OMEGA Int. J. of Mgmt Sci., Vol. 18, No. 2, pp. 171–183, 1990

0305-0483/90 $3.00 + 0.00

Comparing University Departments

JE BEASLEY

Imperial College, UK

(Received February 1989; in revised form September 1989)

In this paper we present a quantitative model for comparing university departments concerned with the same discipline. This model is based upon ideas drawn from data envelopment analysis. Computational results are given for chemistry and physics departments in the United Kingdom.

Key words—data envelopment analysis, education, efficiency

1. INTRODUCTION

DATA ENVELOPMENT ANALYSIS (DEA) was first put forward by Charnes *et al.* [10] in 1978 and is used for evaluating the (relative) efficiency of decision-making units via weights attached to input/output measures. We shall assume throughout this paper some knowledge of DEA. An outline of DEA is presented in Appendix 1 for readers new to the technique.

Since 1978 there has been an increasing amount of interest in DEA and a recent bibliography [22] contains 300 references. However, there appear to have been only a few applications of DEA to the problem of comparing universities or university departments.

Sizer [23] briefly mentioned DEA (using the term "efficient frontier" and referring to [5, 6]) and commented that he remained to be convinced of the practical value of the technique.

Rhodes and Southwick [20] used DEA to compare the efficiencies of 96 public and 54 private universities in the USA. They used five input measures and six output measures and found that the mean efficiency for private institutions was greater than that for public institutions (see also [21]).

Ahn *et al.* [2] used DEA to compare public and private institutions of higher learning (doctoral-granting universities) in the USA. They used three input measures and three output measures and concluded that public universities were more efficient than private universities.

Ahn *et al.* [1] used DEA to compare 33 colleges and universities in Texas. They used four input measures and four output measures and concluded that DEA offered promise as a tool for evaluating the educational performance of such institutions.

Tomkins and Green [25] applied DEA to the problem of comparing university departments of accounting. They presented results for six DEA models defined using varied input/output measures and concluded that there is

> "some hope that DEA, carefully and sensitively used, can offer additional insights on performance which are not available from other methods of assessment".

Kwimbere [18] conducted a similar study to that reported in [25] for university chemical engineering, mathematics and physics departments.

On a more general note readers interested in performance indicators in higher education are referred to [8] and readers interested in the quantitative modelling work that has been presented in the literature with respect to universities to [33, 34].

In this paper we present a model, based upon DEA, for comparing university departments. In the next section we consider the input/output measures that can be used in such a model.

2. INPUT/OUTPUT MEASURES

In deciding the input/output measures that we can use to compare university departments we need first to consider what, conceptually, are the inputs and outputs for a university department and then to consider the data that are actually available.

Conceptual input/output measures

We shall take the approach in this paper of regarding the only relevant input measures for a university department as financial in nature. This is because it is (essentially) the amount of money that a department spends that determines both the equipment it possesses and the number of academic, support and research staff (and their mix with respect to each other and to the various grades available) employed.

Conceptually it is clear that the primary output of a university department is *increased knowledge*. This increased knowledge can be conveniently classified into two types:

(a) person-specific;

(b) general.

Person-specific increased knowledge is, basically, what has been learnt (locked up inside the brain of) people closely associated with a department. Students [undergraduates (UGs), postgraduates—both on taught courses (PGs T) and doing research (PGs R)] are the usual examples of person-specific increased knowledge.

Plainly it is very difficult to, in practice, measure person-specific increased knowledge, e.g. how would we measure the amount learnt in a particular year by a research student?

General increased knowledge is the books, papers, patents, etc. that flow from a university department. Some particular items in this output are more "important" (of higher quality) than other items.

Available input/output measures

The primary available sources of data on input/output measures for all university departments in the United Kingdom in a particular discipline are [11, 27, 28]. Note here that these data are only available for "cost centres" which may not, for a particular university, correspond to a specific department. However we shall continue to use the term department for convenience.

We shall only concern ourselves here with data for the 1986–1987 academic year (the latest year for which data are available at the time of writing).

Although it is clear that university departments should be compared over a number of years (e.g. equipment expenditure in one year will affect research output in future years) our objective in this paper is to develop a model which:

(a) represents an initial attempt to construct a model for the quantitative comparison of university departments;

(b) highlights the data that are needed, but which are not currently available, to make the model more useful;

(c) can be criticised and enlarged upon by others.

Our objective is not to present a model which is a perfect answer to the problem of comparing university departments but to present a model that is an improvement upon previous efforts.

There is some published information as to how departments are rated with respect to their research. This consists of the University Grants Committee (UGC) research ratings which were sent to universities in May 1986 [28]. These ratings classified departments as being either:

(a) star (outstanding), the highest rating; or

(b) A+, above average; or

(c) A, average; or

(d) A−, below average.

Table 1 shows the data that are available for chemistry departments and Table 2 the corresponding data for physics departments. We choose to consider these two disciplines because of the recent reviews [29, 30] that have been carried out into these disciplines.

In the next section we outline the basic DEA model constructed using the data shown in Tables 1 and 2.

Table 1. Data for chemistry departments

University k	Value Weight	General expenditure (£000s) x_{1k} v_1	Equipment expenditure (£000s) x_{2k} v_2	Research income (£000s) x_{3k}, y_{4k} v_3, u_4	UGs y_{1k} u_1	PGs T y_{2k} u_2	PGs R y_{3k} u_3	Research rating: Star y_{5k} u_5	Research rating: A+ y_{6k} u_6	Research rating: A y_{7k} u_7	Research rating: A− y_{8k} u_8
Aston		446	21	183	62	0	37	0	0	0	1
Bath		670	53	288	137	0	43	0	0	1	0
Birmingham		1459	69	288	225	3	63	0	0	1	0
Bradford		613	95	73	92	0	12	0	0	0	1
Bristol		2043	256	1050	253	27	118	1	0	0	0
Brunel		686	46	436	137	18	27	0	0	0	1
Cambridge		2227	620	981	305	0	159	1	0	0	0
City		696	93	354	81	0	31	0	0	0	1
Durham		1027	148	578	187	0	42	0	0	1	0
East Anglia		1155	113	545	126	31	90	0	1	0	0
Essex		620	115	565	76	5	49	0	0	0	1
Exeter		984	138	198	166	0	32	0	0	1	0
Hull		880	78	488	119	9	29	0	0	1	0
Keele		440	51	217	50	0	20	0	0	0	1
Kent		667	281	111	116	0	29	0	0	0	1
Lancaster		685	50	191	92	11	15	0	0	0	1
Leeds		2545	210	763	320	9	82	0	0	1	0
Leicester		919	61	419	173	0	49	0	0	1	0
Liverpool		1259	82	496	195	0	56	0	1	0	0
London—Birkbeck		734	33	142	46	26	48	0	0	0	1
Imperial		1760	742	1061	167	0	141	1	0	0	0
KQC (Kings)		1487	479	521	240	3	42	0	0	1	0
Q Mary		1106	170	430	164	3	37	0	1	0	0
R Hol and Bed		962	131	152	122	0	33	0	0	0	1
Univ Coll		1238	67	490	157	4	60	0	1	0	0
Loughborough		1208	89	397	158	26	49	0	0	0	1
Manchester		1920	191	544	268	0	81	0	1	0	0
UMIST		1758	196	1162	237	9	105	0	0	1	0
Newcastle		1211	79	540	157	0	52	0	0	1	0
Nottingham		1409	122	527	263	0	94	0	1	0	0
Oxford		3337	654	1780	707	0	211	1	0	0	0
Reading		908	120	336	162	5	36	0	0	0	1
Salford		1492	127	613	152	18	102	0	0	0	1
Sheffield		1346	78	250	223	0	64	0	1	0	0
Southampton		1620	420	1224	199	2	124	1	0	0	0
Surrey		691	65	407	122	2	27	0	0	0	1
Sussex		1324	144	565	189	13	104	0	1	0	0
Warwick		927	148	359	147	0	43	0	0	1	0
York		947	146	724	236	7	54	0	0	1	0
Aberystwyth		370	32	102	58	1	10	0	0	0	1
Bangor		360	73	122	89	0	7	0	0	0	1
Cardiff		849	32	258	158	3	53	0	1	0	0
Swansea		764	89	317	132	0	31	0	0	1	0
UWIST		560	99	196	100	0	24	0	0	0	1
Aberdeen		1029	126	391	164	2	39	0	0	0	1
Dundee		619	21	136	73	0	13	0	0	0	1
Edinburgh		1381	254	812	292	0	71	0	1	0	0
Glasgow		2253	131	360	354	7	94	0	0	1	0
Heriot–Watt		768	38	324	142	0	25	0	0	0	1
St Andrews		696	73	408	121	0	29	0	0	0	1
Stirling		421	18	105	57	0	15	0	0	0	1
Strathclyde		1714	112	945	269	15	77	0	1	0	0

Note: See notation in Appendix 1.

3. THE BASIC DEA MODEL

Having regard to Tables 1 and 2 we define our input/output measures as below.

Input measures

For the academic year (1986–1987) we are considering we use three input measures for a university department:

(1) general expenditure (the majority of this expenditure is on salaries);

(2) equipment expenditure;

(3) research income.

Note here that we regard research income as an *input* measure. This contrasts with previous

Table 2. Data for physics departments

University	Value	General expenditure (£000s)	Equipment expenditure (£000s)	Research income (£000s)	UGs	PGs T	PGs R	Research rating Star	A+	A	A−
		x_{1k}	x_{2k}	x_{3k}, y_{4k}	y_{1k}	y_{2k}	y_{3k}	y_{5k}	y_{6k}	y_{7k}	y_{8k}
k	Weight	v_1	v_2	v_3, u_4	u_1	u_2	u_3	u_5	u_6	u_7	u_8
Bath		528	64	254	145	0	26	0	0	0	1
Birmingham		2605	301	1485	381	16	54	0	1	0	0
Bradford		304	23	45	44	3	3	0	0	0	1
Bristol		1620	485	940	287	0	48	0	1	0	0
Brunel		490	90	106	91	8	22	0	0	0	1
Cambridge		2675	767	2967	352	4	166	1	0	0	0
City		422	0	298	70	12	19	0	0	0	1
Durham		986	126	776	203	0	32	0	0	1	0
East Anglia		523	32	39	60	0	17	0	0	0	1
Essex		585	87	353	80	17	27	0	1	0	0
Exeter		931	161	293	191	0	20	0	0	0	1
Hull		1060	91	781	139	0	37	0	0	0	1
Keele		500	109	215	104	0	19	0	0	0	1
Kent		714	77	269	132	0	24	0	0	0	1
Lancaster		923	121	392	135	10	31	0	0	1	0
Leeds		1267	128	546	169	0	31	0	0	1	0
Leicester		891	116	925	125	0	24	0	1	0	0
Liverpool		1395	571	764	176	14	27	0	1	0	0
London—Birkbeck		990	83	615	28	36	57	0	0	0	1
Imperial		3512	267	3182	511	23	153	1	0	0	0
KQC (Kings)		1451	226	791	198	0	53	0	0	1	0
Q Mary		1018	81	741	161	5	29	0	1	0	0
R Hol and Bed		1115	450	347	148	4	32	0	0	0	1
Univ Coll		2055	112	2945	207	1	47	0	1	0	0
Loughborough		440	74	453	115	0	9	0	0	1	0
Manchester		3897	841	2331	353	28	65	1	0	0	0
UMIST		836	81	695	129	0	37	0	0	0	1
Newcastle		1007	50	98	174	7	23	0	0	1	0
Nottingham		1188	170	879	253	0	38	0	0	1	0
Oxford		4630	628	4838	544	0	217	1	0	0	0
Reading		977	77	490	94	26	26	0	0	1	0
Salford		829	61	291	128	17	25	0	0	0	1
Sheffield		898	39	327	190	1	18	0	0	0	1
Southampton		901	131	956	168	9	50	0	1	0	0
Surrey		924	119	512	119	37	48	0	1	0	0
Sussex		1251	62	563	193	13	43	0	0	1	0
Warwick		1011	235	714	217	0	36	0	1	0	0
York		732	94	297	151	3	23	0	0	1	0
Aberystwyth		444	46	277	49	2	19	0	0	1	0
Bangor		308	28	154	57	0	7	0	0	0	1
Cardiff		483	40	531	117	0	23	0	0	0	1
Swansea		515	68	305	79	7	23	0	0	0	1
Aberdeen		593	82	85	101	1	9	0	0	0	1
Dundee		570	26	130	71	20	11	0	0	0	1
Edinburgh		1317	123	1043	293	1	39	0	0	1	0
Glasgow		2013	149	1523	403	2	51	0	1	0	0
Heriot–Watt		992	89	743	161	1	30	0	0	1	0
St Andrews		1038	82	513	151	13	47	0	0	1	0
Stirling		206	1	72	16	0	6	0	0	0	1
Strathclyde		1193	95	485	240	0	32	0	0	0	1

Note: According to [29] the figures given above for the number of students at Q Mary is an underestimate of the actual numbers.

work [18, 25] in which it was considered to be an output measure, although it is admitted [25] that there is some confusion over its role.

Our view is that research income (corrected for size of department, e.g. research income per academic (as in [11, 32]) or research income per £ of general expenditure) is a measure of the *esteem* in which a department, and its work, is held. Whilst evaluating the success of a department in attracting research income may be important, we regard it as equally (or more) important to evaluate how effective a department is at converting this input (money) into outputs (person-specific and general increased knowledge).

Output measures

For the academic year (1986–1987) we are considering we use eight output measures for a university department.

The first three output measures are concerned with the number of students associated with a department and are:

(1) the number of undergraduates (UGs);

(2) the number of postgraduates on taught courses (PGs T); and

(3) the number of postgraduates who are doing research (PGs R).

Having regard to the previous discussion about the output of a university department and person-specific increased knowledge it is clear that we have the implicit assumption that all students in a particular category (UGs/PGs T/PGs R) receive the same amount of increased knowledge irrespective of who they are and irrespective of the department they are associated with. Given the data currently available there appears to be little choice but to make this assumption.

With regard to research output in terms of quantity data concerning publications (and/or citations) has been collected for a number of (relatively small) disciplines (e.g. accounting [15, 16], economics [17], politics [12]) but is not available for chemistry and physics departments (but see [7]). Rather than neglect this factor we decided to use as a *proxy* for research output in terms of quantity the actual amount of research income. Hence the fourth output measure for a department is:

(4) research income.

Let us be clear here, we are not advocating the use of research income as both an input and output measure. Rather we are using research income as a proxy for an output measure which is important but for which no data are available.

Note here however that data relating to publications and citations over a ten year period for all United Kingdom university departments should become available, at least to policy-makers, in the future [26].

With regard to research output in terms of quality the final four output measures for a department are defined by the UGC research ratings and are:

(5) if a department is rated star;

(6) if a department is rated A+;

(7) if a department is rated A;

(8) if a department is rated A−.

Whilst we are aware that these research ratings have been criticised (e.g. see [14, 24]) we are not aware of any other comparative data on quality of research for chemistry and physics departments (but see [7]).

Note here that the work presented in [25] used data relating to publications over a two year timescale to assess research output in terms of quantity (but did not include any assessment of research quality). The work presented in [18] did not include data relating to research quantity or quality.

We would also note here that the 1989 research rating exercise will classify departments into one of five (instead of four) categories [3].

Model

The basic model for comparing a university department with its peers is simply DEA (Appendix 1) with the input/output measures discussed above (where the notation used is as defined in Appendix 1 and Tables 1 and 2 with $s = 8$ and $t = 3$). This model, except for differences in the input/output measures used, is the same as that presented in [18, 25].

One advantage of this model is that it does not require splitting departmental inputs between teaching and research. As the reader will probably appreciate attempting to decide such a split is problematic (e.g. see [4]).

Note here that although this model could be used to compare university departments concerned with *different* disciplines with each other our view would be that, for obvious reasons, it is only valid to use the model to compare departments concerned with the *same* discipline.

We would also take the view that comparing entire universities using DEA (such as in [1, 2, 20, 21]) can be misleading as the results obtained may have nothing to do with efficiency but may be due to the different balance and mix of disciplines present in different universities.

In the next section we illustrate how the basic model can be improved.

4. MODEL IMPROVEMENT

It is possible to improve the basic model to better represent the relative importance of input/output measures. Such an exercise also highlights, for the policy-maker, the issue of what is expected of university departments in a particular discipline. We illustrate this below.

Output improvements

(1) Students. There would probably be general agreement that the weights should satisfy $u_3 \geq u_2 \geq u_1$, i.e. 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 and correspondingly for undergraduates.

As such the constraints $u_3 \geq u_2 \geq u_1$ could be added to the basic model. However, we can go further. To illustrate this we, for the purposes of this study, used:

$$u_3 \geq 1.25u_2 \geq 1.25^2 u_1 \quad (1)$$

$$u_3 \leq 2u_1 \quad (2)$$

Equation (1) ensures that the weight associated with a postgraduate doing research is at least 25% greater than the weight associated with a postgraduate on a taught course and correspondingly for undergraduates. Equation (2) ensures that the weight associated with a postgraduate doing research is at most twice that associated with an undergraduate.

It is clear that similar equations can be constructed to reflect any view that policy-makers might take with respect to the weights (relative importance) of each of these three categories of student.

(2) Numbers. Examination of the Croham review of the University Grants Committee [13] reveals that (ignoring special factors) 63.75% of the total grant made to universities is for teaching, with the remainder (36.25%) being for research. For the purposes of this study this will be taken as indicative of the relative importance attached to teaching (student numbers) output and research output by policy-makers.

Considering student numbers it is clear that we could constrain the proportion of total output associated with student numbers, for a particular department p, to be exactly 0.6375 (this constraint would be $S(1,3,p)/S(1,8,p) = 0.6375$ where $S(i,j,p)$ represents the total output for department p concerned with output measures i to j; see notation in Appendix 1). However it is plainly unrealistic to expect that, for a particular department p, the proportion of its total output related to student numbers should be *exactly* 0.6375. It is clear that, both for a particular department and for the entire set of departments, some flexibility is both necessary and desirable.

For the purposes of this study we take this flexibility as $\pm 20\%$ for both department p (for which we are maximising e_p) *and* for the entire set of departments considered as a whole. This allows the proportion of total output associated with student numbers to be between 0.8(0.6375) and 1.2(0.6375), i.e. between 51.0% and 76.5% of the total output is related to student numbers.

Note here however that we recognise that policy-makers may prefer to allow individual departments greater flexibility than the entire set of departments.

Hence the constraints relating to the proportion of total output associated with student numbers are (see notation in Appendix 1):

$$0.510 \leq S(1,3,p)/S(1,8,p) \leq 0.765 \quad (3)$$

$$0.510 \leq S(1,3,—)/S(1,8,—) \leq 0.765 \quad (4)$$

Note here that it is possible that in choosing the weights to maximise e_p we choose a set of weights which, for some other department q, mean that the proportion of total output associated with student numbers in that department $[S(1,3,q)/S(1,8,q)]$ does not fall within the limits (0.510 to 0.765) defined. If we wish to ensure that this never happens then we merely need to enforce equation (3) for all departments, not just for the department p for which e_p is being maximised.

Equations (3) and (4) implicitly impose constraints upon the proportion of total output not associated with student numbers (essentially research output) and so we shall not explicitly constrain that proportion of total output. Note here that this means that these equations ensure that university departments must do *both* teaching and research—merely doing one or the other is not sufficient.

(3) Research. Also from [13] we have that, of the total grant made to universities based on research, 83.05% is related to a 'floor' level of support for research and to an assessment of the quality of research. Hence, again allowing a flexibility of ±20% we, for the purposes of this study, decided to model this using the constraints:

$$0.8(0.8305) \leq S(5, 8, p)/S(4, 8, p) \leq 1.2(0.8305) \quad (5)$$

$$0.8(0.8305) \leq S(5, 8, —)/S(4, 8, —) \leq 1.2(0.8305) \quad (6)$$

(4) Quality. There would probably be general agreement that the weights should satisfy $u_5 \geq u_6 \geq u_7 \geq u_8$, i.e. that the weight attached to a department with a higher research rating (of higher quality) should be at least that attached to a department with a lower research rating (of lower quality)—irrespective of the size of the departments.

As such the constraints $u_5 \geq u_6 \geq u_7 \geq u_8$ could be added to the basic model. For the purposes of this study we felt that this did not distinguish sufficiently between departments of differing quality and so we used:

$$u_5 \geq 2u_6 \geq 2^2u_7 \geq 2^3u_8 \quad (7)$$

$$u_5 \leq 20u_8 \quad (8)$$

Equation (7) ensures that the weight attached to the research rating of a department is at least twice that attached to a department with a lesser rating. Equation (8) ensures that the weight attached to the research rating of a star department is at most twenty times greater than the weight attached to the research rating of a below average department.

Input improvements

(1) Equipment expenditure. With respect to equipment expenditure we, for the purposes of this study, felt that the weight associated with it should reflect the total amount spent on equipment for the entire set of departments as a fraction of total expenditure (this fraction being F—see Appendix 1). Hence, allowing the usual flexibility of ±20%, we have the constraints:

$$0.8F \leq T(2, 2, —)/T(1, 3, —) \leq 1.2F \quad (9)$$

(2) Research income. With respect to research income we felt that the weight associated with it should be related to the weight associated with general expenditure.

Essentially research income is used to support postgraduates who primarily engage in research but also spend some of their time teaching (e.g. in a laboratory situation for chemistry and physics departments). As such they are analogous to academic staff supported from general expenditure.

For the purposes of this study we took the view that the weight associated with research income should be approximately one half the weight associated with general expenditure. Hence, allowing the usual flexibility of ±20%, we have the constraints:

$$0.8(v_1/2) \leq v_3 \leq 1.2(v_1/2) \quad (10)$$

Complete model

The complete model for comparing university departments consists therefore of the basic model (Appendix 1) together with the constraints [equations (1)–(10)] given above (see Appendix 2).

In the next section we present the results obtained by solving this model using the data shown in Tables 1 and 2.

5. RESULTS

In this section we discuss the results that we have obtained from the complete model and the use to which such results can be put.

Efficiencies

Table 3 shows the efficiencies of university chemistry and physics departments as calculated according to the model (Appendix 2). As both the recent chemistry and physics reviews [29, 30] have recommended a minimum size for departments of 20 academic staff and 200 students we also show in Table 3 the corresponding figures for 1986–1987 for the number of academic staff and the number of students (from [11, 27, 29, 30]).

The linear program associated with the model was solved using a simplex based in-core FORTRAN code [19]. The total computation time required to produce the efficiencies shown in Table 3 was 10.7 Cray X-MP/28 seconds (involving the solution of 102 linear programs).

Table 3. Results

University	Chemistry			Physics		
	Efficiency (%)	Number of academic staff	Number of students	Efficiency (%)	Number of academic staff	Number of students
Aston	81	12	99	—	—	—
Bath	91	19	180	90	16	171
Birmingham	64	41	291	58	69	451
Bradford	54	15	104	79	6	50
Bristol	100	45	398	68	38	335
Brunel	68	19	182	81	13	121
Cambridge	96	42	464	89	49	522
City	51	15	112	89	11	101
Durham	66	26	229	73	26	235
East Anglia	87	23	247	69	12	77
Essex	59	18	130	100	16	124
Exeter	66	22	198	58	24	211
Hull	61	19	157	45	18	176
Keele	58	8	70	77	11	123
Kent	56	17	145	67	15	156
Lancaster	56	20	118	72	25	176
Leeds	44	53	411	59	28	200
Leicester	78	22	222	70	20	149
Liverpool	75	29	251	55	31	217
London—Birkbeck	64	17	120	48	19	121
Imperial	89	34	308	89	88	687
KQC (Kings)	51	32	285	56	37	251
Q Mary	71	26	204	78	35	195
R Hol and Bed	49	21	155	42	27	184
Univ Coll	72	23	221	42	45	255
Loughborough	50	30	233	91	13	124
Manchester	65	41	349	58	76	446
UMIST	53	36	351	54	22	166
Newcastle	58	27	209	81	29	204
Nottingham	88	34	357	68	28	291
Oxford	98	65	918	71	95	761
Reading	57	24	203	63	23	146
Salford	45	36	272	64	19	170
Sheffield	81	33	287	65	21	209
Southampton	100	38	325	90	25	227
Surrey	60	17	151	96	21	204
Sussex	87	31	306	71	30	249
Warwick	67	22	190	86	22	253
York	80	25	297	88	18	177
Aberystwyth	68	9	69	76	10	70
Bangor	80	10	96	79	8	64
Cardiff	100	18	214	79	14	140
Swansea	73	22	163	72	13	109
UWIST	66	12	124	—	—	—
Aberdeen	51	27	205	65	12	111
Dundee	52	10	86	71	15	102
Edinburgh	81	34	363	67	35	333
Glasgow	54	49	455	73	48	456
Heriot-Watt	61	20	167	67	26	192
St Andrews	60	16	150	75	25	211
Stirling	65	12	72	73	5	22
Strathclyde	71	36	361	53	35	272

Discussion

The results from the model as presented in Table 3 need to be considered from two viewpoints:

(1) validity—i.e. is the model a valid way of comparing university departments; and

(2) usefulness—i.e. what use can be made of the results.

We consider each of these viewpoints in turn.

(1) Validity. We would argue that the model presented in this paper is a valid way of comparing university departments for the following reasons:

(a) DEA, the basis of the model developed above, has been applied and found useful in many situations.

(b) Other workers have reported that their view is that DEA has a role to play in comparing university departments (e.g. see [25]).

(c) The constraints presented in Section 4 above clearly enhance the basic DEA approach and enable judgements as to the relative importance of input/output measures to be incorporated into the model.

(d) Alternative methods of comparing university departments, such as those presented in [11], which rely on considering *separately* a number of different performance indicators (such as expenditure per full-time equivalent student and expenditure per full-time equivalent academic) are intuitively easier to understand but frequently fail to give such a clear *overall* picture as the model developed above.

(e) The results from the model appear reasonable (although this is plainly a subjective assessment and relies on detailed consideration of the results for each department).

For example, considering the results for chemistry, four of the five departments (Bristol, Cambridge, Imperial, Oxford and Southampton) with the highest research rating have efficiencies over 95%.

Of the three departments with an efficiency of 100%, two (Bristol and Southampton), have the highest research rating whilst the other (Cardiff) has a lower research rating but has a high efficiency because (in general terms) that research rating is achieved for a much lower cost than for other departments and it has a low cost per student.

Considering the results for physics the four departments (Cambridge, Imperial, Manchester and Oxford) with the highest research rating do not fare so well in terms of efficiency. Manchester, for example, has an efficiency of only 58%. Essentially this is because (in general terms) Manchester has a cost per student some 60% higher on average than Cambridge, Imperial or Oxford, the highest equipment expenditure and the lowest research income.

The department with an efficiency of 100% (Essex) has only an above average research rating but has a high efficiency because (in general terms) that research rating is achieved for a much lower cost than for other departments.

(2) Usefulness. (a) One use of the results given by the model has been considered above in that they provoke insight into exactly why certain departments score well or badly in terms of efficiency.

(b) Further insight, and suggestions as to how efficiency can be improved, can be gained by considering the values of the weights that actually maximise the efficiency for a particular department.

To take just one example, physics at Birmingham, where there is a very large department (69 staff, 451 students) but an efficiency of only 58%. The weights that produce this efficiency are shown in Table 4. With these weights Essex has an efficiency of 100% and so we compare Essex with Birmingham in Table 4.

It is clear from Table 4 that both Birmingham and Essex have a similar input profile but a different output profile. As Essex is a department with an efficiency of 100%, differences between the two output profiles can be used to highlight areas in which Birmingham needs to change to improve its efficiency, specifically:

(1) It is clear that Birmingham needs a higher research rating—in general terms the research rating is not high enough given the size (cost) of the department (cf. Table 2 in which Birmingham is the most expensive department with an above average research rating). If Birmingham were to achieve a star research rating then this would increase its efficiency (given the same weights) to 67%.

(2) It is clear that Birmingham needs to give less emphasis to UGs (by giving more emphasis to some other output). We would suggest that Birmingham gives more emphasis to PGs T and PGs R—for example doubling the number of PGs *R* would increase the efficiency

(given the same weights) to 66%. If this were to be accompanied by a rise in the research rating to star the efficiency (given the same weights) would increase to 75%.

(c) It is clear that knowing the weights enables options (e.g. increasing UGs by 10% and PGs R by 40%) to be easily explored and the effects on efficiency estimated.

(d) One further use of the results of the model presented in this paper concerns investigating the relationship (if any) between size and efficiency.

Considering the results given in Table 3 and comparing the efficiencies as calculated in this study with the minimum size of 20 academic staff and 200 students as recommended by [29, 30] we have:

(1) of the ten departments with the highest efficiencies two chemistry departments (Bath and Cardiff) and five physics departments (Bath, City, Essex, Loughborough and York) are smaller than the recommended minimum size;

(2) of the ten departments with the lowest efficiencies four chemistry departments (Bradford, City, R. Hol and Bed and Dundee) and four physics departments (Hull, Birkbeck, R. Hol and Bed and UMIST) are smaller than the recommended minimum size;

(3) the average efficiency of a chemistry department is 68.8% and of a physics department is 71.0%;

(4) the 24 chemistry departments below the recommended minimum size have an average efficiency of 65.3% whilst the 28 chemistry departments above the recommended minimum size have an average efficiency of 71.9%;

(5) the 27 physics departments below the recommended minimum size have an average efficiency of 72.2% whilst the 23 physics departments above the recommended minimum size have an average efficiency of 69.6%;

(6) the ten departments with the highest number of students (all above the recommended minimum size) have an average efficiency of 75.0% in chemistry and 69.4% in physics;

(7) the ten departments with the smallest number of students (all below the recommended minimum size) have an average efficiency of 62.9% in chemistry and 72.1% in physics.

Our conclusion from this would be that whilst there may be educational reasons for having a recommended minimum size (e.g. breadth of academic expertise available on-site) this study

Table 4. Comparing physics at Birmingham and Essex

Factor		Weight	Relative percentage	
			Birmingham	Essex
Inputs				
General expenditure		0.1000	68.5	66.0
Equipment expenditure		0.1022	8.1	10.0
Research income		0.0600	23.4	23.9
Outputs				
UGs		0.3482	60.1	31.5
PGs T		0.4352	3.2	8.4
PGs R		0.5440	13.3	16.6
Research income		0.0117	7.9	4.7
Research rating	Star	68.9795	—	—
	A+	34.4898	15.6	38.9
	A	10.5785	—	—
	A—	5.2892	—	—

Note: (a) Relative percentage for input factor j for department k is defined as $100[T(j,j,k)/T(1,3,k)]$; (b) relative percentage for output factor i for department k is defined as $100[S(i,i,k)/S(1,8,k)]$; (c) relative percentages may not add to 100% due to rounding errors.

appears to provide little evidence to support such a recommendation from the viewpoint of departmental efficiency.

In the next section we consider how the model presented in this paper can be enhanced with particular reference to the data that could be collected to make the model more useful.

6. MODEL ENHANCEMENT

We stated previously that our objective in this paper was to develop a model which:

(a) represents an initial attempt to construct a model for the quantitative comparison of university departments;

(b) highlights the data that is needed, but which is not currently available, to make the model more useful;

(c) can be criticised and enlarged upon by others.

With respect to the data that could be used to make the model more useful our judgement would be that collecting data relating to publications and citations is of primary importance as such data should be relatively easy to obtain and would provide an immediate payoff in terms of model improvement. Such data could be used (after suitable adjustment if necessary) both to represent research output in terms of quantity (publications) and in terms of quality (citations).

Collecting data relating to person-specific increased knowledge we would regard as being of secondary importance as it is much harder to define appropriate measures and (potentially) much more expensive to obtain such data. For example:

(a) For a particular student, how would we measure his/her increase in knowledge over the course of a year and how expensive would it be to obtain similar data for all such students in the same discipline?

(b) even if we were to restrict ourselves to final year students and define person-specific increased knowledge as the (weighted?) difference between degree class and (relevant?) A-level results this still assumes that all degree classes at all universities are equivalent.

Although it has been suggested [31] that the model should include some measure of the quality of teaching we feel that this suggestion needs to be treated with care. The reason for this is that although quality of teaching is obviously important in a university department it is not a direct output.

Our view would be that some measure of the quality of teaching would best be used in defining a *proxy* for person-specific increased knowledge (e.g. define person-specific increased knowledge for a university department as number of students multiplied by quality of teaching).

Whether measures of the quality of teaching for university departments can realistically be obtained however we would doubt, particularly in the short-term.

7. CONCLUSIONS

In this paper we have presented a model, based upon data envelopment analysis, for comparing university departments. This model is quite flexible and can be used to reflect any view policy-makers might take as to the relative importance of departmental input/output measures. We hope that this model will be enhanced by others and that the merits of a more quantitative approach to comparing university departments will come to be appreciated by policy-makers and used by them.

APPENDIX 1

As stated previously data envelopment analysis (DEA) was first put forward by Charnes *et al.* [10] in 1978 and is used for evaluating the (relative) efficiency of decision-making units (DMUs) via weights attached to input/output measures. Mathematically DEA can be expressed as follows. Let:

s be the number of output measures;
t be the number of input measures;
n be the number of DMUs which are being evaluated with respect to one other;

y_{ik} be the value (≥ 0) of output measure i ($i = 1, \ldots, s$) for DMU k;
x_{jk} be the value (≥ 0) of input measure j ($j = 1, \ldots, t$) for DMU k;
u_i be the weight (≥ 0) to be attached to one unit of output measure i;
v_j be the weight (≥ 0) to be attached to one unit of input measure j;
e_k be the (relative) efficiency of DMU k.

In order to ease the notation in the main text we also define:

$$S(a,b,k) = \left(\sum_{i=a}^{b} u_i y_{ik}\right) \qquad 1 \leq a \leq b \leq s$$

$$S(a,b,—) = \sum_{k=1}^{n} S(a,b,k) \qquad 1 \leq a \leq b \leq s$$

$$T(a,b,k) = \left(\sum_{j=a}^{b} v_j x_{jk}\right) \qquad 1 \leq a \leq b \leq t$$

$$T(a,b,—) = \sum_{k=1}^{n} T(a,b,k) \qquad 1 \leq a \leq b \leq t$$

$$F = \left(\sum_{k=1}^{n} x_{2k}\right) \Big/ \left(\sum_{k=1}^{n}\sum_{j=1}^{3} x_{jk}\right)$$

then we determine the efficiency (e_p) of DMU p using the nonlinear program:

maximise

$$e_p \tag{1}$$

subject to

$$e_k = S(1,s,k)/T(1,t,k) \qquad k = 1, \ldots, n \tag{2}$$

$$0 \leq e_k \leq 1 \qquad k = 1, \ldots, n \tag{3}$$

$$u_i \geq 0 \qquad i = 1, \ldots, s \tag{4}$$

$$v_j \geq 0 \qquad j = 1, \ldots, t \tag{5}$$

Equation (2) defines efficiency as a weighted sum of outputs divided by a weighted sum of inputs, equation (3) ensures that all efficiencies lie between zero and one and equations (4) and (5) ensure that all weights are non-negative. This nonlinear program can be converted into a linear program using an approach due to Charnes and Cooper [9] and hence easily solved.

The key point to note here is that in evaluating the efficiency of DMU p we choose the weights that *maximise* its efficiency. Conceptually we can regard e_p (when maximised) as the efficiency of department p when compared to its peers.

APPENDIX 2

The complete model for comparing university departments is as follows:
maximise

$$e_p$$

subject to

$$e_k = S(1,8,k)/T(1,3,k) \quad k = 1, \ldots, n$$

$$u_3 \geq 1.25u_2 \geq 1.25^2 u_1$$

$$u_3 \leq 2u_1$$

$$0.510 \leq S(1,3,p)/S(1,8,p) \leq 0.765$$

$$0.510 \leq S(1,3,—)/S(1,8,—) \leq 0.765$$

$$0.8(0.8305) \leq S(5,8,p)/S(4,8,p) \leq 1.2(0.8305)$$

$$0.8(0.8305) \leq S(5,8,—)/S(4,8,—) \leq 1.2(0.8305)$$

$$u_5 \geq 2u_6 \geq 2^2 u_7 \geq 2^3 u_8$$

$$u_5 \leq 20u_8$$

$$0.8F \leq T(2,2,—)/T(1,3,—) \leq 1.2F$$

$$0.8(v_1/2) \leq v_3 \leq 1.2(v_1/2)$$

$$0 \leq e_k \leq 1$$

$$u_i \geq 0$$

$$v_j \geq 0$$

where S, T and F are as defined in Appendix 1.

Although this mathematical program is nonlinear it can be easily converted into a linear program. Observe that multiplying all weights (u_i, v_j) by any scaling factor K (>0) leaves the optimal solution to the above nonlinear program unchanged. Since the objective function can be rewritten as

maximise

$$S(1,8,p)/T(1,3,p)$$

we can (implicitly) set a scaling factor by setting the denominator of the objective function ($T(1,3,p)$) equal to one (arbitrarily exclude solutions in which all input weights (v_j) are zero). If this is done then the above nonlinear program becomes a linear program (after elimination of the S, T and e_k variables and rearrangement) and hence is easily solved.

REFERENCES

1. Ahn T, Arnold V, Charnes A and Cooper WW (1988) DEA and ratio efficiency analyses for public institutions of higher learning in Texas. Working paper available from Cooper WW at the Center for Cybernetic Studies, College of Business Administration, The University of Texas at Austin, TX 78712-1177, USA (to appear in *Research in Governmental and Nonprofit Accounting*).
2. Ahn T, Charnes A and Cooper WW (1988) Some statistical and DEA evaluations of relative efficiencies of public and private institutions of higher learning. *Socio-Econ. Plann. Sci.* **22**(6), 259–269.
3. Alexander L (1988) UGC ignores critics in new rankings exercise. *The Times Higher Education Supplement* December 2, 1.
4. Alexander L (1988) Ticking off the clockwatchers. *The Times Higher Education Supplement* December 2, 7.
5. Calvert JR (1978) The measurement of performance in higher education. PhD thesis, Department of Management Studies, Loughborough University of Technology, Loughborough LE11 3TU, UK.
6. Calvert JR (1980) Relative performance. In *Indicators of Performance* (Edited by Billing D), pp. 65–72. Society for Research into Higher Education, Guildford, Surrey, UK.
7. Carpenter MP, Gibb F, Harris M, Irvine J, Martin BR and Narin F (1988) Bibliometric profiles for British academic institutions: An experiment to develop research output indicators. *Scientometrics* **14,** 213–233.
8. Cave M, Hanney S, Kogan M and Trevett G (1988) The use of performance indicators in higher education: A critical analysis of developing practice. Kingsley, London, UK.
9. Charnes A and Cooper WW (1962) Programming with linear fractional functionals. *Nav. Res. Logist. Q.* **9,** 181–186.
10. Charnes A, Cooper WW and Rhodes E (1978) Measuring the efficiency of decision making units. *Eur. J. Opl Res.* **2,** 429–444.
11. Committee of Vice-Chancellors and Principals and University Grants Committee (1988) University management statistics and performance indicators in the UK. Available from the Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom, 29 Tavistock Square, London WC1H 9EZ, UK.
12. Crewe I (1988) Reputation, research and reality: The publication records of UK departments of politics 1978–1984. *Scientometrics* **14,** 235–250.
13. Croham L (1987) Review of the University Grants Committee: Report of a committee under the chairmanship of Lord Croham, GCB. Cm 81. HMSO, London, UK.
14. Gillett R and Aitkenhead M (1987) Rank injustice in academic research. *Nature* **327,** 381–382.
15. Gray RH, Haslam J and Prodhan BK (1987) Academic departments of accounting in the UK: A note on publication output. *British Acctng Rev.* **19,** 53–71.
16. Groves REV and Perks RW (1984) The teaching and researching of accounting in UK universities, a survey. *British Acctng Rev.* **16,** 10–20.
17. Johnes G (1988) Determinants of research output in economics departments in British universities. *Res. Policy* **17,** 171–178.
18. Kwimbere FJ (1987) Measuring efficiency in not-for-profit organisations: An attempt to evaluate efficiency in selected UK university departments using data envelopment analysis (DEA). MSc thesis, School of Management, University of Bath, Claverton Down, Bath BA2 7AY, UK.
19. Marsten RE (1981) The design of the XMP linear programming library. *ACM Trans. Math. Soft.* **7,** 481–497.
20. Rhodes EL and Southwick L (1986) Determinants of efficiency in public and private universities. Working paper, School of Environmental and Public Affairs, Indiana University, Bloomington, IN 47405, USA.
21. Rhodes EL and Southwick L (1988) Relative efficiencies of private and public universities over time. Paper presented at the *TIMS/ORSA Joint National Meeting*, Washington, DC, USA, April 1988.
22. Seiford LM (1989) A bibliography of data envelopment analysis (1978–1989): version 4.0. Working paper available from the author at the Department of Industrial Engineering and Operations Research, The University of Massachusetts, Amherst, MA 01003, USA.
23. Sizer J (1981) Institutional performance assessment, adaptation and change. Working paper available from the author at the Department of Management Studies, Loughborough University of Technology, Loughborough LE11 3TU, UK.
24. Smith T (1987) The UGCs research rankings exercise. *Higher Educ. Q.* **41**(4), 303–316.
25. Tomkins C and Green R (1988) An experiment in the use of data envelopment analysis for evaluating the efficiency of UK university departments of accounting. *Fin. Accountability Mgmt* **4**(2), 147–164.
26. Turney J (1988) ABRC index puts focus on citations. *The Times Higher Education Supplement* October 7, 1.
27. Universities' Statistical Record (1988) *University Statistics 1986–1987*, Vol. 3: *Finance*. Available from Universities' Statistical Record, PO Box 130, Cheltenham GL50 1JW, UK.
28. University Grants Committee (1986) Planning for the late 1980s: Recurrent grant for 1986/1987. University Grants Committee circular letter 4/86 annex C, part 3. Available from the Library, Department of Education and Science, Elizabeth House, York Road, London SE1 7PH, UK.
29. University Grants Committee (1988) The future of university physics: the report of the physics review. HMSO, London, UK.
30. University Grants Committee (1988) University chemistry—the way forward: the report of the chemistry review. HMSO, London, UK.
31. University Grants Committee (1989) Private communication.
32. Ward D (1989) Science review 'wrong about size'. *The Guardian* January 3, 4.
33. White GP (1987) The implementation of management science in higher education administration. *Omega* **15**(4), 283–290.
34. White GP (1987) A survey of recent management science applications in higher education administration. *Interfaces* **17**(2), 97–108.

ADDRESS FOR CORRESPONDENCE: *Dr JE Beasley, The Management School, Imperial College of Science, Technology and Medicine, Exhibition Road, London SW7 2AZ, U.K.*