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

<table>
<thead>
<tr>
<th>University</th>
<th>General expenditure (£000s)</th>
<th>Equipment expenditure (£000s)</th>
<th>Research income (£000s)</th>
<th>UGs</th>
<th>PGs</th>
<th>Research rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>value</td>
<td>$x_{11}$</td>
<td>$x_{12}$</td>
<td>$x_{13}$</td>
<td>$x_{14}$</td>
<td>$y_{11}$</td>
<td>$y_{12}$</td>
</tr>
<tr>
<td>Aston</td>
<td>446</td>
<td>21</td>
<td>183</td>
<td>62</td>
<td>37</td>
<td>0</td>
</tr>
<tr>
<td>Bath</td>
<td>670</td>
<td>53</td>
<td>288</td>
<td>137</td>
<td>43</td>
<td>0</td>
</tr>
<tr>
<td>Birmingham</td>
<td>1459</td>
<td>69</td>
<td>288</td>
<td>225</td>
<td>3</td>
<td>63</td>
</tr>
<tr>
<td>Bradford</td>
<td>613</td>
<td>95</td>
<td>73</td>
<td>92</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>Bristol</td>
<td>2043</td>
<td>236</td>
<td>1050</td>
<td>255</td>
<td>27</td>
<td>118</td>
</tr>
<tr>
<td>Brunel</td>
<td>408</td>
<td>26</td>
<td>137</td>
<td>27</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cambridge</td>
<td>2227</td>
<td>620</td>
<td>981</td>
<td>305</td>
<td>0</td>
<td>159</td>
</tr>
<tr>
<td>City</td>
<td>696</td>
<td>354</td>
<td>81</td>
<td>31</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Durham</td>
<td>1027</td>
<td>148</td>
<td>578</td>
<td>187</td>
<td>0</td>
<td>42</td>
</tr>
<tr>
<td>East Anglia</td>
<td>1155</td>
<td>113</td>
<td>545</td>
<td>126</td>
<td>31</td>
<td>90</td>
</tr>
<tr>
<td>Essex</td>
<td>620</td>
<td>115</td>
<td>565</td>
<td>76</td>
<td>5</td>
<td>49</td>
</tr>
<tr>
<td>Exeter</td>
<td>984</td>
<td>138</td>
<td>198</td>
<td>166</td>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td>Hull</td>
<td>880</td>
<td>78</td>
<td>488</td>
<td>119</td>
<td>9</td>
<td>29</td>
</tr>
<tr>
<td>Keele</td>
<td>440</td>
<td>51</td>
<td>217</td>
<td>50</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Kent</td>
<td>667</td>
<td>281</td>
<td>111</td>
<td>116</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>Lancaster</td>
<td>685</td>
<td>50</td>
<td>191</td>
<td>92</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Leeds</td>
<td>2545</td>
<td>210</td>
<td>763</td>
<td>320</td>
<td>9</td>
<td>82</td>
</tr>
<tr>
<td>Leicester</td>
<td>919</td>
<td>61</td>
<td>419</td>
<td>173</td>
<td>0</td>
<td>49</td>
</tr>
<tr>
<td>Liverpool</td>
<td>1259</td>
<td>52</td>
<td>496</td>
<td>195</td>
<td>0</td>
<td>56</td>
</tr>
<tr>
<td>London—Birkbeck</td>
<td>734</td>
<td>33</td>
<td>142</td>
<td>46</td>
<td>26</td>
<td>48</td>
</tr>
<tr>
<td>Imperial</td>
<td>1760</td>
<td>742</td>
<td>1061</td>
<td>167</td>
<td>0</td>
<td>141</td>
</tr>
<tr>
<td>KQC (Kings)</td>
<td>1487</td>
<td>479</td>
<td>521</td>
<td>240</td>
<td>3</td>
<td>42</td>
</tr>
<tr>
<td>Q Mary</td>
<td>1106</td>
<td>170</td>
<td>430</td>
<td>164</td>
<td>3</td>
<td>37</td>
</tr>
<tr>
<td>R Hol and Bed</td>
<td>962</td>
<td>131</td>
<td>152</td>
<td>122</td>
<td>0</td>
<td>33</td>
</tr>
<tr>
<td>Univ Coll</td>
<td>1238</td>
<td>67</td>
<td>490</td>
<td>157</td>
<td>4</td>
<td>60</td>
</tr>
<tr>
<td>Loughborough</td>
<td>1208</td>
<td>89</td>
<td>397</td>
<td>158</td>
<td>26</td>
<td>49</td>
</tr>
<tr>
<td>Manchester</td>
<td>1920</td>
<td>191</td>
<td>544</td>
<td>268</td>
<td>0</td>
<td>81</td>
</tr>
<tr>
<td>UMIST</td>
<td>1758</td>
<td>196</td>
<td>1162</td>
<td>237</td>
<td>9</td>
<td>105</td>
</tr>
<tr>
<td>Newcastle</td>
<td>1211</td>
<td>79</td>
<td>540</td>
<td>157</td>
<td>0</td>
<td>52</td>
</tr>
<tr>
<td>Nottingham</td>
<td>1409</td>
<td>122</td>
<td>527</td>
<td>263</td>
<td>0</td>
<td>94</td>
</tr>
<tr>
<td>Oxford</td>
<td>3337</td>
<td>654</td>
<td>1780</td>
<td>707</td>
<td>0</td>
<td>211</td>
</tr>
<tr>
<td>Reading</td>
<td>908</td>
<td>120</td>
<td>336</td>
<td>162</td>
<td>5</td>
<td>36</td>
</tr>
<tr>
<td>Salford</td>
<td>1492</td>
<td>127</td>
<td>613</td>
<td>152</td>
<td>18</td>
<td>102</td>
</tr>
<tr>
<td>Sheffield</td>
<td>1346</td>
<td>78</td>
<td>250</td>
<td>223</td>
<td>0</td>
<td>64</td>
</tr>
<tr>
<td>Southampton</td>
<td>1620</td>
<td>420</td>
<td>1224</td>
<td>199</td>
<td>2</td>
<td>124</td>
</tr>
<tr>
<td>Surrey</td>
<td>691</td>
<td>65</td>
<td>407</td>
<td>122</td>
<td>2</td>
<td>27</td>
</tr>
<tr>
<td>Sussex</td>
<td>1324</td>
<td>144</td>
<td>565</td>
<td>189</td>
<td>13</td>
<td>104</td>
</tr>
<tr>
<td>Warwick</td>
<td>927</td>
<td>148</td>
<td>359</td>
<td>147</td>
<td>0</td>
<td>43</td>
</tr>
<tr>
<td>York</td>
<td>947</td>
<td>146</td>
<td>724</td>
<td>236</td>
<td>7</td>
<td>54</td>
</tr>
<tr>
<td>Aberystwyth</td>
<td>370</td>
<td>32</td>
<td>102</td>
<td>58</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Bangor</td>
<td>360</td>
<td>73</td>
<td>122</td>
<td>89</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Cardiff</td>
<td>849</td>
<td>32</td>
<td>258</td>
<td>158</td>
<td>3</td>
<td>53</td>
</tr>
<tr>
<td>Swansea</td>
<td>764</td>
<td>89</td>
<td>317</td>
<td>132</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>UWIST</td>
<td>560</td>
<td>99</td>
<td>196</td>
<td>100</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>Aberdeen</td>
<td>1029</td>
<td>126</td>
<td>391</td>
<td>164</td>
<td>2</td>
<td>39</td>
</tr>
<tr>
<td>Dundee</td>
<td>610</td>
<td>21</td>
<td>136</td>
<td>73</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>1381</td>
<td>254</td>
<td>812</td>
<td>292</td>
<td>0</td>
<td>71</td>
</tr>
<tr>
<td>Glasgow</td>
<td>2253</td>
<td>131</td>
<td>360</td>
<td>354</td>
<td>7</td>
<td>94</td>
</tr>
<tr>
<td>Heriot-Watt</td>
<td>768</td>
<td>38</td>
<td>324</td>
<td>142</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>St Andrews</td>
<td>696</td>
<td>73</td>
<td>408</td>
<td>121</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>Stirling</td>
<td>421</td>
<td>18</td>
<td>105</td>
<td>57</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>Strathclyde</td>
<td>1714</td>
<td>112</td>
<td>945</td>
<td>269</td>
<td>15</td>
<td>77</td>
</tr>
</tbody>
</table>

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
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 policymakers, 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:

\begin{align*}
  u_3 &\geq 1.25u_2 \geq 1.25^2u_1 \\
  u_3 &\leq 2u_1
\end{align*}

(1)

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):

\begin{align*}
  0.510 &\leq S(1,3,p)/S(1,8,p) \leq 0.765 \quad (3) \\
  0.510 &\leq S(1,3,\ldots)/S(1,8,\ldots) \leq 0.765 \quad (4)
\end{align*}

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:

\[ 0.8(0.8305) \leq S(5, 8, \rho)/S(4, 8, \rho) \leq 1.2(0.8305) \]  
\[ 0.8(0.8305) \leq S(5, 8, \rho)/S(4, 8, \rho^-) \leq 1.2(0.8305) \]

(4) Quality. There would probably be general agreement that the weights should satisfy \( u_3 \geq u_4 \geq u_7 \geq u_8 \), 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 \( u_3 \geq u_4 \geq u_7 \geq u_8 \) 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:

\[ u_3 \geq 2u_4 \geq 2^2u_7 \geq 2^3u_8 \]  
\[ u_3 \leq 20u_8 \]

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:

\[ 0.8F \leq T(2, 2, \rho^-)/T(1, 3, \rho^-) \leq 1.2F \]

(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:

\[ 0.8(\epsilon_1/2) \leq \epsilon_1 \leq 1.2(\epsilon_1/2) \]

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

<table>
<thead>
<tr>
<th>University</th>
<th>Chemistry Efficiency (%)</th>
<th>Number of academic staff</th>
<th>Number of students</th>
<th>Physics Efficiency (%)</th>
<th>Number of academic staff</th>
<th>Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aston</td>
<td>81</td>
<td>12</td>
<td>99</td>
<td>90</td>
<td>16</td>
<td>171</td>
</tr>
<tr>
<td>Bath</td>
<td>91</td>
<td>19</td>
<td>180</td>
<td>58</td>
<td>69</td>
<td>451</td>
</tr>
<tr>
<td>Birmingham</td>
<td>64</td>
<td>41</td>
<td>291</td>
<td>79</td>
<td>6</td>
<td>50</td>
</tr>
<tr>
<td>Bradford</td>
<td>54</td>
<td>15</td>
<td>104</td>
<td>68</td>
<td>38</td>
<td>335</td>
</tr>
<tr>
<td>Bristol</td>
<td>100</td>
<td>45</td>
<td>398</td>
<td>81</td>
<td>13</td>
<td>121</td>
</tr>
<tr>
<td>Brunel</td>
<td>68</td>
<td>19</td>
<td>182</td>
<td>89</td>
<td>49</td>
<td>522</td>
</tr>
<tr>
<td>Cambridge</td>
<td>96</td>
<td>42</td>
<td>464</td>
<td>89</td>
<td>11</td>
<td>101</td>
</tr>
<tr>
<td>City</td>
<td>51</td>
<td>15</td>
<td>112</td>
<td>73</td>
<td>26</td>
<td>235</td>
</tr>
<tr>
<td>Durham</td>
<td>66</td>
<td>26</td>
<td>229</td>
<td>69</td>
<td>12</td>
<td>77</td>
</tr>
<tr>
<td>East Anglia</td>
<td>87</td>
<td>23</td>
<td>247</td>
<td>100</td>
<td>16</td>
<td>124</td>
</tr>
<tr>
<td>Essex</td>
<td>59</td>
<td>18</td>
<td>130</td>
<td>58</td>
<td>24</td>
<td>211</td>
</tr>
<tr>
<td>Exeter</td>
<td>66</td>
<td>19</td>
<td>157</td>
<td>45</td>
<td>18</td>
<td>176</td>
</tr>
<tr>
<td>Hull</td>
<td>61</td>
<td>19</td>
<td>70</td>
<td>77</td>
<td>11</td>
<td>123</td>
</tr>
<tr>
<td>Keele</td>
<td>58</td>
<td>8</td>
<td>145</td>
<td>67</td>
<td>15</td>
<td>156</td>
</tr>
<tr>
<td>Kent</td>
<td>56</td>
<td>17</td>
<td>118</td>
<td>72</td>
<td>25</td>
<td>176</td>
</tr>
<tr>
<td>Lancaster</td>
<td>56</td>
<td>20</td>
<td>411</td>
<td>59</td>
<td>28</td>
<td>200</td>
</tr>
<tr>
<td>Leeds</td>
<td>44</td>
<td>33</td>
<td>222</td>
<td>70</td>
<td>20</td>
<td>149</td>
</tr>
<tr>
<td>Leicester</td>
<td>78</td>
<td>22</td>
<td>251</td>
<td>55</td>
<td>31</td>
<td>217</td>
</tr>
<tr>
<td>Liverpool</td>
<td>75</td>
<td>29</td>
<td>120</td>
<td>48</td>
<td>19</td>
<td>121</td>
</tr>
<tr>
<td>London—Birkbeck</td>
<td>64</td>
<td>17</td>
<td>308</td>
<td>89</td>
<td>88</td>
<td>687</td>
</tr>
<tr>
<td>Imperial</td>
<td>89</td>
<td>34</td>
<td>285</td>
<td>56</td>
<td>37</td>
<td>251</td>
</tr>
<tr>
<td>Q Mary</td>
<td>71</td>
<td>26</td>
<td>204</td>
<td>78</td>
<td>35</td>
<td>195</td>
</tr>
<tr>
<td>R Hol and Bed</td>
<td>49</td>
<td>21</td>
<td>155</td>
<td>42</td>
<td>27</td>
<td>184</td>
</tr>
<tr>
<td>Univ Coll</td>
<td>72</td>
<td>23</td>
<td>221</td>
<td>42</td>
<td>45</td>
<td>255</td>
</tr>
<tr>
<td>Loughborough</td>
<td>50</td>
<td>30</td>
<td>233</td>
<td>91</td>
<td>13</td>
<td>124</td>
</tr>
<tr>
<td>Manchester</td>
<td>65</td>
<td>41</td>
<td>349</td>
<td>58</td>
<td>76</td>
<td>446</td>
</tr>
<tr>
<td>UMIST</td>
<td>53</td>
<td>36</td>
<td>351</td>
<td>54</td>
<td>22</td>
<td>166</td>
</tr>
<tr>
<td>Newcastle</td>
<td>58</td>
<td>27</td>
<td>209</td>
<td>81</td>
<td>29</td>
<td>204</td>
</tr>
<tr>
<td>Nottingham</td>
<td>88</td>
<td>34</td>
<td>357</td>
<td>68</td>
<td>28</td>
<td>291</td>
</tr>
<tr>
<td>Oxford</td>
<td>98</td>
<td>65</td>
<td>918</td>
<td>71</td>
<td>95</td>
<td>761</td>
</tr>
<tr>
<td>Reading</td>
<td>57</td>
<td>24</td>
<td>203</td>
<td>63</td>
<td>23</td>
<td>146</td>
</tr>
<tr>
<td>Salford</td>
<td>45</td>
<td>36</td>
<td>272</td>
<td>64</td>
<td>19</td>
<td>170</td>
</tr>
<tr>
<td>Sheffield</td>
<td>81</td>
<td>33</td>
<td>287</td>
<td>65</td>
<td>21</td>
<td>209</td>
</tr>
<tr>
<td>Southampton</td>
<td>100</td>
<td>38</td>
<td>325</td>
<td>90</td>
<td>25</td>
<td>227</td>
</tr>
<tr>
<td>Surrey</td>
<td>60</td>
<td>17</td>
<td>151</td>
<td>96</td>
<td>21</td>
<td>204</td>
</tr>
<tr>
<td>Sussex</td>
<td>87</td>
<td>31</td>
<td>106</td>
<td>71</td>
<td>30</td>
<td>249</td>
</tr>
<tr>
<td>Warwick</td>
<td>67</td>
<td>22</td>
<td>190</td>
<td>86</td>
<td>22</td>
<td>253</td>
</tr>
<tr>
<td>York</td>
<td>80</td>
<td>25</td>
<td>297</td>
<td>88</td>
<td>18</td>
<td>177</td>
</tr>
<tr>
<td>Aberystwyth</td>
<td>68</td>
<td>9</td>
<td>69</td>
<td>76</td>
<td>10</td>
<td>70</td>
</tr>
<tr>
<td>Bangor</td>
<td>80</td>
<td>10</td>
<td>96</td>
<td>79</td>
<td>8</td>
<td>64</td>
</tr>
<tr>
<td>Cardiff</td>
<td>100</td>
<td>18</td>
<td>214</td>
<td>79</td>
<td>14</td>
<td>140</td>
</tr>
<tr>
<td>Swansea</td>
<td>73</td>
<td>22</td>
<td>163</td>
<td>72</td>
<td>13</td>
<td>109</td>
</tr>
<tr>
<td>UWIST</td>
<td>66</td>
<td>12</td>
<td>124</td>
<td>21</td>
<td>21</td>
<td>249</td>
</tr>
<tr>
<td>Aberdeen</td>
<td>51</td>
<td>27</td>
<td>205</td>
<td>65</td>
<td>12</td>
<td>111</td>
</tr>
<tr>
<td>Dundee</td>
<td>52</td>
<td>10</td>
<td>86</td>
<td>71</td>
<td>15</td>
<td>102</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>81</td>
<td>34</td>
<td>363</td>
<td>67</td>
<td>35</td>
<td>333</td>
</tr>
<tr>
<td>Glasgow</td>
<td>54</td>
<td>49</td>
<td>455</td>
<td>73</td>
<td>48</td>
<td>456</td>
</tr>
<tr>
<td>Heriot-Watt</td>
<td>61</td>
<td>20</td>
<td>167</td>
<td>67</td>
<td>26</td>
<td>192</td>
</tr>
<tr>
<td>St Andrews</td>
<td>60</td>
<td>16</td>
<td>150</td>
<td>75</td>
<td>25</td>
<td>211</td>
</tr>
<tr>
<td>Stirling</td>
<td>65</td>
<td>12</td>
<td>72</td>
<td>73</td>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>Strathclyde</td>
<td>71</td>
<td>36</td>
<td>361</td>
<td>53</td>
<td>35</td>
<td>272</td>
</tr>
</tbody>
</table>

**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 |
| 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                           | 0.3440           | 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 $100\frac{T(j, k)}{T(1, k)}$; (b) relative percentage for output factor $i$ for department $k$ is defined as $100\frac{S(i, k)}{S(1, k)}$; (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 ($\geq 0$) of output measure $i$ ($i = 1, \ldots, s$) for DMU $k$;

$x_{jk}$ be the value ($\geq 0$) of input measure $j$ ($j = 1, \ldots, t$) for DMU $k$;

$u_i$ be the weight ($\geq 0$) to be attached to one unit of output measure $i$;

$v_j$ be the weight ($\geq 0$) 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:

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

\[
S(a, b, -) = \sum_{k=1}^{n} S(a, b, k) \quad 1 \leq a \leq b \leq s
\]

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

\[
T(a, b, -) = \sum_{k=1}^{n} T(a, b, k) \quad 1 \leq a \leq b \leq t
\]

\[
F = \left( \sum_{k=1}^{n} x_{k1} \right) \left/ \left( \sum_{i=1}^{s} \sum_{j=1}^{t} x_{ij} \right) \right.
\]

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

maximise

\[
e_p
\]

subject to

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

\[
u_1 \geq 1.25u_5 \geq 1.25^5u_1
\]

\[
u_5 \leq 2u_1
\]

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

\[
0.510 \leq S(1, 3, --)/S(1, 8, --) \leq 0.765
\]

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

\[
0.8(0.8305) \leq S(5, 8, --)/S(4, 8, --) \leq 1.2(0.8305)
\]

\[
u_1 \geq 2u_7 \geq 2^3u_5
\]

\[
u_5 \leq 20u_4
\]

\[
0.8F \leq T(2, 2, --)/T(1, 3, --) \leq 1.2F
\]

\[
0.8(\nu_k/2) \leq v_j \leq 1.2(\nu_k/2)
\]

\[
0 \leq e_k \leq 1
\]

\[
u_i \geq 0
\]

\[
v_j \geq 0
\]

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 $K$ ($\geq 0$) 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).


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.


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