below).

NAME CHANGES

0079 The University of Teesside changed to Teesside University.

0080 Thames Valley University changed to The University of West London.

0089 University of Wales Institute, Cardiff changed to Cardiff Metropolitan University.

0101 The Royal Scottish Academy of Music and Drama changed to Royal Conservatoire of Scotland.

0196 UHI Millennium Institute changed to University of the Highlands and Islands.

0176 Due to University of Wales, Lampeter merging with 0092 Trinity University College, University of Wales, Lampeter changed to The University of Wales Trinity Saint David.

2009/10

NAME CHANGE

0030 Ravensbourne College of Design and Communication changed to Ravensbourne

2008/09

MERGERS

0015 Dartington College of Arts merged with 0017 University College Falmouth in 2007/08, but continued to make separate returns for that collection year. A single return has been made in 2008/09.

Enquiry 32822 Items 3A, 3B and 3C

HE Finance data 2008/09 – 2010/11 (table 7 from Finance Plus Publication for each year)

Notes to tables 2010/11

1. In tables where comparisons are made between HESA FSR figures for the latest financial year 2010/11 and the previous financial year(s), the figures for the previous financial year(s) are those reported in the re-stated financial statements.

2. The list of institutions in the HESA products for 2010/11 has changed:

   The University of Wales Lampeter and Trinity University College have merged to form University of Wales Trinity Saint David.

3. The following institutions have changed their name in the HESA products for 2010/11:

   The Royal Scottish Academy of Music and Drama changed to Royal Conservatoire of Scotland
   The University of Teesside changed to Teesside University

Thames Valley University changed to The University of West London
UHI Millennium Institute changed to University of the Highlands and Islands
University of Wales Institute, Cardiff changed to Cardiff Metropolitan University
The University of Wales Lampeter and Trinity University College have merged to form University of Wales Trinity Saint David.

4. University of London is a confederal organisation. The colleges of the university, shown separately in the HESA reference volumes, are:

Birkbeck College
The Institute of Cancer Research
Central School of Speech and Drama
Courtauld Institute of Art
Goldsmiths College
Heythrop College
Institute of Education
King's College London
London Business School
London School of Economics and Political Science
London School of Hygiene and Tropical Medicine
Queen Mary and Westfield College
Royal Academy of Music
Royal Holloway and Bedford New College
The Royal Veterinary College
St. George's Hospital Medical School
The School of Oriental and African Studies
The School of Pharmacy
University College London.

In addition, the institutes within the umbrella of University of London (Institutes and activities) are:
University of London Institute in Paris
University Marine Biological Station, Millport
School of Advanced Study comprises:

Institute of Advanced Legal Studies
Institute of Classical Studies
Institute of Commonwealth Studies
Institute of English Studies
Institute of Germanic Studies and Romance Studies
Institute of Historical Research
Institute of Musical Research
Institute of Philosophy
Institute for the Study of the Americas
Warburg Institute.

5. The University of Buckingham publishes audited accounts to 31 December each year. Consequently, the income and expenditure in these finance data are prepared from management accounts spanning two financial years, which can be reconciled to the published accounts for 2010 and 2011.

6. The FSR submissions for The University of East Anglia and The University of Essex include figures for income, net return, net assets and deferred capital grants for the University Campus Suffolk, which is a 50:50 joint venture between these two universities.

7. The 2010/11 HESA FSR submission is based on the following institutions' draft consolidated financial statements: Aberystwyth University and Trinity Laban Conservatoire of Music and Dance.

8. Writtle College misreported their capital expenditure data for 2010/11. As a result their data has been suppressed from table 8 and removed from sector totals. The correct figures are shown in Finance_Plus_1011_Table8_Writtle.xls

9. Stranmillis University College misreported their capital expenditure data for 2010/11. As a result their data has been suppressed from table 8 and removed from sector totals. The correct figures are shown in Finance_Plus_1011_Table8_Stranmillis.xls

Notes to tables 2009/10

10. The 2009/10 HESA FSR submission is based on the following institutions' draft consolidated financial statements: Stranmillis University College and Edinburgh College of Art.

11. In 2009/10 The University of Bolton misreported in FSR Table 7 £568K of expenditure in 3a Central administration and services. This should have been reported in 2 Total academic services. The correct figures are as follows: 2 Total academic services: staff costs academic £126K, staff costs other £1,365K, other operating expenses £925K, depreciation £0, Total £2,416K. 3a Central administrations and services: staff costs academic £1,013K, staff costs other £4,040K, other operating expenses £1,945K, depreciation £0, Total £6,998K.

Notes to tables 2008/09

12. In 2008/09 Loughborough University completed both the FSR Table 6a sections 'institutions in England and Northern Ireland only' and head 1d 'EU domicile students' for 'institutions in Wales only' in error. These values have been suppressed from Table 6a, retaining only the HEI's 'home and EU domicile, HE course fees' total figure of £33,602 in 'total HE course fees'. The 2008/09 KFI 9 and KFI 10 calculations for Loughborough University, which include FSR Table 6a 'home and EU domicile students' sub-head figures in the numerator and denominator, and 'non-EU domicile students' figure in the numerator, respectively, have also been suppressed.

Definitions

Coverage

The annual HESA Finance Statistics Return (FSR) is the main source of historical financial information on the total activities of all UK higher education institutions (HEIs). The FSR provides data in respect of the consolidated income and expenditure account, consolidated statement of total recognised gains and losses, consolidated balance sheet and consolidated cash flow statement. The figures recorded for the consolidated income and expenditure account, balance sheet headings, statement of recognised gains and losses and cash flow statement should be the same as those recorded in the HEI's audited/published financial statements. The financial statements should be prepared in accordance with the Statement of recommended practice: accounting for further and higher education (SORP), published by Universities UK (SORP_2007.pdf), and comply with the financial reporting requirements contained in any UK legislation relevant to their constitution, such as the Companies Act and the Charities Act. The FSR uses the principles in the SORP to analyse the financial statements in greater detail than is required for published financial statements.

A copy of the 2010/11 FSR template, used by HEIs to return their data to HESA, can be downloaded from the HESA website. The complete FSR with HE-BCI Survey Collection (2010/11) coding manual can be viewed at www.hesa.ac.uk/C10031. This coding manual contains guidance to HEIs for the return of their finance data, and includes all supporting documentation.

All values in the FSR are returned in units of £1,000 and where necessary rounded to the nearest £1,000.

2010/11 financial data relates to the institutions' financial year 1 August 2010 to 31 July 2011.

2009/10 financial data relates to the institutions' financial year 1 August 2009 to 31 July 2010.

2008/09 financial data relates to the institutions' financial year 1 August 2008 to 31 July 2009.

Region of institution

The allocation of an HEI to a geographical region is done by reference to the administrative centre of that institution. There may be students registered at institutions who are studying in regions other than that of the administrative centre of the institution.

HESA allocates HEIs to Government Office Regions as follows:

North East (NEAS), North West (NWES), Yorkshire and The Humber (YORH), East Midlands (EMID), West Midlands (WMID), East of England (EAST), London (LOND), South East (SEAS), South West (SWES), Scotland (SCOT), Wales (WALE) and Northern Ireland (NIRE).

Although The Open University teaches throughout the UK, its administrative centre is located in South East England, and except where shown separately is counted as a wholly English institution.

Institution identifiers

INSTID - Institution identifier (INSTID) is the unique identifier allocated to institutions by HESA.

Categories of expenditure

Expenditure is analysed by four main categories:

1. Staff costs

This covers the costs of all staff for whom the institution is liable to pay Class 1 National Insurance contributions and/or who have a contract of employment with the institution, and includes any redundancy or restructuring payments (that are not treated as exceptional items) made to these staff.

Academic staff costs includes costs in respect of academic professionals (Standard Occupational Classification Group 2A, as defined in the HESA Staff record), whose primary function is planning, directing and undertaking academic teaching and/or research, paid from within the budgets of academic departments and allocated to the appropriate cost centre. All academic staff are classified to this group regardless of their discipline, and this group includes medical practitioners, dentists, veterinarians and other health care professionals who undertake lecturing or research activities within the HEI.

Other staff costs includes costs in respect of all other staff paid from within the budgets of academic departments and allocated to the appropriate cost centre.

Where the heading **staff costs** appears, this is **academic staff costs** plus **other staff costs** as defined above.

2. Other operating expenses

Other operating expenses includes costs in respect of payments to non-contracted staff or individuals, all other non-staff costs incurred, except for depreciation and interest payable. Equipment that has not been capitalised, expenditure on maintenance contracts and telephone costs (calls, rental and non-capitalised equipment) if not charged to departments are also included in this category.

3. Depreciation

This includes depreciation costs on capitalised equipment according to where the assets being depreciated are located (i.e. academic departments, academic services, administration and central services, premises, research grants and contracts or other expenditure).

4. Interest and other finance costs

This includes costs in respect of interest payable on premises, residences and catering operations (including conferences) and other expenditure.

Academic departments

This includes all expenditure incurred by, or on behalf of, academic departments (including departments of continuing education), and expenditure incurred in connection with special and short courses which is not reimbursable by research councils or other bodies in respect of work carried out on their behalf. There are 34 departmental cost centres to which this expenditure can be attributed:

01 Clinical medicine
02 Clinical dentistry
03 Veterinary science
04 Anatomy & physiology
05 Nursing & paramedical studies

06 Health & community studies

07 Psychology & behavioural sciences

08 Pharmacy & pharmacology

10 Biosciences

11 Chemistry

12 Physics

13 Agriculture & forestry

14 Earth, marine & environmental 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

23 Architecture, built environment & planning

24 Mathematics

25 IT & systems sciences, computer software engineering

26 Catering & hospitality management

27 Business & management studies

28 Geography

29 Social studies

30 Media studies

31 Humanities & language based studies

33 Design & creative arts

34 Education

35 Modern languages

37 Archaeology

38 Sports science & leisure studies

41 Continuing education.

Academic services

This includes expenditure incurred by centralised academic services such as the library and learning resource centres, central computers and computer networks (including maintenance and operating costs), expenditure on centrally run museums, galleries and observatories, and any other general academic services not covered elsewhere.

Administration and central services

This includes expenditure incurred by **Central administration and services**, **General education expenditure**, and **Staff and student facilities**.

(Note: the definitions listed below define the activities for expenditure, against which research grants and contracts income may also be attributed).

Central administration and services includes expenditure in respect of central administrative staff and such payments to Heads of Institutions, Professors, Deans, Tutors, Faculty Officers and the like as are made in respect of central (as distinct from departmental) administrative work. This category also includes expenditure associated with the running costs of an administrative computer system and the following other costs if not charged to their relevant academic cost centre: public relations, advertising, recruitment, removal expenses of all staff, publications (excluding educational publications), rating or council tax advisors, security of wages, bank charges (excluding interest), central postage, superannuation management, expenses of head of institution, legal and audit fees, general insurance costs not included elsewhere and telephone costs where centrally charged.

General education expenditure includes expenditure incurred on examinations, fellowships, scholarships, prizes and other expenditure of a general educational nature. It includes the direct costs of examinations for example of external examiners, salaries, printing, etc. Also included are fee remission and provisions for bad debts in respect of unpaid fees and the following items that cannot be appropriately charged elsewhere: educational publications, public lectures, concerts and exhibitions, subscriptions and contributions to learned societies and similar bodies, contributions to representative bodies and agencies, works of art, contributions to the institution's Press, research projects not returned under other heads, representation at conferences, explorations and expeditions, administration of non-departmental arts centres, widening participation activity and student recruitment costs from home and overseas.

Institutions in Wales are required to separate the expenditure return of **general education expenditure** into **national bursaries**, **institutional specific bursaries and scholarships** and **other general expenditure**. The list of expenditure incurred by these three activities is as detailed under **general education expenditure**.

Staff and student facilities includes expenditure incurred on the provision of facilities and amenities for the use of students and/or staff, e.g. Careers Advisory Service, all grants to student societies, emoluments to wardens of halls of residence, accommodation office, athletic and sporting facilities (excluding maintenance) and the institution's health service.

Premises

This includes all expenditure incurred (whether centrally or departmentally) on the management of premises (including academic buildings, central academic services, art centres, institution’s health service premises, pavilions, sports buildings, etc) and on roads and grounds, except residences and catering.

Repairs and maintenance expenditure includes the maintenance of premises including the pay of staff involved (including estates administrative staff) and maintenance provision charges.

Other expenditure includes rates (the uniform business rate charged by local authorities), payments made for the rental of premises, recurrent costs of energy, water and sewerage, depreciation of all buildings except residential, catering and conference buildings, costs of insuring all premises and their contents, cost of cleaning (i.e. salaries, wages and materials, and payments in respect of contract cleaning) and the cost of portering and security services.

Residences and catering operations (including conferences)

This includes the gross 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. The depreciation costs and financing costs of these operations are included in the appropriate categories of expenditure.

Research grants and contracts

This includes the total of the direct costs attributed to research grants and contracts as detailed for Research grants and contracts income.

Other expenditure

Pension cost adjustment includes any adjustment made to staff pension costs in the income and expenditure account (i.e. the difference between actual contributions made and current service cost figure).

Other includes the total direct costs attributed to other services rendered and all other expenditure not covered above.

This covers all expenditure which increases the value of an institution's (or a subsidiary undertaking's) fixed assets, including the purchase of land, buildings, and those items of equipment which are included in the institution's register of fixed assets and shown in the balance sheet.

Capital expenditure incurred is split into **Residences and catering operations** (including conference operations) and **Other operations** (non-catering and non-conference operations), each sub-divided into expenditure incurred on **Buildings** (land and building projects) and the purchase of **Equipment**.

Sources of capital expenditure funding are categorised as:

Funding body grants includes capital grants allocated by the funding bodies, used to provide assets which have been capitalised.

Retained proceeds of sales includes the contribution from proceeds of sales of exchequer funded properties after surrendering the appropriate amount to the Treasury.

Internal funds includes the amount of internal funds utilised to finance expenditure.

Loans includes all sums borrowed from external sources to fund expenditure.

Other external sources include amounts provided as bequests, donations or all other external sources.