ECONOMIC DISCUSSION PAPERS

Efficiency Series Paper 04/2009

Factors determining the response of hospitals to demand uncertainty

Ana Rodríguez-Álvarez, David Roibás, Alan Wall

Departamento de Economía

Universidad de Oviedo

Available online at: www.uniovi.es/economia/edp.htm

Factors determining the response of hospitals to demand uncertainty*

Ana Rodríguez Álvarez, David Roibás Alonso and Alan Wall

University of Oviedo, Spain.

Abstract
A feature of hospitals is that they face uncertain demand for the services they offer. To cover fluctuations in demand they need to maintain reserve capacity in the form of beds, equipment, personnel etc in order to minimize the probability of excess queuing or turning away patients, creating a trade-off between reserve capacity and economic costs. Using a simple theoretical framework we show how the reserve capacity established can depend on institutional characteristics that can affect the objective of the hospital. In particular, we show that private and public hospitals can provide different levels of reserve capacity. In an empirical application using a panel data set of Spanish hospitals over the period 1996-2006 we model reserve capacity using a distance frontier approach. Our results show that private hospitals do not generally contract as much reserve capacity in response to demand uncertainty, although for some services they contract more capacity than public hospitals.

Keywords: Hospitals, demand uncertainty, technical efficiency, distance function.

JEL Classification: D24, D81, I10.

*Ana Rodriguez acknowledges financial support from the project "MICINN-08-ECO2008-03468/ECON"

1. INTRODUCTION

A salient feature of hospitals is that they face uncertain demand for the services they offer. In order to be able to cover fluctuations in demand they will need to maintain reserve capacity in the form of beds, equipment, personnel etc in order to minimize the probability of excess queuing or turning away patients. As hospitals contract to be able to meet uncertain demand in all but exceptional circumstances, they can be considered as producers of a given provision-of-service probability. As realized demand will therefore generally be less than service capacity, this implies that hospitals operate with some excess capacity.

The influence of demand uncertainty on excess capacity and thus on hospital costs has received a good deal of attention in the literature since the seminal papers of Joskow (1980) and Friedman and Pauly (1981). The idea of service firms producing provision-of-service probabilities, and only incidentally observed output, was first characterized by Duncan (1990). This was first adapted to hospitals in a study by Gaynor and Anderson (1995) which analyzed the effects of demand uncertainty on hospital costs using a cost function, an approach followed in later studies by Carey (1998) and Hughes and McGuire (2003). In a recent paper, Lovell *et al.* (2009) analyzed this issue using a distance function approach which also allowed the possibility of expense preference behaviour to influence costs. All these papers find that demand uncertainty has a significant role to play when defining hospitals' potential output.

This literature shows that reserve capacity is, partially at least, the consequence of rational behaviour on the part of hospital managers to insure against demand fluctuations. However, providing protection against demand variability requires additional inputs to be contracted and this is costly, creating a trade-off between service capacity and economic costs. In the presence of any rational optimizing behaviour, these costs will influence the extent of protection against demand variability and in particular we would expect that reserve service capacity diminishes with the cost of providing the service.

The focus of this paper is on the reaction of hospitals to demand uncertainty and how production costs affect this reaction. To do so, we use tools from the production economics literature. In particular, we adopt an output distance frontier framework to model (unobserved) reserve capacity as a technical inefficiency term and try to determine to what extent it is explained by the existence of demand variability and the economic costs of maintaining this reserve capacity.[1]

Our theoretical setting is one where reserve capacity is a consequence of a rational decision in a context where there is a trade-off between protection against demand fluctuations and economic cost. In this framework, our hypothesis is that decision makers more worried by economic results (private hospitals) will react to a different extent to increases in demand uncertainty than decision makers for whom economic profits are not a primary objective (public hospitals). We analyze therefore whether the reaction to demand uncertainty is different between public and private hospitals. In our empirical work we use a broad dataset of Spanish general hospitals to test, on the one hand, for the influence of the cost of providing the service on the size of the reserve capacity. We then analyze the differences between private and public hospitals in how they react to uncertainty in the demand for different services. To our knowledge, this is the first paper to address both issues.

The paper proceeds as follows. Section 2 deals with the theoretical background to our analysis, where we discuss how reactions to demand uncertainty will have the effect that hospitals will generally produce below their production possibilities frontier and the tool with which we measure differences between potential and actual output, namely the output-oriented distance function. Section 3 discusses the data we use in our empirical analysis. The empirical specification of the distance frontier and the results of the estimation are presented in Section 4. Section 5 concludes.

[1] The technical inefficiency term relates the maximum potential production with the actual one. See Coelli et al (2005) for a good overview of efficiency and productivity analysis.

2. DEMAND UNCERTAINTY, HOSPITAL PRODUCTION AND TECHNICAL INEFFICIENCY

Hospitals will try to ensure that they have sufficient resources available to satisfy random demand at all times, except in extraordinary circumstances. With this in mind, Duncan (1990) provides a stylized analysis of service firm costs in the presence of stochastic demand, in a context where "… *what the firm incurs cost to produce is the option, the capacity, or the readiness to provide service at a certain level and only incidentally the observed output*." In this scenario, hospitals will produce a provision-of-service probability, β , implying that it is prepared to accept all clients up to capacity and produces so that capacity exceeds demand with probability β .

To derive his cost function, Duncan (1990) begins with a firm who faces a production function, $y = f(x)$. Demand for the firm, *d*, is a random variable with conditional distribution function $G(\cdot|d_{-1})$, where d_{-1} represents all relevant information that can be used to predict the probability that demand will be exceeded.

The probability that demand falls below service capacity is:

$$Prob\ (d \leq y|d_{-1}) = Prob\ (d \leq f(x)|d_{-1}) = G(f(x)|d_{-1}) = \beta \tag{1}$$

Now, if *G* is invertible, instead of $y = f(x)$ we can write $G^{-1}(\beta|d_{-1}) = f(x)$. This production relation describes the production behaviour of the firm who produces so that capacity exceeds demand with probability β . Under the assumption that the hospital chooses inputs *before* demand is realized with the constraint that demand is met with probability β , the cost minimization problem of the hospital gives rise to input demand functions of the form

$$x = x(G^{-1}(\beta|d_{-1}), w) \tag{2}$$

where w are the input prices and where it is clear that the inputs are chosen, *ex ante*, as a function of the target service capacity, $G^{-1}(\beta|d_{-1}) = \bar{y}$. Clearly, the greater the target service capacity, the greater the input use.

In the above framework, the provision-of-service probability β and hence the target service capacity, $\bar{y}$, is taken as given. However, the provision-of service probability is a decision taken by a rational agent that follows some objective, which can depend on the institutional nature of the hospital (public or private).

For a given distribution of demand, we assume that hospitals choose capacity as a function of the cost of inputs and the disutility associated with turning away patients. Input costs will influence the hospitals' choice of capacity insofar as they prevent the hospital from choosing unlimited capacity, i.e., the turn-away probability cannot be zero, and this is true regardless of whether the hospital is public or private. The disutility from turning away patients or obliging them to queue for certain services will depend, however, on the type of hospital and the type of service in question. We illustrate this by comparing the optimization problems facing two extreme cases: a private for-profit hospital and a public hospital which maximizes social welfare.

In the case of the private hospital, capacity will be chosen to maximize profits, so its objective function can be expressed as:

$$\max_{\bar{y}} \left[P\left(\int_0^{\bar{y}} y\, f(y)dy + \bar{y} \int_{\bar{y}}^{\infty} f(y)dy \right) - C(\bar{y}) \right] \quad (3)$$

where P is the output price, $f(y)$ is the density of demand and C represents costs. It is assumed that the hospital capacity, $\bar{y}$, is chosen and put into place before the demand for the service is materialized. Thus, the expected output is the expectation of a censored variable (Greene, 2008) because the hospital cannot produce an output level higher than $\bar{y}$.. This gives rise to the following first-order condition:

$$P(1 - F(\bar{y})) = \frac{\partial C(\bar{y})}{\partial \bar{y}} \quad (4)$$

At the other extreme, it is assumed that the public hospital is a social welfare maximizer who chooses capacity to maximize the utility function:

$$\max_{\bar{y}} U\big(Pr(y \leq \bar{y}), C(\bar{y})\big) \qquad (5)$$

where $Pr(y \leq \bar{y})$ is the probability that service capacity exceeds demand. Note that it is assumed in this function that social welfare does not depend on the output level. Given the hospital capacity installed ($\bar{y}$), social welfare is maximized when the demand for hospital services is null (i.e., it is better for citizens to be healthy and not to need hospital services). Citizens' utility thus depends on the probability of being attended if it were necessary. On the other hand, hospital capacity must be paid for and this cost is expected to negatively affect social welfare. The first-order condition to maximize social welfare is:

$$\frac{\partial U\big(Pr(y \leq \bar{y}), C(\bar{y})\big)}{\partial Pr(y \leq \bar{y})} f(\bar{y}) = -\frac{\partial U\big(Pr(y \leq \bar{y}), C(\bar{y})\big)}{\partial C(\bar{y})} \frac{\partial C(\bar{y})}{\partial \bar{y}} \qquad (6)$$

Comparing the first-order conditions, it can be seen that in both cases the choice of capacity depends on the distribution of demand. The private hospital chooses capacity so that the ratio of marginal cost to price equals the turn-away probability $\big(1 - F(\bar{y})\big)$. To see the implications of the first-order condition for the public hospital, we can simplify matters by holding the right-hand side constant i.e., assume that marginal cost and the marginal (dis)utility of cost are constant. Then, as the utility gained from being able to meet demand for a given output, $\partial U(\cdot)/\partial Pr(\cdot)$, increases, capacity rises, corresponding to a lower value (height) of the density function. For a symmetric distribution for demand, say a normal distribution, this means that chosen capacity increases. If utility is not much affected by being able to cover demand, on the other hand, the chosen capacity will be lower. Thus, the provision of service probability β for the public hospital will depend on the parameters of the utility function and in particular on the way the probability of meeting demand for a given service and the cost of providing capacity to meet that demand affect utility. It is worth noting that a comparison of the first order conditions does not permit us to conclude which type of hospital (public or private) provides a higher service capacity when faced with the same demand distribution.

For some services the public hospital may have more capacity than the private one, and for others it may have less. Indeed, this will turn out to be one of the main findings of our empirical work.

The implication of this is that public hospitals may choose a different service capacity than a private hospital when faced with similar costs and distributions of demand for the service. Hospitals that choose smaller turn-away probabilities will use larger quantities of inputs to produce the same expected output and will be operating further below their production possibilities frontier with a higher excess of capacity. If demand uncertainty is not taken into account, these hospitals will seem less efficient than others in providing the services demanded by the citizens. The tools from the production efficiency literature thus turn out to be extremely useful when analysing the effects of demand uncertainty on reserve capacity.

We use the output-oriented distance function, which is a convenient tool with which to measure the difference between potential and observed output, usually denoted as technical inefficiency, in a multi-output context such as hospital production (see Kumbhakar and Lovell, 2000; Coelli et al, 2005).[2] The output distance function is defined as

$$D_O(x,y) = min\left\{\theta : \frac{y}{\theta} \in P(x)\right\} \tag{7}$$

where $P(x)$ is the output set and which represents the set of output vectors, y, that can be produced using the input vector, x. $D_O(x,y)$ is nondecreasing, positively linearly homogeneous and convex in y, and decreasing in x (see, e.g. Färe and Primont, 1995).

Figure 1 below shows the production situation of two hospitals facing the same uncertain demand as represented by the conditional distribution function $G(\cdot|d_{-1})$. The hospitals choose two different probabilities for capacity exceeding demand, β_0

---

[2] Recent papers which have used output distance functions to measure hospitals' efficiency include Ferrari (2006) and Daidone and D'Amico (2009). For surveys on stochastic frontier analysis in health care, see Worthington (2004) and Rosko and Mutter (2008).

and β_1, where $\beta_1 > \beta_0$. to achieve this, the first hospital uses an input vector x_0 that generates a production possibilities frontier PPF_0, while the second hospital chooses the input vector x_1 giving rise to the production possibilities frontier PPF_1. Now, assume expected demand corresponds to point *A*. Given that capacity has been installed in order to be able to produce a level of output greater than expected demand, both hospitals are expected to operate below their PPF, which corresponds to the definition technically inefficiency (Farrell, 1957). The distance functions, which measures the ratio between actual and potential output, will take values less than or equal to 1. The distance function for the first hospital will take an expected value $(= OA/_{OB} < 1$. The second hospital is expected to produce further below its PPF and appears to be more inefficient: the expected value of the distance function is $(= OA/_{OC} < OA/_{OB}$.

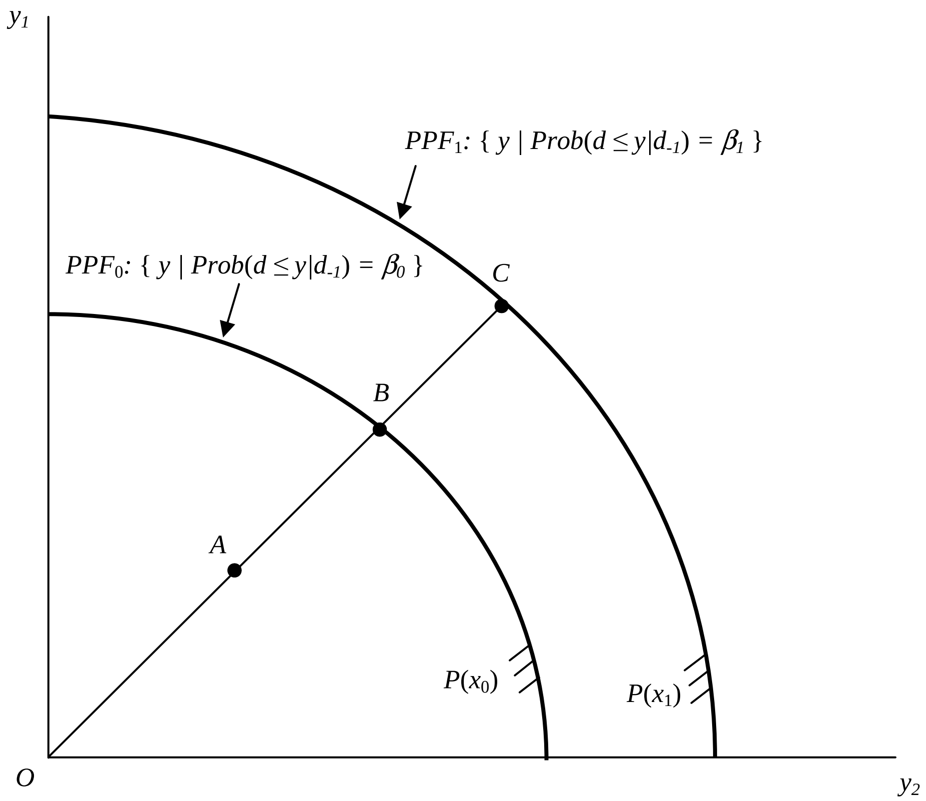

Figure 1. Technical efficiency and uncertain demand

Figure 1 could easily be extended to show the effects of an increase in demand uncertainty. Suppose expected demand remained unchanged but that the variance of demand increased. To maintain the target provision of service probability β the hospital would need a greater capacity, shifting its PPF outwards. This will increase its expected technical inefficiency, reflected by a lower value of the distance function.

In our empirical section we will estimate an output distance function in order to determine the effect of demand uncertainty on technical inefficiency. Output-oriented technical inefficiency can be introduced in (7) by writing:

$$D_O(x, y; \alpha) = e^{-u} \leq 1 \tag{8}$$

where α is a vector of parameters to be estimated and $u \geq 0$ represents output-oriented technical inefficiency: when $u = 0$, the distance function takes the value 1, representing technically efficient production, whereas values of $u > 0$ implies technical inefficiency and the distance function will take values less than 1. To introduce the influence of demand uncertainty and costs on technical inefficiency, we specify an appropriate functional form for $D_O(x, y; \alpha)$ and model u as a one-sided error term with the following distributional assumptions:

$$u \sim iid\ N^+(0, \sigma_u^2)\ ,\ \sigma_u^2 = g(z; \delta) \tag{9}$$

i.e. u is assumed to have a non-negative half-normal distribution with a modal value of technical inefficiency of zero and whose variance depends on a series of explanatory variables, *z*, with δ being a set of parameters to be estimated (see Caudill, Ford and Gropper, 1995). In line with the discussion above, it is expected that the greater the demand uncertainty facing the hospital, the greater the reserve capacity for a given provision-of-service probability and the greater the observed technical inefficiency. On the other hand, for a given level of demand uncertainty, an increase in input costs is expected to reduce the service capacity of the hospital and the subsequent inward shift of the PPF will reduce observed inefficiency. In this model, increases in the variance represent increases in technical inefficiency levels and *vice versa*. The vector of explanatory variables, z will therefore include variables capturing demand uncertainty and economic cost, and to capture possible different objectives between public and private hospitals it will also include dummy variables distinguishing between them. In the next section we describe our data set and how demand uncertainty is estimated.

3. DATA

The data have been obtained from the *"Estadística de Establecimientos Sanitarios en Régimen de Internado*" (EESRI) which have been carried out annually by the Spanish Ministry of Health and Consumption. The sample is an unbalanced panel that includes public and private hospitals observed over the period 1996-2006, with the number of hospitals ranging from 788 in the first year to 746 in the final year, corresponding to a total of 8,414 observations. With the purpose of homogenising the sample, we use only what are categorized as "general hospitals", excluding specialized ones, reducing the sample to 4,841 observations.

We aggregate the production of the hospital into three outputs: non-intensive care discharges (y_{IN})[3], outpatient visits (y_{OUT})[4] and intensive-care discharges (y_{ICU}). These services are provides by using five basic inputs: beds (*BED*), care graduates (*GRAD*), care technicians (*TECH*), expenses on sanitary material and others (*SUPP*) and buildings and equipment (*CAP*). Details of the output and input variables are provided in Table 1.

Table 1. Description of output and input variables

[3] The non-intensive care discharges are weighted by UPAs, or "weighted care units" which depend on the number of days corresponding to a given service.
[4] Second and successive outpatient visits are assigned a weight of 60% that of the first visit.
[4] There are no waiting lists for intensive care units.

Outputs	
y_{IN}	Weighted sum of discharges in general medicine, psychiatry, tuberculosis, long stay, rehabilitation and others, surgery, paediatric, gynaecological surgery, obstetrics, paediatric medicine and neonatology, burns and intensive neo-natals.
y_{OUT}	Weighted sum of first and successive visit and emergencies
y_{ICU}	Discharges in intensive care units
Inputs	
BED	Endowment of beds.
GRAD	Care graduates (doctors, pharmacists and other graduates).
TECH	Care technicians (nurses, matrons and others)
SUPP	Expenditure on sanitary material, food, clothing, fuels and others (euro, year 2006).
CAP	Investments, repairs and amortization of fixed assets (euro, year 2006)

To determine the effect of demand uncertainty on efficiency, we need an estimate of demand variability. We consider that the demands for hospital services are the services effectively provided by the hospital (actual discharges and visits and emergencies) plus the corresponding waiting lists. The demands for each output are denoted y_{IN}^d, y_{OUT}^d and y_{ICU}^d.[5]

When estimating the demand equations, we take into account the fact that demand characteristics may differ from one region to another. We therefore estimate a set of three equations for each Autonomous Community, which is the most disaggregated geographical area in the sample. For each Autonomous Community we estimate separate demand equations for each hospital service, using the lagged values of output to predict demand. The demand equations take the form:

$$ln(y_{ijt}^d) = \delta_j + \delta_1 ln(y_{ijt-1}^d) + \delta_t D_t + e_{itj} \quad (10)$$

where subscript i refers to output, j to the individual hospital and t to the period. D_t are year dummy variables, and the δ *'s* are parameters to be estimated where δ_j is the hospital fixed effect which accounts for unobserved heterogeneity.

Once the demand equations are estimated we use the absolute value of the residuals, $|\hat{\varepsilon}_{ijt}|$, to estimate the standard errors of the demands. As before, the equations are estimated for each Autonomous Community, and we use the same variables as in the demand equation:

$$|\hat{\varepsilon}_{ijt}| = \gamma_j + \gamma_1 ln\left(y^d_{ijt-1}\right) + \gamma_t D_t + \eta_{ijt} \tag{11}$$

where the γ *'s* are parameters to be estimated and η is an error term. Again, γ_j represents a fixed effect which captures unobserved factors which influence the variability of the demand facing each hospital.

The fitted values from the set of equations (11), which we denote σ_{IN}, σ_{OUT} and σ_{ICU}, are our estimates of the standard deviations of the three hospital services and will be included as arguments in the estimation of the hospitals' technical efficiency.

As an alternative specification, we also aggregated the individual demands into an overall demand using the UPA weights and estimated its standard deviation following the procedure outlined above, i.e., estimating a single demand equation for each Autonomous Community. We then re-estimated the distance function modelling the technical inefficiency term as a function of this standard deviation of aggregate or overall demand, σ_y.

To capture the effect of economic costs on technical inefficiency, we use average hospital labour costs, *SAL*. These are calculated by dividing the total salary expenditure by the total number of workers. Finally, to distinguish between public and private hospitals we will use a dummy variable, *PRIV*, which takes the value 1 when the hospital is private.

Regarding the sample used to estimate the distance function, note that when estimating the demand equations we lose the first year's observations for each hospital, a total of 455 observations. A number of hospitals that did not produce all the services corresponding to our three outputs were also eliminated, reducing the sample by a further 2,051 observations. Since we only consider for-profit private hospitals, we eliminated another 140 observations corresponding to private non-profit hospitals. Finally, we eliminated observations with abnormally high or low total salary costs due to misreporting and some hospitals for which the number of care technicians changed drastically in certain periods despite there being little or no change in the remaining inputs, a total of 42 observations.[6] After eliminating these observations, the final sample used for estimating the distance function consisted of 2,136 observations corresponding to the years 1997-2006. The number of public hospitals included ranged from 157 in 1997 to 161 in 2006, while the number of private for-profit hospitals was 39 in 1997 and 70 in 2006. Some descriptive statistics of the output and input variables used in the estimation of the distance function as well as the explanatory variables of technical inefficiency are provided in Table 2.

Table 2. Descriptive statistics of production and efficiency variables

Variable	***Mean***	***Std. Dev.***	***Min***	***Max***
Distance function				
y_{IN}	21788	17233	111	107612
y_{OUT}	43621	41414	58	237297
y_{ICU}	947	1137	1	11803
BED	434	376	7	1927
GRAD	243	207	17	1539
TECH	411	428	10	2305
SUPP	1051	1234	7	16925
CAP	4801	6418	79	119295

[6] However, it should be highlighted that these anomalous observations for salaries and care technicians represent a very small proportion of the sample and that their elimination or inclusion did not change the results of our estimations in any significant way.

Technical efficiency				
σ_{IN}	0.068	0.072	-0.071	0.811
σ_{OUT}	0.104	0.098	-0.079	1.111
σ_{ICU}	0.214	0.219	-0.061	2.529
σ_{y} *(aggregate demand)*	0.074	0.066	-0.029	0.701
SAL	25380	8483	2296	53518
PRIV	0.32	-	-	-
Number of observations	2136			

Table 2 includes the demand uncertainty variables obtained from estimating the demand equations (10-11). The demand equations worked quite well in that they high predictive power. For the equations using aggregate demand, the average R^2 for the 17 different estimations corresponding to each Autonomous Community was 0.98. For the individual demand equations, the average R^2 were 0.98 for both y_{IN}^d and y_{OUT}^d and 0.90 for y_{ICU}^d. Closer inspection of Table 2 reveals that the predicted standard error of demand, obtained from estimating (11), was in negative in some cases. This occurred for 140 observations of the 2139 in the sample. As the estimates of the parameters in the distance function and the technical efficiency term did not change when these observations were excluded, we decided to include them in our final sample.

4. EMPIRICAL SPECIFICATION AND RESULTS

It remains to specify a functional form for the distance function (8) and the inefficiency term (9). We estimate the technology using a (minus) translog output oriented distance function.[7] Linear homogeneity in outputs has been imposed by dividing by y_{IN}, so the distance function is specified as:

[7] That is, the distance function has been multiplied by -1. This is done to give a more intuitive interpretation to the sign of the estimated parameters and is fairly common practice when estimating output distance functions – see, for example, Coelli and Perelman (2000).

$$-\ln y_{INjt} = \alpha_0 + \sum_{k=1}^{5} \alpha_k \ln x_{kjt} + \frac{1}{2} \sum_{k=1}^{5} \sum_{m=1}^{5} \quad (12)$$

where the α 's are parameters to be estimated and $\ln y^*_{ijt} = \ln y_{ijt} - \ln y_{INjt}$. The independent variables were divided by their geometric mean.

The error term is a composed error term where v_{jt} is assumed to be normally distributed. In accordance with (9), u_{jt} is assumed to follow a half-normal distribution where its variance, $\sigma_u^2 = g(z;\delta)$, depends on the demand uncertainty and cost variables. In particular, we model this variance as a linear function of the estimated standard deviations of the demands for the three outputs (σ_{IN}, σ_{OUT} and σ_{ICU},) from equations (11), the average labour costs (*SAL*), and the interaction of the private hospital dummy variable with the standard deviations of the demands, $PRIV \cdot \sigma_{y_i}$:

$$\ln \sigma_{ujt}^2 = \delta_0 + \sum_{i=1}^{3} \delta_1 \, \sigma_{y_i jt} + \delta_3 \, SAL_{jt} + \sum_{i=1}^{3} \delta_4 \, PRIV \cdot \sigma_{y_i jt} \qquad (13)$$

The output distance frontier was estimated using *Stata 10* and the estimated maximum likelihood parameters are presented in Tables 4 below. Table 4 shows the estimates of the frontier distance function using the standard deviations of the demands for each individual service. The estimates from the alternative model where technical inefficiency was modelled using the standard deviation of overall demand were almost identical and are reported in an Appendix.

The model works quite well. In particular, all the first order parameters have the expected sign and are highly significant which implies that the estimated technology complies with the theoretically expected monotonicity conditions.

Table 4. Estimate of distance function: Demand by categories

Variable	*Coeff.*	*S.E.*	*t-stat.*	*Variable*	*Coeff.*	*S.E.*	*t-stat.*
Constant	9.879	0.015	638.5	*ln*(*BED*)·*ln*(*TECH*)	0.067	0.055	1.23

$ln(y_{OUT})$	-0.254	0.011	-23.52	$ln(BED)\cdot ln(SUPP)$	-0.007	0.029	-0.24
$ln(y_{ICU})$	-0.059	0.008	-7.15	$ln(BED)\cdot ln(CAP)$	0.090	0.022	4.12
$ln(BED)$	0.259	0.018	14.69	$ln(GRAD)^2$	0.158	0.047	3.39
$ln(GRAD)$	0.295	0.020	14.87	$ln(GRAD)\cdot ln(TECH)$	0.060	0.039	1.52
$ln(TECH)$	0.192	0.019	10.09	$ln(GRAD)\cdot ln(SUPP)$	-0.136	0.034	-3.99
$ln(SUPP)$	0.040	0.009	4.30	$ln(GRAD)\cdot ln(CAP)$	-0.052	0.020	-2.67
$ln(CAP)$	0.051	0.006	8.00	$ln(TECH)^2$	-0.029	0.060	-0.48
$ln(y_{OUT})^2$	-0.050	0.011	-4.76	$ln(TECH)\cdot ln(SUPP)$	0.037	0.027	1.34
$ln(y_{OUT})\cdot ln(y_{ICU})$	-0.008	0.014	-0.55	$ln(TECH)\cdot ln(CAP)$	-0.085	0.021	-4.13
$ln(y_{ICU})\cdot ln(y_{ICU})$	-0.047	0.010	-4.91	$ln(SUPP)^2$	-0.003	0.009	-0.36
$ln(y_{OUT})\cdot ln(BED)$	0.114	0.024	4.67	$ln(SUPP)\cdot ln(CAP)$	0.060	0.010	5.96
$ln(y_{OUT})\cdot ln(GRAD)$	0.100	0.019	5.22	$ln(CAP)^2$	-0.037	0.010	-3.75
$ln(y_{OUT})\cdot ln(TECH)$	-0.051	0.026	-1.96	D_{1998}	0.004	0.017	0.25
$ln(y_{OUT})\cdot ln(SUPP)$	-0.094	0.020	-4.77	D_{1999}	-0.007	0.017	-0.42
$ln(y_{OUT})\cdot ln(CAP)$	0.035	0.011	3.13	D_{2000}	0.021	0.017	1.21
$ln(y_{ICU})\cdot ln(BED)$	-0.117	0.021	-5.61	D_{2001}	0.040	0.017	2.30
$ln(y_{ICU})\cdot ln(GRAD)$	0.004	0.019	0.23	D_{2002}	0.050	0.017	2.92
$ln(y_{ICU})\cdot ln(TECH)$	0.055	0.019	2.83	D_{2003}	0.025	0.017	1.42
$ln(y_{ICU})\cdot ln(SUPP)$	0.034	0.014	2.37	D_{2004}	0.011	0.017	0.65
$ln(y_{ICU})\cdot ln(CAP)$	0.008	0.010	0.79	D_{2005}	-0.002	0.018	-0.12
$ln(BED)^2$	-0.017	0.075	-0.23	D_{2006}	-0.015	0.018	-0.82
$ln(BED)\cdot ln(GRAD)$	-0.058	0.041	-1.41				

No. observations: 2136

The estimates of the parameters of the variables used to model the variance of the asymmetric error term are presented for both models in Table 5. The results conform to our expectations. Beginning with aggregated demand (Model 1), the coefficient on the demand uncertainty variable is positive and significant. This implies that demand uncertainty positively affects the variance of the asymmetric error term showing the voluntary creation of a "buffer" to deal with demand uncertainty: the higher the demand uncertainty, the larger the buffer needed, and hospitals will be further beneath the PPF on average. Costs, as represented by the salary variable, negatively affect the variance of the asymmetric error term showing that when the hospital services become more expensive, the hospital decision-maker reduces the buffer devoted to dealing with the demand uncertainty. Finally, the coefficient on the interaction of the dummy variable for private hospitals and demand uncertainty is negative and significant, indicating that private hospitals react to demand uncertainty to a lesser degree than public hospitals in the sense that they do not install as much excess service capacity. The fact that adding both coefficients results in a positive significant value implies that private hospitals do install some extra service capacity in reaction to demand uncertainty, but less than the public ones.

Table 5. Determinants of variance of technical efficiency

Model 1 ***Uncertainty in Total Demand***				***Model 2*** ***Uncertainty in Demand by Categories***			
Variable	*Coeff.*	*S.E.*	*t-stat.*	*Variable*	*Coeff.*	*S.E.*	*t-stat.*
Constant	13.715	1.355	10.12	*Constant*	8.811	1.494	5.90
SAL	-1.693	0.135	-12.50	*SAL*	-1.275	0.148	-8.60
σ_y	9.467	0.824	11.49	σ_{IN}	6.916	0.894	7.74
PRIV $\bullet\, \sigma_y$	-4.565	1.055	-4.33	σ_{OUT}	1.560	0.768	2.03
				σ_{ICU}	2.287	0.304	7.53
				PRIV $\bullet\, \sigma_{IN}$	-2.629	1.318	-1.99
				PRIV $\bullet\, \sigma_{OUT}$	2.729	0.906	3.01
				PRIV $\bullet\, \sigma_{ICU}$	-1.402	0.421	-3.33

Turning to the model where demand uncertainty for individual services is specified (Model 2), the cost variable is again negative and significant. Demand uncertainty in each of the three services causes hospitals to increase their service capacity, in line with the previous model, but the interaction terms shed more light on the differences in behaviour of the two types of hospital. Recalling the discussion surrounding the first-order conditions of the public hospital (6), the disutilities associated with turning away patients may be different depending on the service in question and the reserve capacity will vary accordingly. Our results highlight these different disutilities between private and public hospitals. The coefficients on the interaction terms with private hospitals for inpatient and intensive care services are negative and significant, implying that private hospitals do not provide as much extra capacity as public hospitals to deal with uncertainty in these services. However, the interaction term with outpatient services (visits and emergencies) is positive and significant, highlighting that private hospitals install more capacity than public hospitals to deal with outpatients.

These results imply that public hospitals place a much higher value on meeting demand for inpatient and intensive care services than they do on outpatients. Private for-profit hospitals, on the other hand, are shown to be much more preoccupied with being capable of providing outpatient services. This is in accordance with the nature of these hospitals in Spain. Public hospitals provide universal service and from a social welfare perspective they should be equipped to deal with pressing cases as quickly as possible. A private for-profit hospital will not suffer the same outcry as patients can be referred to another hospital (private or public) if they do not have the resources to deal immediately with such cases. Regarding outpatient services, patients have the option of receiving these for free from the public system or paying though the private system. Clearly, the disutility from turning away patients in these cases (consultancies, minor surgery etc.) in the sense of obliging them to join waiting lists is much lower for the public system than for inpatient and intensive care services. One of the main selling points of private for-profit hospitals on the other hand, is that they can attend patients immediately; permitting them to avoid what could be a long wait in a public hospital. It is in their interest to be able to provide such a service for the patient, which, it should be noted, is cheaper for the patient than the other two services and also much less costly for the hospital to produce than the other outputs.

Finally, using the estimated parameters from the models we calculate indices of technical efficiency (percentage of actual to potential output) for public and private hospitals as a function of estimated demand uncertainty. From the estimated distributions of the standard deviations of the demands for the three hospital services we choose three types of demand uncertainty - low, medium and high - corresponding to the first quartile, median, and third quartile of the demand distributions for each service. The indices are presented in Table 6, which shows how the estimated technical efficiency indices for both public and private hospitals decrease as demand becomes more uncertain. The greater reaction of public hospitals to demand uncertainty is reflected in a more pronounced fall in the efficiency indices compared to private hospitals as demand uncertainty increases.

Table 6. Technical efficiency indices according to demand uncertainty

	Degree of demand uncertainty		
	Low.	*Medium*	*High*
Public	86.71	84.04	78.46
Private	87.12	85.01	81.06

Figure 2 below shows this in more detail, tracing the evolution of the efficiency indices as the demand uncertainty increases over all deciles. Again, when demand uncertainty is low the efficiency indices are almost at the same level, and as uncertainty rises, public hospitals react by increasing capacity to a greater extent than private hospitals with a correspondingly greater fall in measured technical efficiency.

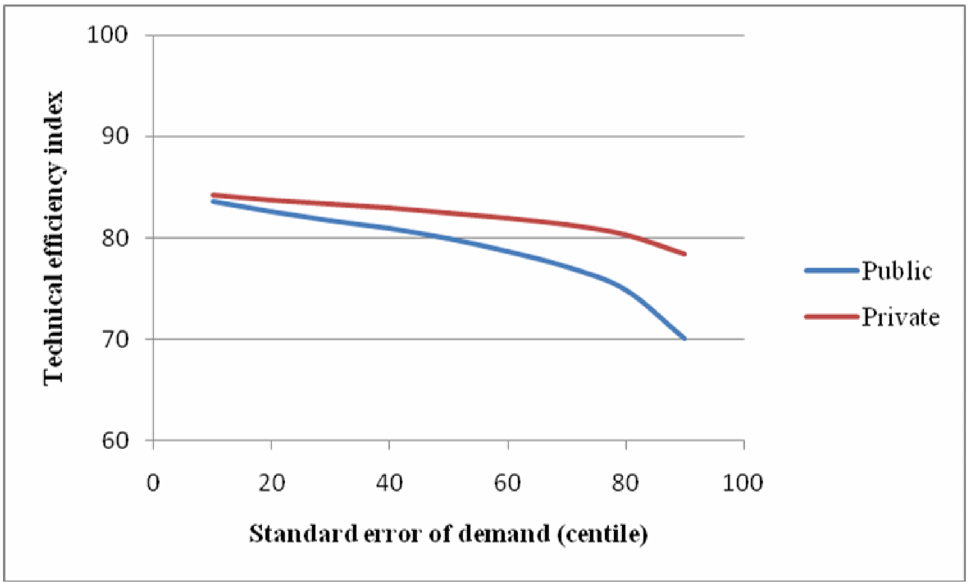

Figure 2. Technical efficiency indices and demand uncertainty

This has important implication for studies which attempt to compare the efficiency performance of public and private for-profit hospitals. A finding that one type of hospital is more efficient than another may be due to a rational decision to provide a different degree of protection against demand fluctuations, and not to pure inefficiency due to suboptimal management.. At a time where there is a heated policy debate over the efficiency of the public system and the idoneity of different management and ownership structures, a comparison of relative efficiencies should take into consideration the objectives of each type of hospital if a fairer and more accurate picture of performance is to be provided.

5. CONCLUSIONS

In this paper we have investigated the effect of demand uncertainty and the cost of the service on hospitals' decisions to contract reserve service capacity. The desire of hospitals to maintain reserve capacity to meet uncertain demand will make them appear technically inefficient as demand generally falls below capacity. Using a simple theoretical framework we show how the disutility associated with turning away patients for a given service can lead public hospitals to provide a different reserve capacity to private for-profit hospitals. We model reserve capacity as the technical inefficiency term of a stochastic output-oriented distance function and compare the reaction to demand uncertainty for a sample of Spanish public and private for-profit general hospitals. The parameters on the variables characterizing the efficiency term show that maintaining some degree of excess capacity is compatible with a rational optimization objective that takes account of the desire to meet uncertain demand, on the one hand, and the cost of providing the corresponding reserve capacity on the other. Our results show that private hospitals react to a lesser extent to demand uncertainty as a whole than public hospitals, which is consistent with the fact that Spanish public hospitals are obliged to provide universal service. When we analyze demand uncertainty in different hospital services, we find that public hospitals react to demand in inpatient and intensive care services by installing more reserve service capacity than private hospitals. For outpatient services, on the other hand, private hospitals react to a greater extent than public hospitals.

Our results show that it may be misleading to attribute a higher measured technical inefficiency of public hospitals to worse technical management if demand uncertainty effects are not taken into account. It is important to take the behavioural objectives of each type of hospital into account as these will affect the reserve capacity decision. Researchers carrying out empirical work comparing the efficiency of public and private hospitals should therefore be careful to take this into account.

Finally, an interesting question which can be raised is whether the health system as a whole has sufficient excess capacity to be able to deal with the closure of hospitals in the short-run due to planned or unplanned "shocks" in the supply of hospital services (see, for example, Ferrier et al, 2009). If citizens are worried about the probability of being attended if necessary but managers are worried about the profits that hospital make by selling services, the reserve capacity provided will generally be different from the optimal one. Policy planners will need to ensure an institutional framework that reconciles managers' objectives with citizens needs.

References

Carey, K. (1998). "Stochastic demand for hospitals and optimizing 'excess' bed capacity." *Journal of Regulatory Economics* 14, 165-187

Caudill, S., Ford, J. and D. Gropper (1995). "Frontier estimation and firm-specific inefficiency measures in the presence of heteroscedasticity." *Journal of Business and Economic Statistics* 13, 105-111.

Coelli, T. and S. Perelman (2000). "Technical efficiency of European railways: A distance function approach." *Applied Economics* 32, 1967-76.

Coelli, T.J., Prasada Rao, D.S., O`Donnell, C.J. and G.E. Battese (2005). *An introduction to efficiency and productivity analysis*. 2nd ed. Springer, New York.

Daidone, S. and F. D'Amico (2009). "Technical efficiency, specialization and ownership form: evidences from a pooling of Italian hospitals." *Journal of Productivity Analysis* 32, 203-216.

Duncan, G.D. (1990). "The effect of probabilistic demands on the structure of cost functions." *Journal of Risk and Uncertainty* 3, 211-220.

Färe, R. and D. Primont (1995). *Multi-output production and duality: Theory and applications*. Kluwer Academic Publishers, Norwell.

Farrell, M.J. (1957). "The measurement of productive efficiency." *Journal of the Royal Statistical Society* Series A, General 120, 253-281.

Ferrari, A. (2006). "The internal market and hospital efficiency: a stochastic distance function approach." *Applied Economics* 38, 2121-2130.

Ferrier, G.D., Leleu, H. and V.G. Valdmanis (2009). "Hospital capacity in large urban areas: is there enough in times of need?" *Journal of Productivity Analysis* 32, 103-117.

Friedman, B. and M.V. Pauly (1981). "Cost functions for a service firm with variable quality and stochastic demand: the case of hospitals." *Review of Economics and Statistics* 63, 620-624.

Friesner, D., Roseman, R. and M.Q. McPherson (2008). "Are hospitals seasonally inefficient? Evidence from Washington State." *Applied Economics* 40, 699-723.

Gaynor, M. and G.F. Anderson (1995). "Uncertain demand, the structure of hospital costs, and the cost of empty hospital beds." *Journal of Health Economics* 14, 291-317.

Greene, W. H. (2008). *Econometric Analysis*. 6th ed. Prentice Hall. New Jersey

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

Joskow, P.L (1980). "The effects of competition and regulation on hospital bed supply and the reservation quality of the hospital." Bell Journal of Economics 11, 421-447.

Lovell, C.A.K., Rodríguez-Álvarez, A. and A. Wall (2009). "The effects of stochastic demand and expense preference behaviour on public hospital costs and excess capacity." *Health Economics* (18), 227-235.

Rosko, M.D. and R.L. Mutter (2008). "Stochastic frontier analysis of hospital inefficiency: A review of empirical issues and an assessment of robustness." *Medical Care Research and Review* 65, 131-166.

Worthington, A.C. (2004). "Frontier efficiency measurement in health care: A review of empirical techniques and selected applications." *Medical Care Research and Review* 61, 135-170.

APPENDIX

Table 4A. Estimate of distance function: Model 1 (overall demand)

Variable	*Coeff.*	*S.E.*	*t-stat.*	*Variable*	*Coeff.*	*S.E.*	*t-stat.*
Constant	9.894	0.016	604.5	$ln(BED)\cdot ln(TECH)$	0.088	0.058	1.53
$ln(y_{OUT})$	-0.258	0.011	-22.93	$ln(BED)\cdot ln(SUPP)$	-0.018	0.030	-0.61
$ln(y_{ICU})$	-0.051	0.009	-5.97	$ln(BED)\cdot ln(CAP)$	0.111	0.023	4.90
$ln(BED)$	0.254	0.019	13.74	$ln(GRAD)^2$	0.185	0.047	3.95
$ln(GRAD)$	0.286	0.021	13.68	$ln(GRAD)\cdot ln(TECH)$	0.046	0.042	1.11
$ln(TECH)$	0.216	0.020	10.75	$ln(GRAD)\cdot ln(SUPP)$	-0.157	0.035	-4.44
$ln(SUPP)$	0.032	0.010	3.30	$ln(GRAD)\cdot ln(CAP)$	-0.061	0.020	-3.02
$ln(CAP)$	0.051	0.007	7.52	$ln(TECH)^2$	-0.022	0.064	-0.34
$ln(y_{OUT})^2$	-0.063	0.010	-6.11	$ln(TECH)\cdot ln(SUPP)$	0.060	0.029	2.05
$ln(y_{OUT})\cdot ln(y_{ICU})$	0.017	0.013	1.27	$ln(TECH)\cdot ln(CAP)$	-0.106	0.022	-4.89
$ln(y_{ICU})\cdot ln(y_{ICU})$	-0.055	0.009	-5.99	$ln(SUPP)^2$	-0.007	0.010	-0.75
$ln(y_{OUT})\cdot ln(BED)$	0.157	0.024	6.67	$ln(SUPP)\cdot ln(CAP)$	0.061	0.011	5.78
$ln(y_{OUT})\cdot ln(GRAD)$	0.061	0.019	3.22	$ln(CAP)^2$	-0.033	0.011	-3.13
$ln(y_{OUT})\cdot ln(TECH)$	-0.067	0.026	-2.59	D_{1998}	0.010	0.019	0.51
$ln(y_{OUT})\cdot ln(SUPP)$	-0.093	0.020	-4.73	D_{1999}	-0.007	0.018	-0.40
$ln(y_{OUT})\cdot ln(CAP)$	0.044	0.011	3.90	D_{2000}	0.018	0.019	0.98
$ln(y_{ICU})\cdot ln(BED)$	-0.109	0.021	-5.31	D_{2001}	0.039	0.019	2.10
$ln(y_{ICU})\cdot ln(GRAD)$	0.020	0.019	1.02	D_{2002}	0.050	0.019	2.68
$ln(y_{ICU})\cdot ln(TECH)$	0.038	0.020	1.97	D_{2003}	0.025	0.019	1.34
$ln(y_{ICU})\cdot ln(SUPP)$	0.043	0.014	3.02	D_{2004}	0.013	0.019	0.71
$ln(y_{ICU})\cdot ln(CAP)$	0.007	0.010	0.71	D_{2005}	-0.002	0.019	-0.12
$ln(BED)^2$	-0.071	0.076	-0.93	D_{2006}	-0.013	0.019	-0.66
$ln(BED)\cdot ln(GRAD)$	-0.049	0.042	-1.17				

No. observations: 2136