

# EFICIÊNCIA TÉCNICA, ECONOMIAS DE ESCALA, ESTRUTURA DA PROPRIEDADE E TIPO DE GESTÃO NO SISTEMA HOSPITALAR BRASILEIRO

*André Proite*

Secretaria do Tesouro Nacional e Universidade de Brasília  
Campus Universitário Darcy Ribeiro S/N, Brasília-DF, 70910-900  
e-mail: [andre.proite@fazenda.gov.br](mailto:andre.proite@fazenda.gov.br)

*Maria da Conceição Sampaio de Sousa*

Departamento de Economia, Universidade de Brasília  
Campus Universitário Darcy Ribeiro S/N, Brasília-DF, 70910-900, e-mail: [mcss@unb.br](mailto:mcss@unb.br)

## RESUMO

Nesse trabalho, computamos fronteiras não paramétricas de eficiência técnica para 1170 hospitais brasileiros da rede SUS, utilizando o método DEA (*Data Envelopment Analysis*), para retornos variáveis de escala, juntamente com o método recentemente proposto que combina as técnicas de reamostragem Bootstrap e Jackknife para eliminar os efeitos de *outliers* e outras discrepâncias estatísticas. Após o cálculo dos índices de eficiência utilizamos técnicas econométricas, em especial, regressão quantílica para investigar os determinantes desses escores. Nossos resultados confirmaram a importância do efeito escala (medido através do total de atendimentos realizados) sobre a eficiência técnica este setor, controlada pelo tempo médio de permanência no hospital, pelo custo médio e pelo capital humano. No que diz respeito à gestão dos hospitais, destacam-se os efeitos negativos das instituições não lucrativas (benéficas e cooperativas) sobre o desempenho, para os quantis maiores de eficiência, resultado que se diferencia da literatura empírica pelo uso de regressão quantílica. A especialização excessiva das instituições também tem efeitos negativos da sobre a eficiência, sugerindo que existe uma combinação ótima entre especialização e generalização dos hospitais. Quanto à propriedade, os hospitais privados tiveram sua performance negativamente afetada para as unidades acima do 30º quantil de eficiência, sugerindo que os efeitos associados às falhas de mercado são importantes para esse setor.

**Palavras-chave:** Análise Envoltória de Dados; Fronteiras Não Paramétricas; Eficiência Técnica; Quantis de Eficiência.

## ABSTRACT

We estimate DEA (Data Envelopment Analysis) technical efficiency scores for 1170 Brazilian hospitals included in the SUS (Central Health System) using a recently proposed method that combines bootstrap and jackknife resampling to eliminate the influence of outliers and possible measurement and recording errors in the data. We use the variable returns to scale variants of the DEA method. After computing the efficiency scores, we use econometric methods, especially quantile regression, to investigate the determinants of those scores. Our results confirm the importance of the scale effect (measured by the total number of services) over the technical efficiency of this sector, controlled by the average permanence in the hospital, its average costs and human capital variables. Regarded to hospital's management, we highlight the negative effects of non-profit organizations over the efficiency, which differs from the empirical literature due the usage of quantile regression in this paper. Excessive specialization also has a negative effect over the efficiency scores, pointing out the existence of an optimal mix between specialization and generalization hospitals' characteristics. As for property, private hospitals have its performance negatively affected for those units over the 30<sup>th</sup> efficiency quantile, suggesting that market failure effects are important in this sector.

**Key Words:** Data Envelopment Analysis; Nonparametric Frontiers; Technical Efficiency; Quantile Regression. **JEL Classification:** C5, C6, C14, H5, I11.

## 1. INTRODUÇÃO

A concepção de que a oferta de serviços de saúde - particularmente aqueles prestados pelos hospitais - não deve se fazer no âmbito do mercado advém das peculiaridades desse setor. A idéia básica é que, no mercado de serviços hospitalares, os compradores estão bem menos informados sobre a qualidade do serviço do que os provedores; conseqüentemente, alguns pesquisadores sugerem que a solução para esse problema de assimetria de informação encontra-se na provisão desses serviços por instituições públicas ou entidades sem fins lucrativos, como cooperativas e beneficências (Arrow (1963)). Juntando-se a isso, o fato de que nesse setor prevalecem as economias de escala, tem-se, então, que os aspectos distributivos e de apreçamento distintos daqueles praticados pela esfera privada abrem espaço para a atuação do governo como agente regulador e, em alguns casos, produtor desses serviços. Não obstante, na maioria dos países, hospitais e outras unidades de prestação de serviços de saúde privados - inclusive beneficências e cooperativas - coexistem com instituições públicas. Nesse contexto, uma questão importante é saber em que medida, a natureza da instituição que presta serviços hospitalares - incluindo-se aí a propriedade e/ou gerenciamento - influencia o desempenho efetivo desses estabelecimentos.

Um amplo debate teórico e empírico, principalmente, no tocante à performance comparada entre hospitais privados - com e sem fins lucrativos - e desses com àqueles assumidos pelo setor público pode ser encontrada na literatura sobre esse tópico. Por sua relevância, essa discussão atraiu inúmeros especialistas, porém, apesar do esforço despendido, está longe de ser conclusiva. Vários autores alegam, por exemplo, que organizações não lucrativas, como as cooperativas e as beneficências, tendem a ser mais eficientes que suas equivalentes privadas e públicas em virtude das assimetrias de informação e da incerteza que caracterizam o mercados de serviços médicos (Weissbrod (1998), Glaeser and Schleifer (2001)). Essa suposta maior eficiência derivaria, em parte, do fato de essas instituições serem legalmente impedidas de distribuir seus lucros (*non distribution constraint*) o que reduziria os incentivos para explorar as vantagens informacionais que elas detêm (Hansmann, Kessler and McClellan (2002)). Ademais, em razão da natureza privada dessas instituições, elas não estão sujeitas às restrições legais dos estabelecimentos públicos, no sentido de prover o mesmo serviço para todos os cidadãos (*categorical constraint*); essa maior flexibilidade lhes permitiria atingir níveis de eficiência mais elevados. Juntos, esses fatores explicariam, em parte, o fato de os estabelecimentos sem fins lucrativos estarem, na maioria dos países, sujeitos a esquemas regulatórios distintos dos demais hospitais privados, particularmente no que concerne à questão fiscal, tornando-os elegíveis para diferentes isenções tributárias. Por outro lado, outros pesquisadores argumentam que as instituições sem fins lucrativos, presumivelmente, são mais ineficientes quando comparados com os demais hospitais privados e/ou públicos. Isto porque a ausência de uma propriedade inequívoca reduziria os incentivos administrativos para a minimização de custos e para a adoção de sistemas eficazes de monitoramento [Fama and Jensen (1983)]. Adicionalmente, de acordo com a teoria padrão da agência, esse tipo de instituição tenderia a ser mais ineficiente em virtude de não estar sujeita à fiscalização dos acionistas (no caso das entidades privadas) e dos contribuintes (no caso das entidades públicas). Daí resultariam o desperdício de recursos - capital e trabalho - o excesso de capacidade e a sobre-utilização de equipamento médico.

No tocante aos hospitais públicos, supõe-se que eles são mais ineficientes, devido, em parte, ao excesso de burocracia, a presença de sistemas hierárquicos rígidos e ao excesso de utilização de mão de obra [Clark (1980), Sloan e Steinwald (1980)]. Isso leva à maximização do orçamento, ao invés da minimização dos custos, de acordo com a teoria da escolha pública proposta por Niskanen (1971). Outros trabalhos afirmam, porém, que os hospitais públicos são menos dispendiosos em razão das

restrições orçamentárias impostas pelo governo e pelo fato de essas instituições, por serem, em média, de maior porte, podem explorar as economias de escala prevalentes na produção reduzindo, assim, o custo médio dos procedimentos [Brown (1983)].

Em síntese, essas visões conflitantes mostram que não existe consenso a respeito da influência da propriedade e/ou gerenciamento sobre os níveis de eficiência do sistema hospitalar. Essa discussão é, pois, basicamente, uma questão empírica. Por essa razão, é fundamental proceder a avaliações dos desempenhos dessas instituições, em situações concretas. Uma maneira apropriada para realizar essa tarefa é estimar fronteiras de eficiência, como aquelas descritas pelo método DEA (*Data Envelopment Analysis*), inicialmente proposta por Charnes, Cooper e Rhodes (1978), a partir da medida elaborada por Farrell (1957). Esse modelo adequa-se bem ao caso dos serviços hospitalares porque é um método relativamente pouco restritivo e capaz de envolver vários aspectos de sua produção, incluindo-se aí a existência de múltiplos insumos e múltiplos produtos [Evans *et al* (1999), Agrell e Bogetoft (1998), Kittlessen e Magnussen (2003)]. Existe uma vasta literatura que aplica esses métodos para computar medidas de eficiência - paramétricas e não paramétricas - para hospitais, incluindo-se aí os trabalhos que examinam o impacto da estrutura de propriedade e da forma de gestão sobre o desempenho dessas instituições [Valdmanis (1992), Mckillop *et al* (1999), Barbeta, Turati e Zago (2001), Dugan (2002)].

No Brasil, os estudos desse tipo são poucos e utilizam uma base de dados limitada [Marinho (2001)]. Considerando-se que, apesar dos elevados custos, nossa extensa rede hospitalar, financiada, em grande parte, pelos recursos dos SUS (Sistema Único de Saúde), atende de forma precária a população, a avaliação da produtividade dos gastos hospitalares no âmbito do SUS é extremamente relevante. Faz-se, pois, necessário proceder à mensuração do desempenho da rede hospitalar e verificar em que medida, esse desempenho é influenciado pelos fatores acima mencionados.

O objetivo deste trabalho é, pois, computar escores de eficiência para os hospitais da rede SUS, que realizaram cirurgias para o ano de 2002, utilizando um método não paramétrico recentemente proposto por Sampaio de Sousa e Stosic (2003), que combina a abordagem DEA com técnicas de reamostragem, como *Bootstrap* e *Jackknife*, para eliminar o efeito dos *outliers* e erros de mensuração na base de dados. Em seguida, utilizaremos técnicas econométricas (incluindo regressão quantílica) para investigar os determinantes dos índices de eficiência calculados. Ênfase especial será dada a investigação (a) das economias de escala que prevalecem no setor e (b) dos impactos da estrutura da propriedade e do tipo de gestão sobre a produtividade hospitalar.

Este trabalho está organizado da seguinte maneira. A Seção 2 apresenta e discute a metodologia utilizada para o cálculo dos níveis de eficiência dos municípios brasileiros. A Seção 3 apresenta a base de dados do estudo, comenta a escolha dos insumos e produtos hospitalares escolhidos e descreve brevemente os aspectos computacionais. A Seção 4 discute os escores de eficiência obtidos e a Seção 5 analisa os determinantes desses escores. Por fim, a Seção 6 reúne algumas conclusões do trabalho e sugere extensões.

## **2. ESTIMAÇÕES ROBUSTAS DE FRONTEIRAS DE EFICIÊNCIA: AS ABORDAGENS DEA E JACKSTRAP**

A teoria econômica da análise de eficiência é baseada no trabalho de Koopmans (1951) e Debreu (1951), sobre análise da produtividade. O trabalho de Farrell (1957) foi o primeiro trabalho empírico onde o problema da mensuração da eficiência para um conjunto de unidades produtivas foi analisado. Os conceitos básicos e a notação do modelo DEA apresentado por Simar e Wilson (2000) são descritos

a seguir. As unidades produtoras se deparam com uma restrição do conjunto de produção  $\Psi$  dos pontos factíveis  $(x,y)$ :

$$[1] \Psi = \{x \in \mathfrak{R}_+^{p+q} | x \text{ pode produzir } y\}.$$

em que  $x \in \mathfrak{R}_+^p$  é um vetor de insumo e  $y \in \mathfrak{R}_+^q$  é um vetor de produto.

Esse conjunto pode ser descrito matematicamente por suas seções. Por exemplo, no espaço dos insumos, temos que os requisitos de insumos para todo produto  $y \in \Psi$ :

$$[2] X(y) = \{x \in \mathfrak{R}_+^p | (x, y) \in \Psi\}.$$

A fronteira de eficiência radial (*input-oriented*)  $\partial X(y)$  é definida por:

$$[3] \partial X(y) = \{x | x \in \mathfrak{R}_+^p, \theta x \notin X(y) \forall 0 < \theta < 1\}.$$

$\partial X(y)$  define os pares “radialmente eficientes”  $(x,y)$  e a medida de Farrel para eficiência para um determinado ponto  $(x,y)$  é definida por:

$$[4] \theta(x, y) = \inf \{\theta | \theta x \in X(y)\}.$$

$\theta(x, y)$  dá a contração radial do uso dos insumos que a firma localizada em  $(x,y)$  deve realizar para se considerada “insumo tecnicamente eficiente”. Dado o nível de insumo  $y$  e a direção (dos insumos) determinada por  $x$ , o nível eficiente de insumos é dado por:

$$[5] x^a(y) = \theta(x, y)x.$$

Como as distâncias radiais são consideradas coordenadas polares de  $x$ , isto pode ser usado para descrever a posição de  $(x,y)$ :

Módulo:  $\omega = \omega(x) \in \mathfrak{R}_+^1$

$$[6] \text{Ângulo: } \eta = \eta(x) \in \left[0, \frac{\pi}{2}\right]^{p-1}.$$

Em situações práticas gerais, quando da análise particular de um setor produtivo, o conjunto  $\Psi$  é desconhecido, assim como  $X(y), \partial X(y)$  e  $\theta(x, y)$ . Tipicamente, somente uma amostra das unidades produtivas é observada:

$$[7] \mathcal{X}_n = \{(x_i, y_i), i = 1, \dots, n\}.$$

A abordagem da “fronteira determinística” assume que  $(x_i, y_i) \in \Psi, i = 1, \dots, n$ . Em outras palavras, todas as unidades observadas são livres de erros de mensuração, de ruídos, etc. O problema então, é estimar os valores desconhecidos listados acima usando  $\mathcal{X}_n$ . Suponha que seja possível obter, através

de uma técnica,  $\hat{\psi}, \hat{X}(y), \hat{\theta}(x, y)$  que são estimadores das quantidades correspondentes de  $\psi, X(y)$  e  $\theta(x, y)$ , baseado na amostra de  $\mathcal{X}_n$ .

Os modelos não paramétricos mais usados são estimadores de  $\psi$  definidos como os mínimos conjuntos que contêm as informações observadas em  $\mathcal{X}_n$ . O estimador de *Free Disposal Hull* (FDH) é baseado em livre disponibilidade de  $\psi$ , enquanto que o *Data Envelopment Analysis* (DEA) utiliza adicionalmente a hipótese de convexidade.

No sentido de apresentar a descrição do modelo DEA, é preciso assumir algumas hipóteses para os dados observados sobre  $\psi$ . A primeira hipótese é que os insumos e produtos contidos em  $\psi$  observam a livre disponibilidade (*free disposal*), i.e., se  $(x, y) \in \psi \Rightarrow (\tilde{x}, \tilde{y}) \in \psi \quad \forall (\tilde{x}, \tilde{y})$  tal que  $\tilde{x} \geq x$  e  $\tilde{y} \leq y$ . Por vezes, essa hipótese será reforçada pela hipótese de convexidade, i.e.,  $\psi$  convexo. Uma outra hipótese é de que as observações amostrais de  $\mathcal{X}_n$  são realizações de variáveis aleatórias iid em  $\psi$  com função de densidade de probabilidade (fdp)  $f(x, y)$ , descrita por coordenadas cilíndricas  $(x, y) \Leftrightarrow (\omega, \eta, y)$  tal que:

$$[8] \quad f(\omega, \eta, y) = f(\omega | \eta, y) f(\eta, y) f(y)$$

na qual todas as densidades condicionais existem. Em particular,  $f(y)$  é definida em  $\mathfrak{R}_+^q$ ,  $f(\eta, y)$  é definida em  $\left[0, \frac{\pi}{2}\right]^{p-1}$  e condicional em  $(\eta, y)$ . Para um dado  $(\eta, y)$  um ponto da fronteira  $x^\partial(y)$  tem módulo

$$[9] \quad \omega(x^\partial(y)) = \inf \left\{ \omega \in \mathfrak{R}_+^1 \mid f(\omega | \eta, y) > 0 \right\}.$$

É importante notar que a medida de eficiência de Farrell (1957) de um ponto  $(x, y)$  é:

$$[10] \quad 0 \leq \theta(x, y) = \frac{\omega(x^\partial(y))}{\omega(x)} \leq 1, \text{ uma função do módulo correspondente.}$$

Por último, para todo  $(x, y)$  no interior de  $\psi$ ,  $\theta(x, y)$  é diferenciável nos dois argumentos.

Conforme anteriormente, o modelo DEA foi proposto por Farrell (1957) e operacionalizado por Charnes et al. (1978) em termos de programação linear. No caso da orientação pelos insumos, assumindo  $\psi$  convexo, o conjunto solução pode ser estimado pelo cone convexo de  $\hat{\psi}$ :

$$[11] \quad \hat{\psi}_{DEA}(\mathcal{X}_n) = \left\{ (x, y) \in \mathfrak{R}_+^{p+q} \mid y \leq \sum_{i=1}^n \gamma_i y_i, x \geq \sum_{i=1}^n \gamma_i x_i, \sum_{i=1}^n \gamma_i = 1, \gamma_i \geq 0 \quad \forall i = 1, \dots, n \right\}.$$

Na abordagem DEA, a eficiência associada ao ponto  $(x_0, y_0)$  é dada pela estimativa relativa à fronteira de  $\hat{\psi}_{DEA}$ :

$$[12] \hat{\theta}_{DEA}(x_0, y_0) = \inf \{ \theta | (\theta x_0, y_0) \in \hat{\psi}_{DEA}(\mathcal{X}_n) \}.$$

Da mesma forma, a estimativa dos níveis eficientes de insumos correspondentes ao ponto  $(x_0, y_0)$  é dada por:

$$[13] \hat{x}^\partial(y_0) = \hat{\theta}_{DEA}(x_0, y_0)x_0.$$

Novamente, por construção, sabe-se que  $\hat{\psi}_{DEA} \subseteq \hat{\psi}$  tal que:

$$[14] \partial \hat{X}_{DEA}(y_0) = \left\{ x \mid x \in \mathfrak{R}_+^p(x, x_0) \in \hat{\psi}_{DEA}(\mathcal{X}_n), (\theta x, y_0) \notin \hat{\psi}_{DEA}(\mathcal{X}_n) \quad \forall \quad 0 < \theta < 1 \right\}.$$

é um estimador que subestima  $\partial X(y_0)$  e  $\hat{\theta}_{DEA}(x_0, y_0)$  é um estimador que superestima  $\theta(x_0, y_0)$ . Além disso,  $\theta(x_0, y_0) \leq \hat{\theta}_{DEA}(x_0, y_0) \leq 1$  para  $\theta(x_0, y_0) \in \hat{\psi}_{DEA}$ .

## 2.1 Estimação Robusta dos Escores de Eficiência: O Método “Jackstrap”

Um problema crítico nas aplicações *naives* do método DEA advém da grande sensibilidade dessa abordagem em presença de *outliers* e de heterogeneidade dos dados. Esse problema afeta, particularmente, a mensuração da eficiência técnicas dos hospitais brasileiros, em razão não somente do grande número e da diversidade dessas instituições, mas também devido a problemas de erros e omissões, inclusive aqueles decorrentes do sistema de incentivos, embutidos nos esquemas de reembolso do SUS (que podem levar ao aumento excessivo do número de procedimentos e/ou dos custos). Esse último ponto é particularmente importante quando se considera que o método DEA é seriamente afetado pela presença dessas observações supereficientes e por erros de mensuração, que tendem a subestimar consideravelmente os escores computados. Como esse método baseia-se no conceito (extremo) de fronteira, composta por um pequeno número de entidades, uma única observação muito discrepante pode deslocar a fronteira aumentando, assim, artificialmente, os requisitos em termos de eficiência para o conjunto dos dados. Portanto, para assegurar a credibilidade dos índices de eficiência calculados é crucial a correção desse tipo de problema. Somente assim poder-se-á obter indicadores que sejam úteis para a avaliação do desempenho dos hospitais brasileiros.

Alguns métodos foram propostos para lidar com essa questão [Banker e Gifford (1998), Banker e Chang (2000), Andersen e Petersen (1993), Simar e Wilson (1995), (2000)]. Essas abordagens têm em comum a utilização de inspeção manual dos dados, o que pode ser considerada inadequada para base de dados como a usada neste trabalho. Recentemente, Sampaio de Sousa e Stošić (2003) propuseram um método baseado em uma combinação de técnicas de amostragem como *Jackknife* e *Bootstrap*<sup>1</sup> para contornar esse problema. A essência dessa abordagem consiste em reduzir estocasticamente o impacto das DMUs mais influentes nas medidas DEA finais, usando o conceito de *leverage* [Cribari e Zarkos, (2003)]. Esse conceito (*leverage*) mede, para cada DMU, o impacto de sua retirada sobre os escores de eficiência computados para as demais observações, isto é, mensura o deslocamento da fronteira de eficiência, quando uma determinada DMU está ausente ao conjunto de dados. A seguir, discutiremos

---

<sup>1</sup> Ver Efron e Tibshirani (1993)

brevemente, esse método. Maiores detalhes podem ser encontrados em Sampaio de Sousa e Stosic (2003).

Considere o conjunto  $\{\theta_k | k = 1, \dots, K\}$ . Removendo-se, sucessivamente, cada DMU, e recalculando-se, para cada remoção, o conjunto de eficiência  $\{\theta_{kj}^* | k = 1, \dots, K; k \neq j\}$ , onde o índice  $j = 1, \dots, K$  representa a DMU retirada, obtém uma medida de influência (*leverage*,  $l_j$ ) para  $j$ -ésima observação:

$$[15] \quad l_j = \sqrt{\frac{\sum_{k=1, k \neq j}^K (\theta_{kj}^* - \theta_k)^2}{K-1}}$$

A aplicação desse procedimento requer uma capacidade computacional muito intensa e pode ser não factível para amostras muito grandes, no nível computacional disponível<sup>2</sup>. Por esse motivo, foi desenvolvido uma forma alternativa, mais eficiente, que combina as técnicas de *Bootstrap* e *Jackknife*. Esse novo procedimento seleciona aleatoriamente um subconjunto de  $L$  DMUs (tipicamente 10% de  $K$ ) e utiliza a expressão [15] para obter os *leverages*  $\tilde{l}_{jl}$ , onde  $j$  assume valores diferentes em  $L$  a partir do conjunto  $\{1, \dots, k\}$ . Repete-se,  $B$  vezes, esse procedimento, acumulando um subconjunto  $\tilde{l}_{jB}$  para todas as DMUs selecionadas aleatoriamente (para  $B$  suficientemente grande, cada DMU deveria ser selecionada  $n_j \approx \frac{BL}{K}$  vezes).

O passo seguinte consiste em calcular o *leverage* médio para cada DMU como  $\tilde{l}_j = \frac{\sum_{b=1}^{n_j} \tilde{l}_{jB}}{n_j}$  e o *leverage*

$$\text{global como } \tilde{l} = \frac{\sum_{j=1}^K \tilde{l}_j}{K}.$$

Pode-se associar várias funções de distribuição de probabilidades para se eliminar as DMUs com elevado *leverage*. Neste trabalho, foi escolhida a função *Heaviside* dada por:

$$[16] \quad P(\tilde{l}_k) = \begin{cases} 1, & \tilde{l}_k \leq \tilde{l} \log K \\ 0, & \tilde{l}_k > \tilde{l} \log K \end{cases}.$$

O critério de corte sugerido por essa função indica que se  $P(\tilde{l}_k) = 0$ , a  $k$ -ésima DMU deve ser retirada da amostra,  $\forall k = 1, \dots, K$ . A aplicação do procedimento acima torna o DEA mais confiável como provedor de medidas de performance<sup>3</sup> provando ser robusto na presença de dados discrepantes e/ou

<sup>2</sup> Sampaio de Sousa e Stosic (2003) argumentam que para 5000 DMUs uma sequência de DEA requer cerca de 1h num Pentium III de 1 Ghz. Todas as seqüências rodadas exigiriam cerca de 7 meses de cálculos contínuos.

<sup>3</sup> Para mais detalhes, ver Sampaio de Sousa e Stošić (2003)<sup>2</sup>

erros de medidas na base de dados, tendo a capacidade de detectar as observações atípicas mesmo se elas não estão na fronteira (ineficiências extremas). Sendo um método completamente automático, não exige a inspeção manual dos dados e é particularmente adequado para grandes bases de dados. Essa técnica é particularmente apropriada para nosso estudo, já que se trata de um grande e heterogêneo conjunto de informações, onde a existência de informações discrepantes restringiria, em muito, a credibilidade da fronteira computada por meio da versão simples do método DEA.

### 3. DADOS E VARIÁVEIS UTILIZADOS

No Brasil, o Sistema Único de Saúde (SUS) é responsável pela maior parte das internações hospitalares. Por essa razão, os dados usados nesse trabalho foram obtidos através de informações do DATASUS, fornecidas pelo Ministério da Saúde. A amostra inclui 1270 hospitais que atendem total ou parcialmente pela rede do Sistema Único de Saúde. Por critério de comparação, foram analisados apenas hospitais que realizaram, dentre outros serviços, procedimentos cirúrgicos de diversas naturezas para o ano de 2002. Essas instituições atualizam seus cadastros junto ao sistema do DATASUS, provendo uma gama de informações detalhadas sobre sua estrutura física e a quantidade de pessoas que lá trabalham. Isto inclui, por exemplo, o número de atendimentos produzidos, o nível e a especialização de seus profissionais, a quantidade de leitos existentes no hospital e a quantidade de leitos contratados pelo SUS, dado que estão incluídos hospitais públicos e privados, que podem ceder uma parte ou toda a sua capacidade para os pacientes usuários do sistema único. Além disso, o banco de dados contém informações qualitativas sobre a natureza da organização do hospital (privado, público, beneficente, universitário, etc).

A função de produção calculada foi composta de três produtos (*outputs*) e seis insumos (*inputs*), conforme pode ser visto na Tabela 1. Todas as variáveis consideradas são quantitativas e sofreram algum tipo de agregação no sentido de captar a maior quantidade de informações disponível, porque nem todas as observações dispunham de informações, quando se considerava uma maior desagregação. Isso elimina o número de zeros existentes no banco de dados, o que é muito importante para o cálculo da fronteira.

**Tabela 1 – Insumos, Produtos e os Correspondentes Serviços Hospitalares**

Indicadores	Sigla	Serviços hospitalares os quais os indicadores servem como <i>proxies</i>
<i>Indicadores de Produtos</i>		
Total de procedimentos não cirúrgicos realizados	TFREQ	Serviços de complexidade média ou baixa
Total de procedimentos cirúrgicos realizados	CFREQ	Serviços de complexidade mais elevada
Inverso do número de óbitos ocorridos	OBITINV	Provisão de saúde
<i>Indicadores de Insumos</i>		
Agentes de saúde, auxiliares e técnicos	AGAUXTEC	Número de pessoal
Enfermeiros, Médicos e Dentistas	ENFDENTMED	Número de pessoal específico da área médica
Profissionais de nível médio e superior	PROFISSIONAIS	Outros tipos de pessoal
Valor médio total dos serviços prestados	TAValue	Valor médio de mercado dos serviços
Permanência média dos procedimentos não cirúrgicos (em dias)	TAVPERM	Tempo que o paciente fica no hospital
Permanência média relativa às cirurgias (em dias)	CAVPERM	Tempo que o paciente fica no hospital

Fonte: DATASUS- Informações do Sistema Único de Saúde- Ministério da Saúde- Governo Federal.

No tocante aos produtos, as duas primeiras variáveis descrevem, respectivamente, a quantidade de serviços que não sejam cirúrgicos prestados à população e a quantidade de cirurgias realizadas pelos hospitais. Conforme já mencionado, essas variáveis foram agregadas e servem, também, como *proxies* para o tamanho do hospital. O terceiro produto considerado é o inverso da taxa de óbito, que procura mensurar todo o aparato físico e profissional do hospital<sup>4</sup>. Pelo lado dos insumos, as três primeiras variáveis agregam os profissionais que constam no cadastro dos hospitais e que envolvidos na produção de serviços. Conforme sugerido pela literatura [Mckillop *et al* (1999)] e [Barbetta, Turati e Zago (2001)], criamos uma variável para o número de médicos, dentistas e enfermeiros. Por fim, a função de produção leva em conta o valor total médio dos atendimentos e a permanência média em dias, controlada para a ocorrência de cirurgias.

#### 4. NÍVEIS DE EFICIÊNCIA DAS UNIDADES DE PRESTAÇÃO DE SAÚDE NO BRASIL: DISCUSSÃO DOS RESULTADOS

Aplicamos, em primeiro lugar, a versão *naïve* do modelo DEA, para retornos variáveis de escala. Verificou-se, de imediato, que a fronteira estimada era fortemente afetada pela presença de *outliers*. Analisando-se, cuidadosamente, essa fronteira, vê-se que algumas DMUs são capazes de transformar poucos insumos em muitos produtos, algo que é muito improvável quando se supõe que o nível tecnológico dessas instituições não está muito distante da média encontrada no Brasil, a menos que essas diferenças representem erros nos dados. Por exemplo, temos os casos de Casa de Saúde Santa Severa (PA) e Unidade Mista de Autazes (AM) que reportaram, respectivamente, uma quantidade de médicos igual a 2 e 3, tendo realizado 1194 e 426 cirurgias, além de 1726 e 406 outros procedimentos médicos. O Consórcio Intermunicipal de Saúde do Alto Sertão, em Sergipe, apresentou apenas um médico realizando em um ano 2621 cirurgias. Exemplos como esses ocorrem em outros estados do Brasil, sugerindo que pode haver sobre-notificação de informações, em uma tentativa da instituição de receber maiores repasses federais, ou simplesmente erro na coleta de dados.

Para corrigir esses problemas, utilizamos o método Jackstrap para detectar *outliers*, associando-os aos maiores valores de *leverage*. Após o cálculo das medidas de *leverage*, foi necessário encontrar um ponto de corte (*cut-off*)  $l_0 \in \mathfrak{R}$  no qual todas as DMUs com *leverage* maiores ou iguais a  $l_0$  seriam excluídas da amostra. O ponto de corte utilizado foi aquele dado pela distribuição *Heaviside*, o que implicou na retirada de 97 observações. O corte indicado por essa regra mostra que o número de DMUs eficientes aumentou. Para concluir se esta distribuição da variável eficiência é diferente entre os grupos, foi feito um teste de Smirnov-Komolgorov<sup>5</sup>, no qual a hipótese nula de igualdade entre a distribuição dos índices de eficiência dados pelo DEA simples e pelo método *Jackstrap* foi rejeitada. O método *Jackstrap* detectou informações distorcíveis do banco de dados, provando ser de grande utilidade analítica por evitar inspeções manuais no banco de dados. Essa vantagem é muito importante para banco de dados muito extensos como o do DATASUS/MS.

Com base no conjunto de observações, resultante da aplicação do método Jackstrap, computamos os escores de eficiências para as 1170 unidades restantes. Esses resultados estão apresentados na Tabela 2. Nota-se, em primeiro lugar, que não somente a grande maioria das instituições analisadas é ineficiente, mas também os níveis médios de eficiência são muito baixos. Isto aplica-se, particularmente, às instituições beneficentes. Em média, os melhores escores são atribuídos às entidades comandadas por

---

<sup>4</sup> Utiliza-se o inverso dos óbitos porque para o DEA *-output oriented-* maximiza o produto e ter o número de óbitos elevados é considerado uma desvantagem.

<sup>5</sup>  $H_0$  testa se a variável  $\theta$  possui a mesma distribuição entre os grupos.

sindicatos; porém a grande dispersão desses índices, mensurada pelo elevado desvio padrão, exige cautela nessa afirmação. No que concerne o tipo de unidade vale ressaltar os baixos índices atribuídos aos postos de saúde. Por fim, embora a percentagem de unidades eficientes nos hospitais privados seja superior àquelas que prevalece nos hospitais públicos, a eficiência média tende a ser mais baixa.

**Tabela 2**  
Caracterização dos Estabelecimentos Segundo a Natureza da Administração, o Tipo de Unidade e a Esfera Administrativa -

Natureza	# de DMUs	Eficientes		Ineficientes				Eficiência Média (todas as DMUs)
		Total	%	Total	%	Mean	Std. Dev.	
<i>Natureza da Adminis.</i>								
Sindicato	25	5	20,00	20	80,00	0,356	0,251	0,484
Cooperativa	30	7	23,33	23	76,67	0,252	0,076	0,426
Beneficente	460	17	3,70	443	96,30	0,283	0,165	0,312
Administração Direta	262	21	8,02	241	91,98	0,357	0,223	0,409
<i>Tipo de Unidade</i>								
Policlínica	22	4	18,18	18	81,81	0,285	0,276	0,415
Pronto-Socorro	24	2	8,33	22	91,66	0,346	0,216	0,401
Postos de Saúde	20	0	0	20	100	0,262	0,120	0,262
Hospital Especializado	93	7	7,53	86	92,47	0,358	0,228	0,406
Hospital Geral	993	51	5,14	942	94,86	0,306	0,188	0,342
<i>Esfera Administrativa</i>								
Pública	319	22	3,89	297	93,10	0,354	0,218	0,399
Privada	852	49	5,75	803	94,25	0,300	0,183	0,340

Várias razões podem ser invocadas para explicar esses resultados. Em primeiro lugar, isto ocorre porque, aparentemente, os hospitais brasileiros não trabalham com a escala mínima exigida e, portanto, não utilizam, de maneira ótima, os recursos disponíveis. Este argumento encontra apoio em estudos empíricos recentes (ver, por exemplo, Grannemann, Randall e Pauly (1986)). As economias de escala ocorrem por várias razões. Em grandes unidades, os custos fixos - tais como equipamentos, mão de obra especializada e espaço físico - são distribuídos entre um grande número de pacientes. Além disso, hospitais maiores reduzem a incerteza inerente à provisão dos serviços de saúde e, portanto, tendem a apresentar maiores taxas de ocupação exigindo, assim, menores reservas para atender as demandas de pico. O maior tamanho leva, também, à eficiência via especialização e os efeitos "*practice makes perfect.*" Por fim, essas unidades, por comprar em grandes quantidades, têm maior poder de barganha junto aos seus fornecedores, o que contribui para reduzir seus gastos. No nosso estudo, esse ponto está bem ilustrado pela Tabela 3, que mostra a distribuição dos escores de eficiência, segundo a dimensão do estabelecimento. Vê-se, claramente, que esses escores são mais elevados para as unidades maiores.

Para complementar a identificação dos retornos de escala, para as instituições analisadas, utilizamos o teste de Färe, Grosskopf e Lovell (FGL), que compara as diferentes hipóteses de retorno de escala, implicadas pelo método DEA - constantes, variáveis e não crescentes. A tabela 3 mostra, para diferentes tamanhos das instituições, medidos pelo número total de procedimentos médicos prestados, a distribuição das economias de escala locais.

Conforme esperado, os retornos crescentes de escala predominam largamente neste tipo de setor, para todos os tamanhos de hospital. A grande maioria das entidades opera na parte decrescente das suas curvas de custo médio. Isso significa que há espaço para amortização dos custos fixos (e portanto para a redução dos custos médios) pela ampliação do número de serviços prestados. Do ponto de vista da

análise de custo-eficiência, um grande número de pequenos hospitais não é desejável para a sociedade como um todo, quando comparado com um número menor de hospitais capazes de realizar mais atendimentos à população. Esse resultado é confirmado tanto pelo teste FGL como pelo teste de Banker que, para esse estudo, apresentaram resultados praticamente idênticos.<sup>6</sup>

Tabela 3 Distribuição dos Índices de Eficiência Segundo o Tamanho do Estabelecimento

Número Total de Atendimentos	Índices de Eficiência						
	média	mediana	desvio	curtose	Assimetria	Max	min
0 a 999	0,145	0,151	0,037	0,42	-0,83	1	0,010
1000 a 1999	0,241	0,242	0,026	-1,19	0,07	1	0,198
2000 a 2999	0,325	0,327	0,022	-1,29	0,04	1	0,289
3000 a 4999	0,460	0,450	0,067	-0,99	0,43	1	0,365
> 5000	0,858	0,950	0,155	-1,45	-0,50	1	0,592

Note-se, por fim, que a existência de retornos crescentes de escala para os hospitais explica também porque o envoltória dos dados pelo método DEA-C é precário. De fato, apenas 84 das 1170 instituições analisadas (7,2%) atingem a fronteira de eficiência, caracterizada por retornos constantes de escala, postulada por esse método (Tabela 3).

Tabela 3 – Rendimentos de Escala para Intervalos de Atendimentos

Número de Atendimentos	# de DMUs	Retornos Constantes de Escala		Retornos Crescentes de Escala		Retornos Decrescentes de Escala	
		#	%	#	%	#	%
0 a 999	338	38	11.24	293	86.69	7	2.07
1000 a 1999	294	13	4.42	281	95.58	0	0
2000 a 2999	169	7	4.14	162	95.86	0	0
3000 a 4999	186	12	6.45	174	93.55	1	0.54
> 5000	183	14	7.65	168	91.80	0	0

Fonte: Nossos Cálculos

## 5. O MODELO ECONOMETRICO

Com o intuito de aprofundar a análise anterior no que diz respeito à natureza da organização, o tipo de unidade e a esfera administrativa foi feita uma análise econométrica dos determinantes dos indicadores de eficiência. Como os hospitais analisados se diferenciam significativamente sobre vários aspectos e apresentaram uma distribuição de eficiência estimada pouco concentrada, o modelo de regressão quantílica (Koenker e Basset, 1978) se faz apropriado para explicar por que os hospitais mais eficientes

<sup>6</sup> No que se refere à mensuração das economias de escala nos modelos DEA, embora Banker, Chang e Cooper (1996) tenham mostrado que a abordagem de Banker (1984) e o enfoque de Färe, Grosskopf e Lovell (1985) produzem resultados equivalentes, pequenas diferenças persistem entre esses métodos deixando alguns pontos em aberto nessa discussão, em particular no que tange à computação das economias de escala.

são afetados diferentemente daqueles de menor eficiência, mesmo quando controlados pelas mesmas características. Para tal regressamos, o logaritmo da eficiência calculada no item anterior pelo modelo DEA-Jackstrap (medida  $\theta$ ). Esta medida é uma combinação linear de um conjunto de variáveis e está restrita ao intervalo fechado  $[0, 1]$ <sup>7</sup>, fazendo com que o estimador de mínimos quadrados ordinários (OLS) do vetor de parâmetros seja inconsistente no sentido de que não converge em probabilidade para o verdadeiro parâmetro desconhecido. No entanto, a literatura mostra que o uso de  $\ln(\theta)$  como variável dependente gera estimadores de mínimos quadrados consistentes e não-viesados, se os valores computados são estritamente positivos<sup>8</sup> [Banker & Natarajan, (2001) e Valdmanis & Blank (2002)].

Os três aspectos considerados consistem em um conjunto de características qualitativas disponíveis para a rede SUS. A natureza da organização, que inclui informações sobre a administração do hospital, indicando se a instituição é beneficente, se pertence à administração direta, ou se é controlada por sindicato ou cooperativa. Essas variáveis são usadas no SUS para avaliar a oferta de serviços de saúde nesses estabelecimentos. Os conceitos das regressoras são, para *Sindicato*, entidade associativa civil de uma ou mais categorias profissionais, com personalidade jurídica de direito privado, que pode desenvolver atividades de assistência a seus associados. Por *Cooperativa*, entende-se instituição civil de direito privado, constituída por membros de determinado grupo social que objetivem atividades em comum. A variável *Entidade Privada* se refere a fundações de personalidade jurídica autônoma de direito privado, sendo atividade pública ou altruísmo. A variável *Administração direta* se refere à administração por órgão governamental de saúde, de administração direta, em qualquer esfera administrativa. Por fim, as *Entidades Benéficas*, incluem as entidades de direito privado sem fins lucrativos, que desenvolvem atividades benéficas ou de assistência social.

O tipo de unidade reúne informações sobre o nível de atuação (abrangência) da instituição. As policlínicas (variável *Policlínica*) são unidades de saúde de atendimento ambulatorial em várias especialidades, incluindo ou não as especialidades básicas, com atendimento 24h. A variável *Pronto-Socorro* representa as unidades destinadas à prestação de atendimento a pacientes com ou sem risco de vida, cujos agravos necessitam atendimento imediato, especializado ou geral. A variável *Postos* corresponde aos postos de saúde, que são unidades destinadas à prestação de serviços a uma população, de forma programada ou não, por profissional de nível médio, com a presença de intermitente ou não do profissional médico. Por *Hospital Especializado*, entende-se hospital destinado à prestação de assistência à saúde em uma única especialidade/área, podendo dispor de serviço de emergência. Finalmente, *Hospital Geral* inclui hospitais destinados à prestação de serviços de especialidade básica, podendo incluir emergência. Essas características foram as presentes em maior número no banco de dados e se originam da classificação do SUS, que possui ainda outras classificações.

A esfera administrativa traz informações sobre a responsabilidade da administração dos recursos da instituição, declarando se o hospital é de administração federal, estadual, municipal ou de esfera privada. Na modelagem que se segue, as informações foram agregadas em esfera pública e esfera privada. Todas as regressões apresentadas nas tabelas seguintes têm a mesma variável dependente  $\ln(\theta)$  e uma variável de *escala* que representa a o logaritmo da quantidade total de serviços produzidos pelo estabelecimento. Esta é uma medida que captura o tamanho absoluto da instituição. Os resultados obtidos por OLS constam na tabela 4 e servem de comparação com os resultados obtidos pelo modelo de regressão quantílica. Com dados em *cross-section*, é comum os erros serem heteroscedásticos e tanto para o modelo OLS descrito abaixo, como para os modelos de regressão quantílica subsequentes,

---

<sup>7</sup> O modelo Tobit não é apropriado. Para detalhes, ver Souza e Tabak (2002).

<sup>8</sup> Outro tratamento dado ao problema é visto em Xue e Harker (1999).

os estimadores utilizaram a matriz de correção de White (1980) e o resultado para a significância dos parâmetros foi satisfatório.

Tabela 4 – Modelos OLS para os Grupos de Informações

Variáveis Explicativas	Variável Dependente: Escores de Eficiência das Unidades de Saúde		
	Natureza da Organização	Tipo de Unidade	Esfera Administrativa
Intercepto	-2.09 (0,17)	-1.85 (0,18)	-2.06 (0,18)
Escala	0.12 (0,02)	0.13 (0,22)	0.12 (0,02)
Sindicato	-0.17 (0,03)	-	-
Cooperativa	-0.24 (0,13)	-	-
Beneficente	-0.14 (0,04)	-	-
Policlínica	-	-0.17 (0,09)	-
Pronto-Socorro	-	-0.42 (0,18)	-
Postos de Saúde	-	-0.58 (0,17)	-
Hospital Especializado	-	-0.42 (0,12)	-
Hospital Geral	-	-0.43 (0,09)	-
Esfera Privada	-	-	-0.095 (0,04)
Número de Observações	1170	1170	1170
Crítério de Akaike	1.932	1.929	1.940
R2-Ajustado	0.055	0.056	0.044

Valores em parênteses são estimativas do desvio-padrão.

A tabela 5 abaixo mostra que os resultados variam, significativamente, com os níveis de eficiência das unidades analisadas. Conforme esperado, em razão da existência de economias de escala, quanto maior o volume de serviços, menor o custo médio do procedimento e, portanto, maior o nível de eficiência. Esse impacto é maior para os quantis inferiores sugerindo que os estabelecimentos ineficientes, o são, em parte, em virtude de trabalharem em uma escala aquém da que seria considerada ótima. Esse resultado difere, porém, para o quantil superior. Nesse caso, o efeito escala não é mais significativo. Isto deve ao fato de que os estabelecimentos mais eficientes são, em sua maioria, hospitais de grande porte, que já amortizaram seus custos fixos e estão trabalhando, provavelmente, na parte crescente da curva de custos médios, onde predominam as deseconomias de escala.

O fato de a gestão de o estabelecimento ser feita por um *sindicato* afeta positivamente os escores dos hospitais mais ineficientes e, negativamente, aqueles onde prevalecem as melhores práticas de gestão de recursos. Os hospitais regidos por sindicatos têm como características a elevada centralização de suas decisões e a atuação em benefício de seus dirigentes, ao invés da comunidade que eles representam. Isto sugere que a sindicalização funciona bem em pequenos hospitais, quando a escala ótima de produção não foi atingida e não existe congestionamento nos serviços prestados. Em grandes estabelecimentos, as restrições associadas à gerência sindical - incluindo-se aí, a questão corporativa - tornam-se mais custosas e, portanto, prejudicam o desempenho dessas unidades.

Fenômeno semelhante acontece com os estabelecimentos regidos por cooperativas. O efeito da gerência por cooperativas é positivo quando os hospitais são ineficientes e negativo quando apresentam melhor desempenho. Novamente, a organização cooperativa parece ser vantajosa para as unidades que se situam abaixo da escala ótima, já que isto reduz os níveis de eficiência. Já os hospitais com melhor

performance perdem com a adoção de práticas cooperativas provavelmente porque seu número de atendimentos de saúde adequa-se melhor a estruturas mais formais de gerenciamento.

No que se refere às instituições *beneficentes*, os resultados obtidos não validam as teses avançadas na literatura teórica de acordo com as quais esse tipo de instituição tenderia a apresentar níveis superiores de eficiência. Eles são corroborados pela evidência empírica existente que tende a demonstrar que as instituições beneficentes não são mais eficientes na provisão de serviços hospitalares quando comparadas com os demais hospitais privados e públicos. [(Duggan, 2000), Barbata, Turati e Zago (2001)]. Note-se, ainda, que enquanto para os quantis inferiores, não parece haver ligação entre o fato de ser beneficente e o nível de eficiência, para os quantis mais elevados, o fato de ser beneficente, quando outros fatores são controlados, reduz claramente os níveis de eficiência.

Finalmente, os hospitais de *administração direta* contribuem para a eficiência para o 75º e o 90º quantis e diminuem a eficiência para o 10º quantil. Isso está de acordo com a idéia de que maior controle do órgão governamental sobre os recursos disponíveis para a prestação de atendimentos de saúde afeta positivamente o número de pessoas atendidas por diversos procedimentos médicos e por cirurgias, além de diminuir o número de óbitos. Esse resultado reforça o papel das agências fiscalizadoras e reguladoras do segmento de saúde na sociedade, como pode ser visto em Kittelsen & Magnussen (2003).

Tabela 5 – Modelo de Regressão Quantílica para a Natureza da Organização

Variáveis Explicativas	Variável Dependente: Escores de Eficiência das Unidades de Saúde					
	5º quantil	10º quantil	25º quantil	50º quantil	75º quantil	90º quantil
<i>Intercepto</i>	-4,5 (0,24)	-4,31 (0,13)	-3,63 (0,11)	-2,63 (0,11)	-1,55 (0,32)	-0,074 (0,31)
<i>Escala</i>	0,32 (0,03)	0,32 (0,02)	0,27 (0,01)	0,18 (0,01)	0,10 (0,04)	-0,007 n/s
<i>Sindicato</i>	0,86 (0,07)	0,57 (0,04)	0,27 (0,03)	-0,09 (0,023)	-0,58 (0,05)	-1,25 (0,04)
<i>Cooperativa</i>	0,51 (0,18)	0,23 (0,13)	-0,08*** (0,17)	-0,07 (0,23)	-0,48 (0,32)	-1,02 (0,16)
<i>Beneficente</i>	0,15 (0,1)	0,024*** (0,05)	-0,04*** (0,04)	-0,15 (0,03)	-0,30 (0,07)	-0,41 (0,06)
<i>Admdireta</i>	0,12*** (0,11)	-0,1 (0,06)	-0,02*** (0,05)	0,03*** (-0,04)	0,14 (0,09)	0,10 (0,07)

Valores em parênteses são estimativas do desvio-padrão. \*, \*\*, \*\*\* significante respectivamente a 1%, 5%, 10%. n/s: não significante.

No que se refere ao estudo para o tipo de unidade, os resultados da tabela 6 mostram que o efeito escala também está presente e perde força à medida que o hospital se torna mais eficiente, isto é, já atingiu a escala ótima de produção. Importante notar que para o 80º quantil o coeficiente de *escala* é estatisticamente diferente de zero apenas a 10% de nível de significância. O quantil mais elevado neste modelo é o que inclui 80% das observações porque para quantis maiores, o modelo não se ajustou, após as iterações computacionais. Enquanto às *policlínicas*, a eficiência é negativamente afetada para os níveis maiores de eficiência (50º, 75º e 80º quantis). Isto pode ser justificado pela pouca especificidade dos tratamentos realizados neste tipo de unidade, porque elas são dotadas, em geral, de profissionais

pouco especializados, que podem se deparar com situações onde são obrigados a encaminharem um grande número de casos, reduzindo o total de atendimentos, o que está de acordo com o resultado do modelo OLS simples da tabela 4. Nos quantis em que se situam os hospitais mais ineficientes (5º, 10º e 25º quantis), o efeito é positivo e decrescente porque trata-se de hospitais de com quantidades menores de atendimentos e de pouco nível de complexidade. Conforme esperado, a variável *pronto-socorro* afeta negativamente a eficiência para todos os quantis (sua magnitude é muito próxima se os quantis mais elevados forem comparados). A justificativa para a esse efeito é que, de regra geral, os casos mais graves de acidentes acabam sendo atendidos em unidades deste tipo, que incorrem num elevado número de óbitos, por maior que seja o número de atendimentos realizados. Os *postos* de saúde, sendo estabelecimentos de menor porte e complexidade, sofrem com o problema do efeito da escala ótima de produção e esse efeito negativo sobre a performance é proporcional ao quantil analisado. Por fim, comparam-se os hospitais gerais e especializados. Os *hospitais especializados* confirmam a suspeita do chamado *overtreatment* já verificado na literatura por Agrell e Bogetoft (2001). Estes hospitais tentem a elevar a permanência do paciente, o que diminui a sua eficiência técnica, principalmente no 25º quantil. Os *hospitais gerais* também tiveram coeficientes negativos para todos os quantis, mas o efeito por trás disso é inverso ao dos hospitais especializados e como no caso das policlínicas, é maior no quantis superiores. Isto porque os hospitais gerais, além de terem um porte maior e de incluírem serviços de urgência em suas atividades, não são capazes de prover tratamento adequado para uma ampla gama de especialidades inerentes ao tipo de demanda com que se defrontam, sendo obrigados a encaminhar muitos casos para outras instituições, como ocorre com as policlínicas. Este resultado sugere que deve haver uma combinação ótima entre a especialização e a generalidade do hospital.

Tabela 6 – Modelo de Regressão Quantílica para o Tipo de Unidade

Variáveis Explicativas	Variável Dependente: Escores de Eficiência das Unidades de Saúde					
	5º quantil	10º quantil	25º quantil	50º quantil	75º quantil	80º quantil
<i>Intercepto</i>	-4.13 (0.26)	-3.77 (0.17)	-3.47 (0.12)	-2.42 (0.14)	-1.16 (0.37)	-0.76 (0.40)
<i>Escala</i>	0.3 (0.03)	0.28 (0.02)	0.28 (0.01)	0.21 (0.01)	0.10 (0.04)	0.091 (0.05)***
<i>Policlínica</i>	1.16 (0.16)	0.89 (0.1)	0.57 (0.07)	-0.05 n/s	-0.7 (0.18)	-1.01 (0.19)
<i>Pronto-socorro</i>	-0.8 (0.32)	-0.41 (0.20)	-0.4 (0.13)	-0.55 (0.13)	-0.44 (0.31)	-0.47 (0.32)
<i>Postos</i>	0.37 (0.22)	0.11 (0.18)***	-0.2 (0.18)	-0.75 (0.3)	-0.8 (0.41)	-1.07 (0.38)
<i>Hospital Especializado</i>	-0.6 (0.21)	-0.72 (0.14)	-0.48 (0.10)	-0.41 (-0.09)	-0.28 (0.22)	-0.46 (0.22)
<i>Hospital Geral</i>	-0.23 (0.16)	-0.30 (0.11)	-0.34 (0.08)	-0.54 (0.07)	-0.55 (0.17)	-0.71 (0.18)

Valores em parênteses são estimativas do desvio-padrão. a) O modelo não se ajustou para quantis maiores do que 80%.

\*, \*\*, \*\*\* significante respectivamente a 1%, 5%, 10%. n/s: não significante.

O último grupo de informações trata da esfera administrativa do estabelecimento. Os resultados constam na tabela 7. Novamente, o efeito escala está presente, exceto para o último quantil, em que o coeficiente estimado não é significativo, cabendo a explicação apresentada anteriormente. Quanto aos hospitais situados nos quantis mais elevados, o fato de serem de esfera privada afeta negativamente sua

eficiência. Isto é consistente com a existência de falhas de mercado no setor de saúde. Além do efeito escala já mencionado, existe informação assimétrica entre pacientes e médicos. Estes afetam a decisão do paciente de consumir um serviço ou um bem, que em geral não são homogêneos, não acontecendo o mesmo em mercados com informação perfeita.

Tabela 7 – Modelo de Regressão Quantílica para a Esfera Administrativa

Variáveis Explicativas	Variável Dependente: Escores de Eficiência das Unidades de Saúde					
	5º quantil	10º quantil	25º quantil	50º quantil	75º quantil	80º quantil
<i>Intercepto</i>	-4.36 (0.29)	-4.44 (0.13)	-3.68 (0.12)	-2.68 (0.13)	-1.29 (0.39)	0.2 n/s
<i>Escala</i>	0.32 (0.03)	0.32 (0.04)	0.27 (0.01)	0.19 (0.03)	0.07 n/s	-0.4 n/s
<i>Esfera Privada</i>	0.15 (0.10)	0.14 (0.41)	-0.06 (0.01)	-0.1 (0.03)	-0.17 (0.09)	-0.23 (0.09)

Valores em parênteses são estimativas do desvio-padrão. n/s: não significante. \*,\*\*,\*\*\* significante respectivamente a 1%, 5%, 10%.

Este setor caracteriza-se, ainda, pela presença de externalidades, dado que alguns serviços de saúde, como por exemplo a vacinação e o tratamento de doenças contagiosas, são bens públicos e os preços relativos de mercado dos produtos ficam distorcidos. Também, é razoável supor que nas instituições privadas o fenômeno da demanda induzida é mais intenso, o que revoga a hipótese de soberania do consumidor, dependendo de como o paciente revela suas informações privadas.<sup>9</sup> Por fim, os hospitais privados têm incentivos a aumentar artificialmente o número de tratamentos oferecidos à população e oferecer tratamentos relativamente mais custosos porque dessa maneira, podem conseguir maiores repasses da rede SUS e aumentarem suas receitas.

À guisa de conclusão, nossos resultados situam-se do lado do debate, baseado em trabalhos empíricos, que mostra que os hospitais públicos tendem a ser mais eficientes que seus equivalentes privados, incluindo-se aí, as instituições sem fins lucrativos [Barbetta, Turago e Zago (2001), Valdmanis (1992), Dugan (2002)].

## 6. CONCLUSÃO

Este trabalho estimou uma fronteira de produção não paramétrica para os hospitais da rede SUS que realizaram cirurgias para o ano de 2002. A partir dessa fronteira, computada mediante o uso do modelo DEA – *Data Envelopment Analysis* com o método “*Jackstrap*”, investigamos o impacto da estrutura da propriedade e do tipo de gestão sobre o desempenho dessas entidades. .

Confirmando as expectativas *a priori*, os estabelecimentos mais eficientes foram aqueles que apresentaram um maior número de cirurgias e outros procedimentos não cirúrgicos para baixo tempo médio de permanência no hospital, controlados pelo custo médio e pelo número de pessoal e quantidade de capital humano qualificado envolvidos nas atividades de cada instituição.

<sup>9</sup> Há ainda a não monotonicidade e a quebra da não saciedade local das preferências. Para mais detalhes, ver Tandon, Murray e Lauer (1999) entre outros.

No sentido de caracterizar os determinantes da eficiência dos hospitais, o passo seguinte foi utilizar a abordagem de regressão quantílica para modelar a eficiência técnica em função de um conjunto de variáveis qualitativas para três grupos distintos de informações, usado pelo sistema SUS sendo estes a natureza da organização, o tipo de unidade e a esfera administrativa. Em todos os modelos foi incluída a variável de escala para captar o efeito das economias de escala, presentes neste setor, caracterizados por elevados custos fixos.

Nossos resultados sugerem que as economias crescentes de escala são prevalentes na maioria das instituições analisadas, independente da natureza, tipo de unidade e esfera administrativa. Isto porque, em virtude do pequeno tamanho, a maioria das unidades não pode se beneficiar das economias de escala inerente à produção dos serviços hospitalares. Isso justifica a política atual de consorciação hospitalar adotada por muitas municipalidades como uma forma de reduzir os custos médios desses os serviços.

No que diz respeito ao debate sobre a influência da natureza destacam-se ainda os efeitos negativos da especialização excessiva dos hospitais sobre a eficiência. As estimativas apontam para a existência de uma combinação ótima entre especialização e generalização dos atendimentos. Além disso, estabelecimentos de esfera privada tiveram sua performance negativamente afetada para os hospitais acima do 30º quantil de eficiência. Isso sugere que os efeitos das falhas de mercado são importantes para o comportamento dos hospitais, isto é, a existência de informação assimétrica nesse mercado, presença de economias de escala e de externalidades na produção fazem com que haja a necessidade de intervenção do setor público como agente complementar onde a provisão de serviços é subótima. Adicionalmente, nesse setor o problema da demanda induzida e o esquema de incentivos do SUS pode fazer com que os estabelecimentos privados passem a aumentar artificialmente o número e o tipo de atendimentos, viesando-os para os de custos mais elevados, a fim de elevar suas receitas.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

- Agrell, J, & Bogetoft, P. (2001). “Should Health Regulators Use DEA?” *Discussion Paper # 32*, Department of Economics, KLV, Denmark.
- Arrow, Kenneth J. (1963) “Uncertainty and the Welfare Economics of Medical Care.” *American Economic Review* 21: 941-973.
- Banker, R.D and Gifford, J.L. (1988) “A Relative Efficiency Method for the Evaluation of Public Health Nurse Productivity.” Mimeo.
- Banker, D.R, Chang, H. “Evaluating the Super-Efficiency Procedure in Data Envelopment Analysis for Outlier Identification and for Ranking Efficient Units.” *Working Paper Series University of Texas at Dallas*, Nov. 2000.
- Banker, R.D., Natarajan, R. “Evaluating Contextual Variables Affecting Productivity Using Data Envelopment Analysis”, *Working Paper Series – University of Texas at Dallas*, Apr. 2001.
- Charnes, A; Cooper, W.W; Rhodes,E. (1978) “Measuring the Efficiency of Decision Making Units”. *European Journal of Operational Research* 2: 429-444.
- Duggan, M. (2000) “Hospital Ownership and Public Medical Spending.” *Quarterly Journal of Economics* 115: 1343–1374.
- Dugan, M. (2002) “Hospital Market Structure and the Behavior of Not-for-Profit Hospitals.” *RAND Journal of Economics* 33: 433–446
- Evans, D.B., Tandon, A., Murray, C., Lauer,J. (1999). “The Comparative Efficiency of National Health Systems in Producing Health: An Analysis of 191 Countries”. *GPE Discussion Paper # 29*, World Health Organization.

- Fama, E. and M. Jensen (1983) "Separation of Ownership and Control." *Journal of Law and Economics* 26:301-325.
- Farrel, M.J. (1957). "The Measurement of Production Efficiency", *Journal of Royal Statistical Society, Series A, CXX, Part 3*, 253-290.
- Glaeser, E. L. and Shleifer, A. (2001) "Not-For-Profit Entrepreneurs." *Journal of Public Economics*, 81: 99-115.
- Grannemann, Thomas, Randall Brown and Mark Pauly (1986) Estimating Hospital Costs: A Multiple-Output Analysis, *Journal of Health Economics* 5: 107-127.
- Hansmann, H., D. P. Kessler and Mark B. McClellan (2002) "Ownership Form and Trapped Capital in the Hospital Industry" Yale Law and Economics research paper No. 266.
- Kittelsen, S.A.C, & Magnussen, J.(2003). "Economies of Scope in Norwegian Hospital Production – A DEA Analysis" *Working Paper # 8, Health Economics Research Programme – University of Oslo*.
- Koenker, R. (1981). "A Note on Studentizing a Test for Heteroskedasticity." *Journal of Econometrics*, 17, 107-112.
- Koenker, R. and Bassett, G. (1978) "Regression Quantiles." *Econometrica* 46: 33-50.
- Long, J.S. and Ervin, L.H. (2000) "Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model." *The American Statistician* 54: 217-224.
- Lovell, C. A K., Walters, L.C. and Wood, L.L. (1994) "Stratified Models of Education Production Using DEA and Regression Analysis." In Charnes, A, Cooper, W.W, A.Y.
- Marinho, A. (2001). "Estudo de Eficiência em alguns Hospitais Públicos e Privados com a Geração de Rankings". *Texto para Discussão 794*, IPEA: Rio de Janeiro, 2001.
- Simar, L e Wilson, P.W. (1997) "Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models". *Management Science* 44, 1, 49-61.
- Simar, L. (2003) "Detecting Outliers in Frontiers Models: A Simple Approach." Discussion Paper 0146, Institut de Statistique, UCL, Louvain-la-Neuve, Belgium, forthcoming (2003) in *Journal of Productivity Analysis*.
- Simar, L. and Wilson, P. (2000) "Statistical Inference in Nonparametric Frontier Models: The State of the Art", *Journal of Productivity Analysis* 13, 49-78.
- Sloan, F., and A. Steinwald (1980). "Effects of Regulation on Hospital Costs and Input Use." *Journal of Law and Economics*, 23:81-109.
- Stosic, B.D. and Sampaio de Sousa, M.C. (2003) "Jackstrapping DEA Scores for Robust Efficiency Measurement." Working Paper N° 291, Department of Economics, University of Brasilia, Brazil
- Souza, G.S., Tabak, B.M. (2002). "Factors Affecting the Technical Efficiency of Production of the Brazilian Banking System: A Comparison of Four Statistical Models in the context of DEA" *Working Paper do Banco Central do Brasil #17*, Nov.2002.
- Valdmanis, V.G, & Blank, J.L.T. (2002) "A Modified Three Stage DEA: An Application to Home for Mentally Disabled in Netherlands", Mimeo.
- Wilson, P. (1993) "Detecting Influential Observations in Data Envelopment Analysis." *Journal of Productivity Analysis* 6: 27-45.
- Weissbrod, B. (1998) *The Non Profit Economy*. Cambridge, Massachusetts: Harvard University Press.
- Xue, M. & Harker, P.T (1999). "Overcoming the Inherent Dependency of DEA Efficiency Scores: A Bootstrap Approach" *Working Paper -University of Pennsylvania - The Wharton School*, 1999. Mimeo.