E Referees' reports

E.1 Referee report of Adjunct Professor Tim Coelli (University of Queensland)

This Productivity Commission (PC) study of hospital performance in Australia sets a new benchmark in terms of the sample coverage and the range of input, output and control variables included in the econometric analysis. The PC team has faced a number of data access challenges, and are to be commended for the quality and breadth of their analysis.

Data sample

The sample data involves 459 hospitals observed over a four-year period from 2003-04 to 2006-/07. Of these hospitals, 343 are public, 99 are private and 17 are contract hospitals. A total of 1806 observations are used to estimate the econometric models. In my assessment, this sample size is more than sufficient to allow the Productivity Commission (PC) to reliably estimate an econometric model that involves a flexible functional form and a number of important input, output and control variables.

One area of concern with the data is the low response rate among not-for-profit hospitals. The PC uses weighted econometric methods to address this issue. They also note that some sample selection bias could remain. For example, if the non-respondents tend to be relatively inefficient. If this was the case, the mean efficiency of the not-for-profit group could be overestimated to some extent. However, the position of the estimated best-practice frontier is unlikely to be notably affected by the omission of inefficient observations, and hence the results for the remaining groups are unlikely to be affected by this sampling issue.

Frontier methodology

There are two frontier estimation methods that are commonly used in the literature: data envelopment analysis (DEA) and stochastic frontier analysis (SFA). DEA is a

linear programming method that has the advantage that no particular functional form needs to be specified. However, SFA is an econometric method that is less susceptible to the effects of data noise and outliers and which also allows one to easily incorporate control variables that involve categorical and ratio data. Hence, in my assessment the choice of SFA is appropriate for this study.

The translog function form has been used. It is a flexible second-order functional form that can accommodate a range of scale and substitution possibilities. This functional form should generally be used in preference to a first-order functional form (such as the Cobb-Douglas) when sufficient sample data is available, as it is in this case.

The PC has chosen to estimate three different types of frontier models: an input-oriented distance function, an output-oriented distance function and a cost function. One can make a case for the use of each of these models on the basis of the management/ownership characteristics of a particular hospital. Given that the PC study pools data from various hospital types (including public, private for-profit and private not-for-profit) the use of the three different models provides a form of a sensitivity analysis, to ensure that no one type of hospital is disadvantaged by the model type that is chosen.

Output measures

The output measures involve a number of categories of admitted and non-admitted separations, with the former casemix-adjusted. The level of detail is substantially better than many past studies of hospital efficiency. The authors emphasize the point that these are measures of intermediate outputs rather than incremental health benefits derived from the services. However, this is standard practice in this literature, given the very substantial challenges that would be involved in attempting to derive these latter output measures.

Input measures

The input measures include three categories of staff members (nursing, diagnostic and other), three monetary measures of non-staff variable inputs (drugs, medical and surgical supplies and other) along with the number of beds. This group of input measures is better than that used in the majority of past studies, but can still be improved upon (given access to better data). In particular, the beds measure treats an intensive care bed no differently to a standard bed, and the staff measures exclude doctors. These issues could introduce some biases in efficiency estimates if the casemix weights (used to define the output measures) include allowances for the

extra capital costs associated with complex cases, and if there are differences among hospitals in the degree to which doctors versus nurses undertake certain "grey area" tasks.

Quality measures

Quality issues have been often overlooked in past studies of health sector efficiency. The PC is to be commended for their efforts in this regard. The inclusion of a mortality rate measure that is adjusted for patient risk characteristics is not a perfect measure, but should go a long way to capturing any notable variations in the effects of service quality upon efficiency potentials.

Control measures

The PC has considered a wide range of exogenous control measures that could potentially be affecting efficiency potentials, including network membership, accident and emergency rates, and so on. These measures help the analyst to avoid labelling a hospital as being 'inefficient' when they may be using more resources per unit output because they face different operating conditions relative to other members of the sample.

Additional comments

In the future, the analysis could be extended to include some estimates of scale economies or scale efficiencies. This information could be particularly valuable to the current public discussion of the effects of casemix funding on small regional hospitals.

The PC has identified two important areas where data is lacking, namely, data on capital costs and medical practitioners. The hospitals sector should be strongly encouraged to collect and then make this type of data available to future studies of this nature.

Finally, I note that the PC have expressed some frustration with the degree to which privacy requirements have constrained what can reported in this document. I would like to add my support here, and also observe that privacy requirements unduly constrain access to data in many regulated sectors in Australia. I find this difficult to comprehend in those situations where public funds and/or monopoly regulation is involved. Even in the USA (where private enterprise is king) there is much more public transparency with regards to data reporting in these situations. As an

example, I encourage the reader to access the Federal Energy Regulatory Commission website (http://www.ferc.gov/) and look at the large amount of detailed firm-level data that is publicly reported there. I am not as familiar with the health sector in the USA, but I understand that similar public data reporting requirements apply there as well.

E.2 Referee report of Professor Jim Butler (Australian National University)

The analyses contained in this report use a large dataset on Australian public and private hospitals covering a period of four financial years (2003-04 to 2006-07). A total of 459 hospitals are included with 1806 observations available for analysis. The compilation of this dataset itself is an impressive achievement, particularly involving as it did the merging of information from ABS collections on private hospitals with data from the AIHW on hospital morbidity and other aspects of public and private hospitals. It is unfortunate that the resulting dataset appears unlikely to be available in the public domain for use by other analysts. Researchers will undoubtedly share the Commission's laments in this regard.

Econometric modelling invariably involves choices of the phenomena to be modelled, the functional form to be estimated, the estimation methods and statistical assumptions to be imposed on the data (to name a few). The report is generally very clear about the choices that have been made. One aspect that could perhaps have been elaborated upon more fully was the decision to treat the data as a pooled cross-section dataset with time trend dummies included rather than as a panel dataset. The data appear to be suitable for analysis using panel methods.

The approaches to estimating hospital quality and hospital efficiency are well documented and have been widely used in the literature. The measure of hospital quality is the hospital-standardised mortality ratio (HSMR), a ratio of actual to expected numbers of in-hospital deaths. Expected numbers of deaths are obtained from an estimated relationship between the number of in-hospital deaths and vectors of patient and hospital characteristics. In principle, these vectors should include factors related to hospital mortality that are beyond the control of the hospital, implying that any remaining unexplained variation in hospital mortality is attributable to factors that are within the control of the hospital. In practice, the distinction between the determinants of mortality that are exogenous and endogenous to the hospital is not always clear cut. The choice of hospital characteristics included in the Commission's analysis is defensible and, as alluded to in the report, may usefully serve as proxies for attributes of patients not captured in the vector of patient characteristics. The use of a negative binomial model to

estimate the relationship is entirely appropriate given that the data are hospital-level and the dependent variable is a non-negative integer (counts of deaths).

The use of stochastic frontier analysis (SFA) with a one-sided error component to measure inefficiency as opposed to deterministic data envelopment analysis (DEA) is a defensible choice. Three sets of efficiency scores are estimated, two relating to technical efficiency and one relating to cost efficiency. The sensitivity of the technical efficiency scores to the assumed distribution of the one-sided error term (but not the cost efficiency scores) is investigated. High correlations (both Pearson's and Spearman's rho) between the technical efficiency scores under two different distributional assumptions are found.

The analyses of hospital quality show private hospitals having significantly lower HSMRs than either public hospitals or public contract hospitals, with variations in this difference occurring across hospital size groupings. This is a very interesting result, but its veracity depends upon the extent to which the vectors of patient and hospital characteristics fully capture differences in the severity of disease between patients. This caveat is recognised in the report.

Two sets of technical efficiency scores are provided based on the estimation of output-oriented and input-oriented distance functions. As both of these are based on the production function, one might expect there to be a high correlation between them. The report investigates this issue (see table 5.5) and, while it indicates that the correlation coefficients are certainly positive, they are not perhaps as high as one would expect. For all public hospitals, the correlation coefficient is 0.486 and for all private hospitals it is 0.539. The two sets of results also show differing effects of hospital ownership on efficiency although the materiality of any differences in efficiency by hospital ownership type is, for pragmatic purposes, debatable. The output-oriented scores show statistically significant differences in efficiency by ownership type with for-profit private hospitals being the most efficient and not-for-profit private hospitals the least, but the difference in average efficiency scores between the most and least efficient groups is less than 10.0 percentage points (94.8 versus 85.6). The input-oriented scores place public contract hospitals as being the most efficient and public hospitals the least, with the difference between the average scores for the most and least efficient being 4.5 percentage points (93.6 versus 89.1).

Regarding cost efficiency, no statistically significant differences in cost efficiency scores by ownership type are found. With the exception of public contract hospitals which have an average cost efficiency score of 90.4, the other three ownership types all have average scores in the range 93.0 - 94.0. Given the duality between cost and production functions, and the statistical significance of ownership type in the

analysis of technical efficiency scores, this result is surprising. Data limitations may be part of the explanation.

Overall, this is a rich report which makes a substantial contribution to the empirical Australian literature on hospital efficiency. The quality of the econometric modelling suggests that the results will be of substantive interest to a range of parties. Nevertheless, given the limitations of the data and the econometric/measurement problems that are inevitable in work of this kind, a robotic interpretation of the results should be resisted.