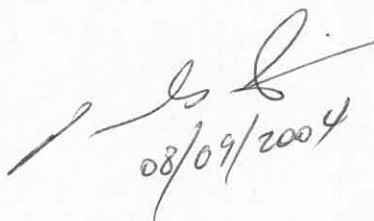


LETÍCIA HARUE ATA

**PRODUTIVIDADE NO SETOR DE SAÚDE PÚBLICA:  
Uma abordagem de fronteira de custos**

Trabalho apresentado ao  
Programa Institucional de  
Bolsas de Iniciação Científica  
(PIBIC).



Handwritten signature and date: 08/09/2004

Orientador:  
Prof. Fernando Garcia

SÃO PAULO  
2004

## RESUMO

O presente trabalho possui como objetivo o estudo da produtividade no setor de saúde pública no Brasil. Além disso, procura identificar as diferenças de eficiência entre as unidades da federação a partir do conceito de fronteira de custos. A metodologia consistiu na coleta de dados que englobam os anos de 1996 a 1998, que foram empregados na estimação da fronteira estocástica de custo. Identificamos uma relação positiva e crescente entre o custo de mão-de-obra dos hospitais públicos e o número de internações, indicando a relação de custo marginal crescente. Também foram identificadas relações positivas do custo com mão-de-obra com dias de internação, equipamentos e participação do salário na renda. Em contrapartida, encontramos uma relação negativa entre o custo de mão-de-obra e o número de leitos e hospitais, indicando a ocorrência de retornos de escala. Além disso, os estados possuem diferenças significativas de eficiência técnica e estão crescentemente ficando menos eficientes tecnicamente.

**Palavras chaves:** Fronteira Estocástica de Custo, Produtividade Total dos Fatores, Saúde Pública, Progresso Técnico, Eficiência Técnica, Eficiência Alocativa.

## SUMÁRIO

INTRODUÇÃO .....	4
1 CARACTERÍSTICAS ESTRUTURAIS E DINÂMICAS DO SISTEMA DE SAÚDE BRASILEIRO.....	6
2 FUNDAMENTOS PARA A ANÁLISE DA PRODUTIVIDADE.....	9
2.1 A FUNÇÃO PRODUÇÃO.....	9
2.2 CUSTOS DE PRODUÇÃO .....	13
2.3 OS CONCEITOS DE INEFICIÊNCIA TÉCNICA E ALOCATIVA.....	17
3 ESTIMAÇÃO DA FRONTEIRA ESTOCÁSTICA .....	25
3.1 CONCEITO DE FRONTEIRA .....	25
3.2 FRONTEIRA DETERMINÍSTICA DE PRODUÇÃO.....	25
3.3 FRONTEIRA ESTOCÁSTICA DE PRODUÇÃO .....	29
3.4 ESTIMAÇÃO DUAL DA EFICIÊNCIA TÉCNICA .....	30
4 ESTIMAÇÃO DA FRONTEIRA ESTOCÁSTICA DE CUSTO PARA SAÚDE PÚBLICA BRASILEIRA .....	34
4.1 DADOS E AMOSTRAS .....	34
4.2 ASPECTOS METODOLÓGICOS ASSOCIADOS A CADA UMA DAS VARIÁVEIS.....	37
4.3 FRONTEIRA DE CUSTOS DE MÃO-DE-OBRA E ANÁLISE DOS RESULTADOS .....	39
CONSIDERAÇÕES FINAIS .....	47
REFERÊNCIAS BIBLIOGRÁFICAS .....	48

## INTRODUÇÃO

O Brasil enfrenta uma crise na saúde pública caracterizado pela existência de uma desigualdade social em saúde favorável aos grupos mais privilegiados economicamente, conforme o artigo "Desigualdade social no acesso aos serviços de saúde no Brasil" de Andrade e Noronha revela. Neste contexto, a distribuição eqüitativa desses serviços visa ao tratamento igualitário entre os indivíduos com as mesmas necessidades médicas, independentemente da sua situação socioeconômica.

A implantação do SUS, Sistema Único de Saúde, tinha como intuito garantir a oferta universal, íntegra e gratuita de assistência médica no Brasil. Entretanto, as restrições de orçamento e a deterioração dos serviços públicos fizeram com que, no mercado de serviços de saúde, outro segmento, a rede privada, cuja oferta ocorre mediante o pagamento direto ou de um plano de saúde privado, tivesse uma demanda progressiva.

A Pesquisa Nacional por Amostra de Domicílios (PNAD) do IBGE estimou em R\$ 37,9 bilhões o gasto privado anual com saúde das famílias em 1998, que inclui gastos com planos de saúde, consultas médicas, consultas com outros profissionais de saúde, exames, medicamentos, artigos ortopédicos e aparelhos médicos, óculos e lentes, odontologia, hospitais, enfermagem domiciliar e outros gastos com saúde. O acesso, portanto, à rede privada de saúde possui como restrição a renda do indivíduo. Quanto maior a renda, maior o acesso aos serviços médicos e hospitalares.



O PNAD, em 1998, agrupou os indivíduos segundo os decis de renda. As pesquisas apontaram, que à medida que sobe a renda familiar per capta, cai a participação do gasto com medicamentos e sobe com planos de saúde e com odontologia no gasto total com saúde das famílias. Outra questão importante abordada diz respeito às pessoas com emprego informais, que normalmente ganham por hora trabalhada e estão menos protegidas por leis trabalhistas. Como resultado, a procura por atendimento médico acarreta uma maior perda na renda que outros vínculos empregatícios. Isto leva a concluir que os mais pobres, que compõe a maior parte da mão-de-obra informal, possuem um custo de oportunidade mais elevado para procurar auxílio médico. Além disso, a probabilidade de internação hospitalar aumenta para os mais pobres. Isso porque eles têm menos acesso aos tratamentos preventivos e consultas rotineiras, procurando assistência em momentos de saúde já bastante precários e graves.

Outro aspecto relevante no que se refere aos serviços de saúde é a questão da inovação tecnológica e a expansão do conhecimento ocorrida no setor, especialmente nas últimas décadas. Tal fenômeno decorre do investimento intensivo em pesquisa para descobrir novos e mais eficientes medicamentos, tratamentos, técnicas, processos e diagnósticos. Fato este ilustrado pelo recente mapeamento do genoma humano. De um lado, a inovação dos processos gera um aumento dos custos na compra de medicamentos ou na realização de exames, por exemplo. De outro, há um barateamento na medida que se pode atingir o mesmo nível de produção economizando fatores produtivos como capital e trabalho.

Dada a relevância dos aspectos citados referentes à saúde, em especial, a pública, procuraremos por intermédio desse trabalho entender de modo mais aprofundado esse assunto, bem como da produtividade nesse setor no período pós-real.

## **1 CARACTERÍSTICAS ESTRUTURAIS E DINÂMICAS DO SISTEMA DE SAÚDE BRASILEIRO**

Nesta sessão estudaremos as características estruturais e dinâmicas do sistema de saúde brasileiro.

Um dos objetivos da Constituição Federal brasileira de 1988 era resgatar e enorme dívida social brasileira herdada do regime militar brasileiro, conforme Marques (1999) aponta. Houve uma mudança de foco e de compreensão do que é saúde pública e como tratá-la. O governo passa a ter como foco a saúde e não na doença. (<http://www.brasil.gov.br>). Assim sendo, a Constituição de 1988 estabeleceu diretrizes gerais para a organização do Sistema Único de Saúde (SUS).

Este visava a amenizar os problemas de saúde no Brasil. Para isso, as seguintes diretrizes foram estabelecidas: descentralização, com direção única em cada esfera do governo; atendimento integral, com prioridade para as atividades preventivas, sem prejuízo dos serviços assistenciais; participação da comunidade. A descentralização estabelece que a execução de serviços seria comandada e realizada pelos Municípios e pelos Estados, cabendo à União as atividades de planejamento, acompanhamento, fiscalização e controle.

A eficácia e eficiência do sistema, entretanto, mostram-se muito questionável principalmente no que diz respeito à garantia de acesso e à qualidade dos serviços de saúde. Além disso, a descentralização ocorre de modo limitado. Há insuficiência de recursos, atrelados, sem dúvida, ao enorme déficit público; bem como na sua aplicação inadequada. Outra razão é a falta de sinergia entre a esfera privada e pública.

Ao invés de haver realmente um sistema único de saúde, observamos basicamente dois pólos: o sistema público de saúde e o sistema privado.

O primeiro, que abrange o SUS, mostra-se atrelado à massa de trabalhadores vinculados aos setores economicamente atrasados, empobrecidos, os socialmente excluídos-desempregados e subempregados; ou seja, aqueles que não possuem renda suficiente para ter acesso ao serviço privado. Têm, pois, de recorrer aos serviços públicos de saúde, financiados historicamente através do fundo público administrado pelo Estado.

O segundo está fortemente associado aos trabalhadores dos setores econômicos de ponta e partícipes do processo de globalização econômica, em conjunto com as elites e os estratos médios e altos da sociedade. O acesso a esse sistema mostra-se limitado para aproximadamente 25% da população nacional, sendo que grande parte desta está concentrada nas regiões sul e sudeste. O financiamento desses serviços decorre de sistemas coletivos de pré-pagamento administrados por empresas privadas a partir do desembolso das empresas. Os recursos destinados para esses fins podem ser ressarcidos através dos mecanismos de renúncia fiscal do Imposto de Renda. Assim sendo, o sistema privado, fundado na renúncia fiscal, impossibilita uma redistribuição dos recursos em pró a rede pública.

O setor privado pode ser dividido em dois principais segmentos: lucrativo e não lucrativo. O lucrativo, com exceção de uma parcela da rede hospitalar, não mantém qualquer tipo de vínculo com o SUS. O não lucrativo abrange instituições filantrópicas, especialmente Santas Casas de Misericórdia, cuja uma parcela significativa é vinculada ao SUS. Com relação às modalidades de produção e gestão de serviços podemos destacar: medicina de grupo, cooperativas médicas, seguro saúde, autogestão e planos de administração.

O sistema privado, além disso, oferece uma grande versatilidade e gama de opções aos seus usuários e aos clientes potenciais como, por exemplo, planos com preços e serviços diferenciados; justificando, em parte, o crescimento desse setor nos últimos anos.

Percebemos, portanto, a existência de uma forte desigualdade e exclusão com relação aos serviços de saúde. O SUS necessitaria para sua implementação, a estruturação de um amplo projeto que contemple a desigualdade e a exclusão social, de maneira que possa orientar a operacionalização da descentralização em moldes diversos daquela baseado apenas no financiamento.

## 2 FUNDAMENTOS PARA A ANÁLISE DA PRODUTIVIDADE

Veremos nessa sessão fundamentos básicos da teoria econômica para a análise da produtividade.

### 2.1 A FUNÇÃO PRODUÇÃO

Ela define a quantidade máxima de um produto  $Y$  que pode ser produzida a partir da cesta de insumos. Por exemplo, descreve a relação entre os insumos do processo produtivo - capital ( $K$ ), trabalho ( $L$ ) e tecnologia ( $A$ ) - e os níveis de produção ( $Y$ ), conforme ilustrado na equação 2.1. A tecnologia, forma pela qual os fatores produtivos são combinados, determina a produtividade dos fatores já que estabelece as relações entre cada fator de produção e o volume produzido.

$$Y = F(A, K, L) \quad (2.1)$$

Conforme observado por Garcia (2003), há três hipóteses referentes a essa função: (i) retornos constantes de escala (ii) ausência de produção livre e (iii) retornos decrescente dos fatores produtivos.

A função de produção de Cobb-Douglas (2.2) é uma das mais utilizadas e satisfaz tais condições.

$$Y = F(K, A \cdot L) = K^{\alpha} \cdot (A \cdot L)^{(1-\alpha)}, \quad 0 < \alpha < 1 \quad (2.2)$$

A primeira premissa (retornos constantes de escala) significa que, caso dobrássemos a quantidade de insumo, a quantidade produzida também dobraria, conforme os cálculos comprovam:

$$\begin{aligned}
 F(\lambda \cdot K, \lambda \cdot A \cdot L) &= (\lambda \cdot K)^\alpha \cdot (\lambda \cdot A \cdot L)^{(1-\alpha)} = \\
 &= \lambda^\alpha \cdot \lambda^{(1-\alpha)} \cdot K^\alpha \cdot (A \cdot L)^{(1-\alpha)} = \lambda \cdot K^\alpha \cdot (A \cdot L)^{(1-\alpha)} = \\
 &= \lambda \cdot F(K, A \cdot L).
 \end{aligned}$$

Sendo  $\alpha$  uma constante e maior que zero.

A ausência de produção livre pode ser comprovada ao supusermos que  $K = L = 0$ . Substituindo os valores, o valor adicionado é nulo ( $Y=0$ )

A terceira hipótese afirma que, conforme aumenta um insumo  $X$ , mantendo-se todos os outros constantes, o produto adicional, ou seja, marginal, diminui. Isto pode ser ilustrado na figura 2.1, que mostra uma função produto total, que nada mais é que a função de produção aplicada a um insumo fixo e que mostra qual será a quantidade produzida (produto total), conforme a quantidade de insumo variável. Do ponto  $G$  ao  $C$ , percebe-se que a inclinação da reta tangente vai diminuindo, ou seja, a unidade adicional de insumo (marginal) aumenta o produto a taxas decrescentes.

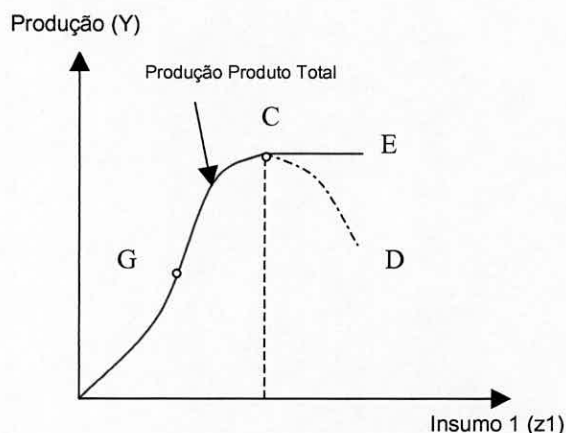


Figura 2.1: Função de produto total

Outra forma de evidenciar essa última premissa é analisando a segunda derivada da função. Caso ela tenha sinal negativo, significa que há retornos

decrecentes dos fatores, conforme observado a seguir na função de Cobb-Douglas para capital e trabalho.

$$\begin{aligned}\frac{\partial F}{\partial K} &= \alpha \cdot K^{(\alpha-1)} \cdot (A \cdot L)^{(1-\alpha)} > 0 & \frac{\partial F}{\partial L} &= (1-\alpha) \cdot K^\alpha \cdot (A \cdot L)^{-\alpha} \cdot A > 0 \\ \frac{\partial^2 F}{\partial^2 K} &= \alpha \cdot (\alpha-1) \cdot K^{(\alpha-2)} \cdot (A \cdot L)^{(1-\alpha)} < 0 & \frac{\partial^2 F}{\partial^2 L} &= -\alpha \cdot (1-\alpha) \cdot K^\alpha \cdot (A \cdot L)^{(-1-\alpha)} \cdot A < 0\end{aligned}$$

Isso aplica-se à questão de saúde. Por exemplo, um médico regular atende 15 pacientes em uma noite. Caso um outro profissional seja contratado, ele poderá atender menos pacientes uma vez que as estruturas fixas do hospital e o número limitado de leitos são fatores restritivos.

É importante ainda observar na função de Cobb – Douglas que o conhecimento (A) aparece multiplicando apenas o fator trabalho (L). Isto revela que um aumento do conhecimento tem efeito apenas sobre a produtividade da mão-de-obra. Tal especificação é denominada Harrod – neutra. Para a equação 2.2, devido à possibilidade de alcançar o mesmo nível de produção Y com uma menor quantidade de trabalho, pode-se afirmar que o processo de inovação tecnológica é essencialmente poupador de trabalho.

Existem ainda outras especificações. A decisão de qual é mais adequada depende de critérios empíricos, cujas informações são dificilmente obtidas. Assim sendo, tal característica na função de produção é deixada de lado.

Existem outras formas da expressão Cobb- Douglas menos restritas, como, por exemplo, a 2.3:

$$Y = A \cdot K^\alpha \cdot L^\beta, \quad \alpha, \beta > 0 \quad (2.3)$$



Sendo  $\alpha$  e  $\beta$  frações positivas que representam elasticidades. Por exemplo, um aumento de 1% em trabalho aumentará  $Y$  em  $\beta$  por cento. As condições de retornos constantes de escala e retornos decrescentes de fator podem ser desrespeitadas, a depender dos valores de  $\alpha$  e  $\beta$ .

A função produção mostra-se como uma restrição no problema na minimização dos custos, meio pelo qual a firma maximiza o seu lucro.

No longo prazo, período de tempo em que todos os insumos são variáveis, a firma pode tanto escolher quantidades ótimas dos insumos  $L$  e  $K$  que minimizem a soma gasta dos fatores. Já no curto prazo, a firma não pode variar todos os seus insumos. Assim, o único meio de minimizar os custos de produzir  $Y$  unidades de produto é escolhendo aumentando  $L$ .

O produto marginal representa a taxa em que a produção muda conforme a quantidade do insumo variável muda. Ele pode ser obtido através da derivada da função produto total em relação a essa variável.

O produto médio é dado pela inclinação da linha determinada pela origem e pelo ponto na função de produção. Ele mostra-se importante, pois permite examinar como a produção varia com mudanças na quantidade do insumo variável.

Quando o  $PMg > PMe$ , o  $PMe$  é crescente.

Quando o  $PMg < PMe$ , o  $PMe$  é decrescente.

Quando o  $PMg = PMe$ , o  $PMe$  é constante.

Na figura 2.2 podemos observar isso. Portanto, a partir do ponto B o produto médio começa a decrescer.



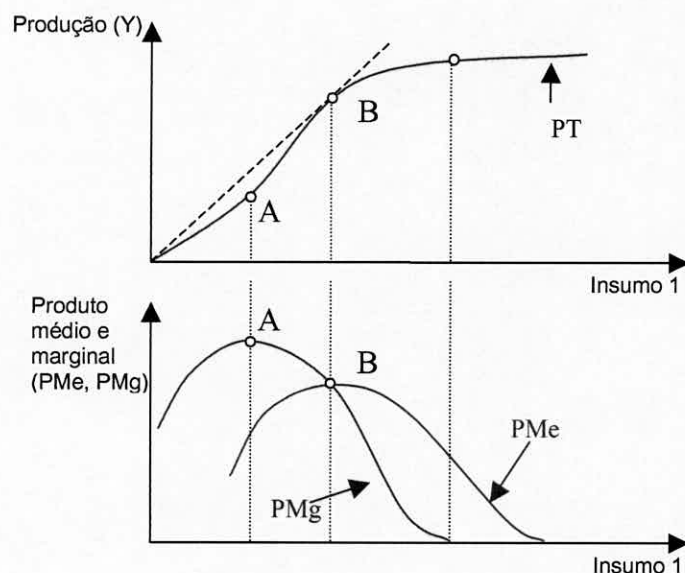


Figura 2.2: Comparações das funções de produto médio e produto marginal

## 2.2 CUSTOS DE PRODUÇÃO

Se soubermos o quanto uma determinada firma deseja produzir, poderemos descobrir a partir do produto total a quantidade mínima do insumo variável requerida para obter tal produção e, por conseguinte, minimizar os custos.

O problema de minimização de custos de curto prazo abrange a redução de  $wL$  pela escolha de  $L$ , respeitando a restrição  $y = PT(L)$ . A expressão (2.4) representa o custo variável, que é o custo variável mínimo necessário para produzir  $y$  unidades de produção.

$$CV(y) = wL \quad (2.4)$$

Analogamente à função produção, podemos derivar o custo variável médio e o custo variável marginal.

$$CVMe(y) = CV(y)/y$$

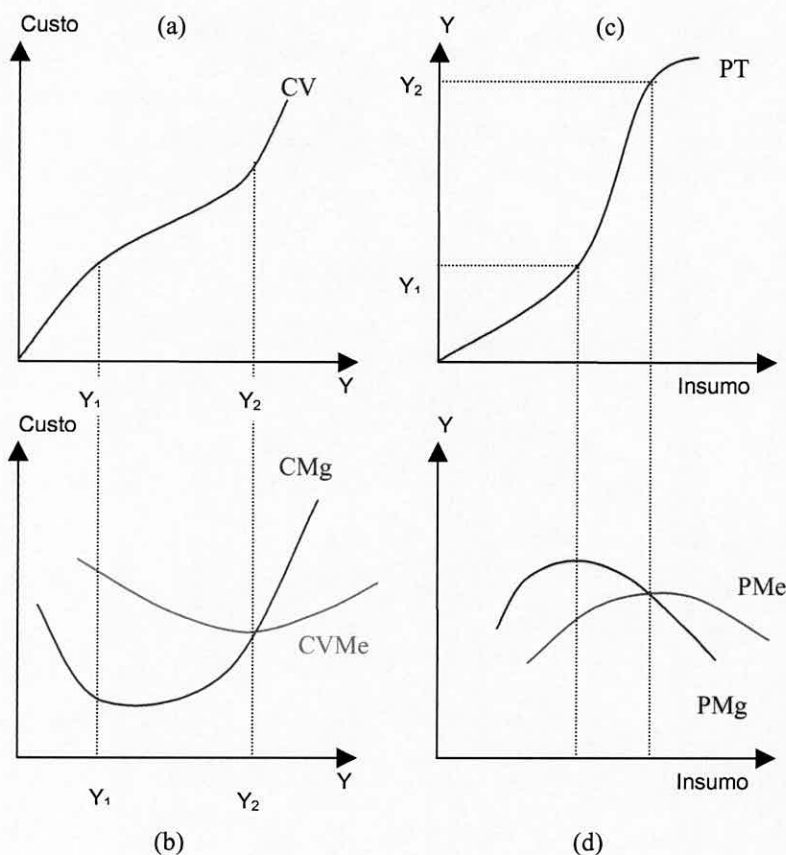
$CMg(y) = \text{inclinação de } CV(y)$

Quando  $CMg < CVMe$ , o  $CVMe$  é decrescente conforme aumenta-se  $y$

Quando o  $CMg > CVMe$ , o  $CVMe$  é crescente

Quando o  $CMg = CVMe$ ,  $CVMe$  é constante

As curvas de produto médio e do custo variável médio, dessa forma, mostram o mesmo tipo de relação bastante similar, mas de forma inversa, conforme percebemos na figura 2.3. Em (d), o  $PMg$  está no máximo quando a produção é  $Y_1$  e diminui depois. Em (b),  $CMg$  nesse ponto está no mínimo e depois aumenta. Portanto, a diminuição do produto marginal parece influir diretamente no aumento do custo marginal.



Além do custo variável, existe o custo fixo, que é o custo do insumo fixo, como por exemplo, aluguel.

A soma do custo fixo ao custo variável resulta no custo total, que representa o custo econômico total da produção, e na expressão (2.5):

$$CT_{cp}(y) = CV(y) + CF \quad (2.5)$$

O custo médio total, ou seja, o custo por unidade produzida é:

$$Cme_{cp}(y) = CT_{cp}(y) / y$$

As figuras 2.4 e 2.5 representam todos os custos citados e vistos de diferentes configurações.

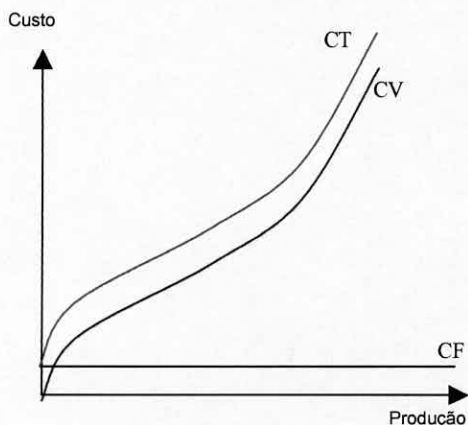


Figura 2.4: custo total como soma do custo variável mais fixo.

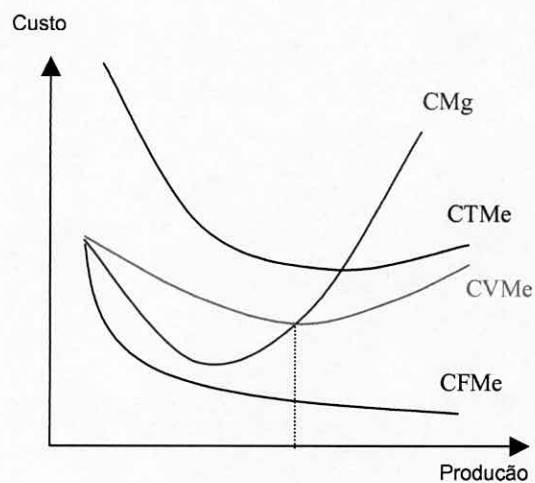


Figura 2.5: curvas de custo.

Dado um custo, podemos realizar diferentes combinações possível de capital (K) e de trabalho (L), formando uma linha de isocusto dada por:

$$C = wL + rK$$

Onde  $w$  e  $r$  são preços do trabalho e do capital respectivamente.

Na figura 2.6 são demonstradas várias linhas de isocustos associadas a isoquantas, que representam todas as combinações possíveis dos insumos K e L que geram o mesmo nível de produção:

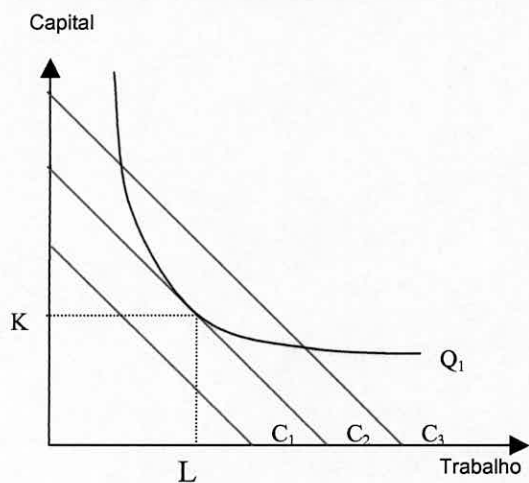


Figura 2.6: linhas de isocusto e isoquanta

O ponto de tangência da isoquanta Q1 com a linha de isocusto C2 representa os custos mínimos de produção. Percebemos que a combinação de minimizadora dos custos encontra-se na isoquanta, e não acima dela. Isto porque para qualquer combinação acima da isoquanta, há combinações mais baratas na isoquanta.

### **2.3 OS CONCEITOS DE INEFICIÊNCIA TÉCNICA E ALOCATIVA**

Conforme Pinilla (2001), a idéia de comparar segundo o comportamento das empresas é de grande interesse para a análise econômica. Desse modo, surgem conceitos como os de produtividade e eficiência e competitividade.

As empresas que maximizam seus benefícios se consideram eficientes. Essa maximização exige que elas decidam sobre:

- 1) a escolha, dentre todos os níveis de produção, do output que maximize o benefício.
- 2) a escolha de uma combinação de inputs que minimiza o custo da produção
- 3) a produção do output com a quantidade mínima de inputs possíveis.

A cada item podemos associar um tipo de eficiência respectivamente:

- 1) Eficiência de escala: quando uma empresa está produzindo em uma escala ótima, na qual permite maximizar o benefício.
- 2) Eficiência alocativa: quando a empresa combina os inputs na proporção que minimize seu custo de produção.
- 3) Eficiência técnica: quando uma empresa obtém o máximo output possível com a combinação de inputs empregada.

Tais tipos de eficiência teoricamente podem ser combinados de várias formas a fim de alcançar a eficiência total.

Apesar de o conceito de produtividade e eficiência serem distintos, elas são constantemente usadas como sinônimos, até mesmo em livros sobre o assunto.

A produtividade média de um fator só seria válida em situações com uma tecnologia de coeficientes fixos, já que de outra forma, não se leva em conta as possibilidades de substituição de inputs, fazendo com que sua utilização seja restrita.

Diversas vezes se faz referência ao conceito de produtividade média de um fator. Entretanto, se utilizarmos esse conceito como critério de comparação entre firmas com combinações diversas de inputs, pode-se gerar conclusões distorcidas. Por exemplo, primeiramente vamos supor que a firma A é intensiva em capital e B, em trabalho, bem como que ambas produzem a mesma quantidade de output. Se compararmos a produtividade média do trabalho de ambas, a firma B seria menos produtiva. Em contrapartida, se tomarmos como referência a produtividade média de capital, a firma A será menos produtiva e B mais.

A fim de corrigir essa distorção, introduziu-se o conceito de *Produtividade Total dos Fatores* (PTF), o qual se pode definir como um coeficiente entre uma soma ponderada de outputs (y) e uma soma ponderada de inputs (x):

$$PTF = \frac{\sum a_i y_i}{\sum b_j x_j}$$

onde  $a_i$  e  $b_j$  são respectivamente as ponderações de outputs e inputs.

Outro engano cometido muitas vezes é usar termos de eficiência e produtividade como sinônimos quando, na verdade, referem-se a aspectos diferentes da produção, sendo que nem sempre é bom um aumento dos mesmos.

Na figura 2.7, em que se representa uma função de produção com rendimentos decrescentes de escala, podemos perceber isso.

As firmas B e C, por estarem contidas na linha que representa a função de produção, são tecnicamente eficientes. Em contrapartida, A mostra-se tecnicamente ineficiente já que produz menos quantidade Y, com o mesmo volume de input X ao compararmos com B. Portanto, de "A" para "B" aumenta-se a eficiência técnica e também a produtividade do fator X.

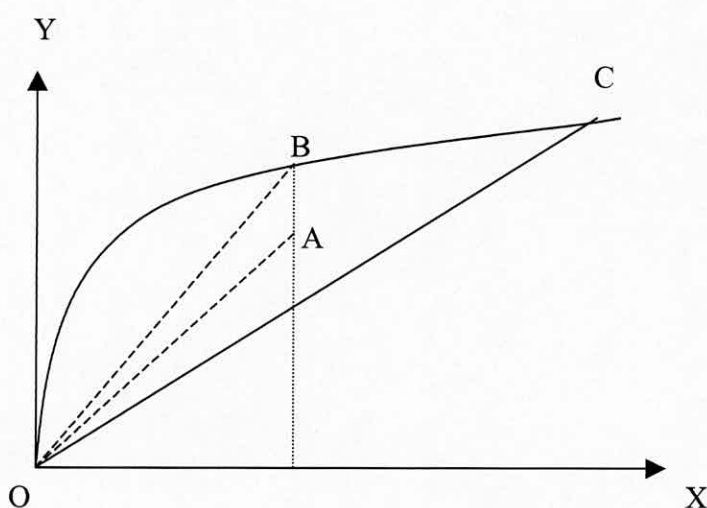


Figura 2.7: Diferenças entre eficiência técnica e produtividade média

Em contrapartida, de "A" a "C" implica um aumento da eficiência, mas uma diminuição da produtividade (razão entre Y e X, ou seja, o coeficiente angular das retas OA e OC).

Da mesma forma, de "B" a "C", ambas tecnicamente eficientes, diminui a produtividade porque o coeficiente angular, que representa a produtividade média, diminui. Portanto, nem sempre uma melhora na eficiência leva a uma melhora na produtividade, e vice – versa.

A medição da eficiência produtiva se baseia na idéia de comparar a atuação real da empresa com relação a um ótimo. Entretanto, um dos problemas é estimar

esse ótimo. Farrell (1957) propôs uma alternativa: a de comparar a firma com aquelas parecidas que atuam no mesmo mercado. Ele destacou-se por determinar empiricamente uma referência para comparar as eficiências. A hipótese fundamental utilizada era a possibilidade de trabalhar com níveis ineficientes, ao contrário das hipóteses clássicas. Para isso, ele usou como instrumental as fronteiras de produção.

Farrell ofereceu uma medida de eficiência global que podia ainda ser decomposta em eficiência técnica e alocativa. A seguir apresentaremos uma abordagem gráfica para dar a intuição sobre esses dois conceitos.

Figura 2.8 Ineficiência técnica

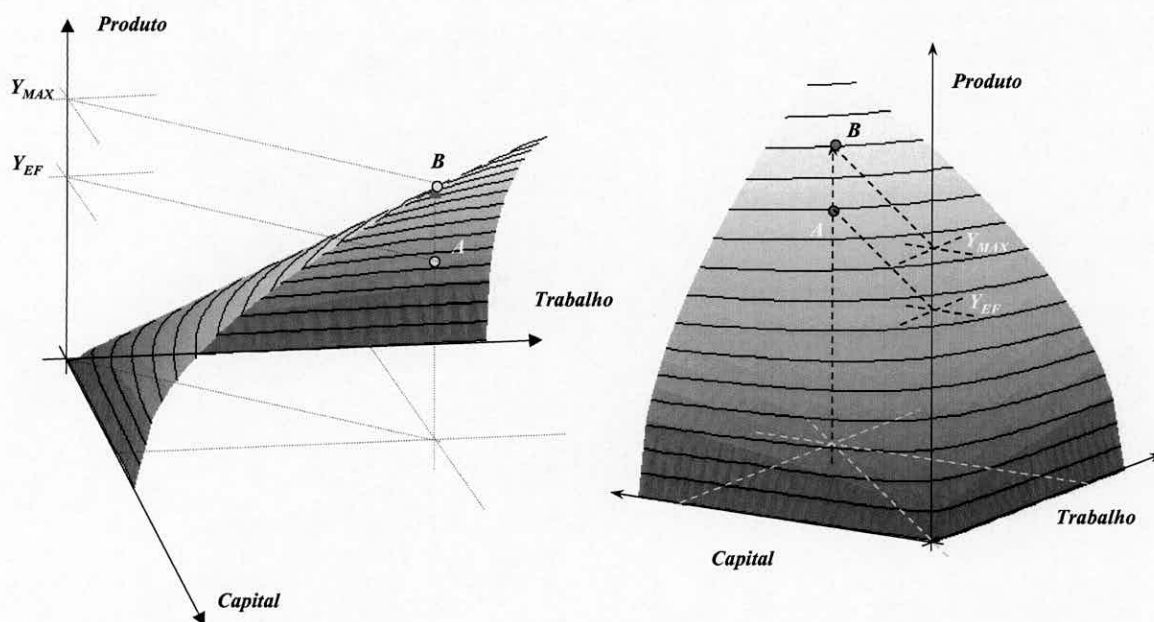
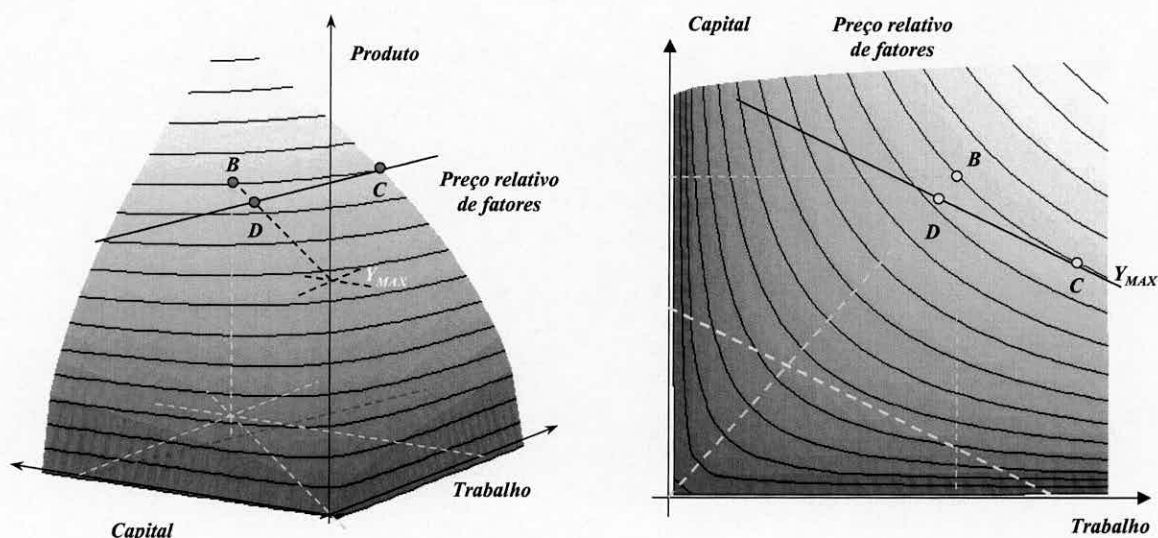




Figura 2.9 Ineficiência alocativa



Analogamente à explicação da figura 2.7, a firma no ponto B na figura 2.8 situa-se na superfície de produção. A firma B ( $Y_{MAX}$ , K, L) é eficiente tecnicamente. Em contrapartida, a firma A ( $Y_{EF}$ , K, L), sob a superfície, é ineficiente tecnicamente. A distância entre os dois pontos é a ineficiência técnica.

Com relação à eficiência alocativa, vimos que ela é alcançada quando a empresa combina os *inputs* na proporção que minimize seu custo de produção. A distância entre B e D representa tal fato. A seguir explicaremos por quê.

As firmas B e C estão sobre a superfície de produção eficiente e sobre a mesma isoquanta. Mostram-se, portanto, tecnicamente eficientes.

As firmas C e D fazem alocação mais eficiente do ponto de vista de econômico por estarem contidas na reta de preços relativos.

Dessa forma, sendo que B está sobre a fronteira de produção, ele é apenas eficiente tecnicamente. O ponto D, em contrapartida, possui apenas eficiência alocativa. Em C, há tanto a eficiência alocativa quanto técnica.

Os pontos D e B combinam a mesma proporção os inputs. Esse fato é comprovado pelo mesmo coeficiente angular das projeções, no plano capital x trabalho, dos vetores dados da origem até B e D.

Assim sendo, a ineficiência alocativa é dada pelas diferenças entre D e B.

Na abordagem de Farrel (1957), as medidas de eficiência são baseadas em distâncias radiais “uniformes” entre as observações ineficientes e a fronteira de produção. Esta é definida como a envoltória linear por partes (piecewise) mais pessimista para os dados observados, isto é, com a maior proximidade possível deles. Seu cálculo é feito através da resolução de sistemas de equações lineares, respeitando as condições que definem a isoquanta unitária. Esta representa combinações eficientes de inputs que permitem produzir um output.

A representação do modelo proposto no diagrama 2.10 é possível dada a hipótese de rendimentos constante de escala.

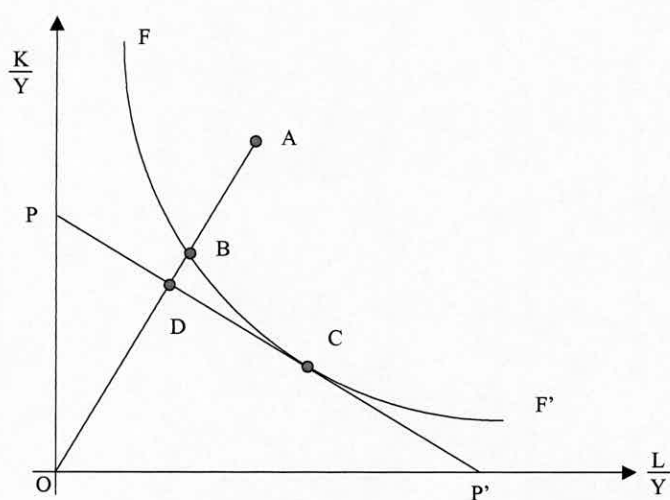


Figura 2.10: Diagrama de Farrel

O ponto A representa a combinação observada de quantidades de insumos para uma determinada firma ineficiente. Dessa forma, a reta OA indica a combinação de fatores empregados pela empresa A.

A isoquanta FF' mostra todas as combinações dos dois insumos que uma firma eficiente poderia empregar para produzir uma unidade de produto, uma vez que nos eixos são medidas as quantidades de insumos por unidade de produto (isoquanta unitária). Por ter inclinação negativa, significa que um aumento da quantidade de um dos insumos (por unidade de produto) implica, tudo mais constante, em uma menor eficiência técnica.

O ponto B representa uma firma eficiente empregando fatores na mesma proporção de fatores de A. Em B a firma produz a mesma quantidade de mercadoria, usando apenas uma fração (OB/OA) da quantidade de cada um dos insumos. Assim, o valor OB/OA pode ser visto como:

$$\frac{OB}{OA} = \frac{\text{combinação de quantidades de insumos necessária para se produzir uma unidade de produto}}{\text{combinação de quantidades de insumos efetivamente empregada}}$$

Esta razão mede a eficiência técnica (ET) de uma empresa.

$$ET = \frac{OB}{OA}$$

Então, a ineficiência técnica (IT) da empresa A pode ser dada por:

$$IT = 1 - \frac{OB}{OA}$$

Por B e C pertencerem à isoquanta unitária, são firmas com alocações tecnicamente eficientes.

A linha PP' é a linha isocusto, que representa as combinações de input que minimizam os custos de produção, dado os preços dos fatores. Ela tem como

inclinação a razão dos preços relativos dos insumos. Por conseguinte, as firmas C e D possuem eficiência alocativa. A eficiência alocativa de D pode ser dada por:

$$EA = \frac{OD}{OB}$$

E a ineficiência alocativa por:

$$IA = 1 - \frac{OD}{OB}$$

Concluimos, dessa maneira, que apenas C representa a combinação ótima de produção (pois ela tangencia a equação PP').

Caso as firmas observadas fossem totalmente eficientes, isto é, tanto sob o ponto de vista técnico como sob o ponto de vista alocativo, então os seus custos de produção seriam uma fração OD/OA do que realmente são. Essa razão é definida então como a eficiência total da firma (EF).

$$EF = \frac{OD}{OA}$$

Como consequência da natureza radial das medidas temos que a eficiência total de uma empresa é o produto da eficiência técnica pela alocativa:

$$\frac{OD}{OA} = \frac{OB}{OA} \times \frac{OD}{OB}$$

Deve-se ressaltar que todas as medidas de eficiência acima (ET, EA e EF) têm valores que oscilam entre 0 e 1. Ademais, caso as quantidades de insumos (por unidade de produto) cresçam indefinidamente, o valor dessas medidas de eficiência tende a zero.

### 3 ESTIMAÇÃO DA FRONTEIRA ESTOCÁSTICA

#### 3.1 CONCEITO DE FRONTEIRA

Um modo de avaliar os efeitos de escala e de eficiência técnica e alocativa, que, por sua vez, permitem avaliar o progresso técnico e conseqüentemente a produtividade, é a fronteira de produção. A função 3.1 define a fronteira de produção, como uma função de produção ajustada por meio de uma medida teórica de ineficiência técnica.

$$Y_i(t) = F(B(t), K_i(t), L_i(t)) \cdot \exp(-u_i(t)), \quad i = 1, 2, \dots, N \text{ e } t = 1, \dots, T \quad (3.1)$$

Sendo que:

- $Y_i(t)$  é a produção da firma  $i$  no instante  $t$ ;
- $F(\bullet)$  é a função de produção;
- $K_i(t)$  e  $L_i(t)$  são as quantidades de capital e trabalho empregadas pela firma  $i$  no instante  $t$ ;
- $B_t$  é o nível de produtividade que espelha a *best practice* no instante  $t$ ; e
- $u_i(t) \geq 0$  é a medida de ineficiência técnica da empresa  $i$  no instante  $t$ .

O nível de produtividade que espelha a *best practice* no instante  $t$  significa que ele representa a tecnologia mais eficiente no mercado em questão. Já  $u_i(t)$  mostra que a ineficiência técnica pode variar no tempo, refletindo que ela não depende apenas das mudanças de produtividade provenientes de acumulação de fatores e do progresso técnico.

#### 3.2 FRONTEIRA DETERMINÍSTICA DE PRODUÇÃO

Uma vez exposto o conceito de fronteira, a seguir abordaremos a análise empírica de ineficiência técnica utilizando o modelo fronteira determinística de Aigner e Chu (1968).

$$Y_i = F(B, K_i, L_i) \cdot \exp(-u_i), \quad i = 1, 2, \dots, N$$

Conforme a expressão acima, percebe-se que agora considera-se apenas um período para a análise da distinção entre as firmas.

Para estimar os parâmetros pode-se minimizar tanto a soma dos valores absolutos dos resíduos quanto a soma dos quadrados dos resíduos. A vantagem desse método é a possibilidade de computar a eficiência técnica de cada firma (observação) diretamente dos resíduos. A desvantagem é que o modelo mostra-se bastante sensível a *outliers*. Além disso, as estimativas obtidas por esse método têm propriedades estatísticas somente se fizermos hipóteses a respeito dos regressores e de “u”. A ausência de propriedades estatísticas provém do fato de a programação matemática do modelo não resultar em erros-padrão. Conseqüentemente, não se pode fazer inferências quanto aos parâmetros, como, por exemplo, testar hipóteses.

Schmidt (1976) mostra que, se os erros são independentes entre si e dos regressores e identicamente distribuídos no tempo, com média e variâncias positivas e finitas, tem-se um modelo de MQO (Mínimos Quadrados Ordinários). A estimação da fronteira por MQO – Schmidt (1976) e Greene (1997) - pode ser representada pela função genérica a seguir:

$$\ln Y = \beta \cdot X - u$$

Em que:

Ln Y: logaritmo natural do vetor de produção

$\beta$  : vetor de parâmetros a serem estimados

X: matriz dos logaritmos das variáveis explicativas

u: vetor de ineficiência técnica, sendo que  $u_i > 0$ , para qualquer i.

É importante ressaltar que o conjunto de variáveis explicativas  $X$  depende da forma funcional da função de produção, ou seja, se ela é Cobb-Douglas, logarítmica-transcendental, entre outras.

Para aplicações práticas, pode-se utilizar a seguinte expressão:

$$\ln y_i = \beta_0 + \beta' \cdot x_i + \varepsilon_i$$

Em que:

$$\varepsilon_i = -u_i$$

$x_i$ : o conjunto de fatores que entram no modelo empírico

As hipóteses são de que  $\varepsilon_i$  é distribuído aleatoriamente entre as firmas e é independente de todas as variáveis do modelo. Assume-se também que  $\varepsilon_i$  tem uma média não-nula, variância constante e que  $E[\varepsilon_i] < 0$ . O modelo pode, então ser escrito como o modelo de regressão clássico:

$$\ln y_i = (\beta_0 + E[\varepsilon_i]) + \beta' \cdot x_i + (\varepsilon_i - E[\varepsilon_i])$$

$$\ln y_i = \beta_0^* + \beta' \cdot x_i + \varepsilon_i^*.$$

Entretanto, somente os parâmetros da função de produção podem ser estimados consistentemente por MQO, com exceção do termo do termo constante. Isto porque não se pode pressupor a normalidade dos distúrbios, visto que  $\varepsilon_i^*$  é a diferença entre uma variável aleatória que sempre é negativa e sua média.

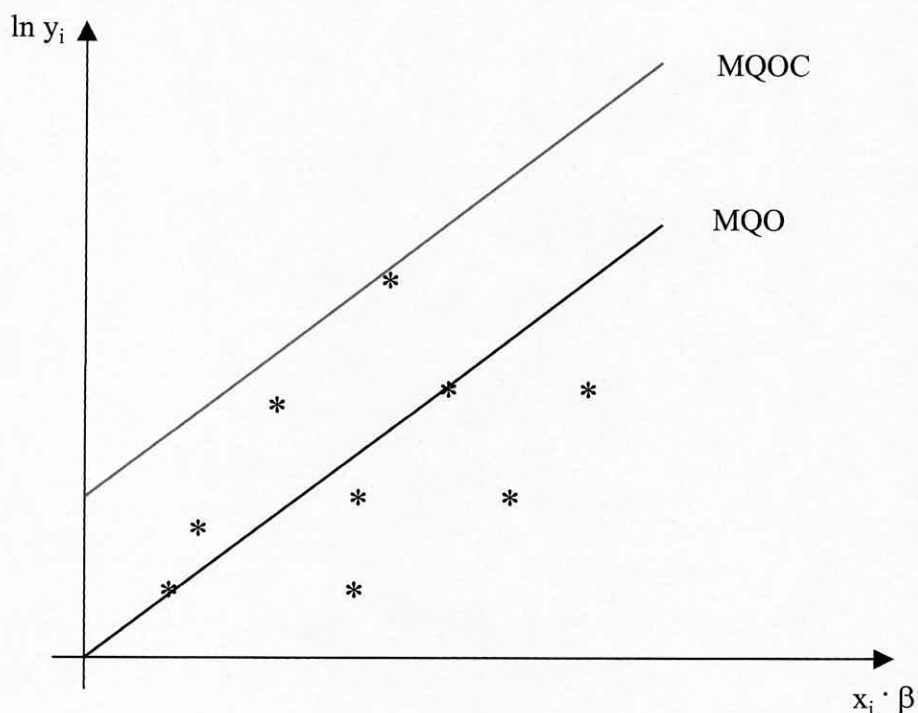
Já que o termo constante é inconsistente por MQO, pode-se obter dos resíduos estimados por MQO as informações para se comparar as ineficiências das firmas.



$$e_i = \ln y_i - b_0^* - b' \cdot x_i = \hat{u}_i + \hat{E}[u_i]$$

Quando se estima por MQO, estima-se uma fronteira “média”, em que uma parcela das firmas com resíduos negativos está abaixo da fronteira e as com positivos, acima da fronteira. No entanto, essa estimativa possui o problema de deslocamento do termo de constante.

Para resolver isto, pode-se utilizar o método de Mínimos Quadrados Ordinários Corrigidos (MQOC), que desloca a função de produção estimada para cima até que todos os resíduos, com exceção de um deles, sejam negativos. Isso corrige a constante tecnológica. O resíduo não negativo, ou seja, o máximo, traduz a idéia de *best practice*.





### 3.3 FRONTEIRA ESTOCÁSTICA DE PRODUÇÃO

Uma vez que fronteira determinística considera todos os desvios da fronteira como provenientes da ineficiência técnica, surgiu uma abordagem alternativa chamada de fronteira estocástica de produção proposta por Aigner, Lovell e Schmidt (1977) e Meeusen e van den Broeck (1977). Esse modelo inclui um erro aleatório  $v_i$  que capta erros de medidas e outros fatores aleatórios, como clima, greves entre outros, que não eram considerados no modelo determinístico.

A expressão geral da fronteira estocástica pode ser dada por:

$$Y_i = F(B, K_i, L_i) \cdot \exp(-u_i) \cdot \exp(v_i), \quad i = 1, 2, \dots, N$$

Nas aplicações empíricas tem-se:

$$\ln y_i = \beta_0 + \beta' \cdot x_i + \varepsilon_i = \beta_0 + \beta' \cdot x_i + (v_i - u_i)$$

Assume-se que  $v_i$  é uma variável aleatória com distribuição normal, independente e identicamente distribuída (i.i.d.), com média zero e variância constante,  $\sigma_v^2$ , e que  $v_i$  é independente de  $u_i$ . Pressupõe-se ainda que  $u_i$  é uma variável aleatória i.i.d., cuja distribuição é exponencial ou meia-normal.

Nesse modelo, os valores do produto estão limitados pelo valor das variáveis estocásticas, ou seja, a parcela determinística mais um componente aleatório:  $\exp(\beta' \cdot x_i + v_i)$ . O erro aleatório  $v_i$  pode ser positivo ou negativo, de modo que o produto das firmas varia em torno da parte determinística do modelo de fronteira,  $\exp(\beta' \cdot x_i)$ .

A principal crítica a esse modelo, conforme exposto por Stevenson (1980) e Greene (1990), corresponde ao fato de não haver a priori justificativa para a escolha de uma forma específica da distribuição de  $u_i$ . As distribuições meia normal e exponencial são escolhas arbitrárias.

Para se estimar os parâmetros, pode-se usar tanto o método dos MQOC ou da máxima verossimilhança (MV).

Nesse modelo, podemos testar a hipótese se há ou não efeitos de ineficiência técnica. A hipótese nula de que não há efeitos de ineficiência técnica significa que  $u_i$  é nulo para toda firma  $i$ , implicando que todas as firmas são eficientes.

### 3.4 ESTIMAÇÃO DUAL DA EFICIÊNCIA TÉCNICA

Conforme a Teoria da Dualidade, existe uma relação bastante forte entre as funções de produção e de custo. Isto porque a solução da maximização da produção, dada as restrições de custo, é a mesma da minimização de custos sujeita a restrições impostas pelo processo produtivo. Dessa forma, podemos estimar a fronteira de custo em vez da fronteira de produção com o intuito de analisar as variações de produtividade do serviço público de acordo com a região. Isso decorre do fato de haver dados melhores sobre custos.

Para esse tipo de análise, Atkinson e Cornewell (1993, 1994b) tomaram como ponto de partida a função produção já discutida anteriormente. Podemos representá-la de forma simplificada por:

$$y = f(x)$$

em que  $y$  é o output e  $x$  é um vetor de insumos e  $f$  é uma função produção que representa a tecnologia.

A ineficiência pode ser representada de duas formas. Podemos introduzir um índice de eficiência técnica orientada ao output  $a$  ou orientada ao input  $b$ , de acordo com as equações abaixo.

$$y = af(x)$$

$$y = f(bx)$$

Se eliminarmos a ineficiência técnica, no primeiro caso, a produção poderia ser aumentada na proporção (1-a). Já no segundo, os inputs poderiam reduzir na proporção (1-b).

A função de custo associada ao primeiro caso seria:

$$C_a\left(w, \frac{y}{a}\right) = \min_x \left\{ wx \mid f(x) \geq \frac{y}{a} \right\}$$

E a do segundo caso:

$$C_b\left(\frac{w}{b}, y\right) = \min_{bx} \left\{ \frac{w}{b} bx \mid f(bx) \geq y \right\}$$

Em que  $w$  é o vetor de preços dos insumos.

Percebemos que, quando a eficiência técnica está orientada ao output, a função de custos depende do quociente entre o output observado e o índice de eficiência técnica. Em contrapartida, quando está orientado ao input, a função de custos depende dos preços dos inputs dividido pelo índice de eficiência técnica. Além disso, concluímos que a eficiência técnica aparece no modelo como um parâmetro.

Entretanto, o modelo apresentado pressupõe que há eficiência alocativa. Para solucionar esse problema, Atkinson e Cornwell (199a) propuseram uma diferenciação de preços: o preço do input observado  $w_i$  e o preço  $w_i^*$  que seria o preço ótimo de minimização de custos. Eles podem ser relacionados através de um parâmetro  $\theta$ :

$$w_i^* = \theta_i w_i$$

A versão empírica da função de custos pode ser dada por uma translog com dados de painel:

$$\ln C_{it} = -\ln b_i + \beta_0^b + \sum_j \beta_j^b \ln w_{jit} + \delta^b \ln y_{it} + \frac{1}{2} \phi^b (\ln y_{it})^2 + \\ + \frac{1}{2} \sum_j \sum_k \gamma_{jk}^b \ln w_{jit} \ln w_{kit} + \sum_j \phi_j^b \ln w_{jit} \ln y_{it} + v_{it}^b$$

Onde  $\beta_0^b, \beta_j^b, \delta^b, \phi^b, \gamma_{jk}^b$  e  $\phi_j^b$  representam os parâmetros comuns a todos os indivíduos,  $v_{it}^b$  é uma perturbação aleatória independente e identicamente distribuída e  $\ln b_i$  é o logaritmo do índice de eficiência técnica orientado ao input.

A função de custo translogarítmica com índice de eficiência técnica orientada ao output pode ser escrita como:

$$\ln C_{it} = \beta_0^a + \sum_j \beta_j^a \ln w_{jit} + \delta^a \ln \left( \frac{y_{it}}{a_i} \right) + \frac{1}{2} \phi^a \left[ \ln \left( \frac{y_{it}}{a_i} \right) \right]^2 + \\ + \frac{1}{2} \sum_j \sum_k \gamma_{jk}^a \ln w_{jit} \ln w_{kit} + \sum_j \phi_j^a \ln w_{jit} \ln \left( \frac{y_{it}}{a_i} \right) y_{it} + v_{it}^a$$

Em que analogamente  $\beta_0^a, \beta_j^a, \delta^a, \phi^a, \gamma_{jk}^a$  e  $\phi_j^a$  representam os parâmetros comuns a todos os indivíduos,  $v_{it}^a$  é uma perturbação aleatória independente e identicamente distribuída e  $a_i$  o índice de eficiência técnica orientado ao output.

Utilizando o LEMA de Shephard, as derivadas logarítmicas são iguais às participações dos inputs nos custos. No caso da função de custos com o índice de eficiência orientado ao input a participação nos custos do input  $j$  pode ser escrito como:

$$S_{it}^j = \beta_j^b \ln w_{jit} + \sum_k \gamma_{jk}^b \ln w_{kit} + \theta_j^b \ln y_{it}$$

em que  $S_{it}^j$  é participação no custo do input  $j$  para o indivíduo  $i$  no momento  $t$ .

No caso em que o índice de eficiência está orientado para o output, a participação nos custos do input  $j$  pode ser escrito como:

$$S_{it}^j = \beta_j^a + \sum_k \gamma_{jk}^a \ln w_{kit} + \theta_j^a \ln y_{it} - \theta_j^a \ln a_i$$

## **4 ESTIMAÇÃO DA FRONTEIRA ESTOCÁSTICA DE CUSTO PARA SAÚDE PÚBLICA BRASILEIRA**

### **4.1 DADOS E AMOSTRAS**

Para o desenvolvimento dessa pesquisa, foram coletados dados do IBGE (Instituto Brasileiro de Geografia e Estatística) de 1996 a 1998, provenientes dos resultados da pesquisa de Despesas Públicas por Funções e do Anuário Estatístico do Brasil.

Com relação à primeira fonte, obtivemos valores referentes às despesas totais, com pessoal e com capital fixo do serviço hospitalar e ambulatorial. Tais despesas abrangem gastos realizados em cada exercício dos governos estaduais, municipais e federal. Além disso, o serviço hospitalar e ambulatorial refere-se às despesas com:

- administração, operação e suporte dos órgãos encarregados dos serviços de atendimento ambulatorial e hospitalar;
- hospitais universitários e militares;
- bancos de sangue;
- programas de alimentação e nutrição de crianças e adolescentes em situação de risco, atendimento aos desnutridos e às gestantes de risco nutricional;
- assistência médico-sanitária e atendimento domiciliar;
- pagamentos realizados pela administração pública por contratos de prestação de serviços de saúde com clínicas e hospitais privados credenciados;

A amostra abrangeu as 27 unidades da federação e 26 capitais ou regiões metropolitanas. A seguir estão as tabelas com a descrição das variáveis:

UF	Variável	Ano	Média	Desvio Padrão	Mínimo	Máximo	Número de observações
	Despesas totais em R\$ mil	1996	235.380,59	363.423,68	16777	1793824	27
		1997	287.185,30	440.666,52	32790	2083948	27
		1998	297.522,78	429.517,00	32783	2166273	27
	Despesas com capital fixo em R\$ mil	1996	9.921,88	10.887,15	5	55364	27
		1997	13.702,31	23.686,32	19	123327	27
		1998	17.045,37	30.691,26	349	163721	27
	Despesas com pessoal em R\$ mil	1996	111.906,78	164.636,06	3723	800550	27
		1997	124.585,04	179.365,40	546	870124	27
		1998	127.109,96	185.235,12	1317	929942	27

CA/RM	Variável	Ano	Média	Desvio Padrão	Mínimo	Máximo	Número de observações
	Despesas totais em R\$ mil	1996	141.903,69	306.843,52	1006	1499975	26
		1997	158.218,50	300.933,66	2051	1414307	26
		1998	218.740,19	378.110,90	2174	1825116	26
	Despesas com capital fixo em R\$ mil	1996	6.911,83	13.922,92	64	63967	26
		1997	33,00	0,00	33	33	26
		1998	5.658,96	6.583,64	1	25013	26
	Despesas com pessoal em R\$ mil	1996	50.862,20	114.651,11	1	558985	26
		1997	58.813,80	113.439,35	601	531035	26
		1998	78.444,42	154.288,07	330	754750	26

Com relação ao Anuário do IBGE, foram coletadas observações sobre hospitalizações pagas pelo SUS, por número de internações, gastos, taxa de mortalidade e média de permanência, segundo as Unidades da Federação e natureza do prestador. A fonte do Anuário provém do Sistema de Informação Hospitalares do SUS (Ministério da Saúde).

	Tipo	Ano	Média	Desvio Padrão	Mínimo	Máximo	Número de observações
número de internações	Privado	1996	307.092,15	347.960,66	967	1.473.380	26
		1997	294.862,19	336.110,93	1.256	1.397.691	26
		1998	282.058,27	319.152,72	1.416	1.307.187	26
	Público	1996	91.460,74	91.631,23	9.995	360.763	27
		1997	97.012,89	94.897,08	2.682	384.637	27
		1998	104.781,52	100.765,21	10.420	426.605	27
	Universitário	1996	67.219,00	97.170,78	746	432.028	22
		1997	67.572,82	95.627,63	924	423.292	22
		1998	67.484,35	100.050,77	289	458.329	23

	Tipo	Ano	Média	Desvio Padrão	Mínimo	Máximo	Número de observações
gastos (R\$)	Privado	1996	73.686.191,37	94.653.522,52	968.122	427.546.186	26
		1997	72.249.163,19	93.344.374,74	888.390	416.218.289	26
		1998	82.202.974,73	104.645.791,80	803.327	458.578.761	26
	Público	1996	17.430.026,04	21.032.023,18	1.362.779	86.950.584	27
		1997	18.765.390,26	22.202.036,11	335.448	93.163.862	27
		1998	24.090.119,41	27.319.914,13	2.643.892	120.528.151	27
	Universitário	1996	36.173.392,97	60.017.732,58	187.026	262.231.927	22
		1997	37.304.919,73	61.169.994,17	227.818	268.428.999	22
		1998	44.415.785,87	74.817.713,14	38.405	338.355.121	23



	Tipo	Ano	Média	Desvio Padrão	Mínimo	Máximo	Número de observações
taxa de mortalidade	Privado	1996	1,47	0,87	0	3	26
		1997	1,48	0,84	0	3	26
		1998	1,60	0,84	0	3	25
	Público	1996	2,17	1,66	0	7	27
		1997	2,17	1,67	0	7	27
		1998	2,26	1,64	0	7	27
	Universitário	1996	4,50	1,36	2	8	22
		1997	4,45	1,40	2	8	22
		1998	4,39	1,46	1	7	23

	Tipo	Ano	Média	Desvio Padrão	Mínimo	Máximo	Número de observações
média de permanência	Privado	1996	7,13	8,24	3	46	26
		1997	6,77	6,97	3	40	26
		1998	6,69	6,61	3	38	26
	Público	1996	5,80	2,30	4	11	27
		1997	5,59	2,21	3	11	27
		1998	5,49	2,09	3	12	27
	Universitário	1996	7,65	1,26	6	11	22
		1997	7,54	1,25	6	10	22
		1998	7,55	1,26	5	10	23

Foram extraídas ainda informações quanto ao número de estabelecimentos hospitalares e leitos existentes, por natureza do prestador e segundo as Unidades da Federação.

	Tipo	Ano	Média	Desvio Padrão	Mínimo	Máximo	Número de observações
hospitais	Privado	1996	162,12	163,19	2	572	26
		1997	159,38	161,12	2	580	26
		1998	157,27	159,93	2	581	26
	Público	1996	74,89	55,49	9	188	27
		1997	78,48	58,21	9	191	27
		1998	79,44	58,55	9	196	27
	Universitário	1996	6,77	7,08	1	26	22
		1997	6,39	7,08	1	26	23
		1998	7,00	7,37	1	27	22

	Tipo	Ano	Média	Desvio Padrão	Mínimo	Máximo	Número de observações
leitos	Privado	1996	13.185,27	16.995,78	186	77.082	26
		1997	12.893,81	16.209,82	280	71.770	26
		1998	12.671,88	15.934,99	186	70.703	26
	Público	1996	4.062,37	4.998,97	547	20.609	27
		1997	4.278,78	5.440,02	513	21.711	27
		1998	4.200,22	5.137,09	604	21.801	27
	Universitário	1996	2.146,27	2.978,17	36	13.268	22
		1997	1.998,87	2.953,96	36	13.335	23
		1998	2.165,77	3.107,22	36	13.964	22



Com relação aos equipamentos, por não variarem bruscamente ao longo do tempo, foram extraídos dados do ano de 1999 com fonte do IBGE (Diretoria de Pesquisas, Departamento de População e Indicadores Sociais, Pesquisa de Assistência Médico-Sanitária 1999). Os equipamentos considerados foram: mamógrafo com comando simples, mamógrafo com estereotaxia, raio X para densitometria óssea, tomógrafo, ressonância magnética, ultrassom doppler colorido, eletrocardiógrafo, eletroencefalógrafo, equipamento de hemodiálise, raio X até 100mA, raio X de 100 a 500mA e raio X mais de 500mA. A seguir apresentamos a tabela descritiva dessa variável.

	Ano	Média	Desvio Padrão	Mínimo	Máximo	Número de Observações
equipamentos	1 999	2 150	3 529	52	16 457	27

## 4.2 ASPECTOS METODOLÓGICOS ASSOCIADOS A CADA UMA DAS VARIÁVEIS

### Número de Estabelecimentos Hospitalares

O número de estabelecimentos hospitalares do setor público, para cada unidade da federação em cada ano, foi obtido pela soma dos hospitais públicos mais os universitários.  $N_{it}$  é o número de hospitais da atividade  $i$  no ano  $t$ , para  $t=1996, 1997$  e  $1998$ .

### Custo de Mão-de-obra dos profissionais da saúde

As despesas com mão-de-obra dos 27 governos federais e DF foram somadas às despesas dos governos das 26 capitais ou regiões metropolitanas correspondentes, resultando em uma despesa global para cada unidade de federação. A fim de conseguir os aumentos reais de despesas, deflacionamos - nas pelo INPC, considerando 1996 como ano base. Além disso, para determinando ano, dividimos as despesas de mão de obra de cada unidade da

federação por número de hospitais. O objetivo disso é encontrar a média de despesas por hospital. Esse, por sua vez, mostra-se como um hospital representativo da unidade da federação.  $C_{it}$  é o custo de mão de obra agregado dos N hospitais que compõem o i-ésimo grupo de atividade econômica no ano t, e  $c_{it}$  representa o custo do hospital representativo do segmento i no instante t:

$$c_{it} = C_{it} / N_{it}$$

#### Número de Internações

O número de internações do setor público, para cada unidade da federação em cada ano, foi obtido pela soma das internações dos hospitais públicos mais os universitários. De maneira análoga ao custo, dividimos o número de internações de uma unidade da federação pelos hospitais correspondentes, para obter um número médio de internações em um hospital representativo no instante t para a observação i.

#### Número de leitos

O número de leitos do setor público foi obtido pela soma dos leitos dos hospitais públicos mais os universitários. Similarmente, dividimos o número de internações da observação i no tempo t pelos hospitais correspondentes, para obter um número médio de leitos em um hospital representativo da unidade da federação.

#### Dias de internação

Os dias de internação, para uma observação i em instante t, foi obtido por:

$$\text{Dias de internação}_{it} = 365 / (\text{internações}_{it} / \text{leitos}_{it})$$

#### Equipamentos

Por falta de informações para os anos abrangidos pela pesquisa, os dados sobre disponibilidade de equipamentos, que se referem ao ano de 1999, foram considerados fixos no período.

#### Participação do salário na renda (SI)

Participação dos salários na despesa (W/C) é uma variável que aproxima o preço relativo da mão-de-obra dos hospitais em cada unidade da federação. Esse procedimento, usual na literatura de fronteiras estocásticas visa minimizar eventuais distorções causadas por diferenças de preços relativos.

### **4.3 FRONTEIRA DE CUSTOS DE MÃO-DE-OBRA E ANÁLISE DOS RESULTADOS**

O modelo de estimação da fronteira de custo foi conduzido utilizando o programa STATA 8, que inclui o modelo de Battese & Coelli (1992).

Não incluímos no modelo o Distrito Federal e o Acre. O Distrito Federal concentra despesas em saúde que são exclusivamente usufruídas pelos moradores do distrito, ao passo que os dados de custos do Acre apresentam variações excessivas entre 1996 e 1998, em função de critérios contábeis: boa parte das despesas empenhadas em 1997 foram liquidadas em 1998, criando uma variação desproporcional ao fluxo de serviços de saúde.

Os parâmetros da tabela 4.1, com exceção da constante, são todos significativos a 5%. A média de ineficiência  $\mu$  também é significativa a 5%, sugerindo que o modelo está apropriado por não rejeitar a hipótese de que a distribuição é uma normal truncada, ou seja, rejeita a hipótese que a distribuição é uma half – normal. Portanto, a hipótese de distribuição é menos restritiva. Além disso, o  $\eta$ , taxa de variação de ineficiência ao longo do tempo, negativo revela que há diminuição de eficiência.

Com relação aos parâmetros, observamos que  $\beta_1$  é positivo. Isso mostra o aumento de internações acarreta um aumento de custo de mão de obra, mantendo as demais variáveis constantes. Esse parâmetro também representa a elasticidade do custo de mão de obra real em relação às internações; isto é, o aumento de 1% de internação, aumentará em 2,9% o custo de mão de obra real.

Outro parâmetro que impacta sensivelmente os custos é o  $\beta_2$ , pois o aumento de 1% de leitos, diminui o custo de mão de obra em 2,4%. A princípio, isso não parece ser muito lógico já que podemos pensar que o aumento de leitos exigirá mais profissionais e, conseqüentemente, haverá maiores custos. Entretanto, mostra-se importante lembrar que os leitos são de um hospital representativo. Dessa forma, o número de leitos indica aproximadamente a magnitude do hospital. Hospital maiores podem ter ganhos de escala ao diluir os custos fixos dos profissionais.

O  $\beta_3$  indica que o acréscimo de um dia de internação aumenta em 0,14% o custo de mão de obra.

O  $\beta_4$  revela que o aumento de equipamentos implica aumento de custo de mão de obra, provavelmente devido aos salários mais altos decorrentes de uma força de trabalho mais qualificada.

Os valores da fronteira estimada seguem a seguir:

$$\begin{aligned} \ln c_i &= \beta_0 + \beta_1 \ln X_1 + \beta_2 \ln X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 \ln X_5 + \beta_6 X_6 + \varepsilon_i = \\ &= \beta_0 + \beta_1 \ln X_1 + \beta_2 \ln X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 \ln X_5 + \beta_6 X_6 + (v_i - u_i) \end{aligned}$$

Tabela 4.1: modelo de ineficiência

Número de observações = 75		Observações por grupo :		min = 3		
Número de grupos = 25				max = 3		
				media = 3		
Log likelihood = -37,9877		Wald chi2(6)		= 143.14		
		Prob > chi2		= 0.0000		
Ln c	Coeficientes	Erro padrão	z	P >  z	Intervalo de Confiança 95%	
					inferior	superior
$\beta_1$	2,9316	0,4290	6,83	0,0000	2,0907	3,7724
$\beta_2$	-2,3722	0,5064	-4,68	0,0000	-3,3646	-1,3797
$\beta_3$	0,1445	0,0223	6,48	0,0000	0,1008	0,1882
$\beta_4$	0,4579	0,2171	2,11	0,0350	0,0324	0,8835
$\beta_5$	-0,4638	0,1469	-3,16	0,0020	-0,0752	-0,1758
$\beta_6$	1,7812	0,3131	5,69	0,0000	1,1674	2,3949
$\beta_0$	-0,2946	2,7876	-0,11	0,9160	-5,7583	5,1691
$\mu$	1,1229	0,4058	2,77	0,0060	0,3275	1,9183
$\eta$	-0,0324	0,0321	-1,01	0,3130	-0,0954	0,0306
$\ln \sigma^2$	-0,8947	0,4028	-2,22	0,0260	-1,6843	-0,1052
$\ln \gamma$	1,6009	0,5647	2,83	0,0050	0,4941	2,7077
$\sigma^2$	0,4087	0,1646			0,1856	0,9001
$\gamma$	0,8321	0,0789			0,6211	0,9375
$\sigma_u^2$	0,3401	0,1666			0,0136	0,6666
$\sigma_v^2$	0,0686	0,0142			0,0407	0,0965

Em que:

c = custo de mão de obra real do hospital representativo da UF

$X_1$  = internações

$X_2$  = leito

$X_3$  = dias

$X_4$  = equipamentos

$X_5$  = número de hospitais

$X_6$  = participação do salário na renda

O  $\gamma$  representa a seguinte razão:

$$\gamma = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$$

Assim sendo, 83% da variação total é proveniente da variação de  $u_{it}$ , ou seja, maior parte da variação total é explicada pela variação na ineficiência representada por  $u_{it}$ .

Tabela 4.2: Eficiência técnica conforme unidade da federação.

Unidade da Federação	1996	1997	1998	Média
Acre	n.d.	n.d.	n.d.	n.d.
Alagoas	0,2868	0,2752	0,2636	0,2752
Amapá	0,3940	0,3820	0,3699	0,3820
Amazonas	0,1613	0,1518	0,1426	0,1519
Bahia	0,2313	0,2203	0,2096	0,2204
Ceará	0,3080	0,2961	0,2844	0,2962
Distrito Federal	n.d.	n.d.	n.d.	n.d.
Espírito Santo	0,1646	0,1550	0,1457	0,1551
Goiás	0,3329	0,3209	0,3090	0,3209
Maranhão	0,7595	0,7524	0,7452	0,7524
Mato Grosso	0,5160	0,5047	0,4933	0,5047
Mato Grosso do Sul	0,8147	0,8091	0,8032	0,8090
Minas Gerais	0,1991	0,1887	0,1786	0,1888
Pará	0,3051	0,2933	0,2816	0,2933
Paraíba	0,4379	0,4260	0,4141	0,4260
Paraná	0,3022	0,2904	0,2787	0,2904
Pernambuco	0,2180	0,2072	0,1967	0,2073
Piauí	0,4516	0,4398	0,4279	0,4398
Rio de Janeiro	0,5640	0,5533	0,5424	0,5533
Rio Grande do Norte	0,2525	0,2412	0,2301	0,2412
Rio Grande do Sul	0,3164	0,3045	0,2927	0,3045
Rondônia	0,6046	0,5945	0,5842	0,5944
Roraima	0,7255	0,7176	0,7096	0,7176
Santa Catarina	0,1296	0,1211	0,1129	0,1212
São Paulo	0,3929	0,3808	0,3688	0,3808
Sergipe	0,1955	0,1852	0,1751	0,1853
Tocantins	0,3450	0,3330	0,3211	0,3331
Brasil	0,3764	0,3658	0,3552	0,3658

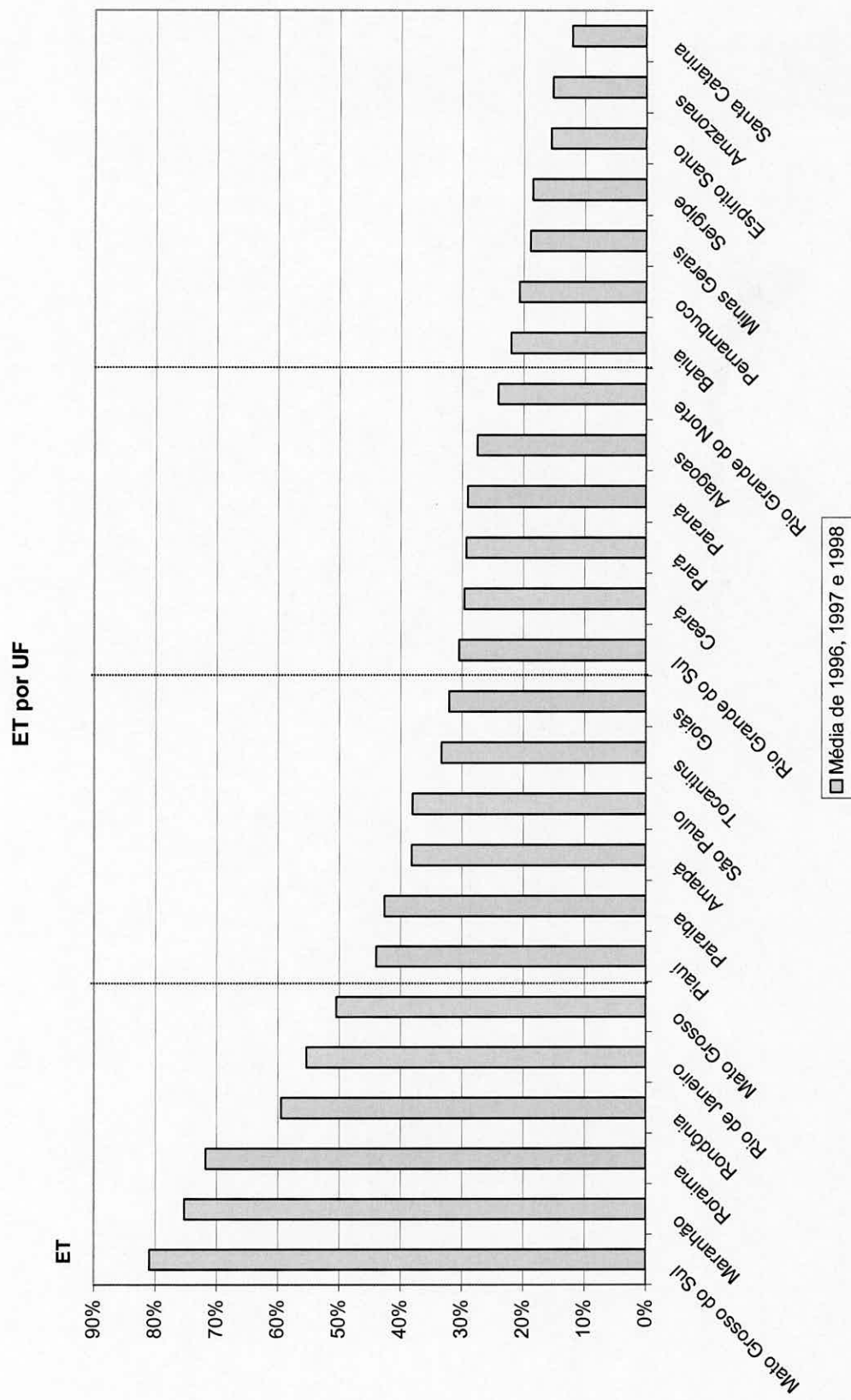


Figura 4.1: Média de eficiência técnica



Na figura 4.1, os estados que se aproximam dos 100% são os mais eficientes tecnicamente pelo modelo. Assim sendo, o Mato Grosso do Sul é o mais eficiente e Santa Catarina, o menos.

Agrupando os estados por região, observamos que o centro oeste é bastante eficiente, já que possui duas unidades da federação (Mato Grosso e Mato Grosso do sul) nos 25% mais eficientes (primeiro quartil de eficiência) e uma unidade do segundo quartil. Com relação à região norte, 67% das unidades da federação (excluindo o Acre), estão nos dois primeiros quartis de eficiência. Em contrapartida, 67% das unidades do nordeste e 100% dos estados do sul estão nos dois últimos quartis de eficiência; ou seja, a região nordeste e sul são as mais ineficientes. Entretanto, como estamos realizando uma análise através dos custos, percebemos que tal diferença se dá provavelmente ao fato dos profissionais da saúde na região sul serem melhor remunerados que os do nordeste, principalmente devido à qualificação. É conhecimento que a região sul apresenta os melhores índices de desenvolvimento humano do país. Já a região Sudeste, apesar de ser a mais economicamente desenvolvida do Brasil, possui dois estados (MG e ES), nos 25% menos eficientes. São Paulo está nos 50% mais eficientes e RJ nos 25%.

Uma ressalva deve ser feita a esse estudo. A eficiência medida nesse estudo não considera possíveis dados de qualidade do serviço, mas sim os custos envolvidos para a prestação dos mesmos.

# Variação de ET entre 1996 e 1998

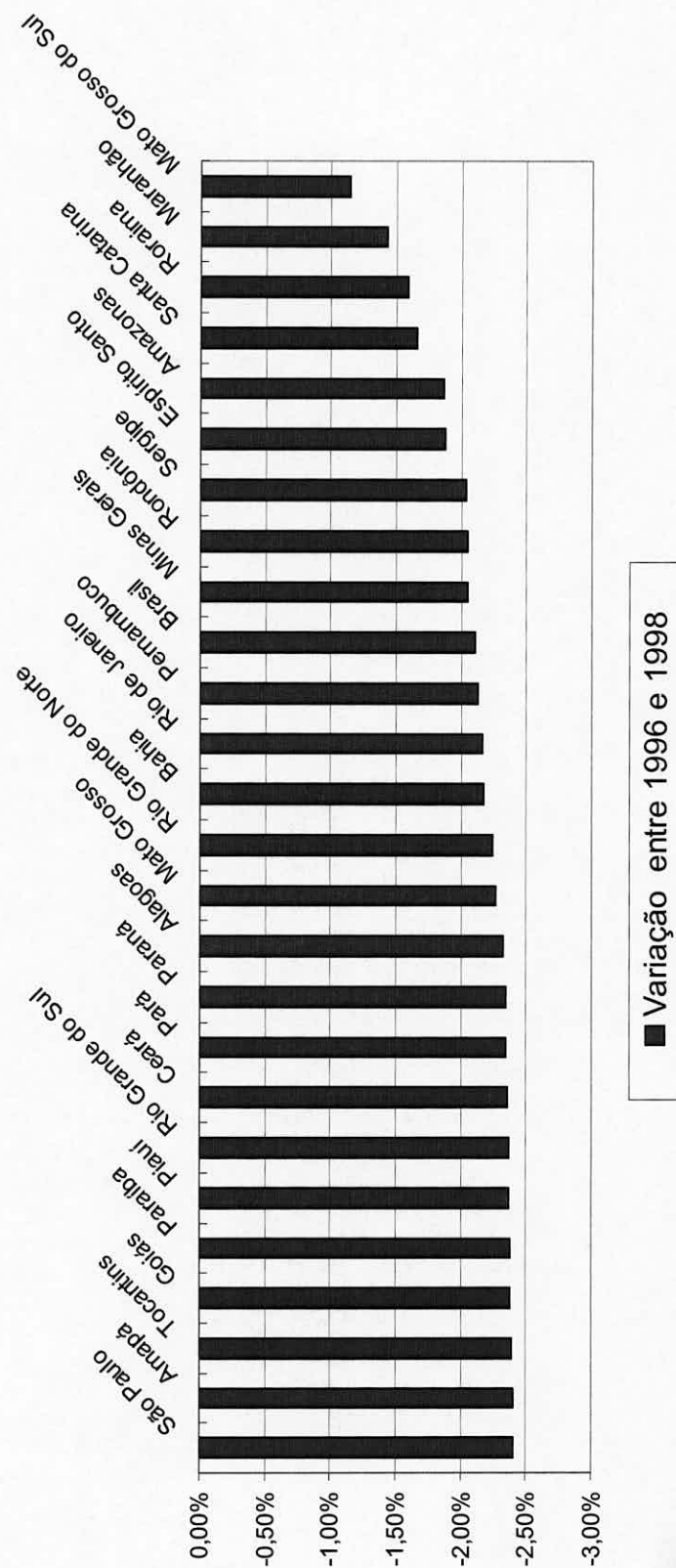


Figura 4.2: Variação de eficiência técnica (ET 1998 – ET 1996)

Na figura 4.2, percebemos que todos os estados estão ficando menos eficientes no período de análise, em especial aqueles que estão situados no segundo quartil de eficiência técnica (São Paulo, Amapá, Tocantins, Goiás, Paraíba e Piauí). Em seguida, a diminuição maior de eficiência são dos estados do terceiro quartil de eficiência técnica, do Rio Grande do Sul até Bahia (figura 4.2), com exceção do Mato Grosso. Assim sendo, notamos que os estados mais e menos eficientes tecnicamente (primeiro e último quartis) apresentam diminuição menor de eficiência; ou seja, os mais eficientes como Mato Grosso do Sul, Maranhão e Roraima, tem tendência de continuarem a ser mais eficientes por terem taxas menores de diminuição de eficiência. Em contrapartida, parece que as políticas públicas nos menos eficientes (BA, Pernambuco, Minas Gerais, Sergipe, Espírito Santo, Amazonas e Santa Catarina), têm sido mais agressivas a fim de diminuir a ineficiência.

## CONSIDERAÇÕES FINAIS

Esse estudo procurou analisar de forma quantitativa a questão dos custos da saúde pública brasileira e sua eficiência conforme as unidades da federação. Identificamos que o custo de mão-de-obra dos profissionais da saúde mostra-se mais sensível às internações e aos leitos, sendo a primeira relação positiva e a segundo negativa. Assim sendo, medidas preventivas de saúde são vitais para diminuir o número de internações. Percebemos ainda que o aumento de equipamentos não diminui o custo de mão-de-obra, mas aumenta-o, por exigir profissionais mais capacitados. Identificamos também que os estados possuem diferenças significativas de eficiência técnica e estão ficando menos eficientes, principalmente aqueles que estão no segundo quartil de eficiência técnica.

A pesquisa nesse campo, entretanto, necessita ainda de mais investigações para gerar idéias e propostas efetivas que possam servir de políticas públicas para amenizar a inequidade entre o setor privado e público de saúde brasileiro e entre os estados.

## REFERÊNCIAS BIBLIOGRÁFICAS

AIGNER, D. J. e CHU, S. F. (1968), On Estimating the Industry Production Function, *American Economic Review*, 58, no.4, 826-839.

AIGNER, D. J., LOVELL, C. A. K. e SCHMIDT, P. (1977), Formulation and Estimation of Stochastic Frontier Production Function Models, *Journal of Econometrics*, 6, 21-37.

ALMEIDA, Célia, TRAVASSOS, Cláudia, PORTO, Silvia and LABRA, Maria Eliana. Healthb sector reform in Brazil: a case study of inequity. *International Journal of Health Services*, vol 30, no 1, 2000.

ÁLVAREZ PINILLA, A.(2001), *La Medición de la eficiencia y la Productividad*, Ediciones Pirâmide, Madrid, Espana.

ANDRADE, M. NORONHA, K. Desigualdade Social no Acesso aos Serviços de Saúde no Brasil.

Banco Mundial (1997): O Estado num mundo em transformação. Relatório sobre *Desenvolvimento Mundial*, Washington, D.C.

BATTESE, G. E. e COELLI, T.J. (1992), Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farms in India, *Journal of Productivity Analysis*, 3, 153-169.

BATTESE, G. E. e COELLI, T.J. (1995), A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data, *Empirical Economics*, 20, 325-332.

COELLI, T. J. (1995), Estimators and Hypothesis Tests for a Stochastic Production Function: A Monte Carlo Analysis, *Journal of Productivity Analysis*, 6, 247-268.

COELLI, T. J. (1996), A Guide to Frontier 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation, *CEPA Working Paper*, Center for Efficiency and Productivity Analysis, University of New England, Armidale, Australia.

EATON, B., EATON D. (1995): *Microeconomia*. Editora Saraiva, São Paulo, Brasil.

FAÇANHA, L.; MARINHO, A. (2000), Hospitais Universitários: Avaliação Comparativa de Eficiência Técnica. *Brazilian Journal of Applied Economics*. v.4, n 2, abr./jun. 2000.

FAÇANHA, L.; MARINHO, A. (2000), Hospitais Universitários: Indicadores de Utilização e Análise de Eficiência. *Brazilian Journal of Applied Economics*. v.6, n 3, jul./set. 2002.

FARREL, M. J. (1957), The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society, Series A (General)*, Vol. 120, issue 3.

GARCIA, F. (2003), A Evolução da Produtividade Total Brasileira: Uma Análise do Período Pós-Real, *Séries Relatório de Pesquisa* nº13,/2003, Núcleo de Pesquisas e Publicações EAESP-FGV.

GARCIA, F; PIRES, J. Productivity of Nations: a stochastic frontier approach to TFP decomposition. Escola de Economia de São Paulo- Fundação Getúlio Vargas- Brazil.

GIAMBIAGI, F., ALÉM, A. (1999): *Finanças Públicas – Teoria e Prática no Brasil*. Editora Campus, Rio de Janeiro, Brasil.

HARBERGER, A. (1978), *Perspectives on Capital Technology in Less Developed Countries* in M.J. Artis and A. R. Nobay (eds.). Contemporary Economic Analysis (London: Croom Helm).

KUMBHAKAR, S. C. (2000), Estimation and Decomposition of Productivity Change When Production is Not Efficient, *Econometric Reviews*, 19, 425-460.

LIMA, M. E. C. M. A Produção e a Estrutura de Custos dos Hospitais Públicos: uma Aplicação de um Modelo Translogarítmico. Universidade do Minho.

MANKIW, N., ROMER, D., WEIL, D. (1992): A Contribution to the Empirics of economic Growth, *The Quarterly Journal of Economics*, (mai).

MARQUES, R. (1999). O Financiamento do Sistema Público de Saúde Brasileiro. Proyecto CEPAL/GTZ "Reformas a los sistemas de salud em América Latina". Chile. Publicação das Nações Unidas.

MEEUSEN, W. e VAN DEN BROECK (1977), Efficiency Estimation From Cobb-Douglas Production Functions With Composed Error, *International Economic Review*, 18, 435-444.

MUSGRAVE R., MUSGRAVE, P. (1980): *Finanças Públicas – teoria e prática*. Editora Campus, Rio de Janeiro, Brasil.

NEHRU, V. & Dhareshwar (1993), A New database on physical capital stock: sources, methodology and results, *Revista de Análisis Económico*, Vol. 08, nº 01, pp 37-59.

PINDICK, R., RUBINFELD, D. (1998): *Econometric Models and Economic Forecasts*. McGraw-Hill.

Pnud, *Informe sobre Desarrollo Humano*, 2000. Eletronic adress:<http://www.undp.org>

REZENDE, f. (2001): *Finanças Públicas*. Editora Atlas, São Paulo, Brasil.

ROMER, D. (1996): *Advanced Macroeconomics*. McGraw-Hill.

SCHMIDT, P. (1976), On the Statistical Estimation of Parametric Frontier Production Functions, *Review of Economics and Statistics*, 58, 238-239.

STEVENSON, R. (1980), Likelihood Functions for Generalized Stochastic Frontier Estimation, *Journal of Econometrics*, 13, 57-66.

TEMPLE, J. (1999): The New Growth Evidence, *Journal of Economic Literature*, (mar).

World Bank –Latin American and Caribbean Studies, *Closing the gap in education and technology*, 2002. Electronic address: <http://www.worldbank.org>. Chapter5

Fontes Primárias de Dados



Anuário estatístico do IBGE (1996-2000), IBGE – Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro.

Pesquisa Nacional por Amostra de Domicílios (1996 -1998), IBGE – Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro.