The joint production of health care and education in UK hospitals

Bruce Hollingsworth
Health Economics Unit
Faculty of Business and Economics
Monash University
Melbourne, Australia
Email: bruce.hollingsworth@buseco.monash.edu.au

David Parkin
City Health Economics Centre
Department of Economics
City University
Northampton Square
London EC1V 0HB, UK
Email: d.parkin@city.ac.uk

Presented to the Joint workshop of the Collège des Economistes de la Santé and the Health Economists' Study Group, Paris, January 2004

Acknowledgements
The research reported in this paper was funded by a grant from the NHS Executive Northern & Yorkshire R& D Committee. We thank our collaborative panel: Bill Kirkup, David Whittaker, Keith Aungiers and Jim Halliday (NHSE); Reg Jordan and Peter Hill (Newcastle University); Hazel Frith (Leeds University); Bernard Chalk (United Leeds Hospital Trust) and; Angela Dragone (Newcastle NHS Trust). We are also grateful to Vicky Spark and Chris Watson (Case Mix Office), Rob White (Newcastle University), Barbara Willard (Leeds University), Catherine Smith and Catherine Graham (NHSE) and John Pearson (Newcastle University) for data and other assistance, and to all those who took the time to respond to our surveys.

Introduction

A major problem in the application of economics to the study of hospitals, including such fundamental issues as reimbursement and the measurement of efficiency, is their multi-product nature. Studies of the production of health care by hospitals therefore have to take account of factors such as case-, specialty- and severity-mix, whose complexity rapidly leads to formidable problems of definition, data collection and analysis. This issue is compounded by the fact that as well as producing health care, hospitals also produce two other major classes of output – medical research and the training of doctors and other health care professionals. Such outputs are fundamentally different to health care, but the production processes of these three classes of output are equally fundamentally linked.

This paper examines the production of teaching, alongside that of health care, in hospitals in the National Health Service (NHS) in the United Kingdom. It does not consider research, although that is of equal importance, because that introduces yet further complexities in data and analysis that have not at present been resolved.

There has been continuing change and development in the structure of the NHS for many years, but there has been little change in reimbursement for teaching (Bevan, 1999). It is dealt with by a formula that allocates funds to hospitals to cover the extra costs of undertaking teaching, arising originally from the work of the 1978 Resource Allocation Working Party (RAWP). This included a Service Increment For Teaching (SIFT) in the more general resource distribution formula, which covered 75% of the observed difference between the average costs of teaching and non-teaching hospitals, a figure based upon research carried out by Culyer *et al* (1978). In 1989 this was increased to 100%. In 1995 the additional funding was split between SIFT and an increment for research and development, but without the previously specified formula.

Despite varying degrees of unhappiness with this resource allocation formula, it has remained in place largely because of the lack of agreed alternatives. Similarly, although there have been huge and continuing efforts to measure and promote efficiency in the production of health care in NHS hospitals, there have been no similar efforts for their teaching activities, presumably because it is viewed as of lower priority. However, three developments have made analysis of reimbursement and efficiency more urgent. First, the number of Medical School places is being expanded considerably, which will lead to a greater volume of teaching activity within hospitals. Secondly, teaching has increasingly been undertaken outside of traditional Teaching Hospitals and this trend will grow with the expansion of student numbers. Thirdly, it is likely that reimbursement of NHS hospitals will, within a few years, be based upon output disaggregated by Healthcare Resource Groups (HRGs), following the “Payment by Results” initiative; as a result, it may no longer be seen as inevitable that teaching should be paid for via an old-style block grant.

The measurement of efficiency and the appropriate method of reimbursement are of course linked, since it is important to ensure that any reimbursement scheme does not encourage inefficiency. As a result, the identification of the impact of teaching on both efficiency and costs is of great importance.

This paper examines the effects of teaching on hospital efficiency and costs, and potential implications for funding mechanisms. We report briefly on a survey of the NHS to establish priorities and issues. We report analyses of efficiency and costs related both to teaching and to Teaching Hospitals. We report briefly on the dissemination of these results to the NHS and discuss the potential implications of our analysis.

Background

Teaching Hospitals and hospitals that teach

An important departure for our analysis is the separation of the issue of teaching from that of the role of designated Teaching Hospitals. It has always been recognised that the label “Teaching Hospital” selects only one of the characteristics of such hospitals, because they in general have a more complex case-mix, provide more high technology services, employ personnel with higher skills and have a role as tertiary referral centres. In recognition of this, there is no longer an official definition of an NHS Teaching Hospital. At the start of SIFT, the problem that higher “Teaching Hospital” costs were not solely because of teaching did not matter, since its origins lay in attempts to deal with the problems of easily identified large hospitals, all of which had that label. However, teaching is now more devolved, so that there are “main” and “associated” teaching hospitals, with only local and rule-of-thumb definitions or agreements of what these mean. So the problems of teaching and of Teaching Hospitals are separable but of course remain highly related.

How do hospitals that teach differ from those that do not?

The two key policy issues that we have identified, performance efficiency and reimbursement, are linked. Differences in overall hospital costs due to teaching seem to have an obvious origin in the resource expended directly on that activity, but the issue is clouded by the likelihood that hospitals that undertake much teaching differ in the health care that they provide. It is also possible that the conditions under which they work affect their efficiency. It is therefore necessary to examine closely the origins of cost differences and how these relate to the outputs of hospitals in terms both of health care and teaching. Possible differences in costs of health care arising directly from teaching include:

- Higher running costs per case. This is due to a number of factors including the need to devote time to teaching as well as care, the necessity to employ more skilled personnel in order to teach and the need to provide, for demonstration purposes, care in excess of the minimum necessary.
- More complex case mix because of the need to demonstrate a wide range of cases to trainees.
- More advanced technology because of the need to demonstrate this to trainees. This may also mean more rapid technological change.
- A lower or higher level of efficiency, due to different pressures on teaching personnel and managers of hospitals that teach.

Of course, some of these differences, especially those due to case-mix and technology, may be not be caused by teaching, rather they may reflect the nature of hospitals that teach. Indeed, it is historically more likely that hospitals were chosen as appropriate places to teach because they possessed these characteristics anyway. Such problems make the analysis of teaching costs even harder to disentangle.

Little work has been undertaken recently concerning possible differences in the running costs of hospitals that teach. It is widely accepted that teaching and other hospital outputs are produced jointly and that the provision of teaching affects the outputs of a hospital (Butler, 1995). In allocating costs in teaching and non-teaching hospitals Sloan *et al* (1983) identified three criteria: costs of medical staff should be included; the sample should contain a number of hospitals and the impact of costs within hospitals should be identified; and case mix and other factors should be accounted for. Evidence on the impact of teaching on costs varies. In an American study Frick *et al* (1985) reported on comparative results from 11 teaching and 20 non-teaching hospitals, with case mix being accounted for in terms of Diagnosis Related Groups (DRGs). Cost per case in the teaching hospitals was 68% higher, with only 23%

of the difference being explained by case mix differences. The residual difference may have been due to indirect costs (e.g. patients receiving more diagnostic tests due to teaching), more complex cases within DRGs, a higher quality of care; or lower efficiency.

Other studies have also found this residual effect on costs of teaching even after case mix adjustment. However, the estimated size of the teaching effect has differed widely. Most studies have been cross sectional in nature with little use being made of panel data, which allows reductions in problems such as omitted variable bias.

The performance and reimbursement of hospitals that teach

As a result of the different factors affecting hospitals that teach, it may be that performance issues are different for hospitals that undertake different levels of teaching, which may impact upon the construction of appropriate performance indicators. This does not simply mean that performance indicators should cover teaching; it may mean that teaching status implies a radically different production technology, which therefore needs to be assessed differently, or at least separately, from other types of hospital.

Similarly, the relative size of the impact of different factors on cost will impact on the appropriate type and level of reimbursement. There are three main alternatives for reimbursement, depending on the source of the higher costs of hospitals that teach:

- Reimbursement by fee-per-student, if higher costs are entirely due to direct teaching costs.
- Reimbursement by increased fee-per-case, if higher costs are entirely due to case mix and technology differences.
- A mixed fee-per-student and increased fee-per-case system, if both factors operate.

Such mechanisms might be supplemented by block grants, as at present, to establish the overall capacity to undertake teaching. However, if differences are in part due to different levels of efficiency, it is important to ensure that any reimbursement scheme does not encourage inefficiency.

Data and Methods

Policy issues survey

We conducted a survey of NHS organisations and Medical Schools, which asked for their views on a number of important policy issues, including the impact of teaching on provision and cost of services, how to take account of teaching in measuring performance in providing health care, how to measure performance in providing teaching, the impact of providing teaching on payment for health care provision and payment mechanisms for teaching. A postal questionnaire was sent to the 25 Medical Schools in the UK and the 322 Trusts and 100 Health Authorities in England and Wales. The response rate was 40%.

Cost, resource and activity data analysis

Data were obtained from the NHS Executive Northern and Yorkshire Office and the CIPFA cost database (CIPFA, 1998). They cover 38 Trusts in the Northern and Yorkshire Region, comprising 13 acute, 12 priority (those specialising in mental health services), and 13 combined (those that undertake both acute and priority activities) for the years 1994/5 to 1997/8. The variables are listed in Table 1, divided into those that are inputs to the production of hospital health care and teaching services and those that are outputs. In addition, we used the case mix index calculated by the National Case Mix Office, which is a national comparison of inpatient and day case length of stay. A value of one is the national average, greater than one means a case mix more complex than the national average, below one means a case mix less complex than the national average. Where it was necessary to convert cost data to a common set of prices, they were converted to 1997/8 levels.

Hospitals were classified using two different criteria: functional, as defined above; and teaching status, defined as:

- Designated Teaching Hospitals (DTHs) - those classified as such by the NHSE
- Other teaching hospitals (OTHs) – those that are not classified as Teaching hospitals, but which undertake a significant amount of teaching
- Non-teaching hospitals (NTHs)- those that undertake little or no teaching

We used three methods of analysing efficiency and costs. The first was Data Envelopment Analysis (DEA), which has been increasingly used in health care, although few have looked at teaching hospitals, with no such applications in the UK (Cooper *et al*, 2000; Hollingsworth *et al,* 1999, Hollingsworth, 2003). It is based on a direct measurement, using linear programming, of a Farrell radial measure of efficiency. DEA produces a number of useful items of information, including a quantified measure of efficiency for each hospital; the identification of hospitals with which inefficient hospitals might compare themselves; information on returns to scale. In addition to the normal method of calculating DEA scores, we make use of "super efficiency", which enables the ranking of efficient units (Anderson and Peterson, 1993). Changes in productivity were also measured using a Malmquist index. As well as changes in overall productivity over time, this method permits such changes to be decomposed into changes in technology - movements in the efficiency frontier –and changes in individual hospital efficiency - the distance that hospitals are from the efficient frontier. For cost models, it is possible to include a third decomposition factor, changes in scale.

The main advantage of DEA is its ability to cope with multiple outputs as well as multiple inputs. Its main problem is that its basis is deterministic, which does not take account of stochastic variation or measurement error; methods for incorporating such issues are not well advanced. As a result, hypothesis testing and model building do not have a firm statistical basis.

Different models were specified based on different aggregation levels for the input and output variables. In addition, case-mix adjusted models were specified, in which FCE outputs were multiplied by the case-mix index. Sensitivity analysis was undertaken to test the robustness of results to changes in the specification of models, according to the methods outlined by Parkin and Hollingsworth (1997) and Hollingsworth and Parkin (1998). In each case, we estimated models using only the health care variables and compared these with similar models including the education variables as well.

The second technique was stochastic frontier analysis (SFA) (Coelli *et al.*, 1998), which has also had a limited number of applications to teaching hospitals, and none in the

UK. It is based on a partition of the regression error into a conventional normally distributed stochastic error term and a random effect term taking a specified distribution. The second of these is interpreted as inefficiency. It also produces efficiency rankings and scores for individual hospitals.

The main advantage of SFA is that it is able to take advantage of statistical inference. Its main disadvantage is that it requires very strong assumptions to be able to produce a multiple output indicator and to interpret the results as being due to inefficiency. A multiple output indicator was calculated using the method described by Gerdtham *et al* (1999), which is the geometric mean of the outputs. This imposes the strong constraint that efficiency affects all outputs in the same proportion. This, along with the assumption that the error terms interpreted as efficiency follow a particular distribution, have no real basis and are justified only by their ability to permit the analysis to be undertaken.

Thirdly, we estimated a cost function which replicated the earlier analyses upon which SIFT allocations were based, using the new data and including panel data methods.

Dissemination survey

We sent a summary of the efficiency and cost results to those in our original sample that were in the Northern and Yorkshire Region, and requested their comments on these. The response rate for this postal questionnaire was 51%.

Results

Issues survey

Detailed results are available in Hollingsworth and Parkin (2000), and here we give a summary of the majority views amongst respondents to the survey:

- Teaching has either no or a negative impact on provision of services in terms of numbers of patients seen and the time that teaching takes up.
- Teaching may improve the quality of care and help encourage retention of higher calibre staff.
- Teaching increases costs of care, although this is uncertain as effects are unknown or unmeasured.
- Teaching costs may be hidden or not sufficiently covered by funding.

- Teaching performance is not measured at present and is difficult to do, but should be measured explicitly and should take account of quality.
- Health service and teaching costs should be measured separately, although this may be difficult and SIFT does not help with this.
- Funding for teaching and health care should be separate; current mechanisms do not help to do this.
- There may be cross-subsidies from teaching to care and *vice versa*.
- There may be inequity between designated Teaching Hospitals and other hospitals that undertake teaching.
- SIFT is a crude instrument, unresponsive, not transparent, out-of-date and inequitable in its distribution of funds.

Data Analysis

Full results for the individual Trusts are available in Hollingsworth and Parkin (2000). Here we report the results overall for the hospitals as a whole and for different types of hospital.

Production efficiency scores from DEA models

Sensitivity analysis was undertaken to test the robustness of results to changes in models used and between time periods (Parkin and Hollingsworth, 1997; Hollingsworth and Parkin, 1998). Overall, there were positive, significant correlations for the relevant models both between different specifications and between different time periods. However, the models with case mix adjusted (CMA) data have much higher correlations and we therefore report here only the results for those models[1].

The best model that looked only at health care included capital, medical staff, nursing staff, other staff and other costs as inputs and total CMA FCEs, Outpatients and A&E attendances as outputs. Table 2 shows the results broken down by functional status. Acute Trusts are the most efficient, followed by Combined Trusts, with Priority Trusts the least efficient. Mean efficiency varied over time, with in general an initial improvement followed by a decline, although the patterns were not identical for the different functional groups. The main observation from these data is that they confirm

[1] The full individual Trust results include bootstrapped 95%confidence intervals for the means, calculated with 1,000 repetitions for all models (Hollingsworth and Parkin, 2000).

earlier analysis on a similar data set (Hollingsworth and Parkin, 1998; Hollingsworth and Parkin, 2003), which increases confidence in the analyses that relate to teaching.

Table 3 shows summary results with the addition of variables to reflect education - SIFT as an input and student weeks as an output. The main observations are that Priority Trusts are more efficient than is suggested by the health care only model, but their efficiency declined over all years. Other patterns over time are also slightly different, but in general still conform to improvements followed by a decline.

Table 4 shows results according to the hospitals' teaching status. Designated Teaching Hospitals (DTHs) are always the most efficient and, except for the first year, Non-Teaching Hospitals (NTHs) are the least efficient. However, apart from the first year, the difference between DTHs and Other Teaching Hospitals (OTHs) is smaller than the difference between both of these and the NTHs. The efficiency of DTHs declined slightly over time, and the efficiency of OTHs improved slightly. However, NTHs appear to be declining rapidly.

Cost efficiency analysis

The cost efficiency model included the same outputs as the health care and teaching production model, but with total costs as the single input. Table 5 shows that, as with the production efficiency results, DTHs are slightly more efficient than OTHs, but NTHs have very low cost efficiency scores. There is very little difference in scores over time for DTHs and OTHs, but NTHs again show a decline, though not as rapid as for the production models.

Super efficiency analysis

Table 6 summarises the results from the "super-efficiency" analysis, which realtes to the cost efficiency models. Because no NTHs are 100% efficient, the results are the same as for the ordinary DEA analyses. In contrast to the results from the ordinary analysis, the mean efficiency of DTHs and OTHs are very similar, and in two of the four years OTHs outperform DTHs. Although variation over time is small for both DTHs and OTHs, their patterns are quite different.

Malmquist index productivity analysis

Table 7 shows the productivity change scores between 1994 and 1998 for each hospital group, using the production models for both health care only and health care and education. A score of greater than one indicates an increase in productivity, a score of less than one indicates a decrease in productivity and a score of one indicates no change.

The pattern of overall change in productivity suggests a general improvement over time, and is consistent with the analyses reported earlier in two ways. First, in looking at health care only, Acute Hospitals have improved the most whilst Priority Trusts have declined. Secondly, taking teaching into account, Priority Trusts appear much better and in fact appear to have improved. However, there is an important difference, in that OTHs have improved productivity at a much faster rate than DTHs, and indeed DTHs have a slightly lower improvement than NTHs. For both DTHs and NTHs, any improvements due to technological change are largely cancelled out by increased relative inefficiency of individual hospitals, though this pattern is far more extreme for NTHs.

For all groups, it appears that overall productivity changes are being driven by changes in technology rather than changes in individual hospital efficiency. There are in general improvements in overall efficiency over time, but only because large technological improvements have outweighed a decline in relative efficiency.

Table 8 shows the productivity change scores between 1994 and 1998 for hospitals according to teaching status, using the cost models for health care and education. As with the production models, OTHs have improved most, but in this case not only have NTHs performed worst, they have actually declined. Again, change is being strongly driven by technological change, with both relative efficiency and scale largely unchanged for DTHs and OTHs but declining markedly for NTHs, which outweighs the very large improvement in overall technology.

Stochastic frontier analysis

As a further means of validation, we used the Coelli panel data stochastic frontier estimator to analyse both production and cost efficiency. The inputs for the production efficiency model were the same as those for the DEA efficiency analysis, but included only one output, the aggregate output variable described earlier. For the cost efficiency analysis, the outputs were the same as for the corresponding DEA analysis, but the only input was total cost. A Cobb-Douglas functional form was specified.

Table 9 shows the efficiency scores derived from the production efficiency models. The closer to one the score is, the more production efficient the hospital is. There appear to be small differences between the DTH and OTH hospital groups and little change over time. DTHs are most efficient and OTHs less efficient, but these differences are not dramatic. The conclusions regarding overall levels of efficiency are similar to those from the DEA analysis, although the inefficiency of NTHs is much more marked. Moreover, the conclusions about trends over time are quite different from each other.It could be that the SFA results are closer to the truth and that the DEA results are artefacts caused by its failure to take account of stochastic variation and measurement error. Alternatively, the DEA results may be closer to reality, and the SFA results are due to the imposition of an arbitrary common distribution for efficiency scores.

Table 10 shows the efficiency scores for the health care and education cost model. In this case, the greater than one the score is, the less cost efficient the hospital is. The OTHs and NTHs appear much more cost efficient than DTHs, which is contrary to findings from the DEA analysis and also to the SFA production analysis. However, efficiency appears to be improving over time for all groups, which is consistent with the findings of the DEA based Malmquist index analysis for DTHs and OTHs, but not for NTHs, and is inconsistent with the SFA production analysis. Moreover, the rate of improvement is in each case constant over time. These results cast some doubt on the SFA analysis; DEA produces findings which are not only more consistent but more intuitively plausible.

Estimating the SIFT cost function

This section is the most preliminary and should in particular be regarded as an early report on work-in-progress. We undertook exploratory analyses to estimate the value of the regression coefficient on the variable student weeks. Initially, we used a primitive specification, the same model as that used by Culyer *et al* (1978), whose results were used in the original SIFT formula. The variables used were those that most closely correspond to the original model, which is:

$$TC = \beta_0 + \beta_1 SW + \beta_2 FCE + \beta_3 OP + \beta_4 AE + \beta_5 CMI$$

where TC = Total Cost; SW = Student weeks; FCE = Total FCEs; OP = Outpatients; AE = A&E attendances; CMI = Case mix index.

Initially, we estimated the model using a simple pooled data set. Table 11 shows that this basic regression model (Model 1) is statistically superior to the original model used to estimate SIFT coefficients. The high R^2 indicates that the model is a good fit to the data, and all the variables are highly statistically significant. However, there is evidence of misspecification, since this fails the RESET test. The conclusions are most obviously that this requires further diagnostic and model-building work, but also that the original model is no longer justifiable.

Dummy variables representing functional status were included, but these were not significant. A dummy representing teaching status was, however, significant and the results are also reported in Table 11 (Model 2). However, the case-mix indicator became insignificant ($p = 0.069$), possibly indicating some collinearity between the variables, though the variance inflation factors for each variable were not large. This may be an important factor, since there is a fairly large effect on the student weeks coefficient.

Not unsurprisingly, given this misspecification, attempts to use the panel data structure of the data to improve the estimates produced results of little value and these are not reported. An obvious approach is to estimate a fixed-effects, or within groups model to take account of individual trust heterogeneity. However, all of the estimated coefficients were insignificant and some were of the wrong sign. The variance estimate was also dominated by the variance of the fixed effects. In fact, a between-groups estimator produced a much better set of estimates. This again suggests a fundamental problem with the specification, since the model works at an aggregate but not at an individual level. This is further confirmed by the fact that a within-years estimate, in which the fixed effects are estimated for years, also produced a better model, one consistent with the simple pooled data model.

There is less justification for using a random-effects estimator, but in any case, again unsurprisingly, such estimates provide results of little value. In looking at Trust heterogeneity the estimates are dominated by the between-groups element, and therefore diverge markedly from the fixed effects and fail the Hausman test. For the within-years effects, the estimates degenerate to the pooled data model.

For illustration, we will use the pooled regression to show how the results might be used; but it should again be emphasised that this is not a reliable model and the calculations should not be regarded as producing meaningful figures. If reimbursement for teaching was based on fee-per-student, the coefficient suggests that this should be at the rate of £14,077 for each student week. However, current allocations are not calculated on this basis. The resource allocation formula in effect assumes joint production, so that most of the cost of teaching are covered by activity-based calculations and SIFT only covers the unaccounted-for excess costs of teaching. It is possible to calculate this for our 1997/8 data using a similar formula to the calculation of SIFT (Bevan, 1999). The excess cost of teaching overall can be estimated using the calculation that the average cost per case mix adjusted FCE is 8.2% higher for DTHs; the same percentage of specific teaching costs should therefore be covered by SIFT. SIFT should therefore be at the rate of £14,077×0.082 = £1,154 per student week. The actual figure for DTHs was £1,156, suggesting that the SIFT formula works well – but since the equation on which it is based is misspecified, this casts some doubt on the underlying basis of SIFT.

Moreover, this does not take account of the fact that OTHs also receive SIFT, although at a much lower rate – in our sample, an average of £215 per student week. This suggests that SIFT may in fact be too low, since it does not fully cover the difference between DTH and OTH costs. However, it is likely that this difference is not fully explained by teaching, in which case SIFT appears much more reasonable.

Dissemination survey

Again, Hollingsworth and Parkin (2000) give detailed results and the following is a summary of the majority views amongst respondents. Respondents were surprised that:

- DTHs are as efficient as OTHs. Explanations included their historically better funding, their ability to use trainees as subsidised labour and their better facilities.
- DTHs are not advancing technologically.

They were not surprised that:

- OTHs are catching up technologically.
- Teaching hospitals have higher costs.

A further point made was that information that would add to the analysis includes data on health outcomes, and on the quality of the teaching that is delivered in hospitals.

Preferences concerning the mechanism for funding teaching were:

- Reimbursement per student: preferred by 58% of Trusts, 25% of HAs and all Medical Schools.
- Block budgets: preferred by 25% of Trusts and 25% of HAs.
- A mixture of the two: preferred by 17% of Trusts and 50% of HAs.

The main stated objective for any reimbursement mechanisms was that funding should follow quality teaching, and be spread more evenly.

Discussion and conclusions

This paper demonstrates that the quantitative approaches of DEA and SFA can be used on routine data to produce useful information on the efficiency and productivity of NHS hospitals taking into account teaching. The results from these two methods are, however, not consistent and further work is needed to validate them, in order to see which method produces analyses that are closest to the truth. The authors' preferences are for DEA, with its admitted faults, and the rest of these conclusions do not draw upon the SFA results.

A major problem with all of the analyses reported here is the deficiency of the data in two very important respects. There are no data on the quality of teaching, and although there are data on research activities, they are of very dubious value.

The DEA results dispel to some extent any belief that DTHs are inefficient. However, OTHs are catching up rapidly in terms of productivity, and in production efficiency terms there is little to choose between the two hospital groups. In cost terms DTHs are less cost efficient than other Trusts. However, this may simply mean that DTHs operate with respect to a different cost technology and their performance should be assessed in a different manner to OTHs. NTHs are unambiguously the least efficient group, but this is of less relevance to the issue of teaching, since such hospitals may be entirely unsuitable for it.

Clearly, much more work is required to obtain estimates that would be of value in calculating the appropriate value of any cost-per-student fee or the SIFT allocation factor. There is, however, a hint that current SIFT factors may be based on a misspecified model, though the evidence is at present weak.

These findings have implications for the funding mechanism. Teaching has an impact on efficiency and costs independent of case mix. A cost-per-student element to funding is therefore justified to accompany a case-mix based reimbursement; this was the mechanism most respondents to our second survey favoured. This would suggest a much bigger role for OTHs, who would be better reimbursed for teaching. Since OTHs compare favourably with DTHs in terms of efficiency, this would lead to a better allocation of resources for teaching. It would also ensure that DTHs were fully reimbursed and funding was transparent. However, the weak evidence that DTHs operate with reference to a different cost technology suggests that there might be different levels of case-mix payment, based on the amount of teaching. Further research is needed to establish this key point.

References

Anderson P. and Peterson N.C. A procedure for ranking efficient units in data envelopment analysis. *Management Science;* 1993: **39(10)** 1261-1264.

Bevan, G. The medical service increment for teaching (SIFT): a £400m anachronism for the English NHS? *British Medical Journal*; 1999: **319** 908-911.

Butler, J.R.G. *Hospital Cost Analysis* Kluwer Academic Publishers, 1995.

Chartered Institute of Public Finance and Accountancy (CIPFA) *The Health Service Financial Database.* London: CIPFA, 1998.

Coelli, T., Rao, D.S.P., Battese, G.E. *An introduction to efficiency and productivity analysis.* London: Kluwer Academic Publishers, 1998.

Cooper W, Seiford LM, Tone K. Data Envelopment Analysis. Dordrecht: Kluwer Academic Publishers; 2000.

Culyer, A.J., Wiseman, J., Drummond, M.F., West, P.A. What accounts for the higher costs of teaching models? *Social Economic Administration;* 1978: **12** 20-30.

Department of Health *Planning the Medical Workforce.* Medical Workforce Standing Advisory Committee: Third Report. December, 1997.

Department of Health *Our Healthier Nation. A Contract for Health.* Green Paper Cm 3854, February, 1998.

Frick, A.P., Martin, S.G., Schwartz, M. Case-mix and cost differences between teaching and non-teaching hospitals. *Medical Care;* 1985 **23(4)** 283-295.

Gerdtham, U-G, Löthgren, M., Tambour, M., Rehnberg, C. Internal markets and health care efficiency: a multi-output stochastic frontier analysis. *Health Economics*; 1999: **8** 151 - 164.

Hollingsworth, B. and Parkin, D. Efficiency and Productivity Change in the English National Health Service: Can Data Envelopment Analysis provide a robust and useful measure? *Journal of Health Services Research and Policy;* 2003: **8** 230-236.

Hollingsworth, B. Non-parametric and parametric applications measuring efficiency in health care. *Health Care Management Science*; 2003: **6(4)** 203-218.

Hollingsworth, B., Dawson, P., Maniadakis, N. Efficiency Measurement of Health Care: A review of non-parametric methods and applications. *Health Care Management Science;* 1999: **2**(3) 161-172.

Hollingsworth, B. and Parkin, D. *The effects of teaching on hospital costs and efficiency in the UK NHS*. A Report to the NHS Executive Northern & Yorkshire Region R&D Directorate. Health Economics Group, School of Health Sciences, University of Newcastle, 2000.

Hollingsworth, B. and Parkin, D. *Developing Efficiency Measures for use in the NHS.* A Report to the NHS Executive Northern & Yorkshire Region R&D Directorate. Health Economics Group, School of Health Sciences, University of Newcastle, 1998.

Parkin, D. and Hollingsworth, B. Measuring Productive Efficiency of Acute Hospitals in Scotland, 1991-1994: Validity Issues in Data Envelopment Analysis. *Applied Economics*; 1997: **29(11)** 1425-1438.

Sloan, F.A., Feldman, R.D., Steinwald, A.B. Effects of teaching on hospital costs. *Journal of Health Economics;* 1983: **2** 1-28.

Table 1: Data used in analysis

Inputs:
Medical and dental staff numbers
Administration & clerical staff numbers
Professions allied to medicine staff numbers
Nursing & midwifery staff numbers
Ancillaries staff numbers
Other staff numbers
Total staff numbers
Staff costs*
Clinical Supplies and Services costs*
General supplies costs*
Establishment costs*
Transport costs*
Premises costs*
Capital (K) charge§*
Service increment for teaching payment

§ A proxy for capital - actually a charge levied for use of capital assets and new capital investment.

Outputs:
Leaning disability occupied bed days
Mental illness occupied bed days
Mental illness (MI) & learning disability (LD) FCEs
Maternity FCEs
General & Acute (G&A) FCEs - Elective
General & Acute FCEs - Non-elective
General & Acute FCEs - Day Cases
Total elective General & Acute
Total General & Acute FCEs
Day Cases
Total FCEs
Day / night attenders
Mental illness & leaning disability 1st attendances
Maternity outpatient 1st attendances
General & Acute outpatient 1st attendances
Accident & Emergency (A&E) attendances
Student undergraduate weeks

Source: * = CIPFA database, otherwise NHSE N&Y

FCE = Finished Consultant Episode

Table 2: Summary of health care production efficiency DEA models by functional status.

Health Care	**Median**	**Mean**	**Minimum**	**Standard Deviation**	**Number Efficient**
All					
1994/5	79.27	70.69	19.07	27.8	7
1995/6	83.28	77.63	3.86	25.7	12
1996/7	85.11	72.60	9.91	32.3	14
1997/8	77.88	67.42	8.22	33.2	10
Acute					
1994/5	86.94	84.98	56.0	13.78	2
1995/6	100	91.28	63.47	13.75	8
1996/7	100	92.23	64.59	12.78	8
1997/8	97.44	90.6	63.0	13.29	6
Priority					
1994/5	30.76	43.8	19.07	27.03	1
1995/6	66.74	60.42	23.68	22.52	0
1996/7	21.90	36.14	9.91	29.12	1
1997/8	18.73	28.64	8.22	21.58	0
Combined					
1994/5	86.57	81.22	23.64	21.23	4
1995/6	90.77	79.87	3.86	29.51	4
1996/7	89.52	86.62	35.94	17.8	5
1997/8	83.39	80.04	17.83	22.97	4

Table 3: Summary of health care and education production efficiency DEA models by functional status.

Health and Education	**Median**	**Mean**	**Minimum**	**Standard Deviation**	**Number Efficient**
All					
1994/5	100	90.12	51.4	14.3	20
1995/6	100	90.19	18.7	17.9	21
1996/7	100	87.19	11.6	21.6	22
1997/8	100	85.14	25.1	21.9	21
Acute					
1994/5	100	93.55	51.4	13.72	9
1995/6	100	99.23	89.9	2.8	12
1996/7	100	98.83	84.8	4.2	12
1997/8	100	98.76	83.9	4.4	12
Priority					
1994/5	96.24	88.26	52.9	16.5	5
1995/6	86.53	77.54	18.7	26.6	2
1996/7	72.48	68.31	11.6	29.8	4
1997/8	66.37	64.42	25.1	27.2	3
Combined					
1994/5	89.72	88.4	60.5	13.2	6
1995/6	100	92.82	77.4	9.1	7
1996/7	99.58	92.99	81.1	8.0	6
1997/8	94.1	90.64	71.2	10.9	6

Table 4: Summary of health care and education production efficiency DEA models by teaching status.

Health and Education	**Median**	**Mean**	**Minimum**	**Standard Deviation**	**Number Efficient**
Designated Teaching Hospitals					
1994/5	100	96.16	85.8	6.6	5
1995/6	100	95.24	79.6	8.5	5
1996/7	100	94.19	77.4	9.9	4
1997/8	100	93.57	74.3	11.1	5
Other Teaching Hospitals					
1994/5	91.17	87.24	51.4	15	10
1995/6	100	93.34	59.5	10.6	13
1996/7	100	91.74	47.0	13.3	14
1997/8	100	91.11	43.9	14.8	14
Non-Teaching Hospitals					
1994/5	100	93.09	52.9	16.4	5
1995/6	89.06	76.56	18.7	18.7	3
1996/7	73.84	68.0	11.6	11.7	4
1997/8	60.86	60.59	25.1	25.1	2

Table 5: Summary of cost efficiency DEA models by teaching status

Health and Education	**Median**	**Mean**	**Minimum**	**Standard Deviation**	**Number Efficient**
Designated Teaching Hospitals					
1994/5	87.46	74.85	28.3	29.4	2
1995/6	93.42	76.80	39.7	27.1	2
1996/7	91.54	74.86	18.5	32.2	2
1997/8	92.49	75.44	27.1	29.8	2
Other Teaching Hospitals					
1994/5	73.85	70.82	12.2	27.2	4
1995/6	72.81	66.64	10.4	27.5	1
1996/7	77.62	70.22	8.0	32.7	5
1997/8	76.45	70.01	7.8	30.3	3
Non-Teaching Hospitals					
1994/5	28.89	34.48	5.3	28.2	0
1995/6	32.23	34.62	3.7	26.2	0
1996/7	24.05	28.20	3.8	21.9	0
1997/8	9.16	26.45	2.8	29.7	0

Table 6: Summary of super efficiency, cost efficiency DEA models by teaching status

Health and Education	**Median**	**Mean**	**Minimum**	**Standard Deviation**	**Number Efficient**
Designated Teaching Hospitals					
1994/5	87.46	77.34	28.3	32.5	2
1995/6	93.42	82.41	39.7	33.6	2
1996/7	91.54	81.27	18.5	39.2	2
1997/8	92.49	80.74	27.1	35.7	2
Other teaching hospitals					
1994/5	73.85	79.19	12.2	46.0	4
1995/6	72.81	70.22	10.4	36.0	1
1996/7	77.62	72.32	8.0	35.0	5
1997/8	76.45	88.52	7.8	95.7	3
Non teaching hospitals					
1994/5	28.89	34.48	5.4	28.2	0
1995/6	32.23	34.62	3.7	26.2	0
1996/7	24.05	28.20	3.8	21.9	0
1997/8	9.16	26.45	2.8	29.7	0

Table 7: Malmquist indexes for DEA production models of health care by functional status and health care and education by functional status and teaching status

	Efficiency Change	***Technological Change***	***Total Change***
Health Care			
All	0.95	1.1	1.05
Acute	1.03	1.1	1.13
Priority	0.86	1.12	0.97
Combined	0.99	1.1	1.09
Health care & Education			
All	0.97	1.08	1.05
Acute	1.02	1.04	1.07
Priority	0.89	1.18	1.04
Combined	1.01	1.05	1.06
Designated Teaching	0.99	1.02	1.01
Other teaching	1.02	1.06	1.08
Non teaching	0.92	1.12	1.02

Table 8: Malmquist indexes for DEA cost models by teaching status.

	Efficiency Change	***Technological Change***	***Scale Change***	***Total Change***
Designated Teaching	1	1.05	0.99	1.05
Other teaching	0.99	1.09	0.98	1.08
Non teaching	0.84	1.13	0.86	0.94

Table 9: Production efficiency SFA model.

Average efficiency scores

	1994/5	*1995/6*	*1996/7*	*1997/8*
All Trusts	0.46	0.45	0.45	0.45
Designated Teaching	0.54	0.54	0.54	0.53
Other teaching	0.53	0.53	0.52	0.52
Non teaching	0.17	0.17	0.16	0.16

Table 10: Summary of cost efficiency SFA model.

Average efficiency scores

	1994/5	*1995/6*	*1996/7*	*1997/8*
All hospitals	3.02	2.93	2.84	2.76
Designated Teaching	5.18	4.96	4.76	4.57
Other teaching	2.63	2.56	2.49	2.43
Non teaching	2.26	2.21	2.17	2.12

Table 11: SIFT cost function models.

	Model 1		Model 2	
	Coefficient	Standard Error	Coefficient	Standard Error
Constant	20,400,000	3,518,469	21,600,000	3,341,710
Student weeks	14,076.64	1,001.46	10,866.23	1,215.68
Total FCEs	177.39	56.2	161.96	53.3
Outpatients	3,934,836	1,618,278	2,701,639	1,559,610
A&E attends	430.59	70.81	438.75	67.06
Case mix index	1,373,570	584,730	1,024,975	559,677
Teaching			17,900,000	4,234,034
F (Prob > F)	201.08	(0.000)	189.95	(0.000)
Adj R^2	0.870		0.883	