HEDG Health, Econometrics and Data Group
THE UNIVERSITY of York

HEDG Working Paper 09/26

The costs of new organisational and financial freedom: The case of English NHS trusts

Giorgia Marini

Marisa Miraldo

July 2009

ISSN 1751-1976

york.ac.uk/res/herc/hedgwp

The costs of new organisational and financial freedom: The case of English NHS trusts*

Giorgia Marini† University of York

Marisa Miraldo Imperial College London

July 6, 2009

Abstract

In this paper we estimate hospital costs and evaluate economies of scale and scope using a generalised multiproduct cost function and a sample of English NHS trusts with different types of ownership, namely Foundation Trusts and non Foundation Trusts. Evaluating the behaviour of different types of hospitals separately might be particularly helpful for the design, and future developments, of the optimal provider reimbursement tariff. Also it might shed some light on the ability of different types of hospitals to profit from the existence of economies of scale and scope.

Results show that, even though these two group of providers do not exhibit differences regarding economies of scale, Foundation Trusts exhibit global diseconomies of scope while non Foundation Trusts exhibit global scope economies.

KEYWORDS: Economies of Scale, Economies of Scope, Foundation Trusts, Payment by Results

JEL CLASSIFICATION: I11, O33

*A previous version of this paper was presented at the 10th Portuguese Conference on Health Economics (November 2007), at the Health Econometrics Data Group Seminar Series, University of York (January 2008), at the DTE Seminar Series, University of Rome La Sapienza (February 2008), and at the 7th European Conference on Health Economics hosted by the University of Rome Tor Vergata (July 2008).

†Correspondence to: Giorgia Marini, Centre for Health Economics, Alcuin A Block, University of York, Heslington, York YO10 5DD, UK. Telephone: 0044 (0)1904 321421. Fax: 0044 (0)1904 321454. E-mail: gm509@york.ac.uk.

1 Introduction

Since 2000 the English NHS has been the target of a major reform outlined by a ten-year policy programme announced by the NHS Plan (Department of Health, 2000) and reinforced by the Wanless Report (Wanless, 2002). With this reform, the Labour party aimed at bringing England to EU spending average: "We decided to make an historic commitment to a sustained increase in NHS spending. Over five years it amounts to an increase of a third in real terms. Over time, we aim to bring it up to the EU average." (Foreword by the Prime Minister Tony Blair, the NHS Plan, Department of Health, 2000, p. 8). This investment would have been implemented through a series of policy reforms that, to ensure the taxpayer value for money, would have been shaped around patient needs, with patient choice playing a crucial role.

A package of demand- and supply-side reforms and the areas where these reforms should apply was set in 2005 in the "framework for health reform in England" (Department of Health, 2005). The package included increased patient choice, new systems of information, a new system of financial flows (Payment by Results), increased provider diversity with the introduction of a new legal status for NHS trusts (Foundation Trusts) and new commissioning frameworks with commissioning of health care services from the private sector.

In the context of this paper, we focus our attention mainly on Foundation Trusts (FTs). In particular we are interested in analysing whether providers that have acquired this new legal status are significantly diverse from the remaining with respect to their cost structures and whether FTs perform differently under the newly implemented provider reimbursement system, Payment by Results (PbR).

In 2003 the UK Parliament passed the Health and Social Care Act that transformed the best performing English NHS trusts into NHS FTs (Health and Social Care Act, 2003), with the first wave of FTs being authorised in 2004 and followed by further waves in subsequent years.

These new trusts are given greater autonomy in order to allow the development of business and strategies that better coordinate their financial and operating structure with local needs.

With increased planning flexibility, extended borrowing freedom and ability to retain surpluses, FTs are expected to channel investment to local needs and to improve national targets (Eaton, 2005; Healthcare Commission, 2005). In particular, compared to non-FTs, FTs face a different set of constraints (limited borrowing from private sector under the Prudential Borrowing Limit set by FTs new regulatory body - Monitor, binding contracts with the organisations commissioning services from them, use of national tariffs to price their activities) and incentives (more control over their own future activity, more control over appointing directors, more and quicker access to capital investment, more local control over setting priorities and more freedom in employment of new staff) that may encourage them to change their behaviour (Healthcare Commission, 2005).

Even though the Foundation policy has drastically re-shaped the hospital sector in England, the most important reform has been the change in the financial incentives in secondary care with the introduction of an activity-based reimbursement system, i.e. PbR.

Also implemented by the government in 2003, the PbR is a prospective payment system that rewards hospitals and other providers on the basis of their casemix adjusted activity: hospitals receive a fixed payment - the national tariff - for each type of treatment supplied (Department of Health, 2002b).

The PbR system is targeted at overcoming the deficiencies of both cost-based reimbursement and negotiated budgets. Incentives for cost control and efficient behaviour are introduced by relating payment directly to activity and by ensuring that hospitals cannot influence the price they face. Indeed, under this financial arrangement, the price paid for each treatment is independent of the hospital actual costs and it is fixed in advance as the average cost of all hospitals for each health resource group (HRG).[1]

As this new activity-based reimbursement policy rewards providers on the basis of their relative performance, it therefore generates incentives for promoting efficiency and consequently lower costs, it overcomes inefficient spending arising from the open-ended budgets and it creates incentives for improving performance. It is then expected that providers reduce slack, increase effort, reduce lengths of stay, reduce rates of use of ancillary services, reduce the total ratio of personnel to patients, "unbundle" services, invest in technology and acquire hospital supplies more prudently and invest less on clinical research, continuing education and other discretionary activities (Department of Health, 2002b).

The PbR system assumes that cost differences between providers are exclusively efficiency related and therefore it penalises providers with costs above the average. Nevertheless, we argue that hospitals might face exogenous constraints, that are not efficiency related and that hinder hospitals from improving relative performance with respect to their peers.

The ability to improve relative performance will depend on the internal structure of different providers, with scale and scope of the services provided assuming an important role. For example, hospitals with a higher installed capacity and guaranteed levels of activity will, *ceteris paribus*, benefit from scale economies and will therefore be able to produce at lower costs. Also, hospitals with a higher installed capacity might offer a broader mix of services enabling them to benefit from scope economies. On the other hand, providers obliged to invest in services with a volatile demand and locked in their own capacity, might either supply care far below their capacity and incur into high costs for unused capacity, or produce unnecessary care by inducing demand.

The hypothesis we want to test is whether these structural differences between providers are mainly due to technological differences arising from ownership, management and financial structures: the lower the ability of a provider to reshape its internal structure and the range and amount of services to be

[1] HRGs are standard groupings of clinically similar treatments, which use similar levels of healthcare resources.

supplied, the higher the relevance of these structural differences.

These issues are particularly relevant for the analysis of the performance of FTs and non-FTs. Indeed, we would expect that, under new financial and managerial freedoms, FTs will invest mainly in services where they are able to extract economies of scale and scope, with a consequent cost advantage for FTs with respect to non-FTs that cannot benefit from such flexibility.

A way to verify the existence of such opportunistic behaviour of FTs is to estimate a cost function for each type of provider, FTs and non-FTs. Evaluating the behaviour of different typologies of providers separately might be particularly relevant for the design of their optimal reimbursement and might also have an impact on the future developments of PbR.

The literature on estimation of hospital scale and scope economies using cost functions is vaste (Bartlett and Le Grand, 1994; Carey, 1997; Dor and Farley, 1996; Fournier and Michell, 1992; Given, 1996; Hughes and McGuire, 2003; Preyra and Pink, 2006; Scuffham et al., 1996; Smet, 2002; Vita, 1990; Vitaliano, 1987; Wholey et al. 1996) but, to the best of our knowledge, only Bartlett and Le Grand (1994) focus on the English case. Using data from the 1990s for English hospitals the authors show that ward costs rose with increasing scale. Our analysis departs from theirs on scope, data and methodology.

The remaining of the paper is organised as follows. The next section describes the empirical methodology adopted in the analysis. Section 3 describes the data, section 4 the results and section 5 concludes.

2 Empirical methodology

We split our panel of 171 trusts into two groups (54 FTs and 117 non-FTs) in order to test whether there are any differences in the cost structure between FTs and non-FTs following the introduction of the Foundation policy which has made FTs more likely than non-FTs to exploit economies of scale and scope in the take-up of the whole reform agenda. We use panel data methods (Carey, 1997; Dor and Farley, 1996) rather than cross-section models (Cowing and Holtmann, 1983; Fournier and Mitchell, 1992; Given, 1996; Preyra and Pink, 2006; Vita, 1990; Vitaliano, 1987; Wholey et al., 1996) because panel data methods capture hospital-specific effects, such as quality of services, severity of illness and managerial ability, which are likely to be omitted by cross-section models that therefore result to be biased.

We therefore track all variables involved in the analysis over time for FTs before they actually became FTs and investigate how they compare to non-FTs over time. In doing so, however, we do not intend to evaluate the Foundation policy using a natural experiment that compares the change in costs for FTs before and after the policy intervention with the change in costs for non-FTs that are not undergoing the intervention, over the same period. Instead, we *assume* that FTs and non-FTs have different cost structures and thus we estimate two separate generalised multiproduct cost functions (Caves et al., 1980), one for FTs and the other for non-FTs. The advantage of the generalised multiproduct cost function with respect to a basic translog multiproduct cost function (Berndt and Christensen, 1973) is that it allows for the value zero in the permissible domain of outputs.

For both hospital types, the general form of the generalised multiproduct cost function is

$$\begin{aligned}\ln C_{it}^{v} \quad = \quad & \alpha_0 + \sum_{j=1}^{n} \alpha_j Q_{jit}^{(\pi)} + \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_{jk} Q_{jit}^{(\pi)} Q_{kit}^{(\pi)} + \sum_{l=1}^{p} \beta_l \ln P_{lit} + \\ & + \frac{1}{2} \sum_{l=1}^{p} \sum_{m=1}^{p} \beta_{lm} \ln P_{lit} \ln P_{mit} + \sum_{j=1}^{n} \sum_{l=1}^{p} \delta_{jl} Q_{jit}^{(\pi)} \ln P_{lit} + \gamma_1 \ln K_{it} + \frac{1}{2} \gamma_2 (\ln K_{it})^2 + \\ & + \sum_{j=1}^{n} \eta_j Q_{jit}^{(\pi)} \ln K_{it} + \sum_{l=1}^{p} \theta_l \ln P_{lit} \ln K_{it} + \sum_{r=1}^{s} \mu_r X_{rit} + \sum_{t=1}^{13} \nu_t D_t + u_i + e_{it} \qquad (1)\end{aligned}$$

in which C_{it}^{v} equals the total variable costs for trust i in year t where i refers either to 54 FTs or to 117 non-FTs and t covers 13 years from 1994/95 to 2006/07, $Q_{jit}^{(\pi)}$ is the Box-Cox transformed output j for trust i in year t, P_{lit} is the input price l for trust i in year t, K_{it} is the fixed capital for trust i in year t and X_{it} is a vector of r observable factors affecting the dependent variable for trust i in year t. In all models we also include 13 time dummies in order to capture unobserved heterogeneity due to unavailable variables and potential time trends present in the regressors: D_t is the year dummy for $t = 1$ (1994/95), $t = 2$ (1995/96), and so on. Finally, u_i is the trust-specific effect and e_{it} is the heteroskedastic component of the error. It is reasonable to think that the error structure of the model is heteroskedastic because "small trusts are more likely to have their average costs substantially affected by the lack of presence of extreme clinical cases" (Wholey et al., 1996, p.670). A more complete cost model would take into account also demand uncertainty: "The demand uncertainty represents the level of unexpected emergency arrivals, which is the difference between actual emergency demand and the forecast demand" (Hughes and McGuire, 2003, pp.1001-1002). However, we were not able to estimate this model because of lack of monthly data on emergency demand, necessary to estimate the unpredicted demand.

The cost function (1) must be non-decreasing, homogeneous of degree one and concave in input prices. In particular, linear homogeneity in input prices requires that, for $l, m = 1, ..., p$,

$$\sum_{l=1}^{p} \beta_l = 1, \; \sum_{l=1}^{p} \beta_{lm} = \sum_{m=1}^{p} \beta_{lm} = 0, \; \sum_{l=1}^{p} \delta_{jl} = 0, \; \sum_{l=1}^{p} \theta_l = 0$$

where the second constraint holds because of symmetry, $\beta_{lm} = \beta_{ml}$.

Symmetry also requires that $\alpha_{jk} = \alpha_{kj}$. Therefore, with 3 outputs ($n = 3$), 2 input prices ($p = 2$) and 4 observable factors ($r = 4$), equation (1) can be re-written as

$$\begin{aligned}\overline{C}_{it} &= \alpha_0 + \sum_{j=1}^{3}\alpha_j Q_{jit}^{(\pi)} + \frac{1}{2}\sum_{j=1}^{3}\sum_{k=1}^{3}\alpha_{jk} Q_{jit}^{(\pi)} Q_{kit}^{(\pi)} + \beta_1 \overline{P}_{it} + \beta_2 \overline{P}_{it}^2 + \sum_{j=1}^{3}\delta_j Q_{jit}^{(\pi)}\overline{P}_{it} + \\ &+ \gamma_1 \ln K_{it} + \frac{1}{2}\gamma_2 \left(\ln K_{it}\right)^2 + \sum_{j=1}^{3}\eta_j Q_{jit}^{(\pi)} \ln K_{it} + \theta \overline{P}_{it} \ln K_{it} + \sum_{r=1}^{4}\mu_r X_{rit} + \\ &+ \sum_{t=1}^{13}\nu_t D_t + u_i + e_{it} \qquad (2)\end{aligned}$$

where

$$\begin{aligned}\overline{C}_{it} &\equiv \ln C_{it}^{v} - \ln P_{2it} \\ \overline{P}_{it} &\equiv \ln P_{1it} - \ln P_{2it}\end{aligned}$$

We explore the robustness of our results comparing different estimation methods, different measures of the fixed input and different combinations of the control variables.

In order to choose the appropriate estimation technique, we perform the Hausman's specification test for both FTs and non-FTs. According to our results, we cannot reject the hypothesis that the trust-level effects are modelled by a random effects model for both FTs and non-FTs. This result may be due to the fact that, when the number of periods is much smaller than the number of individuals, the random effects model tends to perform better than the fixed effect one (Carey, 1997).

Once we have identified the appropriate estimation method, we can estimate two cost functions, one for FTs and the other for non-FTs. Estimating two separate cost functions has the advantage of estimating different values of the Box-Cox parameter π and therefore finding the most suitable cost structure for each type of hospital.

In Stata 10 (Stata, 2007) we run the random effects model using the GLS estimator (xtreg, re). The model allows the option of clustering on trusts and the calculation of robust standard errors.

The value of the parameter π is found from an iterative grid search. The criteria for selecting optimal values of π is minimising the sum of squared errors of the random effects model. The search ranges between -10 and 10, with an initial gradient of 0.2 and a 0.00001 gradient around the conversion point.

The least squares estimates and the estimated value of π constitute the least squares estimates of all the parameters involved in the empirical analysis. However, as the optimal value of π is an estimate, the least squares standard errors underestimate the correct asymptotic standard errors (Greene, 2003, p.174). To get the appropriate values of the standard errors, we therefore need the derivative of the right-hand side of (2) with respect to all parameters and π. We then use these derivatives to estimate the variance matrix of the estimated parameters. Further details on Box-Cox transformation and estimates of parameters can be found in Greene (2003), pp.173-175, and Cameron and Trivedi (2005), pp. 726-729 and 734-736. The Stata program developed for FTs and non-FTs is available on request.

The least squares estimates of the parameters are then used to calculate economies of scale and to test for the existence of economies of scope.

Once we have the estimated costs for FTs and non-FTs, we are able to check for the existence of economies of scale and scope and for weak cost complementarities, for each type of hospital.

3 Data description

Our data are annual and cover all acute and specialist trusts in England for a period of 13 years starting in 1994/95.

The dependent variable C^v is defined as the annual operating expenses, excluding land, buildings and dwellings.

Output is measured by the total number of inpatient spells (Q_1), the total number of outpatient attendances (Q_2), including first attendance and subsequent ones, and the total number of A&E attendances (Q_3), including first attendance and subsequent ones.

As clinical and general supply prices (including drugs) are uniformly set by the Purchasing and Supply Agency, an NHS regulatory agency (http://www.pasa.nhs.uk/pasaweb/productsandservices/pharmaceuticals/landingpage.htm), we only consider in the analysis variable input prices relating to medical and non-medical staff, including consultants, nurses, administrative staff, housekeeping, kitchen, laundry and maintenance staff. Input prices for medical and non-medical staff (labelled P_1 and P_2, respectively) are derived from annual data on total salaries and wages for each group measured as a proportion of the respective WTE staff employed (see also Scuffham et al., 1996).

The fixed input is measured by either the annual amount of fixed assets employed (assets, K), the most appropriate -but not always available- measure of capital input, or the average number of available beds (beds, K), in order to have results comparable with those of the existing literature on hospital costs.

Some control variables are included in the model to assess trust-specific characteristics such as the average length of stay, the patient load, the competition between trusts and the teaching status. The average length of stay (X_1) is calculated as the average number of days spent by each inpatient in hospital. This variable is included in the empirical specification to control for the outpatient variation among inpatients not captured by the number of inpatient spells. The trust caseload severity is measured by the casemix index (X_2). The casemix index is a weighted average of the procedures performed at each trust. The higher the index, the more ill, on average, the patients treated at the hospital. As the average length of stay, the casemix index is also used to control for outpatient variation among inpatients not captured by the number of inpatient spells. Competition between trusts is measured by the number of trusts within a 20km range of each trust (X_3). This is a plain measure of competition defined on the simple number of neighbour competitors and used to control for non-price competition (e.g., quality and/or demand competition), instead of price competition (e.g., technical efficiency). The teaching status is identified by a dummy variable (X_4) which takes value 1 if the trust is a teaching hospital and value 0 otherwise.

4 Empirical results

Table 1 shows the descriptive statistics for the dependent variable and the explanatory variables included in the model.

[Table 1 about here]

In Table 2 we present some t-tests on the differences between FTs and non-FTs. Given the degrees of freedom, $(N_{non-FTs} - 1) + (N_{FTs} - 1)$, the critical value of the 2-tailed t-statistic is equal to 1.960. If the t-statistic is lower than the critical value 1.960, we fail to reject the null and we conclude that the different between FTs and non-FTs is significantly not different from zero. Therefore, differences between FTs and non-FTs are significantly not different from zero regarding number of A&E attendances, number of available beds, the casemix index and the competition index. FTs and non-FTs are instead different in terms of number of inpatient spells (because FTs are paid according to the number of patients they treat), number of outpatient attendances (because FTs can control and organise their own elective activity), input prices (because FTs can freely set different salaries, within the national grade salary scheme), amount of assets employed (because FTs have more freedom in investment), average length of stay (because FTs may discharge patients quicker than non-FTs) and teaching status. In particular, we look at the test on the number of available beds, used as test on the size of FTs compared to non-FTs. We fail to reject the null and we conclude that the different in size between FTs and non-FTs is significantly not different from zero. In other words, differences in costs between FTs and non-FTs are not due to the fact that they operate at different points on the same cost function.

[Table 2 about here]

Table 3 presents the least square estimates of the parameters of the generalised translog multiproduct cost function.

[Table 3 about here]

The coefficients of the output parameters (α_1 to α_3) are always positive, if significant. Therefore, the cost function is non-decreasing in the levels of output. In particular, an increase of 1% in the inpatient activity implies that variable costs increase between 27% and 61% for FTs and between 18% and 38% for non-FTs. The impact of outpatient activity is not significant for FTs (and anyway small: limited to a decrease of $1 - 2\%$ of the variable costs increase). On the other hand, an increase of 1% in the outpatient activity of non-FTs implies an increase of $23\% - 27\%$ in their variable costs. The impact of A&E activity on total variable costs is not significant for both FTs and non-FTs (and anyway small: up to 5% for FTs and $1-2\%$ for non-FTs).

The normalisation of the right-hand side variables results in the first-order input price parameter β_1 having coefficient equal to the cost elasticity at the mean for the corresponding variable: β_1 is equal to the average cost share for the medical staff with respect to the non-medical staff.

The coefficients of the parameters β_1 and β_2 are always positive and negative, respectively, implying that the cost function is non-decreasing and concave in the input prices. Linear homogeneity is always verified as it has been imposed by the constraint $\sum_{l=1}^{p} \beta_l = 1$.

The coefficient of the parameter γ_1 is always positive and significant. This result implies that both FTs and non-FTs are not in the long-run equilibrium (see the Appendix A for a discussion on equilibrium and a test for it).

The coefficients of the parameters μ_1 to μ_4 are always positive. An increase in the average length of stay (μ_1) implies an increase in the costs associated to inpatient activity and therefore an increase in total variable costs. Total variable costs increase up to 19% for FTs and between 3% and 21% for non-FTs. An increase in the casemix index (μ_2) implies that patients treated are, on average, more severly ill and therefore that total variable costs increase, between 34% and 70% for FTs and around $12-13\%$ for non-FTs. An increase in competition (μ_3) implies that trusts have to spend more in order to attract more patients and consequently total variable costs increase, between up to 8% for FTs and around $2-3\%$ for non-FTs.[2] Finally, the presence of a teaching unit within the trust implies an increase in total variable costs by $41-48\%$ for FTs and by $32-38\%$ for non-FTs.

The coefficient of the intercept is ambiguous: the dependent variable will be positive for positive values of outputs and input prices, when α_0 is positive; negative for positive values of outputs and input prices, when α_0 is negative.[3]

Finally, it seems that, while the cost structure of FTs is well represented by a generalised cost function ($\pi > 1$ either if we use assets as a proxy for the fixed input or if we use beds), the cost structure of non-FTs is closer to a translog cost function as the parameter π is closer to zero (0.54808 when we use assets and 0.60999 when we use beds).

Using the parameters of the variable cost function (2), economies of scale are calculated applying the following formula

$$ES_R(K) = \frac{1-\gamma_1}{\sum_{j=1}^{3} \alpha_j} \tag{3}$$

[2] Note that, being health care free at the point of use in England, trusts compete for patients on the quality of the services provided: increased competition implies higher quality investment.

[3] Remember that our dependent variable is $\overline{C}_{it} \equiv \ln C_{it}^{v} - \ln P_{2it}$, which explains why costs may be negative.

(see Appendix A.1 for further details on derivation of the formula).

[Table 4 about here]

Table 4 presents estimates of the degree of scale economies at the mean outputs. Both FTs and non-FTs face economies of scale, with non-FTs having a greater advantage in increasing the provision of offered services when using fixed assets employed as a proxy for capital (1.22 versus 1.18) and FTs having a greater advantage in increasing the provision of offered services when using available beds as a proxy for capital (1.51 versus 1.30). These figures imply that increasing all of the average hospital's outputs by 1% induce a total variable cost increase between 0.66% and 0.85% for FTs and between 0.77% and 0.82% for non-FTs.

In order to check the robustness of this result, we conduct a test for all the three output measures (see Mester, 1987). Doubling or halving the levels of output activities one by one while keeping all the other variables constant, does not really make any difference in terms of economies of scale for either FTs or non-FTs, with two exceptions. Doubling the level of inpatient activities, both FTs and non-FTs experience diseconomies of scale, with an increase in total variable costs between 1.24% and 1.82% for FTs and between 1.10% and 1.29% for non-FTs; doubling the level of outpatient visits, only non-FTs experience diseconomies of scale, with an increase in total variable costs between 1.16% and 1.19%. These findings are consistent with the literature (Mester, 1987).

In the case of three outputs, global scope economies exist if

$$SCOPE = \frac{C^v(Q_1,0,0) + C^v(0,Q_2,0) + C^v(0,0,Q_3) - C^v(Q_1,Q_2,Q_3)}{C^v(Q_1,Q_2,Q_3)} > 0 \tag{4}$$

where $C^v(Q_1,0,0)$, $C^v(0,Q_2,0)$ and $C^v(0,0,Q_3)$ are the estimated variable costs associated to the exclusive production of Q_1, Q_2 and Q_3 respectively and $C^v(Q_1,Q_2,Q_3)$ is the estimated variable cost associated to the joint production of Q_1, Q_2 and Q_3.

Since FTs and non-FTs exhibit different global scope economies, we also look at product-specific economies of scope, in order to check choices relative to each service provided. Product-specific economies of scope exist if

$$SCOPE_j = \frac{C^v(Q_j) + C^v(Q_{-j}) - C^v(Q_1,Q_2,Q_3)}{C^v(Q_1,Q_2,Q_3)} > 0 \tag{5}$$

where $C^v(Q_j)$ is the cost of producing only output j and $C^v(Q_{-j})$ is the cost of producing all the other outputs different from output j.

[Table 5 about here]

Table 5 presents the estimates of global and product-specific scope economies. While FTs exhibit global diseconomies of scope, non-FTs exhibit strong global economies of scope. If we look at product-specific economies of scope, FTs always exhibit diseconomies of scope (with the exception of inpatient activity when the fixed input is measured by the annual amount of fixed assets employed), while non-FTs always exhibit economies of scope. These results are more informative when analysing weak cost complementarities between pairs of output. Weak cost complementarities allow to infer whether an increase in the level of one service has an impact on the marginal cost of providing another. Unlike scope economies these pairwise comparisons only give insights on how the cost functions behave on a local neighbourhood and therefore do not allow to infer on global cost savings (Fraquelli et al., 2004).

Weak cost complementarities exist if

$$\frac{\partial^2 \overline{C}}{\partial Q_j \partial Q_k} = \alpha_j \alpha_k + \alpha_{jk} \leq 0 \text{ and } \frac{\partial^2 \overline{C}}{\partial Q_j \partial K} = \eta_j \leq 0 \tag{6}$$

(see Appendix A.2 for more details on how these conditions are derived).
Since symmetry holds, $\alpha_{jk} = \alpha_{kj}$. Thus condition (6) implies that a trust may face a cost advantage if it decides to diversify the provision of services by carrying on service j joint with service k and service k joint with service j.

[Table 6 about here]

Looking at Table 6, FTs face a cost disadvantage if they decide to carry on inpatient and outpatient services (when using available beds as a proxy for the stock of capital), while non-FTs face a cost advantage if they decide to carry on A&E and outpatient services (independently of the variable used as a proxy for the stock of capital).

5 Discussion and conclusions

In the previous section we tested whether we could detect any differences between FTs and non-FTs in terms of cost saving arising from increasing the provision of offered services (economies of scale) or from diversifying the provision of services by carrying on various services (economies of scope).

In order to evaluate economies of scale and scope, we estimate two separate cost functions, one for each type of provider, using a generalised translog multiproduct cost function on a panel of English hospitals. Once we have the estimated costs for FTs and non-FTs, we are able to check for the existence of economies of scale and scope for each type of hospital.

While we cannot detect any difference between FTs and non-FTs in terms of economies of scale, we find different patterns between FTs and non-FTs in terms of economies of scope. Results show that FTs' cost function exhibits global diseconomies of scope, while non-FTs can profit from global economies of scope.

These results have significant policy implications, especially since both the Foundation policy and the PbR push FTs against diversification and therefore in favour of separate production of alternative services (specialisation). Given the extended new freedoms, our results show that FTs could achieve efficiency gains if they offered different services on different sites.

On the other side of the spectrum, the fact that non-FTs are locked in their own capacity and constrained to provide a broad range of services, does not seem to be problematic to non-FTs as their technology exhibits efficiency gains from joint production of outpatient, inpatient and A&E services. However, the extent to which they can effectively profit from this technology feature, will depend on whether local demand characteristics, as well as PCT commissioning practices, allow them to deliver an efficient mix of inpatient, outpatient and A&E services.

Our results on cost complementarities shed some light on the optimal service portfolio for FTs and non-FTs. FTs face a cost disadvantage if they decide to carry on inpatient and outpatient services: the joint production of inpatient and outpatient services have negative effects on cost savings. On the other hand, non-FTs face a cost advantage if they decide to carry on A&E and outpatient services: any potential cost savings from joint production can be realised only through a limited range of hospital services. However, these pairwise comparisons only give insights on how the cost functions behave on a local neighbourhood and therefore do not allow to infer on global cost savings.

Still, on the grounds of efficiency, concentrating the delivery of services will lead to cost savings, with larger hospitals achieving economies of scale. Results also show that there are strong arguments for creating larger hospital units to facilitate links between departments, to foster multidisciplinary teams and to ensure optimal use of expensive resources. These can be achieved either by investing on additional capacity for each provider or through hospital mergers, being the former more likely to lead to less caveats.

Indeed the literature on hospital scale and scope economies shows that mergers do not necessarily undertake to concentrate facilities in order to attain reduced total costs. This might be due to the possibility of increased management costs in the long-run and to the difficulty of integrating staff and systems.

Although efficiency gains and better competition may be achieved through specialisation and single unit capacity growing, these savings can come at the cost of loosing the "localness" of health care provision and therefore the direct contact with the community. Indeed, given the geographical distribution of providers, hospital mergers are likely to reduce access to health care, increasing inequalities. This is especially the case of providers in rural areas. In this context, it is crucial that PCTs balance the needs of both demand and supply of secondary care, by ensuring their local population access to key services whilst encouraging hospitals' efficiency gains.

Given that, under a fixed price regime (such as the PbR), the revenue stream depends on relative efficiency performance and therefore only the most efficient trusts will flourish, in the long-run such a payment system might discriminate in favour of those providers that are in a better position to exploit scale and scope economies by reorganising the offer of their services (such as FTs), or might even discriminate against providers that either face capacity constraints or are limited on the mix of services commissioned by the purchasers (such as non-FTs).

Our analysis has therefore implications for the design of the tariff payable to different types of providers in order, not only to better achieve PbR efficiency targets, but also to avoid discrimination and inequitable treatment between providers. If these differences across providers are indeed exogenous and not efficiency related, then a unique price system applied to different typologies of providers -as PbR is- will discriminate against non-FTs, widening the performance gap between FTs and non-FTs. Indeed, both the Foundation status and the PbR policies have been frequently seen as policies that encourage a two tier system, locking up non-FTs in a sort of "poverty trap" (Marini et al., 2008).

In conclusion, a combination of Foundation policy and PbR might reinforce the gap between FTs and non-FTs and could be particularly lethal for non-FTs, leading to a two-tier systems with FTs outperforming non-FTs.

6 Acknowledgements

We would like to thank Pedro Pita Barros, Hugh Gravelle, Arne Hole, Andrew Jones, Guido Pellegrini, Nigel Rice and Giacomo Rodano for their useful comments on preliminary versions of this paper and for discussions on our STATA routine. The usual disclaimer applies.

A Appendix

A.1 Economies of scale

The use of a multiproduct cost function gives rise to scale economies. Ray scale economies are defined as the proportional increase in total costs that would result from a proportional increase in all outputs.

If it is believed that hospitals operate in equilibrium, it is appropriate to estimate a long-run cost function

$$C = C\left(Q, P\right) \tag{A.1}$$

Baumol et al. (1982) define the degree of ray economies of scale at output vector Q in a multiproduct firm as

$$ES_R = \frac{1}{\sum_{j=1}^{n} \frac{\partial C(Q,P)}{\partial Q_j} \frac{Q_j}{C(Q,P)}} \tag{A.2}$$

under the assumption that all of the firm's inputs can be freely varied to minimise total costs, i.e. under the assumption that the firm is operating in equilibrium. Ray economies of scale measure overall economies of scale and assume that the composition of the output bundle remains fixed while the size of the composite output bundle can vary. The production is characterised by constant returns to scale if ES_R equals one. Ray economies or diseconomies of scale are said to exist if ES_R is greater than or less than one, respectively. ES_R can be interpreted as "the elasticity of the output of the relevant composite commodity with respect to the cost needed to produce it" (Baumol et al., 1982, p. 51).

However, if not all inputs can be adjusted quickly in response to changing output levels or changing input prices, hospitals will only employ optimal quantities of the adjustable variable inputs (e.g., labour), given the non-optimal levels of the fixed inputs (e.g., capital stock). In this case, it is more appropriate to estimate a short-run cost function

$$C^s = C^v\left(Q, P_L, K\right) + P_K K \tag{A.3}$$

in which $C^v\left(Q, P_L, K\right)$ are the variable costs and $P_K K$ are the (short-run) fixed costs, P_L is the price vector of variable inputs and P_K is the price of the fixed input K.[4]

[4] A well known empirical test for short- versus long-run equilibrium consists in estimating the variable costs $C^v\left(Q, P_L, K\right)$ and verifying that the envelope condition, $\partial C^v\left(Q, P_L, K\right)/\partial K + P_K = 0$, holds (Vita, 1990; Cowing & Holtmann, 1983). If the envelope condition holds, hospitals operate in equilibrium. Solving the envelope condition for K^*, the cost-minimising level of K, and substituting K^* into the short-run cost function (A.3) yields the long-run cost function: $C^v\left(Q, P_L, K^*\right) + P_K K^* = C\left(Q, P\right)$. On the other hand, if the envelope condition does not hold, hospitals only employ optimal quantities of the adjustable variable inputs and the appropriate cost function to be estimated is the (A.3). Further details and a survey of this test can also be found in Smet (2002).

When the structure of production is analysed via the estimation of a variable cost function, two methods can used to infer the degree of ray economies of scale. One consists of using the estimated parameters and the price of the fixed input, applying the envelope condition, $\partial C^v(Q, P_L, K)/\partial K + P_K = 0$, to solve for the optimal level of the fixed factor, and deriving the long-run cost function. Long-run scale economies can then be derived by evaluating the long-run cost function (Braeutigam and Daughety, 1983; Vita, 1990). Using the parameters from the variable cost function, the formula for ray economies of scale becomes

$$ES_R(K^*) = \frac{1 - \frac{\partial C^v(Q,P_L,K)}{\partial K} \frac{K}{C^v(Q,P_L,K)}\Big|_{K=K^*}}{\sum_{j=1}^{n} \frac{\partial C^v(Q,P_L,K)}{\partial Q_j} \frac{Q_j}{C^v(Q,P_L,K)}\Big|_{K=K^*}} \tag{A.4}$$

Evaluated at the optimal levels of the fixed inputs, K^*, equation (A.4) is equivalent to the Baumol et al. definition (A.2). At $K = K^*$, $\frac{\partial C^v(Q,P_L,K^*)}{\partial K^*} = 0$ (because the fixed input is set at its cost-minimising optimal level) and the (A.4) coincides with the (A.2).

An alternative approach proposed by Caves et al. (1981) calculates a measure of scale economies using the parameters of the variable cost function, without reference to the prices of the fixed factors. Instead of the optimal value of the fixed factor, the estimated value is used in the calculation of the ray economies of scale. The formula for ray economies of scale becomes

$$ES_R(K) = \frac{1 - \frac{\partial C^v(Q,P_L,K)}{\partial K} \frac{K}{C^v(Q,P_L,K)}}{\sum_{j=1}^{n} \frac{\partial C^v(Q,P_L,K)}{\partial Q_j} \frac{Q_j}{C^v(Q,P_L,K)}} \tag{A.5}$$

When applying the method proposed by Caves et al. (1981), it should be kept in mind that it will generally not yield the same estimate of scale economies as would be obtained by enveloping the variable cost function (see Braeutigam and Daughety, 1983). The latter method evaluates scale economies along a ray from the origin that passes through the actual point of operation observed in the sample. Since the motivation for adopting the variable cost framework is the belief that the firms being studied are not necessarily operating on their efficient expansion path, scale economy estimates computed using the second method would rarely be expected to coincide with those derived using the first.

As most studies argue that there is no evidence that hospitals operate in their long-run equilibrium, to avoid estimation bias, we estimate a variable cost function in which hospitals minimise variable costs, given the stock of capital and exogenous input prices and output quantities, then check the envelope condition and proceed accordingly, as explained previously.

If a variable cost function is estimated with the generalised multiproduct cost function,

$$\ln C^v_{it} = \alpha_0 + \sum_{j=1}^{n} \alpha_j Q^{(\pi)}_{jit} + \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_{jk} Q^{(\pi)}_{jit} Q^{(\pi)}_{kit} + \sum_{l=1}^{p} \beta_l \ln P_{lit} + \\ + \frac{1}{2} \sum_{l=1}^{p} \sum_{m=1}^{p} \beta_{lm} \ln P_{lit} \ln P_{mit} + \sum_{j=1}^{n} \sum_{l=1}^{p} \delta_{jl} Q^{(\pi)}_{jit} \ln P_{lit} + \\ + \gamma_1 \ln K_{it} + \frac{1}{2} \gamma_2 (\ln K_{it})^2 + \sum_{j=1}^{n} \eta_j Q^{(\pi)}_{jit} \ln K_{it} + \sum_{l=1}^{p} \theta_l \ln P_{lit} \ln K_{it} \tag{A.6}$$

the elasticity of cost with respect to any individual output (e.g., Q_j) is given by

$$\begin{aligned} \rho_{C,Q_j} &= \frac{\partial C^v (Q, P_L, K)}{\partial Q_j} \frac{Q_j}{C^v (Q, P_L, K)} = \exp (RHS) \times \\ &\times \left[\alpha_j Q_j^{\pi - 1} + \frac{1}{2} \sum_{k=1}^{n} \alpha_{jk} Q_j^{\pi - 1} Q_k^{(\pi)} + \sum_{l=1}^{p} \delta_{jl} Q_j^{\pi - 1} \ln P_l + \eta_j Q_j^{\pi - 1} \ln K \right] \times \frac{Q_j}{C^v (Q, P_L, K)} \\ &= C^v (Q, P_L, K) \times \left[\alpha_j + \frac{1}{2} \sum_{k=1}^{n} \alpha_{jk} Q_k^{(\pi)} + \sum_{l=1}^{p} \delta_{jl} \ln P_l + \eta_j \ln K \right] \times Q_j^{\pi - 1} \times \frac{Q_j}{C^v (Q, P_L, K)} \\ &= \left[\alpha_j + \frac{1}{2} \sum_{k=1}^{n} \alpha_{jk} Q_k^{(\pi)} + \sum_{l=1}^{p} \delta_{jl} \ln P_l + \eta_j \ln K \right] \times Q_j^{\pi} = \alpha_j \end{aligned}$$

in which RHS identifies the right-hand side of equation (A.6). At the sample mean, the normalisation of the right-hand-side variables of equation (A.6) causes the cost elasticity for each output Q_j to equal α_j, the first-order output parameter from the estimated cost function.

On the other hand, the elasticity of cost with respect to the fixed input K is given by

$$\begin{aligned} \rho_{C,K} &= \frac{\partial C^v (Q, P_L, K)}{\partial K} \frac{K}{C^v (Q, P_L, K)} = \exp (RHS) \\ &\times \left[\gamma_1 \frac{1}{K} + \gamma_2 \frac{\ln K}{K} + \sum_{j=1}^{n} \eta_j \frac{Q_j^{(\pi)}}{K} + \sum_{l=1}^{p} \theta_l \frac{\ln P_l}{K} \right] \times \frac{K}{C^v (Q, P_L, K)} \\ &= C^v (Q, P_L, K) \times \left[\gamma_1 + \gamma_2 \ln K + \sum_{j=1}^{n} \eta_j Q_j^{(\pi)} + \sum_{l=1}^{p} \theta_l \ln P_l \right] \times \frac{1}{K} \times \frac{K}{C^v (Q, P_L, K)} \\ &= \left[\gamma_1 + \gamma_2 \ln K + \sum_{j=1}^{n} \eta_j Q_j^{(\pi)} + \sum_{l=1}^{p} \theta_l \ln P_l \right] = \gamma_1 \end{aligned}$$

At the sample mean, the normalisation of the right-hand-side variables of equation (A.6) causes the cost elasticity for the fixed input K to equal γ_1, the first-order output parameter from the estimated cost function.

Therefore, when the variable cost function is estimated with a generalised multiproduct cost function (A.6), equations (A.4) and (A.5) become

$$ES_R (K^*) = \frac{1}{\sum_{j=1}^{n} \alpha_j}$$

and

$$ES_R(K) = \frac{1-\gamma_1}{\sum_{j=1}^{n} \alpha_j}$$

A.2 Weak cost complementarities

Weak cost complementarities exist if

$$\frac{\partial^2 C}{\partial Q_j \partial Q_k} \leq 0 \tag{A.7}$$

where C is the long-run cost function (A.1). This condition can be related to the variable cost function according to the following decomposition

$$\frac{\partial^2 C}{\partial Q_j \partial Q_k} = \frac{\partial^2 C^v}{\partial Q_j \partial Q_k} + \frac{\partial^2 C^v}{\partial Q_j \partial K} \frac{\partial K^*}{\partial Q_k} \leq 0$$

If we assume $\partial K^*/\partial Q_k > 0$, then a sufficient condition for the validity of the (B.3) is

$$\frac{\partial^2 C^v}{\partial Q_j \partial Q_k} \leq 0 \text{ and } \frac{\partial^2 C^v}{\partial Q_j \partial K} \leq 0$$

a necessary condition is

$$\frac{\partial^2 C^v}{\partial Q_j \partial Q_k} \leq 0 \text{ or } \frac{\partial^2 C^v}{\partial Q_j \partial K} \leq 0$$

When the variable cost function is estimated with equation (A.6), at the sample mean, the above conditions reduce to

$$\begin{aligned}
\frac{\partial^2 C^v}{\partial Q_j \partial Q_k} &= \exp(RHS) \left\{ \begin{array}{c} \left[\alpha_k Q_k^{\pi-1} + \frac{1}{2}\sum_{j=1}^{n} \alpha_{jk} Q_j^{(\pi)} Q_k^{\pi-1} + \sum_{l=1}^{p} \delta_{kl} Q_k^{\pi-1} \ln P_l + \eta_k Q_k^{\pi-1} \ln K\right] \times \\ \times \left[\alpha_j Q_j^{\pi-1} + \frac{1}{2}\sum_{k=1}^{n} \alpha_{jk} Q_j^{\pi-1} Q_k^{(\pi)} + \sum_{l=1}^{p} \delta_{jl} Q_j^{\pi-1} \ln P_l + \eta_j Q_j^{\pi-1} \ln K\right] + \\ + \left[\alpha_{jk} Q_j^{\pi-1} Q_k^{\pi-1}\right] \end{array} \right\} = \\
&= C^v(Q, P_L, K) Q_j^{\pi-1} Q_k^{\pi-1} \left\{ \begin{array}{c} \left[\alpha_k + \frac{1}{2}\sum_{j=1}^{n} \alpha_{jk} Q_j^{(\pi)} + \sum_{l=1}^{p} \delta_{kl} \ln P_l + \eta_k \ln K\right] \times \\ \times \left[\alpha_j + \frac{1}{2}\sum_{k=1}^{n} \alpha_{jk} Q_k^{(\pi)} + \sum_{l=1}^{p} \delta_{jl} \ln P_l + \eta_j \ln K\right] + \alpha_{jk} \end{array} \right\} = \\
&= C^v(Q, P_L, K) Q_j^{\pi-1} Q_k^{\pi-1} \left(\alpha_j \alpha_k + \alpha_{jk}\right) = \alpha_j \alpha_k + \alpha_{jk}
\end{aligned}$$

and

$$\frac{\partial^2 C^v}{\partial Q_j \partial K} = \exp(RHS) \left(\eta_j Q_j^{\pi-1} \frac{1}{K}\right) = C^v(Q, P_L, K) Q_j^{\pi-1} \frac{1}{K} \eta_j = \eta_j$$

since $\frac{\partial \exp(RHS)}{\partial K^*} = \frac{\partial C^v(Q, P_L, K^*)}{\partial K^*} = 0$.

Table 1. Descriptive statistics and variable definitions, pooled data 1994/95-2006/07

Variable name	Definition (Source)	Non-FTs					FTs				
		N	Mean	S.D.	Min	Max	N	Mean	S.D.	Min	Max
Dependent variable											
C^v	Annual operating expenses (000000) (CIPFA)	2057	112	87.30	3.32	7.36	627	133	107	6.82	664
Outputs											
Q_1	Total number of inpatient spells (HES)	2086	50720.39	32317.79	222	232033	632	56214.18	33703.08	1040	188551
Q_2	Total number of outpatient attendances (HES)	2048	200713.20	151212.30	0	3667170	632	223657.80	147291.50	0	1333426
Q_3	Total number of emergency attendances (HES)	2045	64143.96	44616.53	0	426295	626	63338.30	41984.36	0	197398
Input prices											
P_1	Medical staff price (derived from CIPFA and IC)	1653	70021.52	25198.79	9826.47	683548.90	547	65560.61	22818.72	12437.64	240313.90
P_2	Non-medical staff price (CIPFA and Information Centre)	987	27896.21	5592.16	16913.59	49775.40	414	24215.23	7925.98	4325.25	45281.29
Capital input											
K (assets)	Annual amount of fixed assets employed (000000) (CIPFA)	2031	92.20	67	5.39	500	626	108	95.90	5.76	756
K (beds)	Average number of available beds (HAS)	2118	673.04	364.50	4	2838.14	634	696.02	402.88	13.54	2058.78
Other explanatory variables											
X_1	Average length of stay (HES)	1862	6.43	7.58	1.10	136.81	589	5.33	2.54	0.61	23.10
X_2	HRG casemix index based on Reference Costs (NHSIA)	2074	100.87	24.37	68.54	306.13	622	100.86	22.75	67.11	256.02
X_3	Number of trusts within a 20km range of each trust (derived)	1620	9.98	13.60	0	62	584	10.55	14.03	0	59
X_4	Dummy variable =1 if Trust is teaching hospital; =0 otherwise (derived)	2131	0.10	0.30	0	1	637	0.19	0.40	0	1

CIPFA - Chartered Institute of Public Finance and Accountancy, http://www.cipfastats.net/health/default.asp
HES - Hospital Episodes Statistics, http://www.hesonline.nhs.uk
IC - Information Centre, http://www.ic.nhs.uk
HAS - Hospital Activity Statistics, http://www.performance.doh.gov.uk/hospital activity
NHSIA - NHS Information Authority, http://www.connectingforhealth.nhs.uk and http://www.ic.nhs.uk

Table 2. t-tests on the differences between FTs and non-FTs

	$N_{non-FTs}$ N_{FTs}	$\mu_{non-FTs}$ μ_{FTs}	$\mu_{non-FTs} - \mu_{FTs}$	$\sigma^2_{non-FTs}$ σ^2_{FTs}	$\sigma^2_{non-FTs} / N_{non-FTs}$ σ^2_{FTs} / N_{FTs}	$\sqrt{\frac{\sigma^2_{non-FTs}}{N_{non-FTs}} + \frac{\sigma^2_{FTs}}{N_{FTs}}}$	$t = \frac{\lvert \mu_{non-FTs} - \mu_{FTs} \rvert}{\sqrt{\frac{\sigma^2_{non-FTs}}{N_{non-FTs}} + \frac{\sigma^2_{FTs}}{N_{FTs}}}}$	d.f.
C								
non-FTs	2057	112	21	7621.29	3.71	4.69	4.48	2682
FTs	627	133		11449	18.26			
Q_1								
non-FTs	2086	50720.39	5493.79	1044439550	500690.10	1515.91	3.62	2716
FTs	632	56214.18		1135897601	1797306.33			
Q_2								
non-FTs	2048	200713.20	22944.6	22865159671	11164628.75	6744.76	3.40	2678
FTs	632	223657.80		21694785972	34327192.99			
Q_3								
non-FTs	2045	64143.96	805.66	1990634749	973415.53	1946.59	0.41	2669
FTs	626	63338.30		1762686485	2815793.11			
P_1								
non-FTs	1653	70021.52	4460.91	634979017	384137.34	1155.87	3.86	2198
FTs	547	65560.61		520693982	951908.56			
P_2								
non-FTs	987	27896.21	3680.98	31272253	31684.15	428.28	8.59	1399
FTs	414	24215.23		62821159	151741.93			
K (assets)								
non-FTs	2031	92.20	15.8	4489	2.21	4.11	3.84	2655
FTs	626	108		9197	14.69			
K (beds)								
non-FTs	2118	673.04	22.9756	132857.99	62.73	17.85	1.29	2750
FTs	634	696.02		162314	256.02			
X_1								
non-FTs	1862	6.43	1.1	57.39	0.03	0.20	5.38	2449
FTs	589	5.33		6	0.01			
X_2								
non-FTs	2074	100.87	0.01	593.8000	0.29	1.06	0.01	2694
FTs	622	100.86		518.0000	0.83			
X_3								
non-FTs	1620	9.98	0.57	184.87	0.11	0.67	0.85	2202
FTs	584	10.55		197	0.34			
X_4								
non-FTs	2131	0.1	0.09	0.0900	0.00	0.02	5.25	2766
FTs	637	0.19		0.1600	0.00			

Table 3. Estimates for the random effects model

Variable name	Parameter name	K (fixed assets employed)		K (available beds)	
		FTs	non-FTs	FTs	non-FTs
Q_1	α_1	0.612***	0.377***	0.268**	0.180***
		(0.142)	(0.0264)	(0.106)	(0.0276)
Q_2	α_2	-0.0109	0.272***	-0.0162	0.228***
		(0.103)	(0.0260)	(0.0508)	(0.0275)
Q_3	α_3	-0.0669	0.0157	0.0537	0.0129
		(0.100)	(0.0148)	(0.0647)	(0.0155)
Q_1^2	α_{11}	-0.385**	0.243***	-0.0252	0.185***
		(0.189)	(0.0498)	(0.01783)	(0.0618)
Q_2^2	α_{22}	-0.00366	-0.109***	-0.000146	-0.0878***
		(0.014)	(0.0205)	(0.00053)	(0.0191)
Q_3^2	α_{33}	0.0177	-0.00223	0.0274	0.00994
		(0.178)	(0.0237)	(0.0287)	(0.0207)
Q_1Q_2	α_{12}	-0.0320	-0.0779**	0.00328	-0.108**
		(0.036)	(0.0309)	(0.00300)	(0.0518)
Q_1Q_3	α_{13}	-0.0183	-0.0705**	-0.00193	-0.0750**
		(0.116)	(0.0327)	(0.0142)	(0.0353)
Q_2Q_3	α_{23}	0.121	0.0229	0.0123	0.0185
		(0.080)	(0.0323)	(0.0113)	(0.0329)
$Q_1\overline{P}$	δ_1	0.939**	0.120	0.399**	0.226**
		(0.431)	(0.086)	(0.162)	(0.104)
$Q_2\overline{P}$	δ_2	-0.346	-0.170*	-0.176**	-0.0858
		(0.300)	(0.092)	(0.0806)	(0.0912)
$Q_3\overline{P}$	δ_3	-0.0789	-0.0135	-0.309*	0.0342
		(0.311)	(0.0575)	(0.162)	(0.0584)
$Q_1\ln K$	η_1	0.464**	-0.161***	-0.00657	-0.0780
		(0.192)	(0.0358)	(0.093)	(0.092)
$Q_2\ln K$	η_2	0.00167	0.00367	-0.0608	0.0548
		(0.097)	(0.0338)	(0.0570)	(0.0727)
$Q_3\ln K$	η_3	-0.261**	0.0603**	-0.290**	0.0482
		(0.129)	(0.0246)	(0.126)	(0.0526)
$\overline{P}$	β_1	0.303	0.460***	0.319	0.430***
		(0.193)	(0.0420)	(0.211)	(0.0440)
$\overline{P}^2$	β_2	-1.719***	-0.702***	-1.467**	-0.628***
		(0.635)	(0.151)	(0.627)	(0.147)
$\ln K$	γ_1	0.369***	0.190***	0.539***	0.453***
		(0.088)	(0.0199)	(0.135)	(0.0330)
$\ln K^2$	γ_2	-0.0440	0.104***	0.181**	0.0733
		(0.161)	(0.0258)	(0.080)	(0.138)
$\overline{P}\ln K$	θ	-0.140	0.0314	0.119	-0.273**
		(0.223)	(0.0522)	(0.175)	(0.119)
X_1	μ_1	0.190*	0.210***	0.00228	0.0257
		(0.103)	(0.0214)	(0.1250)	(0.0213)
X_2	μ_2	0.344	0.126**	0.702**	0.117*
		(0.332)	(0.0610)	(0.357)	(0.0611)
X_3	μ_3	0.000593	0.0179**	0.0834**	0.0255***
		(0.0370)	(0.0089)	(0.0391)	(0.0090)
X_4	μ_4	0.413***	0.320***	0.483***	0.381***
		(0.094)	(0.0442)	(0.109)	(0.0427)

Intercept	α_0	0.0551	-0.207***	-0.0184	0.0732
		(0.335)	(0.0688)	(0.370)	(0.0695)
Box-Cox parameter	π	2.0387321***	0.54807633***	3.814533***	0.60999477***
		(0.319)	(0.101)	(0.419)	(0.110)
σ_u		0.094	0.112	0.142	0.122
σ_e		0.334	0.081	0.330	0.079
ρ		0.074	0.657	0.155	0.707
R-squared		0.834	0.930	0.825	0.929
Observations		336	801	336	802
Number of hospitals		47	126	47	126

Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.
Time dummies not reported.

Table 4. Ray economies of scale

Parameter name	K (fixed assets employed)		K (available beds)	
	FTs	non-FTs	FTs	non-FTs
α_1	0.612	0.377	0.268	0.180
α_2	-0.0109	0.272	-0.0162	0.228
α_3	-0.0669	0.0157	0.0537	0.0129
γ_1	0.369	0.190	0.539	0.453
$ES_R(K)$	1.18121	1.21859	1.50900	1.29960

Table 5. Global and product-specific economies of scope

	K (fixed assets employed)		K (available beds)	
	FTs	non-FTs	FTs	non-FTs
Global	-0.572	1.439	-0.310	1.724
Q_1	0.081	0.821	-0.189	0.989
Q_2	-0.379	0.726	-0.627	0.891
Q_3	-0.592	0.689	-0.071	0.814

Table 6. Test for Weak Cost Complementarities

Parameter name	K (fixed assets employed)		K (available beds)	
	FTs	non-FTs	FTs	non-FTs
$\alpha_1\alpha_2+\alpha_{12}$	-0.03867	0.02464	-0.00106	-0.06696
$\alpha_1\alpha_3+\alpha_{13}$	-0.05924	-0.06458	0.01246	-0.07268
$\alpha_2\alpha_3+\alpha_{23}$	0.12173	0.02717	0.01143	0.02144
η_1	0.464	-0.161	-0.00657	-0.078
η_2	0.00167	0.00367	-0.0608	0.0548
η_3	-0.261	0.0603	-0.290	0.0482

References

Bartlett W, Le Grand J. Costs and trusts (Studies in decentralization & quasi-markets). Policy Press; 1994.

Baumol WJ, Panzar JC, Willig RD. Contestable markets and the theory of industry structure. Harcourt Brace Jovanovich: New York; 1982.

Berndt ER, Christensen LR. The translog function and the substitution of equipment, structures, and labor in U.S. manufacturing 1929-68. Journal of Econometrics 1973; 1; 81-113.

Box GEP, Cox DR. An analysis of transformations. Journal of the Royal Statistical Society, Series B 1964; 26; 211-246.

Braeutigam R, Daughety A. On the estimation of returns to scale using variable cost functions. Economics Letters 1983; 11; 25-31.

Cameron AC, Trivedi PK. Microeconometrics. Methods and applications. Cambridge University Press: New York; 2005.

Carey K. A panel data design for estimation of hospital cost functions. Review of Economics and Statistics 1997; 79; 443-453.

Caves D, Christensen LR, Tretheway MW. Flexible cost functions for multiproduct firms. Review of Economics and Statistics 1980; 62; 477-481.

Caves D, Christensen LR, Swanson JA. Productivity growth, scale economies, and capacity of utilization in U.S. railroads, 1955-74. American Economic Review 1981; 71; 994-1002.

Cowing TG, Holtmann AG. Multiproduct short-run hospital cost functions: Empirical evidence and policy implications from cross-section data. Southern Economic Journal 1983; 49; 637-653.

Department of Health. The NHS plan: A plan for investment, a plan for reform. Department of Health: London; 2000. Available at http://www.dh.gov.uk /en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH _4002960?IdcService=GET_FILE&dID=23612&Rendition=Web

Department of Health. A guide to NHS Foundation Trusts. Department of Health: London; 2002a. Also available at http://www.dh.gov.uk/assetRoot/04 /06/04/80/04060480.pdf

Department of Health. Reforming NHS Financial Flows: Introducing Payment by Results. Department of Health: London; 2002b. Also available at http://www.dh.gov.uk/assetRoot/04/06/04/76/04060476.pdf

Department of Health. Health Reform in England: Update and Next Steps. Department of Health: London; 2005. Available at http://www.dh.gov.uk/en/ Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4 124723

Dor A, Farley DE. Payment source and the cost of hospital care: Evidence from a multiproduct cost function with multiple payers. Journal of Health Economics 1996; 15; 1-21.

Eaton L. Foundation Trusts argue they really do make a difference. British Medical Journal 2005; 331; 864.

Fournier GM, Mitchell JM. Hospital costs and competition for services: A multiproduct analysis. Review of Economics and Statistics 1992; 74; 627-634.

Fraquelli G, Piacenza M, Vannoni D. Scope and scale economies in multi-utilities: evidence from gas, water and electricity combinations. Applied Economics 2004; 36; 2045-2057.

Given RS. Economies of scale and scope as an explanation of merger and output diversification activities in the health maintenance organization industry. Journal of Health Economics 1996; 15; 685-713.

Greene WH. Econometric analysis (fifth edition). Prentice Hall International Inc., Upper Saddle River: New Jersey; 2003.

Health and Social Care Act. The Health and Social Care (Community Health and Standards) Act; 2003. Available at http://www.opsi.gov.uk/acts/acts2003 /20030043.htm

Healthcare Commission. The Healthcare Commission's Review of NHS Foundation Trusts. Healthcare Commission: London; 2005. Available at http:// www.healthcarecommission.org.uk/ db/ documents/04018494.pdf.

Hughes D, McGuire A. Stochastic demand, production responses and hospital costs. Journal of Health Economics 2003; 22; 999-1010.

Mester LJ. A Multiproduct cost study of savings and loans. Journal of Finance 2003; 42; 423-445.

Preyra C, Pink G. Scale and scope efficiencies through hospital consolidations. Journal of Health Economics 2006; 25; 1049-1068.

Scuffham PA, Devlin NJ, Jaforullah M. The structure of costs and production in New Zealand public hospitals: an application of the transcendental logarithmic variable cost function. Applied Economics 1996; 28; 75-85.

Smet M. Cost characteristics of hospitals. Social Science & Medicine 2002; 55; 895-906.

Stata. Stata Reference Manual Release 10. Stata Press: College Station, Texas; 2007.

Vita MG. Exploring hospital production relationships with flexible functional forms. Journal of Health Economics 1990; 9; 1-21.

Vitaliano DF. On the estimation of hospital cost functions. Journal of Health Economics 1987; 6; 305-318.

Wanless D. Securing our future health: Taking a long-term view. Final report. HM Treasury: London; 2002. Available at http://www.hm-treasury.gov. uk/wanless

Wholey D, Feldman R, Christianson JB, Engber J. Scale and scope economies among health maintenance organizations. Journal of Health Economics 1996; 15; 657-684.