Issues in cost function specification for neonatal care: the Fordham case

C. O’Neill and A. Largey

Summary

Background Econometric techniques have been used to examine the relationship between costs of provision, case mix and unit size for health care providers. Estimation of cost functions in a health care context is complicated by poor understanding of the underlying production relationship and the constraints under which production takes place. Different results and policy implications can follow from different model specifications. This underscores the need for care in the construction of such functions and the interpretation of their results.

Methods Cost and activity data from a study of neonatal care for an English Regional Health Authority are re-examined. Cost functions are estimated using alternative functional forms, and average cost per day is estimated and compared for two of these functions [one produced by Fordham et al. (J Publ Hlth Med 1992; 14(2): 127–130) and the other based more explicitly in economic and econometric theory].

Results It is shown that estimates of average cost per day are sensitive to model specification. Such are the differences in cost that significantly different policy implications could follow from the different models.

Conclusion We conclude that care must be taken in the construction and estimation of cost functions and that the assumptions upon which they rest be made explicit so that results can be properly interpreted.

Keywords: neonatal, costs, specification

Introduction

A cost function summarizes the relationship between output and what is normally assumed to be the minimum cost of producing that output. It shows, for example, how the cost of producing a good varies with the quantity of the good produced, thus allowing the researcher to establish if per unit costs rise or fall as output rises, or if the use of a certain combination of inputs is more efficient than another. Implicit in such a function is the existence of an underlying production relationship. That is, implicit in the relationship between output and cost is a relationship between output and inputs which give rise to costs. A cost function can be derived (through simultaneous equations) from a fully specified production function and indeed can ultimately only be properly interpreted if such a production function exists.

In health care provision the production relationship is often poorly understood. The objective functions of practitioners may differ, as could the constraints within which they operate in their particular institutions. For example, one practitioner may be interested in developing innovative techniques to a much greater degree than another. That is, part of their objective may be to be at the forefront of their chosen specialty. Another practitioner may prefer to have a greater role in administration within their setting and devote relatively more effort to this activity. With clinical freedom this could see different solutions adopted for the same problem and, depending on the degree to which practitioners devote themselves to different tasks, different throughputs of patients per year between practitioners. By the same token, different hospital administrations may suggest different drug regimes or the same drug regime with differing degrees of force for the treatment of a particular complaint. That is, the constraints within which practitioners operate may differ.

Finally, here, even if all practitioners shared the same goal and operated within the same constraints, the fact that the production process for health care is not of an engineering type, in the sense that heterogeneous patients may respond differently to the same treatment, may account for differences between practitioners in terms of output. These and the difficulties associated with defining and measuring output, accounting for case mix and the heterogeneous nature of inputs make the concept of a production frontier, i.e. defining the maximum output that can be obtained from a given amount and combination of inputs, fuzzy.

The corollary of this is that a definitive approach to the specification of production and cost functions in the context of health care does not exist, and a variety of approaches have been used in practice to investigate the relationship between costs and outputs. Fordham et al. presented a function for neonatal care and used it to examine the relationship between
costs and scale of operations and between cost and activity mix. This paper argues that the Fordham model may be mis-specified and compares the original results with those obtained using a more appropriate functional form. (The interested reader is referred to Folland et al.\textsuperscript{10} for a fuller discussion of the theory and practice of cost function analysis in health care.) We begin by outlining the approach adopted by Fordham et al.\textsuperscript{9} and providing a brief critique of this. A more rigorous approach to finding the appropriate functional form is presented and used to generate a function. Results obtained using this and the Fordham model are compared and discussed.

**The Fordham function**

Fordham et al.\textsuperscript{9} presented a function in which costs are determined by the size of the unit (measured using the total number of days of care it provides) and the intensity of the care it provides (measured using the proportion of intensive care days in the total number of days it provides). The economic theory on which the function is based is not made explicit but can be inferred. An assumption of cost minimization is implicit which (the previous reservations concerning practitioner objective functions and the constraints operating on them not withstanding) seems plausible\textsuperscript{11} especially given the units involved were from one region.

Output is defined in terms of activity, i.e. days of care provided with account taken of the level of care provision. This circumvents the difficulties associated with defining or measuring output in terms of health outcomes or of taking further account of differences in the complexity of cases presenting for treatment (i.e. the need to measure value added is obviated). Care is warranted in the interpretation of results. Results from the analysis can be interpreted as showing the relationship between cost and outputs if output can be assumed to be some constant function of activity. What is more realistic is to assume that the function shows the relationship between cost and activity. However, this is not to discount its usefulness to policy makers. Should scale economies or efficiency gains associated with specialization exist in the production of activity, this should be discernible from the function, and the activity level and mix planned for appropriately. That is, the function should allow us to identify whether larger units are more efficient than smaller ones and to establish if specialized units are more efficient than less specialized units. The results have, indeed, been referred to in debates concerning the provision of neonatal care in the United Kingdom,\textsuperscript{12,13} even though some misgivings have been expressed concerning them.\textsuperscript{14}

Fordham et al.\textsuperscript{9} suggested that plausible a priori reasons exist for expecting economies of scale in the production of care for neonates over a range of output at least. They also suggested that a positive correlation should be expected between intensity of care and costs. Using two functional forms - linear and quadratic - they produced results apparently supporting these contentions. (A fuller discussion of the data has been given by Fordham et al.\textsuperscript{9}) Simplicity and plausibility were cited as justification for their use of these forms. No tests were reported on the forms used. The authors referred to the size of the sample as further justification for their lack of sophistication. Neither model takes account of possible interaction effects between variables.

**A brief critique of the Fordham approach**

The approach adopted by Fordham et al. can be criticized on two fronts. First, as no reference is made in their model to interaction terms between regressors, i.e. the combined effect of two variables together changing, the potential for mis-specification of the function owing to variable omission exists.\textsuperscript{15,16} A priori, such interactions could be expected and indeed the authors acknowledge this, though they nevertheless fail to take account of it. For example, there appears no reason to assume that any increased efficiency associated with increased scale will be uniform across the types of care offered. Thus, as scale and the proportion of intensive care of total care provided by neonatal units increase so units may become more (or less) efficient in the production of intensive care owing to greater specialization; such an effect can be accounted for by including the product of scale and case mix variables. The omission of interaction terms would imply that were they significant the function would produce biased coefficient estimates.

The failure to formally test alternative functional forms provides the second source of difficulty with the results of Fordham et al. As shown by O'Neill and Davis,\textsuperscript{17} albeit in a different context, results can often be highly sensitive to choice of functional form. Plausibility and simplicity are valid criteria by which to assess functional form, but are not sufficient, especially if results are thought to be sensitive to the form used.

**Results**

The original Fordham et al. results are given in Equations (1) and (2) of Table 1, and Equation (3) shows results for a linear function including an interaction term. Using the same data set as Fordham et al., a Box–Cox transformation was used to find the appropriate functional form. The use of log transformations to both independent and dependent variables was suggested (see Table 2). (A fuller discussion of the use of the Box–Cox model in assessing appropriate functional forms has been given by Greene\textsuperscript{18}. Results obtained using this form are given in Equation (4) of Table 1. The cost function was augmented to include an interaction term between scale and the proportion of intensive care days. Results obtained using a double-log function with this additional variable are presented in Table 1, Equation (5). Equation (6) shows the results of a double-log function with an interaction term, where the Fordham case mix variable is redefined as the percentage of intensive care days.
Table 1 Parameter estimates for cost functions under different functional forms

<table>
<thead>
<tr>
<th>Equation</th>
<th>PIC</th>
<th>TD</th>
<th>$R^2$</th>
<th>$F$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cons</td>
<td>PIC</td>
<td>TD</td>
<td>$-0.03279^*$</td>
</tr>
<tr>
<td>2</td>
<td>Cons</td>
<td>PIC</td>
<td>TD</td>
<td>$-0.06554^*$</td>
</tr>
<tr>
<td>3</td>
<td>Cons</td>
<td>PIC</td>
<td>TD</td>
<td>$-0.0407^*$</td>
</tr>
<tr>
<td>4</td>
<td>Cons</td>
<td>ln(PIC)</td>
<td>ln(TD)</td>
<td>$-0.3998^*$</td>
</tr>
<tr>
<td>5</td>
<td>Cons</td>
<td>ln(PIC)</td>
<td>ln(TD)</td>
<td>0.09818</td>
</tr>
<tr>
<td>6</td>
<td>Cons</td>
<td>ln(PIC)</td>
<td>ln(TD)</td>
<td>$-0.91397^*$</td>
</tr>
</tbody>
</table>

*Significant at $\alpha = 0.05$. All models were tested for heteroscedasticity using the Breusch-Pagan test. In no instance was it found to be significant at $\alpha = 0.05$.

The dependent variable is total cost divided by total days or, in Equations (4), (5) and (6), ln(ACPD). IC is the total number of IC days produced. PIC = IC/TD. Inter = PIC x TD ln(lnter) = ln(PIC) x ln(TD).

Table 2 Results from a Box–Cox transformation of the model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of IC days</td>
<td>0.2948</td>
<td>0.109</td>
</tr>
<tr>
<td>Total days</td>
<td>-76728</td>
<td>-0.029</td>
</tr>
<tr>
<td>Constant</td>
<td>76865</td>
<td>0.032</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>-1.0000</td>
<td>-0.308</td>
</tr>
<tr>
<td>$\theta$</td>
<td>1.0006</td>
<td>0.441</td>
</tr>
</tbody>
</table>

(PRIC) rather than the proportion (PIC). The inclusion of Equation (6) in the analysis illustrates the importance of correct interpretation of the regression results. [When checking for scale effects in the presence of the interaction term, it is not enough to check the sign and significance of the coefficient on ln(TD), where TD represents total number of days of care provided. This coefficient is positive and insignificantly different from zero in Equation (5), whereas it is significant and negative in Equation (6). In the Appendix, however, it is shown that Equations (5) and (6) in fact give equivalent results.]

Table 3a Estimates of average cost per day obtained using Equation (2)

<table>
<thead>
<tr>
<th>Total no. of days</th>
<th>Proportion of days which were IC</th>
<th>0.05</th>
<th>0.1</th>
<th>0.4</th>
<th>0.6</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>216.30</td>
<td>229.62</td>
<td>309.51</td>
<td>362.78</td>
<td>416.04</td>
<td></td>
</tr>
<tr>
<td>1391</td>
<td>195.53</td>
<td>208.84</td>
<td>288.74</td>
<td>342.00</td>
<td>395.26</td>
<td></td>
</tr>
<tr>
<td>1392</td>
<td>195.47</td>
<td>208.79</td>
<td>288.69</td>
<td>341.95</td>
<td>395.21</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>166.33</td>
<td>179.64</td>
<td>259.54</td>
<td>312.80</td>
<td>366.06</td>
<td></td>
</tr>
<tr>
<td>3500</td>
<td>110.82</td>
<td>124.14</td>
<td>204.03</td>
<td>257.30</td>
<td>310.56</td>
<td></td>
</tr>
<tr>
<td>5000</td>
<td>78.66</td>
<td>91.98</td>
<td>171.88</td>
<td>225.14</td>
<td>278.40</td>
<td></td>
</tr>
</tbody>
</table>

Table 3b Estimates of average cost per day obtained using Equation (6)

<table>
<thead>
<tr>
<th>Total no. of days</th>
<th>Proportion of days which were IC</th>
<th>0.05</th>
<th>0.1</th>
<th>0.4</th>
<th>0.6</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>236.45</td>
<td>224.84</td>
<td>203.31</td>
<td>197.41</td>
<td>193.33</td>
<td></td>
</tr>
<tr>
<td>1391</td>
<td>196.54</td>
<td>196.52</td>
<td>196.50</td>
<td>196.50</td>
<td>196.49</td>
<td></td>
</tr>
<tr>
<td>1392</td>
<td>196.46</td>
<td>196.47</td>
<td>196.49</td>
<td>196.50</td>
<td>196.50</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>160.36</td>
<td>169.47</td>
<td>189.28</td>
<td>195.49</td>
<td>200.03</td>
<td></td>
</tr>
<tr>
<td>3500</td>
<td>117.20</td>
<td>134.89</td>
<td>178.65</td>
<td>193.96</td>
<td>205.61</td>
<td></td>
</tr>
<tr>
<td>5000</td>
<td>95.97</td>
<td>116.62</td>
<td>172.20</td>
<td>193.00</td>
<td>209.24</td>
<td></td>
</tr>
</tbody>
</table>
Discussion

A comparison of the various equations shows the sensitivity of results to model specification. Comparing estimates of average cost per day obtained using Equation (6) (functional form suggested by Box-Cox with interaction term included) with those from Equation (2), the Fordham equation, for example, shows stark differences in results and the policy implications of those results. Table 3a, derived from Equation (2), shows average cost per day falling as scale increases, regardless of the proportion of IC days the unit is producing. (Average cost per day is, however, higher for units producing a greater proportion of IC days, \textit{ceteris paribus}, as can be seen from the table.) In Table 3b, the results obtained using Equation (6) show average cost per day falling for units with a percentage of IC days below 60 per cent. Equation (6), in fact, predicts that this will be the case for the percentage of IC days up to approximately 64 per cent. For units with higher percentages of IC days, average cost per day rises with scale. That is, the scale effects are not independent of case mix. As units become more specialized, so scale economies become less significant, and for units with a percentage of IC days above 64 per cent these are replaced by scale diseconomies. As can also be seen from a comparison of the two tables, estimates of average cost per day for small units with a high percentage of IC days are very much lower than those obtained using Equation (2). For example, Fordham et al.\textsuperscript{9} would predict costs for a 1000 bed unit with 40 per cent of cases being IC to be over 52 per cent higher than suggested by Equation (6) (£309.51 and £203.31, respectively) and for 1000 bed units with 80 per cent of cases being IC to be over 115 per cent higher (£416.04 and £193.33, respectively). Such are these differences in magnitude, it could be expected, that the results of Equation (2) would prompt very different policy responses from government considering changes in the provision of services, compared with those obtained using Equation (6).

This serves to underline the need for care in the development of cost functions of this type. That the results obtained from the original Fordham et al. paper have been used in debate concerning the provision of services, as noted,\textsuperscript{11,12} despite misgivings expressed concerning their reliability,\textsuperscript{13} further underscores this. Although it is unlikely that policies would be based on the results from one study using just 17 observations, such findings could, in the absence of other information, be seized upon and used in public debate. Given the relationship between public opinion and government allocation decisions this could have adverse consequences on efficiency. We do not suggest that the results obtained using Equation (6) are definitive, though they are arrived at in a fashion more consistent with economic theory and econometric practice, and in this sense are superior to those of Fordham et al. However, given our limited understanding of cost functions in health care and the paucity of data the researcher often has available, clearly there is a need for care in interpreting results from analysis such as this, and to resist the temptation to seize upon results or use them in a oversimplified fashion in debate.

References


Accepted on 6 August 1996
Appendix

The general form of equation (5) is

\[
\ln (ACPD) = a_0 + a_1 \ln (PIC) + a_2 \ln (TD) + [a_3 \ln (PIC) \times \ln (TD)]
\]

Setting PRIC = PIC \times 100, the general form of Equation (6) is

\[
\ln (ACPD) = b_0 + b_1 \ln (PIC \times 100) + b_2 \ln (TD) + b_3 [\ln (PIC \times 100) \times \ln (TD)]
\]

Using common rules for manipulation of logs, this can be reduced to

\[
\ln (ACPD) = [b_0 + b_1 \ln (100)] + b_1 \ln (PIC)
+ [b_2 + b_3 \ln (100)] \ln (TD)
+ b_3 [\ln (PIC) \times \ln (TD)]
\]

Thus comparing the two equations, we should have:

\[
a_0 = b_0 + b_1 \ln (100), \quad a_1 = b_1, \quad a_2 = b_2 + b_3 \ln (100) \quad \text{and} \quad a_3 = b_3
\]

It can be verified that these are the relationships which hold between the estimated parameters in Equations (5) and (6) (Table 1).