Economic Issues, Vol. 12, Part 2, 2007

Using Cooper's Approach to Explore the Extent of Congestion in the New British Universities

A T Flegg and D O Allen[1]

ABSTRACT

This paper uses data envelopment analysis (DEA) to explore the issue of congestion in British universities. The focus is on 41 former polytechnics that became universities in 1992, and the analysis covers the period 1995/6 to 2003/4. These new universities differ from the older universities in many ways, especially in terms of their far higher student:staff ratios and much lower research funding per member of staff. The primary aim is to examine whether this under-resourcing of the new universities has led to 'congestion', in the sense that their output has been decreased as a result of having too many students. This phenomenon is measured using the method proposed by Cooper et al. *in a series of articles. To check the sensitivity of the results to different specifications, three alternative DEA models are formulated. The results reveal that a substantial amount of congestion was present throughout the period under review, and in a wide range of universities. An overabundance of undergraduate students is identified as the largest single cause of congestion in the former polytechnics. Less plausibly, the results also suggest that academic overstaffing was a major cause of congestion. By contrast, postgraduates and 'other expenditure' are found to play a noticeably smaller role in generating congestion.*

1. INTRODUCTION

There has been a rapid expansion of higher education in the United Kingdom in recent decades. This growth has taken place both in the older universities (those created before 1992) and in other higher education institutions. The latter include the former polytechnics that were granted university status in 1992, along with university colleges, institutes of higher education, and so forth. Here we have chosen to review the experience of the former polytechnics in the period 1995/6 to 2003/4. These institutions form a relatively homogeneous group, sharing a common history and facing similar opportunities and problems. As far as we are aware, this is the first study to employ data envelopment analysis (DEA) to examine the efficiency of the former poly-technics as a separate group.[2]

Figure 1: Students and staff: older universities and former polytechnics

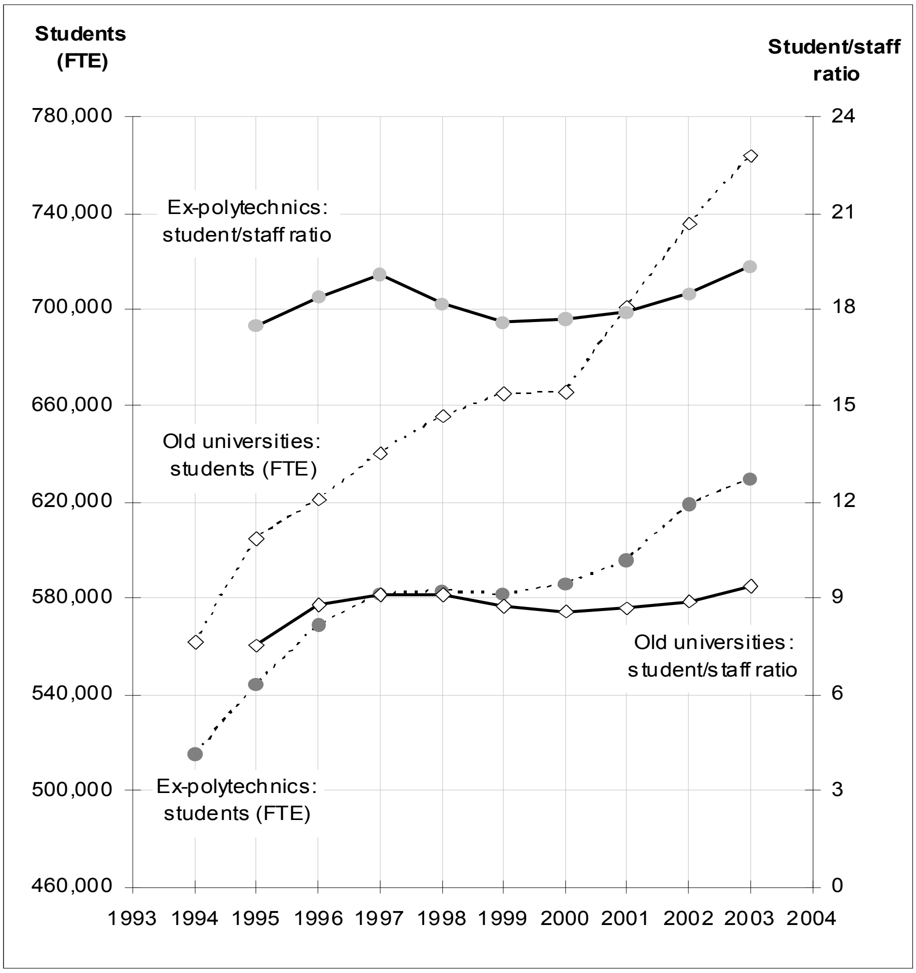


Figure 1 illustrates the point that the former polytechnics operate under much higher student : staff ratios than do the older universities. In addition, the older universities typically receive considerably more research funding per member of staff, and a higher proportion of their undergraduate students gain first-class degrees and upper seconds.[3] It is also interesting to observe from Figure 1 that there has been a much smaller rise in the number of students in the former polytechnics than in the older universities.[4] In view of these clear dissimilarities, it seems right to analyse the older universities and the former polytechnics separately.[5]

The particular issue we wish to address is whether the number of students in the former polytechnics has grown to such an extent that it has caused output to fall. This inverse relationship between inputs and outputs is commonly referred to as *congestion*.

2. The problem of congestion

Congestion of inputs occurs when their use has increased by so much that output actually falls. A more comprehensive definition is given by Cooper *et al.* (2001a, p. 228):

> **Definition 1.** Evidence of congestion is present when reductions in one or more inputs can be associated with increases in one or more outputs — without worsening any other input or output. Proceeding in reverse, congestion occurs when increases in one or more inputs can be associated with decreases in one or more outputs — without improving any other input or output.

Cooper *et al.* then describe congestion of inputs as 'an especially severe form of inefficiency (or waste)' (*ibid.*, p. 229).

However, a problem with the above definition is that it does not shed any light on what could be the underlying cause of the congestion. It might be more appropriate, therefore, to redefine this concept as follows:

> **Definition 2.** Congestion is evident whenever more (less) of any input is employed, with all other inputs held constant, and there is a concomitant fall (rise) in output.

This alternative definition is grounded in the hypothesis of diminishing marginal returns, with the added feature that congestion requires the marginal product of the input under consideration to be negative beyond a certain point.

With regard to universities, it seems fair to assume that an excessive number of students could lead to congestion. For instance, Figure 1 shows that the number of full-time equivalent students in the former polytechnics rose substantially during the period under review; as a result, the marginal product of students might have become negative in some of these new universities. This argument suggests that their output — measured in terms of research and the number of undergraduate and postgraduate qualifications awarded — could have been higher if they had admitted fewer students. Some doubt is cast on this argument, however, by the fact that this growth in the number of students was accompanied by only a moderate rise in the student: staff ratio for the period as a whole.[6]

3. DEA models

DEA makes use of linear programming techniques to construct an 'efficiency frontier', with the most efficient organizations within a group being used to define the standard against which the performance of the other organizations is evaluated. The concept of efficiency is thus relative rather than absolute. The organizations being evaluated are known as decision-making units (DMUs).

The starting point for our analysis is the Charnes-Cooper-Rhodes (CCR) model, which assumes constant returns to scale (CRS) and no congestion. In its output-oriented form, this model can be specified as follows:

$$\theta^* = \max \theta \tag{1a}$$

subject to:

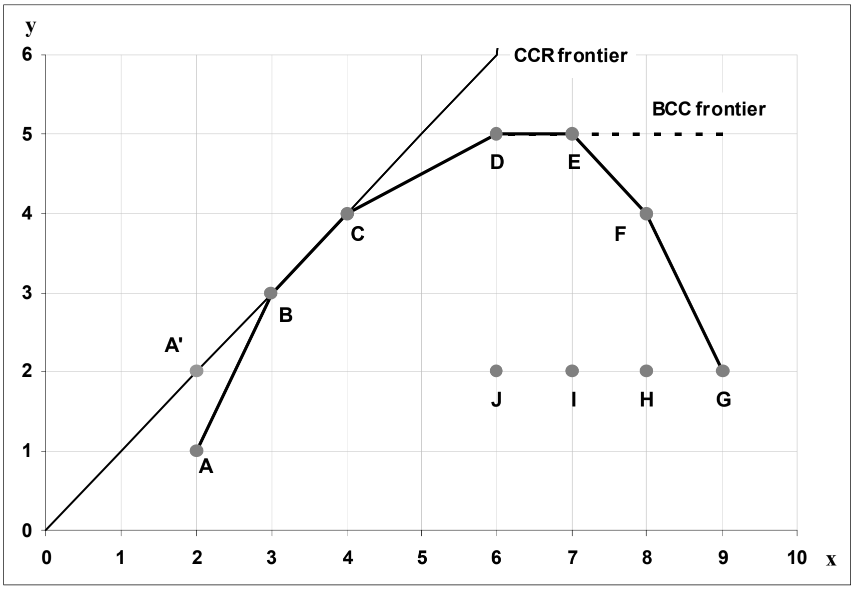
$$\sum_j \lambda_j x_{ij} \le x_{ik} \qquad i = 1, 2, \ldots, m \tag{1b}$$

$$\sum_j \lambda_j y_{rj} \ge \theta y_{rk} \qquad r = 1, 2, \ldots, s \tag{1c}$$

$$\lambda_j \ge 0 \qquad j = 1, 2, \ldots, n \tag{1d}$$

where x_{ij} and y_{rj} are the quantities of input i and output r produced by DMU j, and the λ_j are a set of weights with values to be determined. The model is solved for each DMU k, and an efficiency score, $\theta^* \ge 1$, is thereby produced. It is more convenient, however, to define a new measure of *technical efficiency*, $TE \equiv 1/\theta^*$, so that efficient DMUs have $TE = 1$, whereas inefficient DMUs have $TE < 1$.

Figure 2: DEA models and congestion



It is worth noting that the weights are chosen in such a way that the efficiency of each DMU is maximized. Essentially, the DEA software computes a weighted sum of outputs and a weighted sum of inputs, and then calculates their ratio, subject to the restriction TE $\leq$ 1. The weights are chosen so that each DMU is able to appear in the most favourable light possible. It can play to its strengths and ignore its weaknesses! For instance, if a university decides to specialize in teaching or research and does exceptionally well at its chosen activity, it could end up on the frontier, notwithstanding mediocre performance elsewhere.

The CCR model is illustrated in Figure 2.[7] For simplicity, it is assumed that each DMU employs a single input, x, to produce a single output, y. DMUs B and C operate under CRS and hence are located on the CCR frontier; both have TE = 1. The other DMUs are deemed to be inefficient. For example, A has TE = 0.5, showing that it is producing only half of its potential output; to be efficient, it would need to move to point A' on the frontier.

To capture possible scale effects, we need to modify the CCR model to produce the following Banker-Charnes-Cooper (BCC) model.[8]

$$\phi^* = \max \phi \tag{2a}$$

subject to:

$$\sum_j \lambda_j x_{ij} \leq x_{ik} \qquad i = 1,2,\ldots,m \tag{2b}$$

$$\sum_j \lambda_j y_{rj} \geq \phi y_{rk} \qquad r = 1,2,\ldots,s \tag{2c}$$

$$\sum_j \lambda_j = 1 \tag{2d}$$

$$\lambda_j \geq 0 \qquad j = 1,2,\ldots,n \tag{2e}$$

The crucial difference between these two models is the addition of the convexity constraint $\sum_j \lambda_j = 1$. In Figure 2, this constraint generates a new frontier ABCDE and its horizontal extension from E. This BCC frontier exhibits variable returns to scale (VRS). The BCC model is solved in two stages. In the first stage, ϕ^* is evaluated for each DMU k, while the second stage involves maximizing the sum of the slacks, conditional on this value of ϕ^* (cf. Cooper *et al.*, 2000b, pp. 3-5).

In terms of the new model, A and D are now regarded as being efficient. However, even though E has $\phi^* = 1$, it is still deemed to be inefficient owing to the *slack* of one unit in x. Notice that x can be reduced by one unit without affecting y. Identifying inefficiencies of this kind is the aim of the second stage of the BCC model.

To measure *scale efficiency*, we can define a new ratio, $SE \equiv \phi^* / \theta^*$. This yields SE = 1 for B and C but values of 0.5 for A, 0.833 for D and 0.714 for E. The diagram shows that A is subject to increasing returns to scale, whereas D and E are subject to decreasing returns. What is more, all of the inefficiency of these three DMUs can be attributed to the fact that they are operating at an inappropriate scale.

As regards F and G, it is clear from Definition 2 above that both DMUs would be regarded as being congested. This is because y and x are inversely related over the relevant part of the frontier. By contrast, E would be held to be technically inefficient rather than congested. This is because y is constant over the range $x = 6$ to $x = 7$. Classifying the remaining three DMUs is a little more complicated but their situation becomes clearer once we project them onto the frontier: J to D, I to E and H to F. Once this is done, it is evident that only H suffers from congestion. Even so, all three DMUs do suffer from pure technical inefficiency. This is because they are located beneath the frontier *ABCDEFG*.

4. Cooper's measure of congestion

Although congestion can be measured in several different ways, we have opted in this paper to use the method proposed by Cooper *et al.* (2000b, 2001a, 2001b) as the basis for our analysis. For simplicity, this procedure is referred to hereafter as Cooper's method. An alternative approach is considered in the next section and reasons are given there for our choice of Cooper's method.

Cooper's measure of congestion, denoted here by C_C, is calculated from the results of the BCC model. It involves a straightforward decomposition of the slacks from this model. The first step is to specify a formula for calculating the amount of congestion:

$$c_i = s_i^* - \delta_i^* \tag{3}$$

where c_i is the amount of congestion associated with input i, s_i* is the total amount of slack in input i and δ_i* is the amount of slack attributable to technical inefficiency (cf. Cooper *et al.*, 2001b, p. 69). The asterisks denote optimal values generated by the DEA software. The measured amount of congestion is thus a residual derived from the DEA results. We can then rewrite equation (3) as follows:

$$c_i / x_i = s_i^* / x_i - \delta_i^* / x_i \tag{4}$$

where c_i/x_i is the proportion of congestion in input i, $s_i{}^*/x_i$ is the proportion of slack in input i and $\delta_i{}^*/x_i$ is the proportion of technical inefficiency in input i. The final step is to take arithmetic means over all inputs to get:

$$\mathrm{C_C} = \overline{s/x} - \overline{\delta/x} \tag{5}$$

Hence C_C measures the average proportion of congestion in the inputs used by a particular DMU. It has the property $0 \le C_C \le 1$. See Cooper *et al.* (2001b, p. 73).

To illustrate the meaning of Cooper's measure, let us return to Figure 2. It was noted earlier that DMUs F and G were both congested but we now need to measure the extent of this congestion. G will be examined first. The diagram reveals that there are two DMUs that could be used for evaluating G, viz *D* and *E*. However, although both would yield $\phi^* = 2.5$, D is the one that would maximize the slack in input x (giving three units rather than two). Hence *D* is the DMU picked out by Cooper's model for the purpose of evaluating *G*. In this instance, the three units of slack in input x obtained from the BCC model would be divided into two units of congestion and one unit of technical inefficiency. In terms of equation (5), we would have $\overline{s/x} = 3/9$ and $\overline{\delta/x} = 1/9$, giving $C_C = 2/9 \approx 0.222$. Similarly, we can calculate $C_C = (2/8 - 1/8) = 0.125$ for *F* (and likewise for *H*). *I* and *J* would be free from congestion.

It is worth noting that, in real data sets, horizontal segments such as *DE* in Figure 2 are rare and, in our own data set of former polytechnics, we found no instance where slack existed, yet $\phi^* = 1$. If the data set does not have any DMUs such as *E*, then the amount of congestion for each input equals the BCC slack for this input. This greatly simplifies the computation of C_C because the second stage of Cooper's procedure is redundant.[9]

5. An alternative approach

The most widely used method of measuring congestion is not, in fact, the procedure outlined above; instead, most researchers have followed the approach associated with Färe and Grosskopf.[10] An attractive feature of this approach is that its underlying concepts are firmly grounded in economic theory. In addition, as a *radial* method, Färe and Grosskopf's approach has the considerable advantage that a straightforward decomposition of overall technical efficiency into scale, congestion and purely technical components can be effected via the identity:

$$TE \equiv PTE \times SE \times CE \qquad (6)$$

where *PTE* is pure technical efficiency, *SE* is scale efficiency and *CE* is congestion efficiency.

Nonetheless, for several reasons, we opted instead to use Cooper's procedure. The first reason was that this approach made it possible to employ an output orientation to measure congestion of inputs.[11] Here we felt that an objective of maximizing output from given resources would be much closer to the aims of British universities than the alternative of minimizing the resources used to produce a given output. Secondly, as demonstrated elsewhere (Flegg and Allen, 2006b), only certain instances of negative marginal productivity are deemed to constitute congestion under Färe and Grosskopf's approach. Hence this approach would not be consistent with the definitions of congestion given earlier in this paper. Thirdly, Cooper's procedure makes use of concepts that can easily be identified and measured in a set of data. Moreover, his measure of congestion is easy to grasp and one can immediately see which factors are apparently causing the problem and to what extent. This is more difficult to establish from Färe and Grosskopf's procedure.[12]

Having provided a rationale for using Cooper's procedure, we can now consider the outputs and inputs to be used in the DEA.

6. Output variables

Following previous research (Flegg and Allen, 2007), the output of each university will be measured using the following variables:

- the number of undergraduate qualifications awarded (y_1);
- the number of postgraduate qualifications awarded (y_2);
- income from research grants and contracts in £ thousands (y_3).

Sources of data and other details are given in Appendix A.

Undergraduate qualifications

The total number of undergraduate qualifications awarded is undeniably a key component of the output of any university but how should one take account of the *quality* of such qualifications? In Flegg and Allen (2007), we examined this issue by formulating two alternative models: *Model 1* used the sum of first-class honours degrees and upper seconds, whereas *Model 2* used the sum of all undergraduate qualifications, including all degrees irrespective of classification, as well as diplomas and certificates. The latter type of output has become increasingly important in the new universities.[13] To facilitate comparisons with our earlier study of the older universities, we have used the same two models here.

A potential problem with Model 1 is that a university could improve its relative performance by deliberately raising the proportion of firsts and upper seconds awarded. However, if *all* universities did so and to the same extent over time, this would not affect the DEA results because each year is analysed separately.[14] Furthermore, any tendency towards awarding too many firsts and upper seconds would not benefit a university in terms of our Model 2.

Another concern is that students' achievements depend not only on the quality of teaching but also on their effort, ability and initial qualifications. Although we were unable to adjust for any of these attributes, this may not be too serious in the case of effort. This is because universities with well-designed and stimulating degree courses are likely to evoke more effort by their students and hence end up with higher efficiency scores.

The failure to adjust for ability and initial qualifications is potentially much more serious. However, Rodgers (2007, p. 56) notes that the correlation between 'A' level points and degree class is generally weak. What is more, the statistical relationship between these two variables may be nonlinear (*ibid.*, p. 57). His study suggests that a student's academic entry qualifications are but one of many possible determinants of degree class and that one needs to consider factors such as gender, ethnic origin, age and social background as well. For these reasons, along with a desire to maintain the simplicity of our models, we decided not to include a separate input variable measuring entry qualifications.

A final caveat is worth mentioning: the only output recognized is qualifications awarded to students in their final year, even though all undergraduates are used as an input. However, this should not be problematic except where there are marked differences in the rate of growth of the number of undergraduates in different institutions.

Postgraduate qualifications
Figures were available for four separate categories of postgraduate qualification: doctorates, other higher degrees, postgraduate certificates of education, and other postgraduate qualifications. However, for simplicity, and in order to avoid artificially raising the efficiency scores, these categories were combined into a single variable. A drawback of this aggregation is, of course, that it masks any differences in the mix of postgraduate qualifications across institutions. This issue is examined later in the paper. Variable y_2 also fails to allow for any differences in the quality of postgraduate qualifications.

Income from research grants and contracts
Variable y_3 includes, inter alia, income received from research councils, charities, central government, local authorities, health authorities, industry, commerce and public corporations. It includes income from both UK and overseas sources, although income from 'other services rendered' was excluded owing to concerns about the comparability of some of the data.

The importance of research as an output is incontrovertible, yet its measurement is awkward. Indeed, the use of research income as a proxy for research output has often been criticized on the basis that such income should be treated as an input into the research process rather than as an output. Research income may also be distorted by disparities in research costs across subject areas. However, in a study of this nature, one has little option but to use research income as a proxy for research output since annual data for most alternative variables are unavailable. In fact, the use of research income as a proxy for research output is widespread; notable recent examples include Izadi *et al.* (2002), Stevens (2005) and Johnes (2006).

In defence of our use of income from research grants and contracts as an output, we would maintain that such income is likely to reflect the perceived quality, as well as quantity, of research output; it should also offer a more up-to-date picture of such output than, for example, the scores in some earlier research assessment exercise (cf. Stevens, 2005, p. 357).

In similar vein, Izadi *et al.* (2002, p. 66) argue that research grants may be regarded 'as a measure of the market value of the research being undertaken' and that their award is indicative of 'the grantee's strong research performance in the recent past'. A problem with this argument is that much research is speculative in nature; such research does not have an immediate market value, although it may yield valuable 'spin offs' at a later stage. Furthermore, one might argue that an important function of a university is to carry out research of uncertain market value.

7. Input variables

The following inputs are used in the DEA analysis:

- the number of full-time equivalent undergraduate students (x_1);
- the number of full-time equivalent postgraduate students (x_2);
- academic staff expenditure in £ thousands (x_3);
- other expenditure in £ thousands (x_4).

See Appendix A for sources of data and other details. While the first two inputs are self-explanatory, some remarks on the last two are needed.

Academic staff expenditure
This input measures a university's total expenditure on academic staff. It has the advantage, therefore, of being measured in the same units as 'other expenditure'. A possible shortcoming of variable x_3 is that staffing expenditure will vary with the proportion of staff on different grades and only approximately with the number of staff hours available for teaching, research, administration, etc. For this reason, an alternative variable — the full-time equivalent number of academic staff — is examined in our third DEA model.

Other expenditure
Variable x_4 measures a university's total expenditure minus its academic staff expenditure. It encompasses expenditure on academic cost centres, academic services, administration and central services, premises, residences and catering, and on research grants and contracts.

8. Technical efficiency

At the outset, let us examine the overall *technical efficiency* (*TE*) of the former polytechnics in the period 1995/6 to 2003/4. Table 1 displays the findings from three alternative models.[15] *Model 1* is the one outlined earlier, in which the output of undergraduate qualifications is measured by the sum of firsts and upper seconds. This output variable is replaced in *Model 2* by the sum of all undergraduate awards.[16] *Model 3* is a modified version of Model 2, whereby the number of full-time equivalent staff takes the place of expenditure on academic staff. All models contain three outputs and four inputs. The sample comprises 41 institutions up to 2001/2 but 40 thereafter. This is due to the merger of London Guildhall University and the University of North London to form London Metropolitan University.[17]

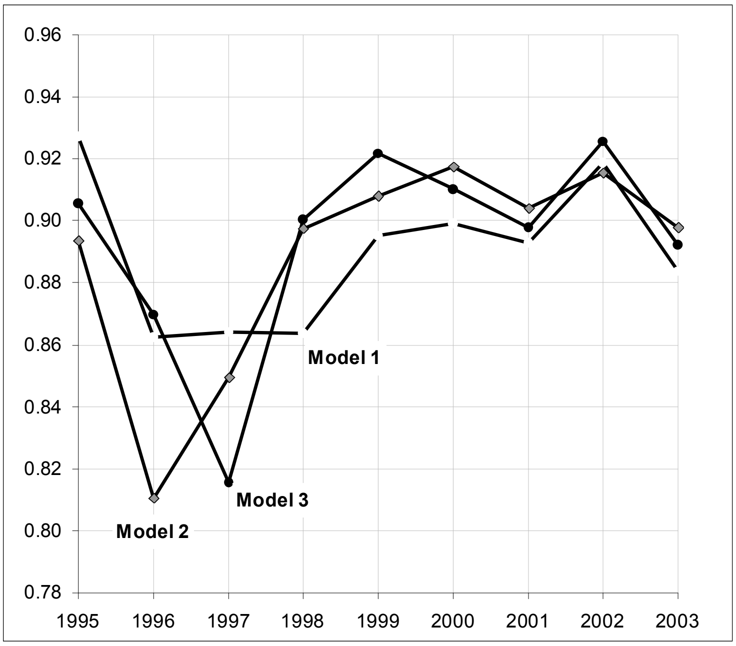
Along with the annual unweighted arithmetic mean *TE* scores for each model, Table 1 also shows the corresponding weighted arithmetic mean scores, which were calculated using each university's share of the total number of students as a weight. This was done to take account of the unequal size of universities (see Table 8). The unweighted results, which are also illustrated in Figure 3, will be considered first.

Table 1. Annual mean technical efficiency (*TE*) scores for alternative models

	Unweighted mean TE score	*Weighted mean TE score*	*Difference*	*Minimum*	*Standard deviation*	*Number on frontier*
Model 1						
1995/6	0.927	0.938	0.011	0.493	0.118	19
1996/7	0.862	0.875	0.012	0.487	0.144	13
1997/8	0.864	0.875	0.011	0.527	0.134	13
1998/9	0.864	0.885	0.021	0.428	0.147	13
1999/0	0.895	0.908	0.013	0.577	0.129	20
2000/1	0.899	0.919	0.020	0.430	0.146	20
2001/2	0.893	0.906	0.014	0.483	0.132	14
2002/3	0.919	0.927	0.009	0.597	0.099	16
2003/4	0.884	0.887	0.003	0.531	0.118	11
Model 2						
1995/6	0.894	0.899	0.006	0.586	0.123	12
1996/7	0.811	0.813	0.002	0.497	0.164	10
1997/8	0.849	0.853	0.004	0.580	0.132	9
1998/9	0.897	0.904	0.006	0.477	0.107	9
1999/0	0.908	0.910	0.002	0.722	0.100	17
2000/1	0.918	0.925	0.008	0.526	0.103	15
2001/2	0.904	0.909	0.005	0.536	0.111	14
2002/3	0.915	0.916	0.001	0.616	0.096	15
2003/4	0.898	0.894	-0.004	0.511	0.104	11
Model 3						
1995/6	0.906	0.910	0.004	0.594	0.111	14
1996/7	0.870	0.873	0.004	0.501	0.137	11
1997/8	0.816	0.816	0.000	0.580	0.140	7
1998/9	0.900	0.906	0.006	0.546	0.099	8
1999/0	0.922	0.923	0.002	0.743	0.083	17
2000/1	0.910	0.916	0.006	0.519	0.107	15
2001/2	0.898	0.903	0.005	0.545	0.111	14
2002/3	0.926	0.926	0.000	0.602	0.091	15
2003/4	0.892	0.888	-0.004	0.485	0.108	11

If we ignore the erratic results for the first three years, then the unweighted mean *TE* scores from Model 2 exceed those from Model 1 in five years out of six. This is evident from both Table 1 and Figure 3. This outcome probably reflects the fact that it is possible, with Model 2, to substitute one type of undergraduate award for another, while keeping the overall number of awards constant, e.g. an upper second could be replaced by a lower second. This would tend to lessen the intertemporal fluctuations in output and reduce the dispersion in *TE* scores across universities. This, in turn, would tend to raise the mean *TE* scores.

Figure 3: Mean technical efficiency



If we again ignore the first three years, then Table 1 and Figure 3 also reveal that the unweighted mean *TE* scores from Models 2 and 3 are not that different. What is more, there is no tendency for these results to diverge in a systematic way. This suggests that it may not make much difference to the conclusions whether one measures the input of academic staff in terms of full-time equivalents or expenditure. The close relationship between Models 2 and 3 for the last six years was confirmed by the finding of a strong positive correlation of 0.943 between the 244 individual *TE* scores generated by each model. By contrast, $r = 0.739$ for Models 1 and 2.

If we now look at the first three years, it is surprising that the minima of the graphs for Models 2 and 3 occur in different years. However, this may merely reflect possible errors in the data for full-time equivalent academic staff. This series is much more erratic than the corresponding one for academic staff expenditure, and we observed some very large annual changes in the FTE figures for some institutions, especially in the earlier years.

With regard to weighting, the 'Difference' column in Table 1 shows that this procedure enhances the mean scores for Model 1, albeit by a modest amount in most cases. For Models 2 and 3, the weighting slightly raises the mean scores in all years apart from 2003/4. Taking the results as a whole, however, there is a clear tendency for the scores from the different models to converge during the period under review.

Whilst the mean levels of technical efficiency are generally fairly high, there is no evidence of an upward trend, especially from 1998/9 onwards. Indeed, all of the models show that the rise in 2002/3 was offset by a downturn in the final year. This is true for both weighted and unweighted scores. It is worth noting too that all models record a fall in the number of frontier universities in the final year. However, it should be borne in mind that the *TE* scores do not measure technical efficiency in an absolute sense but instead measure it relative to the frontier in each year. Hence the drop in the mean *TE* scores in 2003/4 could mean that the universities were moving further away from a static frontier or, alternatively, that the frontier had shifted outwards.[18] It may be noted, finally, that the mean *TE* scores being discussed here are somewhat lower than the comparable scores we obtained for 45 older British universities over the same period (Flegg and Allen, 2007). This suggests a greater degree of heterogeneity in the sample of former polytechnics.

9. Subject mix

Before examining the extent of congestion in the former polytechnics, it is worth considering the possible impact of a university's subject mix on its efficiency. This mix could affect efficiency in two main ways: by changing a university's costs and income, and by influencing the classification of its undergraduate awards. For instance, a science-oriented university might incur higher costs and receive more research income than an arts-oriented university. In addition, the proportion of first-class degrees and upper seconds might vary with a university's subject mix.

To explore this issue, we first calculated the percentage, π, of 'science' students in each university in 2003/4, using data for all students in the following ten subject areas: medicine and dentistry, subjects allied to medicine, biological sciences, veterinary science, agriculture and related subjects, physical sciences, mathematical sciences, computer science, engineering and technology, and architecture, building and planning.[19]

We then measured the correlation between π and undergraduate results. This analysis yielded a correlation of -0.327 (p = 0.040) between π and the proportion of first-class degrees and upper seconds in 2003/4.[20] This negative correlation, which is significant at the five per cent level, may reflect lower entry requirements on many science-based courses in the former polytechnics, as compared with arts-based courses. The next task was to examine the relationship between π and efficiency scores.

Table 2 presents the values of π, along with two sets of efficiency scores obtained from Models 1 and 2. For simplicity, the discussion is confined to those two models and to 2003/4. The BCC efficiency scores, referred to hereafter as β scores, measure technical efficiency after removing any scale effects. The 'super-efficiency' scores will be discussed later.

Table 2. Efficiency scores and universities' 'science' orientation in 2003/4

	Percentage of 'science' students (π)	Model 1		Model 2	
		BCC efficiency score	Super-efficiency score	BCC efficiency score	Super-efficiency score
Abertay Dundee	63.6	1	1	1	1
Anglia Polytechnic	42.8	0.8647	0.8372	0.7980	0.7935
Bournemouth	39.1	1	1.0445	0.8688	0.7170
Brighton	46.6	0.9443	0.8902	0.8515	0.8271
Central England	52.2	1	1.0063	0.9613	0.7689
Central Lancashire	48.8	1	1.0268	1	1.0262
Coventry	54.2	1	1.0263	1	1.0203
De Montfort	43.7	1	1.1103	1	1.1131
Derby	37.5	0.9902	0.7618	0.9560	0.6886
East London	45.0	0.7914	0.5922	0.8424	0.6118
Glamorgan	44.3	0.7802	0.6733	0.7707	0.6317
Glasgow Caledonian	57.2	0.8027	0.5449	0.9607	0.6021
Greenwich	46.6	1	1.0496	1	1.0496
Hertfordshire	52.5	0.7777	0.6865	0.8245	0.6958
Huddersfield	37.7	0.9821	0.7532	1	1.0131
Kingston	46.8	0.9057	0.6170	0.9081	0.5999
Leeds Metropolitan	38.4	1	1.0075	1	1.0172
Lincoln	34.5	1	1.0636	1	1.0286
Liverpool J Moores	52.2	1	1.1213	1	1.0752
London Metropolitan	26.0	0.8319	0.5851	0.9219	0.6418
London South Bank	64.4	0.7577	0.7248	0.7361	0.6919
Luton	43.7	1	1.3382	1	1.1448
Manchester Met	35.6	0.9966	0.9163	1	1.0364
Middlesex	38.3	0.9842	0.4610	1	1.0077
Napier	57.7	0.7104	0.6219	0.9187	0.6905
Northumbria	47.4	0.8785	0.7143	0.9099	0.7226
Nottingham Trent	25.3	1	1.0103	0.9917	0.6234
Oxford Brookes	39.7	1	1.2003	1	1.2074
Paisley	50.3	1	1.0204	1	1.0697
Plymouth	49.3	0.9266	0.8711	0.9015	0.8482
Portsmouth	41.5	0.8167	0.7752	0.9527	0.8133
Robert Gordon	54.3	1	1.0764	1	1.0764
Sheffield Hallam	45.0	1	1.2609	1	1.1922
Staffordshire	52.0	1	1.1133	1	1.0780
Sunderland	32.2	0.9923	0.9799	0.9582	0.8742
Teesside	66.1	0.8378	0.6144	0.8269	0.6070
Thames Valley	50.2	0.5319	0.2556	0.5248	0.2444
West of England	46.0	0.8971	0.7670	0.8537	0.7448
Westminster	37.1	0.9334	0.8935	0.9185	0.8731
Wolverhampton	41.0	1	1.0028	0.9901	0.7925
Mean	45.7	0.9234	0.8754	0.9287	0.8565
Standard deviation	9.2	0.1074	0.2316	0.0983	0.2110
Correlation with π		-0.302	-0.124	-0.245	-0.082

The BCC results play a key role in Cooper's analysis of congestion. β = 1 implies that a university is located on the BCC frontier and hence is uncongested by definition, whereas β < 1 opens up the possibility of congestion. In the case of Model 1, the BCC method identifies 18 efficient and 22 inefficient universities. It is evident that the non-frontier institutions vary greatly in terms of performance. For instance, β = 0.982 for Huddersfield, which indicates that it was producing 98.2 per cent of its potential output; it was, therefore, close to being efficient. A striking contrast is offered by Thames Valley. This university has β = 0.532, which suggests that it was producing only 53.2 per cent of its potential output, the lowest percentage in the sample. Switching to Model 2 results in some changes in the frontier institutions, with three promotions and four demotions. However, the mean value of β rises only slightly, from 0.9234 to 0.9287.

Using the data in Table 2 for Model 1, we found a correlation of -0.302 between β and π (p = 0.059). Model 2 yielded a somewhat weaker correlation of -0.245 (p = 0.129); this outcome can be explained by the use of a broader measure of undergraduate output. The fact that both correlations are negative suggests that science-oriented universities tend to be less efficient than arts-oriented universities. However, neither correlation is significant at the five per cent level. Even so, it is interesting that some of the most science-oriented universities have relatively low efficiency scores, e.g. London South Bank and Teesside.

A shortcoming of the analysis hitherto is that the BCC model does not discriminate between efficient DMUs; all are assigned unitary scores. This makes it difficult to measure the true relationship between subject mix and efficiency. For this reason, we also computed the 'super-efficiency' scores shown in Table 2, which are referred to hereafter as σ scores.[21] The higher the value of σ, the more efficient a university is considered to be.

The values of σ are calculated by deleting each frontier university in turn from the data set and then examining what effect this has on the frontier. If this frontier is unchanged, then $\sigma = 1$. Looking at the results for Model 1, Abertay Dundee is an example of a university whose presence or absence has no effect on the scores for the remaining universities. By contrast, universities like Oxford Brookes, Sheffield Hallam and Luton have a big impact on the position of the frontier and hence have correspondingly large values of σ. Deleting such universities would cause noticeable changes in the scores for several of the remaining universities.

The BCC and super-efficiency methods are in accord in determining which universities are inefficient but in discord about the extent of this inefficiency; what is more, σ < β in all cases but one. For instance, β = 0.982 for Huddersfield, yet σ = 0.753, which indicates a great deal more inefficiency. Another striking contrast is offered by Thames Valley, which has β = 0.532 but σ = 0.256, even though both methods find it to be the least efficient university of all. The reason for the differences in the values of β and σ is that the two methods use different formulae to project inefficient DMUs onto the efficiency frontier. The BCC projection is a radial (proportional) method, whereas the super-efficiency method employs a slacks-based measure, which is non-radial (Tone, 2002).

Using the data in Table 2 for Model 1, we found a correlation of -0.124 between σ and π (p = 0.449). Model 2 generated an even weaker correlation of only -0.082 (p = 0.615). The weakness of these correlations is an important finding because, by allowing for variation in performance within the group of frontier universities, the super-efficiency scores are arguably the most reliable way of measuring the association between subject mix and efficiency.

The absence of a strong association between subject mix and efficiency can probably be ascribed to the fact that we included income from research grants and contracts as an output, and research expenditure as one component of 'other expenditure' (our input x_4). Thus there is no in-built bias in our study in favour of any particular type of university.[22]

In the light of these findings, it seems unnecessary to complicate the analysis of congestion by including a measure of subject mix as a separate input variable.[23] Indeed, by adding another variable in this way, we would most likely have impaired the discriminatory power of the DEA models, while adding little of benefit in terms of additional information.[24]

10. Results from Cooper's procedure

The first step in implementing Cooper's procedure was to work out C_C, the average proportion of congestion in the inputs employed by each institution in each year. This was done by using the *DEA-Solver-Pro* software (www.saitech-inc.com) to derive the BCC slacks and *Excel* to perform some supplementary calculations. These congestion scores were then averaged, first over all universities, and then over the congested universities alone. Both weighted and unweighted means were computed for the whole sample. However, Table 3 shows that the differences between the weighted and unweighted means are typically very small and, for simplicity, only the latter will be discussed here. These unweighted mean scores are denoted by the symbol $\overline{C_C}$.

Figure 4 reveals many similarities, along with some noteworthy differences, in the results for the three models. For instance, the values of $\overline{C_C}$ from Models 2 and 3 follow essentially the same path from 1998/9 onwards, and both models suggest a decline in congestion over the period as a whole. What is more, again from 1998/9 onwards, the values of $\overline{C_C}$ from Model 1 are much the same as those from the other two models, although 2000/1 is an anomaly here. All three models also indicate that, in the final two years, congestion first fell and then rose, whereas technical efficiency did the opposite.

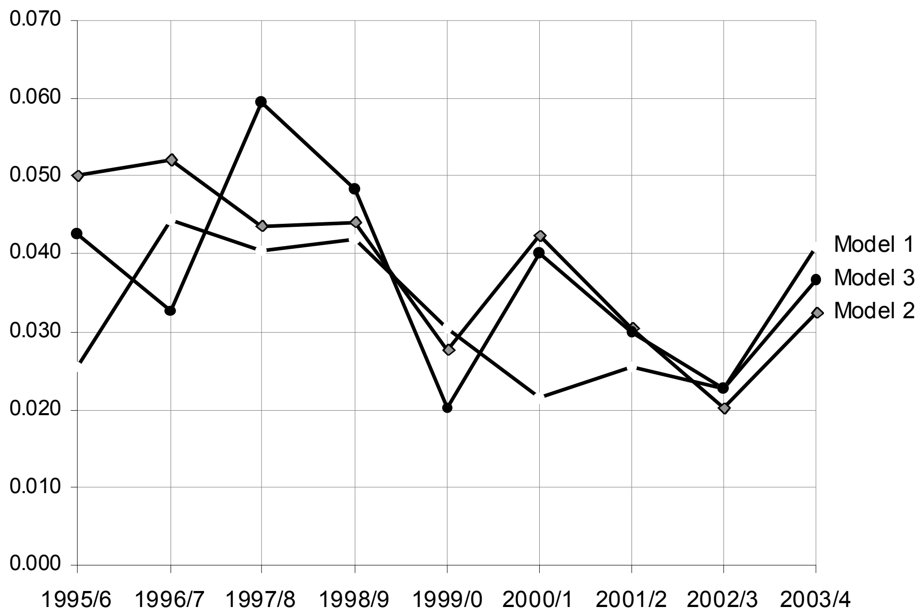
Although there is a large gap between the graphs for Models 2 and 3 in both 1996/7 and 1997/8, it is worth recalling that the mean *TE* scores from these two models also fluctuated erratically in these two years (see Figure 3), possibly as a result of errors in the data. Hence we are loath to pay too much attention to the outcomes in these two years. Indeed, for the period as a whole, the similarities in the values of $\overline{C_C}$ from the three models are more striking than the differences. For this reason, it seems sensible to base most of the subsequent discussion on the findings from a single model. Model 2 was chosen for this purpose because it yields the most stable results in terms of congestion; it also appears to be less affected than Model 1 by differences in subject mix.[25]

Table 3. Annual mean congestion scores for alternative models

	All universities				*Congested universities*	
	Unweighted mean score	*Weighted mean score*	*Difference*	*Standard deviation*	*Number*	*Unweighted mean score*
Model 1						
1995/6	0.0254	0.0255	0.0001	0.164	17	0.0612
1996/7	0.0443	0.0467	0.0025	0.284	24	0.0756
1997/8	0.0403	0.0403	0.0000	0.258	23	0.0719
1998/9	0.0418	0.0405	-0.0013	0.267	25	0.0686
1999/0	0.0304	0.0296	-0.0007	0.194	19	0.0655
2000/1	0.0214	0.0192	-0.0022	0.138	16	0.0549
2001/2	0.0254	0.0251	-0.0004	0.162	21	0.0497
2002/3	0.0226	0.0204	-0.0023	0.144	18	0.0503
2003/4	0.0409	0.0465	0.0056	0.259	22	0.0744
Model 2						
1995/6	0.0501	0.0483	-0.0018	0.065	24	0.0857
1996/7	0.0521	0.0536	0.0015	0.060	28	0.0763
1997/8	0.0435	0.0455	0.0020	0.062	24	0.0743
1998/9	0.0441	0.0476	0.0035	0.056	27	0.0670
1999/0	0.0275	0.0276	0.0001	0.039	22	0.0513
2000/1	0.0423	0.0427	0.0004	0.053	23	0.0754
2001/2	0.0304	0.0301	-0.0003	0.035	22	0.0566
2002/3	0.0201	0.0178	-0.0023	0.032	20	0.0401
2003/4	0.0325	0.0334	0.0010	0.040	23	0.0564
Model 3						
1995/6	0.0425	0.0434	0.0009	0.272	23	0.0758
1996/7	0.0327	0.0362	0.0036	0.210	23	0.0582
1997/8	0.0595	0.0584	-0.0011	0.378	26	0.0938
1998/9	0.0484	0.0522	0.0038	0.308	27	0.0735
1999/0	0.0203	0.0220	0.0017	0.131	21	0.0396
2000/1	0.0401	0.0395	-0.0005	0.255	22	0.0747
2001/2	0.0299	0.0301	0.0002	0.190	24	0.0511
2002/3	0.0226	0.0220	-0.0006	0.143	19	0.0477
2003/4	0.0366	0.0379	0.0013	0.230	23	0.0636

Model 2 signifies that congestion for the whole sample fell from an average of five per cent of inputs in 1995/6 to a more modest 3.25 per cent in 2003/4. However, in assessing the degree of congestion, it may be more appropriate to focus on the congested universities alone. For instance, congestion averaged 5.64 per cent for the 23 congested universities in 2003/4, which suggests that congestion was a more serious problem than it might initially have appeared to be.

Figure 4: Mean congestion scores



11. Sensitivity analysis: postgraduate qualifications

The postgraduate output variable used thus far, y_2, is very broadly defined, so it is of interest to see whether this breadth has an undue influence on the results. More specifically, y_2 is the sum of doctorates (p_1), other higher degrees (p_2), postgraduate certificates of education or PGCEs (p_3) and other postgraduate qualifications (p_4). Some information on the components of y_2 in 2003/4 is shown in Table 4. One can see from the summary statistics that there is quite a lot of variation across universities in the mix of postgraduate qualifications awarded. This variation is clearly evident in Figure 5, which shows this mix plotted by institution. For ease of interpretation, the universities have been ordered in terms of the proportion of 'other postgraduate qualifications'. The figure is largely self-explanatory, although it may be worth noting that 12 of the 40 former polytechnics do not offer PGCEs; typically, this gap is filled by additional 'other higher degrees'.

Figure 5: Mix of postgraduate qualifications in 2003/4

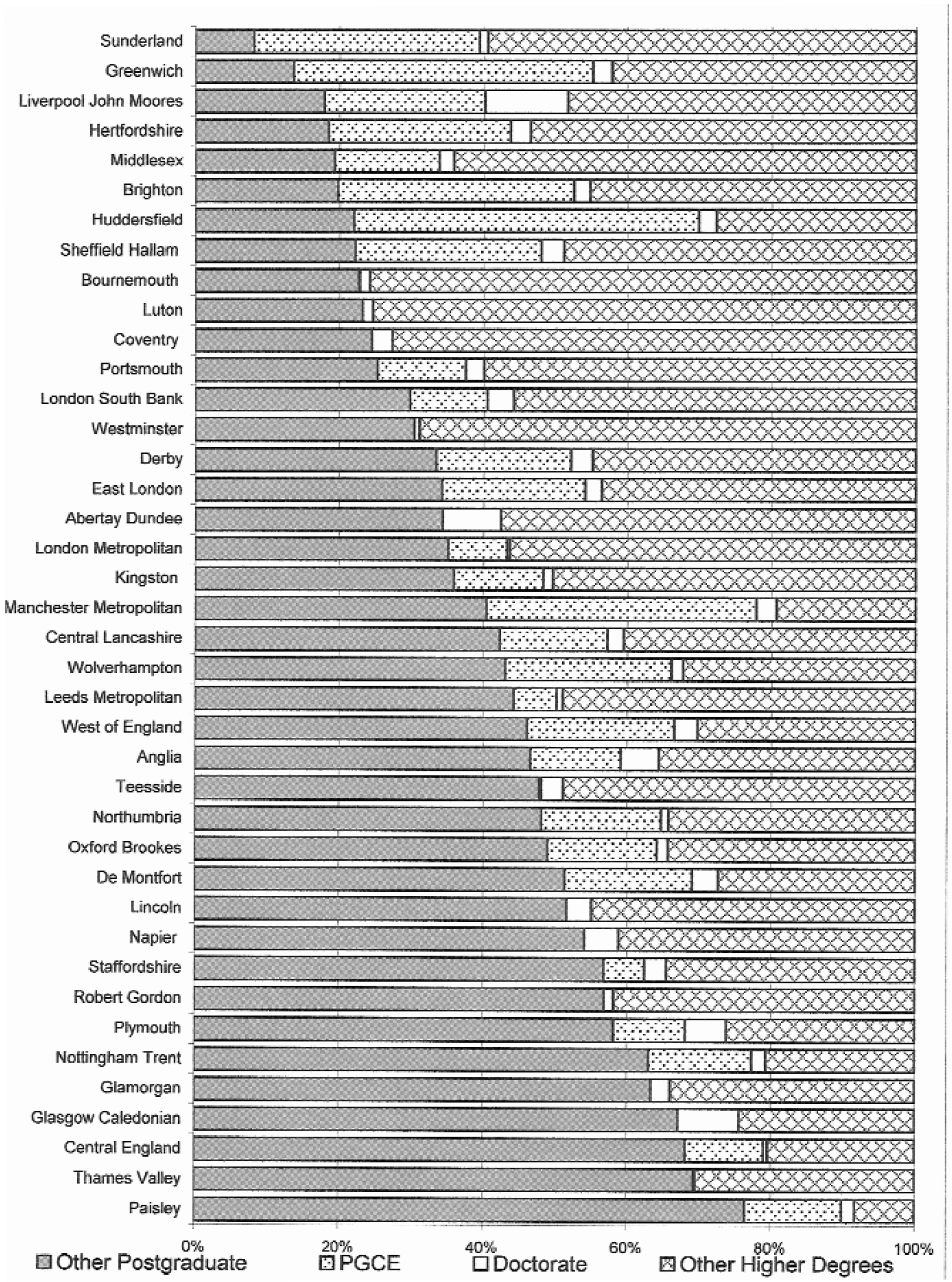


Table 4. Summary statistics for postgraduate qualifications in 2003/4

	Doctorates	*Other higher degrees*	*PGCEs*	*Other postgraduate qualifications*	*All postgraduate qualifications*
Mean	34	594	215	539	1382
Standard deviation	25	404	226	380	734
Coefficient of variation	0.735	0.680	1.051	0.705	0.531
Median	28	506	182	433	1349
Minimum	0	55	0	35	100
Maximum	135	2065	885	1700	3475

Note: Minima and maxima have been rounded in accordance with HESA's policy.

To test for sensitivity, we specified three new postgraduate output variables:

$$y_{2a} \equiv p_1 + p_2; \; y_{2b} \equiv p_3 + p_4; \; y_{2c} \equiv 4p_1 + 2p_2 + p_3 + p_4$$

The rationale for splitting y_2 into y_{2a} and y_{2b} was to take account of the possibility that a university might wish to specialize in producing a particular type of postgraduate qualification. The weighted output, y_{2c}, is an attempt to recognize the enhanced academic status attached to higher degrees, along with the associated resource costs, e.g. in supervising and examining dissertations. Giving a doctorate twice the weight of a masters degree and four times that of a PGCE or professional qualification is, of course, arbitrary but so too is the unitary weighting scheme attached to the original variable, y_2. The weights underlying y_{2c} are not unreasonable and experiments with different weights did not yield very different results.[26]

Table 5 shows that the aggregated postgraduate output variable, y_2, and the weighted variant, y_{2c}, yield very similar BCC efficiency scores and the mean hardly changes. Indeed, over half of the scores are identical or very similar. The number of efficient universities remains the same at 17, although this outcome hides the fact that Manchester Metropolitan switches from efficient to inefficient, whereas Westminster does the opposite. This occurs because Manchester Metropolitan has a relatively low proportion of higher degrees, whereas Westminster's postgraduate provision is skewed the other way. The congestion scores are also very similar on the whole, although the mean does rise from 0.0325 to 0.0375.

By contrast, disaggregating the postgraduate output variable has a more noticeable impact; in particular, the mean BCC score rises from 0.9287 to 0.9416 and four more universities become efficient. The latter outcome occurs because the output of Central England and Nottingham Trent is skewed towards 'other postgraduate qualifications', Westminster's provision is skewed towards 'other higher degrees' and Wolverhampton has a high proportion of non-degree postgraduate output. Nonetheless, the congestion scores are, on average, less affected than the BCC scores; the mean value of C_C drops

Table 5. Sensitivity analysis: postgraduate output measures (Model 2, 2003/4)

	BCC efficiency scores			Congestion scores		
	Single unweighted variable	*Separate variables*	*Single weighted variable*	*Single unweighted variable*	*Separate variables*	*Single weighted variable*
Abertay Dundee	1	1	1	0	0	0
Anglia Polytechnic	0.7980	0.8019	0.8090	0.0624	0.0673	0.0639
Bournemouth	0.8688	0.9573	0.9261	0.0320	0.1128	0.0792
Brighton	0.8515	0.8559	0.8510	0.0088	0.0237	0.0296
Central England	0.9613	1	0.8827	0.0844	0	0.0573
Central Lancashire	1	1	1	0	0	0
Coventry	1	1	1	0	0	0
De Montfort	1	1	1	0	0	0
Derby	0.9560	0.9665	0.9302	0.0184	0.0086	0.0065
East London	0.8424	0.8524	0.8563	0.0920	0.1071	0.1088
Glamorgan	0.7707	0.7707	0.7707	0.0224	0.0224	0.0224
Glasgow Caledonian	0.9607	0.9607	0.9607	0.1370	0.1370	0.1370
Greenwich	1	1	1	0	0	0
Hertfordshire	0.8245	0.8246	0.8244	0.0389	0.0389	0.0389
Huddersfield	1	1	1	0	0	0
Kingston	0.9081	0.9085	0.8890	0.0100	0.0090	0.0236
Leeds Metropolitan	1	1	1	0	0	0
Lincoln	1	1	1	0	0	0
Liverpool J Moores	1	1	1	0	0	0
London Metropolitan	0.9219	0.9899	0.9769	0.0584	0.0757	0.1225
London South Bank	0.7361	0.7999	0.7941	0.0474	0.0748	0.0850
Luton	1	1	1	0	0	0
Manchester Met	1	1	0.9825	0	0	0.1501
Middlesex	1	1	1	0	0	0
Napier	0.9187	0.9187	0.9187	0.0444	0.0444	0.0444
Northumbria	0.9099	0.9114	0.9099	0.0766	0.0725	0.0766
Nottingham Trent	0.9917	1	0.9299	0.0961	0	0.0705
Oxford Brookes	1	1	1	0	0	0
Paisley	1	1	1	0	0	0
Plymouth	0.9015	0.9481	0.8800	0.0558	0.0355	0.0484
Portsmouth	0.9527	0.9647	0.9583	0.0253	0.0231	0.0206
Robert Gordon	1	1	1	0	0	0
Sheffield Hallam	1	1	1	0	0	0
Staffordshire	1	1	1	0	0	0
Sunderland	0.9582	0.9728	0.9582	0.1284	0.1235	0.1284
Teesside	0.8269	0.8339	0.7976	0.1107	0.1142	0.0906
Thames Valley	0.5248	0.5493	0.5174	0.0324	0.0761	0.0446
West of England	0.8537	0.8777	0.8272	0.0275	0.0124	0.0403
Westminster	0.9185	1	1	0.0692	0	0
Wolverhampton	0.9901	1	0.9615	0.0197	0	0.0121
Mean	0.9287	0.9416	0.9278	0.0325	0.0295	0.0375
Standard deviation	0.0983	0.0931	0.0972	0.0393	0.0419	0.0448
Number on frontier	17	21	17	17	21	17

only marginally from 0.0325 to 0.0295 when the postgraduate output variable is disaggregated.

It seems fair to conclude that the results are not very sensitive to changes in the postgraduate output variable. For this reason, we decided that it would not be worthwhile to forego the simplicity of our original model by incorporating either of the modifications discussed above.

12. Decomposing congestion

An advantage of Cooper's approach is that it is possible to measure, for each congested university, the contribution of each input to the observed amount of congestion. Table 6 takes a closer look at this feature of his approach, using annual means to summarize the data. The table shows a decomposition by input of the annual unweighted mean value of C_C. To explain these calculations, consider the results for Model 2 in 2003/4. Table 6 shows $\overline{C_C}$ = 0.0564. In the case of undergraduates, congestion averaged 0.0954 (see Table 7), which is 42.3 per cent of 4 × 0.0564. The other percentage contributions shown in Table 6 were computed in the same way. It is worth noting that these contributions would have been exactly the same if all, rather than just congested, universities had been considered.

The results for *Model 1* indicate that an overabundance of undergraduate students was the largest single cause of congestion in the former polytechnics during the period under review. On average, such students accounted for 34.5 per cent of the value of $\overline{C_C}$. However, the results suggest that academic overstaffing was also a major cause of congestion in these new universities! Indeed, at 30.8 per cent, the average share of academic staff is not far behind that of undergraduates. By contrast, the results suggest that postgraduates and 'other expenditure' played a noticeably smaller role in generating congestion.

The pre-eminence of undergraduates in generating congestion is confirmed by the results from Model 2. Indeed, there is now a noticeably wider gap between the average shares of undergraduates and academic staff. With an average share of 26.0 per cent, academic staff are now clearly in second place. What is surprising is that the switch from a narrower to a broader measure of undergraduate output has had so little impact on the share of undergraduates. As regards postgraduates and 'other expenditure', the results show that these two inputs have gained in importance, although their respective shares are still similar. It is worth noting that these various changes in shares have little impact on the overall mean value of $\overline{C_C}$.

As expected, the results for Models 2 and 3 are broadly similar and there is again hardly any change in the overall mean value of $\overline{C_C}$. Undergraduates are shown once more to be the largest single factor underlying congestion, with an average share that is only slightly lower than before. Nonetheless, some changes are worth noting. In particular, as a result of using full-time equivalents rather than expenditure, there is a further appreciable fall in the average share of academic staff and concomitant rise in the shares of postgraduates and 'other expenditure'. The average shares of these three inputs are now of roughly comparable size.

Table 6. Percentage contribution of each input to the annual unweighted mean congestion score, $\overline{C_C}$, in congested universities

	Other expenditure	*Academic staff*	*Post-graduates*	*Under-graduates*	*Number congested*	$\overline{C_C}$
Model 1						
1995/6	19.7	33.3	13.1	33.9	17	0.0612
1996/7	17.0	33.9	20.0	29.1	24	0.0756
1997/8	11.6	31.8	36.5	20.1	23	0.0719
1998/9	5.9	41.9	23.7	28.5	25	0.0686
1999/0	17.6	29.2	7.2	46.0	19	0.0655
2000/1	16.6	24.9	10.8	47.7	16	0.0549
2001/2	25.7	18.7	19.0	36.6	21	0.0497
2002/3	20.1	35.7	8.5	35.7	18	0.0503
2003/4	22.2	27.8	17.4	32.6	22	0.0744
Mean	17.4	30.8	17.4	34.5		0.0636
Model 2						
1995/6	14.0	39.5	15.0	31.5	24	0.0857
1996/7	20.5	27.7	14.9	36.9	28	0.0763
1997/8	14.3	26.1	35.6	24.0	24	0.0743
1998/9	8.9	36.1	34.0	21.0	27	0.0670
1999/0	30.9	23.8	15.6	29.7	22	0.0513
2000/1	18.5	20.6	26.3	34.6	23	0.0754
2001/2	26.8	15.4	20.5	37.3	22	0.0566
2002/3	28.8	15.0	13.7	42.5	20	0.0401
2003/4	16.0	29.5	12.3	42.3	23	0.0564
Mean	19.8	26.0	20.9	33.3		0.0648
Model 3						
1995/6	11.8	33.9	19.1	35.2	23	0.0758
1996/7	21.8	20.0	20.9	37.3	23	0.0582
1997/8	15.7	21.5	34.6	28.2	26	0.0938
1998/9	14.4	29.4	33.8	22.4	27	0.0735
1999/0	43.4	11.4	18.3	26.9	21	0.0396
2000/1	17.2	21.7	26.1	34.9	22	0.0747
2001/2	33.6	11.6	18.3	36.4	24	0.0511
2002/3	32.6	21.9	11.5	33.9	19	0.0477
2003/4	14.6	40.3	11.1	34.0	23	0.0636
Mean	22.8	23.5	21.5	32.1		0.0642

To put these findings into context, let us consider the results for 2003/4 from Model 2. Here the figure of 42.3 per cent for congestion due to undergraduates is equivalent to 751 'surplus' undergraduates, on average, for all universities, or an average of 1306 for the 23 congested universities. The comparable figures for postgraduates are 53 and 91.5, respectively. The mean expenditure on 'surplus' academic staff was £1,437,000 for all universities or £2,499,000 for the congested universities alone. Finally, for 'other expenditure', the relevant figures are £1,495,000 and £2,600,000, respectively.

13. Disaggregated results

Some additional insights can be gleaned from the results for individual universities, which are exhibited in Table 7. These results are, once again, based on *Model 2* and relate to 2003/4.

In terms of Cooper's measure, Glasgow Caledonian and Sunderland are clearly the most congested universities, with Teesside not far behind. However, the underlying causes are rather different in each case. For instance, whereas Glasgow Caledonian has an overabundance of academic staff, Sunderland has excessive 'other expenditure'. In addition, both have too many undergraduates. In the case of Teesside, the salient factors are academic overstaffing and too many undergraduates.

Of the four factors underlying congestion, an excessive number of undergraduates is undeniably the pre-eminent one, affecting all but four congested universities. This problem is especially serious in Teesside, Central England, Nottingham Trent, Anglia and Glasgow Caledonian. By contrast, academic overstaffing, whilst still a cause for concern, is both less prevalent and less acute in most cases.

Earlier in the paper, it was noted that Thames Valley had by far the lowest BCC and super-efficiency scores in the sample, yet Table 7 shows that it was only moderately congested in terms of Cooper's measure. One can see that its C_C score is below average for the congested universities; this occurs because its congestion in terms of undergraduates, which is above average, is outweighed by negligible congestion elsewhere.

It is interesting that only five of the new universities are congested in terms of postgraduates. Here East London and, to a lesser extent, London Metropolitan are conspicuous in terms of having too many postgraduates. 'Other expenditure' is a determinant of congestion in nine of the new universities. Sunderland and, to a lesser extent, Northumbria stand out as having particular problems in this respect.

Several cases of more moderate congestion are also shown in Table 7. For example, both London South Bank and West of England have below-average congestion. In both cases, the congestion can be attributed to academic overstaffing and, to a lesser extent, to having too many undergraduates.

Table 7. Disaggregation of Cooper's congestion score for each congested university (Model 2, 2003/4)

	Other expenditure	*Academic staff*	*Post-graduates*	*Under-graduates*	C_C
Anglia	0.0321			0.2175	0.0624
Bournemouth				0.1279	0.0320
Brighton	0.0177			0.0176	0.0088
Central England		0.0954		0.2423	0.0844
Derby				0.0734	0.0184
East London			0.3681		0.0920
Glamorgan				0.0651	0.0224
Glasgow Caledonian		0.3351		0.2128	0.1370
Hertfordshire	0.0924			0.0633	0.0389
Kingston				0.0399	0.0100
London Metropolitan		0.0495	0.1443	0.0397	0.0584
London South Bank		0.1498		0.0400	0.0474
Napier		0.1776			0.0444
Northumbria	0.1761	0.1304			0.0766
Nottingham Trent		0.0984	0.0679	0.2181	0.0961
Plymouth	0.1226			0.1006	0.0558
Portsmouth		0.1012			0.0253
Sunderland	0.3026		0.0300	0.1811	0.1284
Teesside	0.0132	0.1712		0.2584	0.1107
Thames Valley	0.0027			0.1267	0.0324
West of England		0.0794		0.0306	0.0275
Westminster	0.0691	0.1197	0.0274	0.0607	0.0692
Wolverhampton				0.0786	0.0197
Mean	0.0360	0.0666	0.0277	0.0954	0.0564

14. Scale inefficiency, congestion and size

Table 8 shows a set of *TE* and *SE* (scale efficiency) scores for individual universities in 2003/4, based on Model 2. The congestion scores are also reproduced from Table 7. The *SE* scores were calculated by taking the ratio of the efficiency scores from the CCR and BCC models.

Thames Valley is an interesting case. This university's *TE* score suggests that it was producing only 51 per cent of its potential output in 2003/4, yet Cooper's measure indicates only a moderate amount of congestion. What is more, Thames Valley's *SE* score of 0.9729 reveals that it was operating at a high level of scale efficiency, with only 2.7 per cent of potential output being lost as a result of its failure to achieve full scale efficiency. The implication of these results is that much of Thames Valley's inefficiency must have been purely technical in origin.

Table 8. Technical and scale efficiency, congestion and size (*Model 2*, 2003/4)

	Share of students	*Technical efficiency*	*Rank*	*Scale efficiency*	*Rank*	*Congestion score*	*Rank*
Abertay Dundee	0.007	1	1	1	1	0	1
Anglia Polytechnic	0.028	0.7886	35	0.9883	17	0.0624	32
Bournemouth	0.020	0.8520	28	0.9806	20	0.0320	25
Brighton	0.024	0.8305	31	0.9753	21	0.0088	18
Central England	0.029	0.9203	19	0.9573	26	0.0844	35
Central Lancashire	0.033	0.9510	16	0.9510	28	0	1
Coventry	0.022	0.9700	14	0.9700	25	0	1
De Montfort	0.030	1	1	1	1	0	1
Derby	0.018	0.9088	23	0.9506	29	0.0184	20
East London	0.019	0.8373	29	0.9939	16	0.0920	36
Glamorgan	0.022	0.7696	38	0.9987	15	0.0224	22
Glasgow Caledonian	0.022	0.9599	15	0.9991	13	0.1370	40
Greenwich	0.025	1	1	1	1	0	1
Hertfordshire	0.030	0.7736	37	0.9384	31	0.0389	27
Huddersfield	0.021	1	1	1	1	0	1
Kingston	0.027	0.8122	33	0.8943	39	0.0100	19
Leeds Metropolitan	0.034	0.9089	22	0.9089	37	0	1
Lincoln	0.017	1	1	1	1	0	1
Liverpool J Moores	0.027	1	1	1	1	0	1
London Metropolitan	0.036	0.8319	30	0.9023	38	0.0584	31
London South Bank	0.022	0.7353	39	0.9988	14	0.0474	29
Luton	0.013	1	1	1	1	0	1
Manchester Met	0.045	0.8547	27	0.8547	40	0	1
Middlesex	0.027	0.9720	13	0.9720	24	0	1
Napier	0.015	0.8946	24	0.9738	22	0.0444	28
Northumbria	0.032	0.8637	25	0.9492	30	0.0766	34
Nottingham Trent	0.039	0.9917	12	0.9999	12	0.0961	37
Oxford Brookes	0.023	1	1	1	1	0	1
Paisley	0.013	1	1	1	1	0	1
Plymouth	0.034	0.8256	32	0.9157	35	0.0558	30
Portsmouth	0.028	0.9362	18	0.9827	18	0.0253	23
Robert Gordon	0.014	0.9101	21	0.9101	36	0	1
Sheffield Hallam	0.037	1	1	1	1	0	1
Staffordshire	0.019	1	1	1	1	0	1
Sunderland	0.019	0.9398	17	0.9808	19	0.1284	39
Teesside	0.021	0.7879	36	0.9528	27	0.1107	38
Thames Valley	0.019	0.5106	40	0.9729	23	0.0324	26
West of England	0.037	0.8007	34	0.9380	32	0.0275	24
Westminster	0.026	0.8613	26	0.9377	33	0.0692	33
Wolverhampton	0.028	0.9177	20	0.9269	34	0.0197	21
Mean	0.025	0.8979		0.9669		0.0325	
Number on frontier		11		11		17	
Correlation with *TE*						-0.2795	

Another example worth considering is Manchester Metropolitan. This university has $TE = SE = 0.8547$. The fact that its TE and SE scores are identical indicates that it was operating on the BCC frontier ($\phi^* = 1$). According to Cooper's measure, all of its inefficiency would be attributed to its inappropriate scale, so that any congestion would be ruled out. There are five other universities in a similar situation.

Scale efficiency in 2003/4 was generally fairly high. The mean of 0.9669 implies that, on average, 3.3 per cent of output was sacrificed across the 40 former polytechnics because of a failure of certain universities to operate at an appropriate scale.

Table 8 also shows each institution's share of all higher education students studying in the former polytechnics. Here it is worth noting that this measure of relative size is very weakly correlated with the various measures of efficiency and congestion.[27]

15. Reappraisal of findings

The sizable role attributed to academic staff in generating congestion in the former polytechnics is perplexing. What the findings suggest is that, other things being equal, this overabundance of academic staff actually *lowered* the output of congested universities in terms of earnings from research and consultancy, as well as undergraduate and postgraduate qualifications obtained. One possible explanation of this seemingly perverse finding is that the academic staff variable might be picking up the effects of an omitted variable such as the quality of management. It is also conceivable that the presence of 'surplus' academic staff in the congested universities could be indicative of institutional inefficiency in a broader sense. A final possibility is that the results may reflect heterogeneity of both staff and students. For instance, institutions are likely to differ in terms of their capacity to recruit and retain the suitably qualified and experienced staff needed to supervise research degrees and to develop masters programmes. Differences are also bound to exist in terms of the ability to recruit suitable students onto such programmes.

The congesting role attributed to 'other expenditure' in all three models is just as puzzling. What it suggests is that, beyond a certain point, extra expenditure actually lowered congested universities' output. However, 'other expenditure' is a very broadly defined input variable, comprising expenditure on academic cost centres, academic services, administration and central services, premises, residences and catering, and on research grants and contracts, and it is conceivable that one or more of these components could be congesting. For instance, by creating an unwieldy bureaucracy, overspending on administration might reduce a university's efficiency and hence output in terms of research and qualifications awarded. It is also possible that a rise in the proportion of 'other expenditure' devoted to research could decrease the output of undergraduate and postgraduate qualifications, if it meant lower spending on teaching-related activities. Hence higher total expenditure could lead to a fall in one or more outputs and a concomitant rise in others, with the overall impact on efficiency uncertain.

On a deeper level, the implausible findings with respect to academic staff and 'other expenditure' might indicate a shortcoming in Cooper's methodology. More specifically, the calculations could be attributing too high a proportion of the BCC input slacks to congestion and too low a proportion to purely technical inefficiency. As regards academic staff, we find it hard to understand why an increase in staffing should lead to a fall in output or, indeed, why a decrease should lead to a rise in output. It seems much more likely that the BCC slacks are capturing technical inefficiency rather than congestion. If so, the way to augment output would be to endeavour to raise the productivity of the existing academic workforce. Likewise, in the case of 'other expenditure', it is hard to see why a rise in spending should lead to a fall in output or, indeed, why a cut should lead to a rise in output. Here too, it seems reasonable to reinterpret most of the 'congestion' as representing technical inefficiency.

16. Conclusion

This paper has used data envelopment analysis (DEA) to examine the performance of 41 former British polytechnics that became universities in 1992, using annual data for the period 1995/6 to 2003/4. These new universities differ from the older universities in many ways, especially in terms of their far higher student:staff ratios and substantially lower research funding per member of staff. Moreover, this under-resourcing increased during the period under review, as exemplified by a further rise in the student:staff ratio from 17.5 to 19.3.[28]

In view of this under-resourcing of the new universities, we fully expected to find evidence of congestion. Congestion would mean that the output of the new universities — as measured by the number of undergraduate and postgraduate awards, along with earnings from research grants and contracts — would tend to be lower than it might otherwise have been. We used the procedure proposed by Cooper *et al.* to measure the extent of this problem.

The results suggested that congestion was a widespread phenomenon in the former polytechnics. For instance, based on our preferred Model 2, 23 out of 40 institutions were found to be congested in 2003/4, with a mean congestion score equal to 5.64 per cent of inputs (when averaged over the whole group, the mean score was 3.25 per cent).[29] Here it is interesting to note the corresponding results we obtained for 45 older universities: a mean of 6.17 per cent for the 17 congested universities but 2.33 per cent for the group as a whole (Flegg and Allen, 2007, Table 2). These findings indicate that the problem of congestion was more widespread in the former polytechnics than in the older universities, yet somewhat less severe in individual institutions.

The underlying causes of congestion were explored via a decomposition analysis based on Cooper's procedure. This revealed that an overabundance of undergraduate students was the largest single cause of congestion in the former polytechnics during the period under review. On average, based on our Model 2, such students accounted for 42.3 per cent of the value of Cooper's congestion score in 2003/4. Less plausibly, the results suggested that academic overstaffing was also a major cause of congestion in the new universities! Here the results indicated a share of 29.5 per cent. In contrast, the results suggested that postgraduates (12.3 per cent) and 'other expenditure' (16.0 per cent) played a noticeably smaller role in generating congestion. These findings are rather different to those we obtained for the older universities (Flegg and Allen, 2007, Table 6), which gave a less prominent role to academic overstaffing (15.2 per cent) but a more prominent role to postgraduates (44.8 per cent).

Some doubts were raised about the credibility of the findings with respect to academic overstaffing and excessive 'other expenditure'. It was suggested that Cooper's procedure might be attributing too high a proportion of the BCC input slacks to congestion and too low a proportion to purely technical inefficiency. We would recommend, therefore, that analysts should always decompose Cooper's congestion score and carefully examine its components. Only those components where an *a priori* explanation can be offered for the existence of congestion should be interpreted as such.

Nonetheless, it is important to note that Cooper's method typically generated mean congestion scores that were noticeably lower than those from alternative methods (see Flegg and Allen, 2006a), so the results reported here may well represent minima rather than maxima. It is worth mentioning too that the findings were not greatly affected by changes in the input and output variables used in the DEA models. Furthermore, a sensitivity analysis revealed that differences in universities' subject mix had minimal impact on efficiency scores.

Cooper's measure indicated a clear but rather bumpy downward trend in mean congestion scores over the period as a whole. However, this finding was not corroborated by the other approaches considered in Flegg and Allen (2006a), although all methods recorded a rise in congestion and fall in technical efficiency in the final year.

In terms of implementing the findings of this study, one important caveat needs to be stated: it may well be much easier to comprehend the causes of congestion than to realize the potential gains in output from eliminating such congestion. Also, since the different methods of identifying and measuring congestion have their respective merits and demerits, yet produce different results, it would be prudent not to rely exclusively on a single method or on the rankings derived therefrom (see Flegg and Allen, 2006a).

There are several areas where this study could usefully be built upon. The first is that a more detailed investigation could be carried out into how one might interpret the BCC input slacks in situations where the inputs in question are unlikely to have been affected by congestion. Secondly, it would be interesting to experiment with a weighted undergraduate qualifications variable to mirror that used for postgraduates (see Johnes, 2006). Thirdly, one could explore what effect lagging some of the inputs (e.g. research expenditure) would have on the findings. Fourthly, a Malmquist analysis could be used to discriminate between changes in congestion caused by shifts in the efficiency frontier, as opposed to movements towards or away from this frontier. Finally, it would be interesting to investigate the reasons for the differences between the findings reported here and those we obtained for 45 older British universities over the same period (Flegg and Allen, 2007).

Accepted for publication: 13 July 2007

Appendix A: Sources and definitions

Most of the data employed in this study were taken directly from various issues of the following publications of the Higher Education Statistics Agency (HESA):

- *Resources of Higher Education Institutions*
- *Students in Higher Education Institutions*

See HESA (various years). In certain cases, noted below, data were obtained directly from HESA under contract. We decided to omit 1994/5 from our study because of missing data for Luton and Robert Gordon universities. The results for 1995/6 should be treated cautiously owing to possible problems with the data on full-time equivalent numbers of students and staff.

Some essential information on the variables used in this study is given below. More detailed information is given in the HESA publications mentioned above.

- *Income from research grants and contracts*

  Because of concerns about the comparability of some of the data, this variable excludes data on what HESA defines as income from 'other services rendered'. Source: *Resources of Higher Education Institutions*, Table 3 up to 2001/2, Table 1c thereafter.

- *Number of undergraduate and postgraduate qualifications awarded*

  The qualifications data published in *Students in Higher Education Institutions* could not be used for two reasons:

  (i) the severe rounding of the published data from 1999/2000 onwards;

  (ii) the unspecified qualifications of 'dormant students' from 1995/6 to 1999/2000.

  Fortunately, we were able to obtain the necessary data directly from HESA. Even so, it is worth noting the following caveat with regard to the 1996/97 data (*Students in Higher Education Institutions*, 1996/97, p. 6): 'It is known that the qualifications obtained were under-reported for the following institutions: De Montfort University, Sheffield Hallam University, South Bank University and the University of Paisley.' There were also problems relating to the treatment of 'dormant' students. In addition, in the case of Luton in 1997/8, the figures for undergraduate degrees awarded were not separated into classes, so we used interpolation to estimate the missing figures for use in Model 1.

- *Full-time equivalent undergraduate and postgraduate students* (x_1 and x_2)
  HESA did not publish full-time equivalent numbers for 1994/5 and 1995/6 because of concerns about the quality of the data. Although we were able to obtain the unpublished data directly from HESA, we have used the figures for 1995/6 in our study with some reservations. Data from 1996/7 onwards were taken from *Students in Higher Education Institutions*, Table 0b.
- *Academic staff expenditure* (x_3)
  Source: *Resources of Higher Education Institutions*, Table 7 up to 2001/2, Table 2b thereafter.
- *Other expenditure* (x_4)
  Variable x_4 was calculated by subtracting what HESA defines as 'other expenditure' from each university's total expenditure and then deducting academic staff expenditure (x_3). HESA's 'other expenditure' was not included, as we had concerns about the comparability of some of the data. Source: *Resources of Higher Education Institutions*, Tables 6 and 7 up to 2001/2, Tables 2a and 2b thereafter.
- *Full-time equivalent number of academic staff*
  The HESA data on this variable were downloaded from http://www.data-archive.ac.uk. It should be noted that we have some concerns about the reliability of the data for 1995/6. In particular, the aggregate student: staff ratio for that year looks unrealistically high.

Appendix B: Summary statistics for 2003/4

	Inputs				*Outputs*		
	Undergraduate students	*Postgraduate students*	*Academic staff expenditure*	*Other expenditure*	*Research income*	*Undergraduate awards*	*Postgraduate awards*
Mean	13007	2252	33239	61907	3783	4378	1382
Std. deviation	4172	972	11051	18621	2363	1517	734
Coeff. of variation	0.321	0.431	0.332	0.301	0.625	0.346	0.531
Median	12197	2111	33811	64318	3085	4116	1349
Minimum	3450	610	8320	20105	420	1240	100
Maximum	23430	4285	57945	101035	10830	8095	3475

Notes: Expenditure is in £1000; students are full-time equivalents. Minima and maxima have been rounded in accordance with HESA's policy.

ENDNOTES

1. School of Economics, Bristol Business School, University of the West of England, Coldharbour Lane, Bristol BS16 1QY. Tony.Flegg@uwe.ac.uk; David.Allen@uwe.ac.uk. We are most grateful to two anonymous referees, whose helpful comments led to numerous improvements in this paper. We would also like to thank our colleagues, Peter Howells, John Sloman and Chris Webber, and members of the Editorial Board, for some very helpful suggestions. Finally, we would like to thank Kate Lang of the Higher Education Statistics Agency (HESA) for the efficient way in which she answered our many queries and produced the data we needed.

2. The basic data used in this study were obtained, either directly or indirectly, from the Higher Education Statistics Agency (HESA). See Appendix A for details.

3. In 2003/4, for example, 13.2 per cent of undergraduates in the 45 older universities gained first-class degrees and 49.3 per cent gained upper seconds, whereas the proportions in the ex-polytechnics were 7.7 per cent and 39.0 per cent, respectively. Source: Authors' own calculations using HESA data.

4. The number of full-time equivalent students in the ex-polytechnics rose by 15.1 per cent between 1995/6 and 2003/4, compared with a rise of 26.4 per cent in the 45 older universities.

5. Congestion in the older universities is examined in Flegg and Allen (2007). Another reason for analysing the two groups of universities separately is that DEA requires all inputs and outputs to be homogeneous. This assumption is most unlikely to hold for a combined sample, especially in terms of the intake of students.

6. The student : staff ratio in the ex-polytechnics rose from 17.5 in 1995/6 to 19.3 in 2003/4.

7. Figure 2 is adapted from Tone and Sahoo (2004, Figure 2).

8. For a detailed discussion of the properties of the CCR and BCC models, see Cooper *et al.* (2000a).

9. In fact, it is also necessary to ensure that all frontier DMUs are located at extreme points on the BCC frontier — such as A, B, C and D in Figure 2 — rather than on the line segments connecting such points. Such cases are rare in real data sets.

10. See, for example, Flegg *et al.* (2004).

11. *OnFront* (www.emq.com), the software underpinning Färe and Grosskopf's approach, uses an input orientation to measure congestion of inputs.

12. For a detailed discussion of the pros and cons of alternative procedures for identifying and measuring congestion, see Flegg and Allen (2006b). Also see the exchange between Cherchye *et al.* (2001) and Cooper *et al.* (2001b, 2001c).

13. For the ex-polytechnics, 'other undergraduate awards' such as certificates and diplomas have gained in importance, rising from 27.7 per cent of all undergraduate awards in 1995/6 to 34.3 per cent in 2003/4.

14. It is worth noting here that, in 2003/4, 7.7 per cent of undergraduates in the ex-polytechnics gained first-class degrees and 39.0 per cent gained upper seconds, compared with 4.3 per cent and 37.8 per cent, respectively, in 1995/6. Source: Authors' own calculations using HESA data.

15. These *TE* scores were obtained (using *DEA-Solver-Pro*) from the CCR model, which assumes constant returns to scale (CRS) and no congestion. The orientation of the model has no effect on the *TE* scores under CRS. *OnFront* (www.emq.com) generated identical results.

16. This broader variable encompasses all undergraduate degrees, as well as 'other undergraduate awards' such as certificates and diplomas in business, computing, engineering, medicine, nursing and technology, along with higher national diplomas, certificates and diplomas of higher education, etc. For the ex-polytechnics, these 'other awards' have gained in importance, rising from 27.7 per cent of all undergraduate awards in 1995/6 to 34.3 per cent in 2003/4. In some cases, these other awards are a default qualification rather than one that would be sought in its own right.

17. Although DEA does not require a balanced panel of DMUs, we nonetheless experimented with models in which the data for London Guildhall and North London were pooled to form a single entity in the first seven years. However, it made little difference whether these two universities were combined into a single DMU or analysed separately. The explanation for this is that neither university appeared on the frontier in any year.

18. To discriminate between these two possibilities would require a Malmquist analysis (see Flegg *et al.*, 2004). However, an analysis of this kind is beyond the scope of the present paper.

19. Source: HESA, *Students in Higher Education Institutions*, Table 8j.

20. The correlation with firsts alone was 0.063 (p = 0.701). All tests used are two-tailed.

21. The calculations were done using the routine in *DEA-Solver-Pro* for estimating a VRS-based, output-oriented, super-efficiency model. This approach is explained in Appendix A of the user's manual (www.saitech-inc.com).

22. Athanassopoulos and Shale (1997) constructed a rather different DEA model in order to measure universities' cost efficiency. The inputs used were general academic expenditure and research income, while the outputs were the number of successful leavers, the number of higher degrees awarded, and a weighted research rating derived from Research Assessment Exercise (RAE) scores. This type of model would tend to exaggerate the efficiency of an arts-oriented university relative to a science-oriented university with similar RAE scores. This is because the science-oriented university would tend to have higher research income.

23. An alternative approach would have been to disaggregate the undergraduate data into arts and sciences, broadly defined. This route was taken by both Izadi *et al.* (2002) and Stevens (2005) in their stochastic frontier analyses of British universities' cost functions. However, this approach is not so appropriate in a DEA study such as the

present one. It would mean adding an extra input and output, which would introduce greater complexity into the modelling and also reduce the discriminatory power of the analysis. The latter point arises from the fact that the number of frontier DMUs tends to rise as the number of variables employed increases.

24. A good illustration of this feature of DEA is provided in a paper by Athanassopoulos and Shale (1997), who examined data relating to 1992/3 for 45 'older' British universities and university colleges. Using three outputs and six inputs, they obtained a mean *TE* score of 0.9716. Furthermore, when VRS was assumed, the mean efficiency score rose to 0.9823, and the number of efficient institutions increased from 27 to 31. In such circumstances, a DEA model has limited discriminatory power.

25. Using data for 2003/4 and the results for *Model 2*, we found a correlation of only 0.0235 between the percentage of science students, π, and a university's congestion score, C_C.

26. The alternatives we tried were $y_{2d} = 4.5p_1 + 3p_2 + 1.5p_3 + p_4$ and $y_{2e} = 4.5p_1 + 3p_2 + p_3 + 1.5p_4$.

27. Using data for 2003/4 and the results for Model 2, we found correlations of -0.106, 0.066, 0.072 and 0.115, respectively, between a university's share of students and its *TE*, β (BCC), C_C and σ (super-efficiency) scores.

28. In the older universities, the student : staff ratio rose from 7.5 in 1995/6 to 9.4 in 2003/4.

29. n = 40 in the final two years owing to the merger between London Guildhall and North London.

References

Athanassopoulos A D and Shale E (1997) 'Assessing the comparative efficiency of higher education institutions in the UK by means of data envelopment analysis', *Education Economics*, 5, 117–134.

Cherchye L, Kuosmanen T and Post T (2001) 'Alternative treatments of congestion in DEA: a rejoinder to Cooper, Gu, and Li', *European Journal of Operational Research*, 132, 75–80.

Cooper W W, Deng H, Gu B, Li S and Thrall R M (2001a) 'Using DEA to improve the management of congestion in Chinese industries (1981–1997)', *Socio-Economic Planning Sciences*, 35, 227–242.

Cooper W W, Gu B and Li S (2001b) 'Comparisons and evaluations of alternative approaches to the treatment of congestion in DEA', *European Journal of Operational Research*, 132, 62–74.

Cooper W W, Gu B and Li S (2001c) 'Note: alternative treatments of congestion in DEA – a response to the Cherchye, Kuosmanen and Post critique', *European Journal of Operational Research*, 132, 81–87.

Cooper W W, Seiford L M and Tone K (2000a) *Data Envelopment Analysis*, Boston: Kluwer.

Cooper W W, Seiford L M and Zhu J (2000b) 'A unified additive model approach for evaluating inefficiency and congestion with associated measures in DEA', *Socio-Economic Planning Sciences*, 34, 1–25.

Flegg A T, Allen D O, Field K and Thurlow T W (2004) 'Measuring the efficiency of British universities: a multi-period data envelopment analysis', *Education Economics*, 12, 231–249.

Flegg A T and Allen D O (2006a) 'Are the new British universities congested?', Discussion Paper No. 06/10, School of Economics, Bristol Business School, University of the West of England, Bristol (http://carecon.org.uk/DPs/). Paper presented at the 48th Annual Conference of the Operational Research Society, Bath, September 2006.

Flegg A T and Allen D O (2006b) 'Does it matter how we measure congestion?', Discussion Paper No. 06/14, School of Economics, Bristol Business School, University of the West of England, Bristol (http://carecon.org.uk/DPs/).

Flegg A T and Allen D O (2007) 'Does expansion cause congestion? The case of the older British universities, 1994–2004', *Education Economics*, 15, 75–102.

HESA (various years) *Resources of Higher Education Institutions* and *Students in Higher Education Institutions*, Cheltenham: Higher Education Statistics Agency.

Izadi H, Johnes G, Oskrochi R and Crouchley R (2002) 'Stochastic frontier estimation of a CES cost function: the case of higher education in Britain', *Economics of Education Review*, 21, 63–71.

Johnes J (2006) 'Data envelopment analysis and its application to the measurement of efficiency in higher education', *Economics of Education Review*, 25, 273–288.

Rodgers T (2007) 'Measuring value added in higher education: a proposed methodology for developing a performance indicator based on the economic value added to graduates', *Education Economics*, 15, 55–74.

Stevens P A (2005) 'A stochastic frontier analysis of English and Welsh universities', *Education Economics*, 13, 355–374.

Tone K (2002) 'A slacks-based measure of super-efficiency in data envelopment analysis', *European Journal of Operational Research*, 143, 32–41.

Tone K and Sahoo B K (2004) 'Degree of scale economies and congestion: a unified DEA approach', *European Journal of Operational Research*, 158, 755–772.