

A MEDIÇÃO DA EFICIÊNCIA NO SECTOR HOSPITALAR

O caso português

Rui B. Leote de Paiva

Tese dirigida pelo Prof.
Doutor Carlos Silva Ribeiro e
apresentada na Universidade
Técnica de Lisboa

Universidade Técnica de Lisboa
Instituto Superior de Economia e Gestão



AGRADECIMENTOS

Quero agradecer em primeiro lugar ao Professor Carlos Murillo que aceitou acompanhar os trabalhos desta tese durante a minha permanência na Universidade de Barcelona que durou aproximadamente dois anos lectivos. O Professor Carlos Murillo estabeleceu uma forte relação de trabalho, habituando-me a recorrer a ele sempre que julguei necessário.

Ao Professor Carlos Silva Ribeiro que acompanhou em Lisboa a elaboração da tese, agradeço-lhe a disponibilidade que sempre manifestou e a compreensão pelas vicissitudes por que este trabalho passou.

Tendo esta tese como pressuposto a dedicação em tempo integral à sua elaboração, devo agradecer ao Departamento de Matemática de ISEG pela dispensa de serviço docente durante três anos lectivos, e à INVOTAN pela bolsa concedida para a minha permanência a Barcelona. Às Universidades de Barcelona e Técnica de Lisboa devo também agradecer os meios materiais postos à minha disposição.

A parte empírica desta tese, feita com dados portugueses, foi possível graças ao apoio do Ministério da Saúde e concretamente a Direcção Geral dos Hospitais e o Departamento de Gestão Financeira dos Serviços de Saúde, onde, respectivamente o Dr. André Gíria e a Dra. Suzette Tranquada intervieram no sentido de se disponibilizarem os dados necessários. Os dados referentes aos recursos humanos dos hospitais foram obtidos graças à intervenção do Professor Correia de Campos; para estes dados obtive um financiamento do Departamento de Matemática do ISEG.

Por todo o seu apoio e sacrifício agradeço a minha mulher. A todos os meus amigos que quizeram ter uma palavra de encorajamento, o meu muito obrigado.

Í N D I C E

Capítulo 1 - Introdução	1
Capítulo 2 - O sistema sanitário português e a evolução recente dos serviços de saúde na Europa ...	9
2.1 - O sistema sanitário português	11
2.2 - Os países da OCDE nos anos 80 e o seu sistema sanitário	19
Capítulo 3 - Funções fronteira e medida da eficiência ..	30
3.1 - Funções fronteira e eficiência	32
3.1.1 - A eficiência	35
3.1.2 - Produtividade e eficiência	38
3.1.3 - A evolução tecnológica.....	40
3.2 - Especificação e estimação de funções fronteira com dados seccionais	44
3.2.1 - Modelos com fronteiras não paramétricas ...	45
3.2.2 - Modelos não estatísticos de fronteira paramétrica e determinística	48
3.2.3 - Modelos estatísticos de uma fronteira paramétrica e determinística	49
3.2.4 - Fronteiras estocásticas	53
3.2.5 - Factores-x e factores-z	56

3.3	- Fronteiras estocásticas e dados de painel .	60
3.3.1	- Estimaco com dados de painel	61
3.3.2	- Estimaco com dados de painel e funoes de fronteira	73
3.4	- Funoes de custo de fronteira	75
Capítulo 4	- A eficincia no sistema hospitalar	79
4.1	- A unidade hospitalar num contexto no lucrativo	81
4.2	- Funoes de custo hospitalares	94
4.3	- A medio do output dos hospitais	103
4.4	- A medio do output e a estimaco de funoes de custo e de produo	112
Capítulo 5	- A eficincia no sistema hospitalar portugus	116
5.1	- Os dados	118
5.1.1	- Descrio dos dados da amostra	120
5.1.2	- Algumas medidas de output no caso portugus	122
5.2	- Obteno da ineficincia de custo	127
5.3	- Funo de produo hospitalar e separaco da ineficincia de custo em ineficincincia tcnica e de afectaco ...	141
5.4	- Estimaco de uma forma funcional flexvel: a translog	157
Capítulo 6	- Conclusoes	167
Bibliografia	173

R E S U M O

Neste trabalho pretendemos aplicar funções fronteira na medição da eficiência do sector hospitalar. Assim, depois de um capítulo de introdução, o capítulo 2 refere-se à actualidade do problema da medição da eficiência no sistema sanitário e em particular no sistema hospitalar.

O capítulo 3 aborda da medição do desempenho das empresas e mais especificamente as funções de fronteira. Estas funções são estudadas nas suas implicações econométricas e também como instrumentos de medição da eficiência das empresas.

Seguidamente, no capítulo 4, pretendemos diagnosticar alguns problemas suscitados pela aplicação de funções fronteira ao sector hospitalar.

A segunda parte do trabalho é a aplicação ao caso português dos instrumentos desenvolvidos anteriormente. Os resultados são constituídos pela análise do sector com estes instrumentos e as consequentes orientações de política sanitária.

CAPÍTULO 1

INTRODUÇÃO

A despesa pública com a saúde cresceu a um ritmo muito superior ao do PIB durante as décadas de 60 e 70, na generalidade dos países da OCDE, mas nos anos 80, numa época de "frugalidade fiscal", muitos países questionaram o esforço financeiro do estado nos níveis que eram praticados naquela época. Não se punha em questão o papel das administrações públicas na consecução de objectivos de equidade dos cidadãos perante a doença, ou a filosofia de grande intervenção no sector, que em muitos países se materializou num Serviço Nacional de Saúde, mas, pelo contrário, pretendia-se conseguir uma maior "eficiência" com a despesa pública no sector. Deste modo, e seguindo os princípios dominantes no período, afirma-se a confiança no papel do mercado, dos preços e da busca de incentivos não muito diferentes dos que são próprios das organizações privadas para os agentes que actuam no sector sanitário.

Esta busca da eficiência passa em pouco tempo ao sector hospitalar. Este sector é, na generalidade dos países europeus da OCDE, o responsável pela maior parte da despesa pública com a saúde. Não admira portanto que o problema da medição da eficiência económica das unidades hospitalares que apresentamos nesta tese, seja um tema de actualidade para a avaliação da actuação no sector.

Igualmente não estranha que os instrumentos de medição utilizados não sejam muito diferentes dos utilizados em empresas privadas actuando em mercados competitivos.

É conveniente esclarecer desde o início o que entendemos por eficiência. Dizemos que um hospital é (tecnicamente) eficiente quando o seu output coincide com o output máximo atingível com a dotação de inputs disponível. Dizemos igualmente, que um hospital tem eficiência de custo, quando o custo observado coincide com o custo mínimo atingível quando se produz um determinado volume de output e se enfrenta um determinado vector de preços para os inputs. Fala-se de ineficiência quando se compara aqueles níveis máximo ou mínimo com os níveis efectivamente observados, e estes são diferentes dos óptimos. As diferenças existentes proporcionam medidas úteis da ineficiência.

Este conceito de eficiência só é bem compreendido quando posto no contexto da função de produção e da função de custo hospitalares. O output máximo e o custo mínimo, tal como o definimos, coincidem com os proporcionados pela função de produção e pela função de custo. Dispondo-se, por um lado, só de observações sobre unidades hospitalares cujo output é inferior aquele máximo ou, na melhor das hipóteses, coincide com o máximo atingível e, por outro lado, de observações sobre os custos que são superiores ao mínimo atingível ou, na melhor das hipóteses, coincide com aquele mínimo, teremos então perturbações aleatórias menores ou iguais a zero, no caso da função de produção, ou positivas ou nulas no caso

da função de custo, quando se estima estas funções de explicação das condutas optimizadoras. Estas funções recebem na literatura a designação de "funções de fronteira".

Esta característica das funções fronteira afasta as hipóteses habituais de média nula e normalidade das perturbações aleatórias e afasta também estas funções, das funções (de produção e de custo) estimadas por MQO. Estas funções, a que chamaremos funções médias, pressupõem perturbações aleatórias com função de densidade de probabilidade simétrica em relação à origem. As funções médias admitem que os valores assumidos pelo output (ou pelos custos) podem ser maiores (ou menores) que os que são previstos pela função de produção (ou de custo), facto que contradiz a teoria microeconómica que define aquelas funções como funções de fronteira.

Como veremos mais adiante, o conceito de eficiência tal como o introduzimos aqui, com o uso das funções fronteira, revela-se superior a outros conceitos desenvolvidos também com a finalidade de medir o desempenho das unidades produtivas. É o caso do conceito de produtividade (pelo facto de não se referir ao conceito de função de produção), ou a medida do deslocamento da função de produção no tempo (por não se referir ao conceito de fronteira).

De entre os problemas que surgem na estimação de fronteiras no contexto do sistema hospitalar, sobressai o relativo à ineficiência estrutural. Se pelo seu estatuto de propriedade ou pelas relações

que os agentes mantêm entre si as unidades hospitalares se afastam do comportamento que a teoria da produção considera paradigmático, então este sector poderá incorrer em ineficiência estrutural. Este problema tem consequências na estimação de fronteiras: não poderemos associar à fronteira estimada o sentido de fronteira absoluta. Contudo, é possível outorgar um sentido relativo à fronteira estimada. Este aspecto relativo pretende significar que a actividade produtiva está associada a um determinado quadro institucional do sistema.

Outro problema empírico é que as funções de custo habituais na literatura da economia da saúde são do tipo "behavioral" ou "ad hoc". Quer dizer, são funções que incluem como variáveis explicativas os indicadores da actividade hospitalar (número médio de casos por cama, taxa de ocupação, ou a duração média da estadia hospitalar) além da variável de dimensão que o número de camas constitui. Este tipo de funções é pouco apropriado à estimação de uma fronteira de custo que em rigor e enquanto estimativa de custos mínimos, deve ser uma função dual de uma função de produção, tal como a definimos anteriormente e na forma que se especifica habitualmente.

Finalmente, um terceiro problema na estimação de fronteiras num contexto hospitalar é o relativo à medição do output. Lopez e Wagstaff (1988) e Wagstaff (1989) encontraram sérios problemas de multicolinearidade na estimação, respectivamente, da função de produção e da função de custo. Concluimos, com um diagnóstico

apropriado, que aquela colinearidade resultava da forma como mediam o output estes autores. De facto, consideravam em ambas as funções o vector de case-mix (ou seja, a percentagem das altas em cada uma das especialidades médicas) como regressor, o que era justificado pela necessidade de ter em conta a heterogeneidade do número de casos como medida do output. A sua aproximação era uma abordagem multi-produto transformada. Contudo, no caso espanhol e também no caso português, o vector de case-mix é altamente colinear, pelo que as estimativas onde se inclui este grupo de variáveis como regressores, são muito instáveis. Deveríamos, então adoptar uma medida de output que reduzisse drasticamente a dimensão do vector de case-mix, mantendo ao mesmo tempo a propriedade de ter em conta a heterogeneidade das altas como medida do output.

O trabalho está organizado da seguinte forma. O capítulo 2 inclui um breve apontamento sobre a pertinência da abordagem em termos de eficiência económica do sistema hospitalar, comparando-se a evolução dos sistemas e políticas sanitárias nos países europeus da OCDE. Este capítulo inclui também uma caracterização sintética do sistema hospitalar português.

O capítulo 3 refere-se ao estudo das funções de fronteira, pretendendo-se a sua caracterização e o estudo dos problemas econométricos que suscitam, assim como a comparação com outras medidas do desempenho das empresas. Dêmos uma atenção especial ao problema da estimação com dados de painel e à sua relação com a estimação de fronteiras.

No capítulo 4 abordam-se alguns problemas suscitados pela estimação de fronteiras no contexto hospitalar. Fundamentalmente o tema tratado refere-se à ineficiência estrutural no sistema hospitalar público ou não lucrativo, e, igualmente, os problemas da especificação de funções de custo e medição do output hospitalar.

O capítulo 5 trata da medição da eficiência no sistema hospitalar português. A primeira medição da eficiência de custo realiza-se com uma função de custo com uma especificação intermédia entre as especificações "behavioural" e as funções de custo duais de uma função de produção: a função de custo é especificada como sendo do tipo Cobb-Douglas, mas é posteriormente transformada, decompondo-se o output no produto de três variáveis de dimensão (camas, taxa de ocupação e duração média). Encontrámos uma elevada ineficiência média e também uma grande dispersão de aquela ineficiência de custo. Explorando mais este facto, encontrámos como variáveis explicativas destas variações entre centros a sua dimensão e uma variável artificial representando o tipo de centro. Juntando estas duas relações pode-se efectuar uma simulação do custo médio como função do número de camas. É possível encontrar um mínimo perto das 400 camas, valor que se situa ligeiramente acima da dimensão média dos centros da amostra.

Atendendo ao facto de que a ineficiência de custo é composta por ineficiência técnica e por ineficiência de afectação, procurámos numa segunda estimação a separação de aqueles dois tipos

de ineficiência. Para tal foi necessário a especificação de uma função de custo do tipo Cobb-Douglas e ter em conta os preços dos diferentes factores produtivos. Esta decomposição, que explora plenamente a dualidade entre a função de custo e a função de produção, exige a estimação desta última função. Com a função de produção pudémos comparar as proporções das despesas com os diferentes factores produtivos e concluir sobre a direcção de uma eventual reafecção de recursos.

Chegados a este ponto considerámos que era importante averiguar da possibilidade da reafecção de factores produtivos: era necessário ver quais eram os sinais e magnitudes das elasticidades de substituição entre factores. Como a função de produção Cobb-Douglas restringe ao valor 1 estas elasticidades, convém passar a formas funcionais flexíveis que não ponham restrições àqueles valores. Esta é a razão pela qual se estimou uma função de custo do tipo "translog". O trabalho completa-se com um capítulo de síntese e conclusões.

C A P Í T U L O 2

**O SISTEMA SANITÁRIO PORTUGUÊS E A EVOLUÇÃO RECENTE DOS SERVIÇOS
DE SAÚDE NA EUROPA**

A despesa sanitária cresceu na generalidade dos países da OCDE, nas últimas décadas, a um ritmo superior ao do PIB. Ao mesmo tempo, o financiamento da despesa sanitária era, numa grande proporção, realizada pelo Estado. Este facto tinha como consequência uma pressão crescente sobre o orçamento do Estado, num período em que os cidadãos sentiam os níveis como sendo excessivamente elevados . Isto contribuiu para a procura de incentivos para uma maior "eficiência" nos sistemas sanitários. Esta procura confundiu-se com a adopção de incentivos de tipo empresarial no sistema sanitário público com o objectivo da contenção das despesas. Não surpreende portanto, que no sistema hospitalar, onde são empregues a maior parte dos recursos públicos do sector, essa busca da eficiência tenha marcado a política do sector, assemelhando além disso a unidade hospitalar a uma empresa. Em Portugal essas tendências da política sanitária, ou mais especificamente da política hospitalar, reflectiram-se nomeadamente em novos sistemas de financiamento do sector, pretendendo medir o output hospitalar de uma forma mais rigorosa e associar o nível de financiamento ao nível de output.

Neste capítulo e numa primeira secção, pretendemos fazer uma breve caracterização do sistema sanitário português, com especial

atenção ao sistema hospitalar. Numa segunda secção aborda-se a evolução recente dos sistemas sanitários dos países europeus da OCDE, pretendendo-se conhecer quais as linhas fundamentais daquela evolução.

2.1. - O SISTEMA SANITÁRIO PORTUGUÊS

Desde 1979 que o Serviço Nacional de Saúde garante a toda a população portuguesa o acesso gratuito ao sistema de cuidados de saúde. Este facto é resultante da sucessiva extensão da cobertura sanitária, começada em 1971, quando se reconhece o papel do Estado como garante do direito à saúde de toda a população. A reforma de 1971 foi importante por diversos motivos que vão desde a reorganização dos serviços de saúde, à criação de serviços de coordenação e administração do sistema, passando pela reestruturação da organização e funções do pessoal sanitário, mas sobretudo pela extensão da população coberta pelo sistema sanitário público e por uma nova filosofia dos cuidados de saúde onde a assistência na doença pretendia ser tendencialmente universal e onde os preços não constituíssem uma barreira de acesso ao sistema.

Não é senão depois da revolução de 1974 que se generaliza na população a consciência de que o acesso à assistência sanitária é

um direito de todos. É no período posterior à revolução que se populariza o recurso ao sistema sanitário público. A partir dessa altura todos os indicadores de utilização registam incrementos notáveis, reconhecendo-se constitucionalmente que o direito à saúde da população deve materializar-se na existência de um sistema sanitário público gratuito e com cobertura universal (1976). A expansão do sistema foi gradual até à criação do Serviço Nacional de Saúde (SNS) em 1979.

Apesar da existência do SNS universal e gratuito, cerca de 24% da população portuguesa está coberta actualmente por esquemas de assistência sanitária próprios de certos grupos profissionais (os funcionários públicos, os militares, os funcionários bancários) e, em consequência, não recorrem sistematicamente ao sistema sanitário público. Além disso, só 2% da população está coberta por um seguro de doença privado. (Correia de Campos - 1991).

A distribuição funcional da despesa sanitária do Serviço Nacional de Saúde é tal que, aproximadamente 43% é destinada aos cuidados hospitalares, 33% aos cuidados primários e 21% a despesa com medicamentos. Esta despesa é, em todo o caso, integralmente financiada pelo Orçamento do Estado (Correia de Campos - 1991).

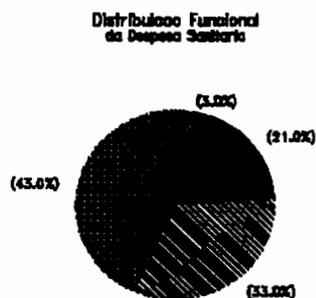


Figura 2.1.

Apesar da constituição de 1976 ter como objectivo a socialização da medicina, o SNS não oferece todos os tipos de cuidados de saúde. Cerca de 36% do seu orçamento destina-se ao financiamento de serviços prestados pelo sector privado ("DRG's in Portugal" - 1992).

Portugal gastava no sector da saúde em 1987 cerca de 6.4% do seu PIB. Este valor representa um esforço superior ao realizado por outros países com serviços nacionais de saúde como, por exemplo, a Espanha e a Inglaterra. Contudo, a despesa pública com o sector era só de 3.9% do PIB o que é, por seu lado, a menor percentagem de todos os países da CEE (Correia de Campos - 1991).

O esforço de financiamento privado das despesas de saúde (cerca de 40% do total) destina-se principalmente à despesa com medicamentos (54% daquele valor) e a consultas médicas. Contudo, 73% das hospitalizações e 67% das consultas médicas eram realizadas pelo SNS; mas só 48% dos exames radiológicos e cerca de 30% das

análises clínicas são feitas no SNS. Igualmente só 48% das consultas a especialistas e 16% do tratamento de estomatologia são realizadas no sector público (Correia de Campos - 1991). Outros cuidados ambulatoriais especializados, como é o caso, por exemplo, das diálises e da fisioterapia, efectuam-se predominantemente no sector privado, em centros cujos profissionais costumam exercer também no SNS. Com efeito, a maioria dos médicos (e outros profissionais de saúde) que garantem o funcionamento dos centros privados, exercem habitualmente também no SNS. Esta simultaneidade de funções conduz à sub-utilização dos meios e equipamentos dos hospitais públicos.

Os hospitais que não são do sector público (4% das camas no sector com fins lucrativos e 18% das camas no sector não lucrativo) concentram a sua actividade na cirurgia de baixa complexidade e em partos (Correia de Campos et al. (1990)). Assim é possível dizer que a maior parte dos meios auxiliares de diagnóstico e tratamentos ambulatoriais especializados são realizados pelo sector privado, enquanto no que respeita à hospitalização em regime de internamento o sector público é predominante.

Em Portugal existem 87 hospitais no sector público, 24 dos quais são Hospitais Centrais (que incluem 7 hospitais especializados, 3 maternidades e 3 centros de oncologia) e 63 são Hospitais Distritais, dos quais 24 eram anteriormente Hospitais Concelhios ("DRG's in Portugal" - 1992).

Enquanto quase todos os hospitais repartem a sua actividade entre internamento, consultas externas e urgências, os Hospitais Centrais diferenciam-se dos Distritais porque estão situados em grandes centros urbanos. Os Hospitais Centrais têm um número de especialidades médicas (e os correspondentes serviços clínicos) superior a 30, enquanto que os Hospitais Distritais raramente excedem os 15 serviços especializados.

Os Hospitais Centrais e Distritais constituem uma hierarquia de serviços entre os níveis regional e nacional; mas a esta hierarquia não corresponde sempre a localização mais eficiente dos serviços médicos, pela dificuldade de persuadir os médicos de certas especialidades a estabelecerem-se nas regiões periféricas.

A administração e o planeamento dos serviços de saúde estão organizados também em dois níveis: o nível nacional, do qual dependem os Hospitais Centrais e Distritais, e o nível regional (as Administrações Regionais de Saúde - ARS's- coordenadas centralmente pela Direcção Geral dos Cuidados de Saúde Primários) do qual dependem os 343 Centros de Saúde que têm a responsabilidade dos cuidados primários, embora fora dos grande centros urbanos tenham também, por vezes, internamento.

Esta dupla estrutura de poderes faz com que exista pouca coordenação entre cuidados hospitalares e assistência primária, apesar da sua necessidade ser reconhecida cada vez com maior intensidade dado o impacto da medicina comunitária e preventiva.

Esta dualidade do sistema tinha como consequência que, nas áreas urbanas, o internamento nos grandes hospitais, passa quase exclusivamente pelas urgências (Correia de Campos et al.(1990)).

Até ao final da década de 70 o financiamento do sistema hospitalar baseava-se na análise pormenorizada da contabilidade dos hospitais e no crescimento do capítulo correspondente do Orçamento do Estado. Durante a década de 80 utilizou-se um novo sistema de financiamento com o objectivo de melhorar a produtividade e gestão dos centros. As principais características deste sistema são o pagamento por acto clínico com preços baseados nos custos directos de cada um dos dois grupos de hospitais, quer dizer, Hospitais Centrais e Distritais.

As unidades de pagamento são as seguintes:

- . Em relação ao internamento as altas são pagas de acordo com os custos directos e com um preço por cada especialidade médica de acordo com estimativas da duração média da estadia e do índice de ocupação, dentro de intervalos estabelecidos antecipadamente.
- . Nas consultas, urgências e serviços complementares de diagnóstico, o pagamento é por acto clínico mediante um sistema de preços fixados anualmente de acordo com os custos directos.
- . A alimentação é paga por cada dia de internamento para os doentes e por refeição para o pessoal, com preços de acordo com os custos directos.

- . Os serviços administrativos não têm unidade de medida específica. O pagamento efectua-se em função do número de dias de internamento, convertendo-se naquela unidade as consultas externas e as urgências, de acordo com a seguinte fórmula:

$$N = \text{"estadias"} + a \cdot \text{"urgencias"} + b \cdot \text{"consultas"}$$

onde a e b são os parâmetros de conversão a estimar.

- . Nos outros serviços, os pagamentos realizam-se em função dos custos apresentados pelos hospitais.

O financiamento de cada unidade está limitada superior e inferiormente, como consequência dos limites fixados pelo Orçamento do Estado para o conjunto do sistema. O financiamento retido para cada hospital é calculado da forma exposta quando o valor que proporciona está compreendido dentro daqueles limites. Caso contrário o financiamento resultante coincide com aquele limite.

Estes sistemas de financiamento enfrentaram vários tipos de críticas. Assim, o primeiro é um sistema pelo qual os hospitais têm um financiamento baseado nos custos observados, o que é criticável por não terem incentivos para uma maior eficiência na gestão: não se penalizam os custos unitários elevados, antes, pelo contrário, estes contribuem para financiamentos mais elevados.

O segundo sistema tem a vantagem de vincular o financiamento aos resultados (output), mas tem ainda alguns inconvenientes. Com efeito, a medição do output é pouco precisa por ter uma

classificação do case-mix baseada em categorias pouco homogéneas. O sistema de pagamento por acto clínico é apontado como estimulando a sobre-utilização dos serviços auxiliares, e incluir incentivos para a substituição de meios de diagnóstico de baixo custo por meios caros, por existirem meios com custos diferentes pagos ao mesmo preço.

Este tipo de críticas contribuiu para que se considerasse um novo sistema de financiamento baseado no output, mas melhorando a forma da sua medição. Trata-se de estabelecer como categoria básica do case-mix o conceito de Grupo de Diagnóstico Homogéneo (DRG) (Fetter, 1980) que é uma categoria homogénea em relação às estadias e consumo de serviços. Este sistema de financiamento, a que voltaremos mais adiante, começou a aplicar-se de forma experimental em 1990.

Apesar dos diferentes modos de financiamento, existe o problema do sub-financiamento crónico do sistema, que contribui para limitar o poder negocial das administrações dos hospitais, para aumentar os custos médios devido ao atraso no cumprimento dos compromissos, diminuir a confiança no sistema por parte de utentes e fornecedores.

A análise do sistema sanitário português pode ser encontrada em mais detalhe em :

- . Correia de Campos, A. (1983, 1990, 1991(a), 1991(b)).
- . Correia de Campos, A., Caldeira da Silva, Costa, C. (1990)

. "DRGs in Portugal"- Relatório sobre a aplicação dos DRG's ao financiamento dos hospitais portugueses.

. Mantas, A., et al.(1989).

2.2. - OS PAÍSES DA OCDE NOS ANOS 80 E O SEU SISTEMA SANITÁRIO

Nas duas últimas décadas assistiu-se a um significativo aumento das despesas com a saúde, que cresceram a um ritmo superior ao do PIB na generalidade dos países da OCDE :

QUADRO 1 - DESPESAS COM A SAÚDE E PIB PER CAPITA 1970-1987⁽¹⁾

PAÍS	1970			1987			TAXA DE CREC.	
	DESP. SAÚDE	PIB	DSPSS/PIB	DESP. SAÚDE	PIB	DSPSS/PIB	DESP. SAÚDE	PIB
ALEMAN.	220	3 993	.055	1 072	13 308	.081	9.8	7.3
AUSTRIA	163	3 056	.053	988	11 710	.084	11.2	8.2
BELGICA	147	3 652	.040	881	12 183	.072	11.1	7.3
DINAMA.	252	4 147	.061	784	13 129	.060	6.9	7.0
ESPAÑA	102	2 473	.041	521	8 676	.060	10.1	7.7
FINLAN.	183	3 280	.056	970	13 061	.074	10.3	8.5
FRANÇA	223	3 685	.061	1 117	12 849	.087	9.9	7.6
GRECIA	70	1 756	.040	337	6 410	.053	9.7	7.9
HOLAND.	232	3 881	.060	1 041	12 263	.085	9.2	7.0
IRLAND.	122	2 196	.056	553	7 446	.074	9.3	7.4
ISLAND.	288	3 382	.085	1 205	15 566	.077	8.8	9.4
ITALIA	171	3 093	.055	837	12 190	.069	9.8	8.4
NORUEG.	191	3 083	.062	1 149	15 495	.074	11.1	10.0
PORTU.*	105	2 240	.046	385	5 979	.064	10.5	7.8
RE.UNI.	161	3 562	.045	763	12 414	.061	9.6	7.6
SUECIA	359	4 976	.072	1 233	13 770	.090	7.5	6.2
MEDIA**	192	3 347	.057	896	12 031	.073	9.6	7.8

(1) A unidade monetária retida é o dólar; a taxa de crescimento é a taxa de crescimento anual média.

* Os dados de Portugal referem-se aos anos 1974 e 1987

** Média não ponderada dos países do quadro excepto Portugal

FONTE: OECD - "Health Care Systems in Transition - The Search for Efficiency"-Paris 1990

As despesas com a saúde têm uma participação crescente no PIB, sendo a elasticidade das despesas com a saúde em relação ao PIB maior do que a unidade, apesar desta elasticidade ter diminuído nos últimos anos, como se pode ver no quadro 2.

QUADRO 2 - ELASTICIDADES DA DESPESA SANITÁRIA EM RELAÇÃO AO PIB REAL

PAÍS	1960-1975	1975-1984
ALEMANHA	1.2	0.9
AUSTRIA	0.7	0.7
BELGICA	1.3	1.5
DINAMARCA	1.9	1.4
ESPAÑA	1.7	2.1
FINLANDIA	2.0	0.9
FRANÇA	1.6	2.6
GRECIA	1.8	1.8
HOLANDA	1.5	0.5
IRLANDA	2.3	0.9
ITALIA	0.9	1.3
NORUEGA	1.7	1.5
PORTUGAL	-	1.1
REINO UNIDO	2.1	1.0
SUECIA	2.4	1.6
MEDIA*	1.7	1.3

* Não ponderada e sem incluir Portugal

FONTE: OECD - "Health Care Systems in Transition - The Search for Efficiency"-Paris 1990

Poullier (1990) calculou a partir da regressão linear simples das despesas com a saúde sobre o PIB e com dados de tipo "cross-section" para os países da OCDE, a elasticidade da despesa sanitária, encontrando valores superiores à unidade, resultado que concorda com o apresentado no quadro anterior. Estas regressões

apresentam R^2 (próximos de 0.9) e raios t elevados.

QUADRO 3 - DECOMPOSIÇÃO DA TAXA DE CRESCIMENTO DA DESPESA SANITÁRIA E TAXA DE CRESCIMENTO DO DEFLACTOR DO PIB (1960-84)

PAÍSES	TAXA DE CRESCIMENTO ANUAL				
	CRESCIMENTO DA DESPESA SANITÁRIA				DEFLAC- TOR DO PIB
	TOTAL	PREÇO	POPULA.	SERVIÇOS	
ALEMANHA	10.1	5.6	0.4	4.1	4.3
AUSTRIA	11.3	8.3	0.3	2.7	5.1
BELGICA	11.8	6.3	0.3	5.2	5.4
DINAMARCA	14.1	8.2	0.5	5.4	8.2
ESPAÑA	21.8	13.0	1.0	7.8	12.1
FINLANDIA	15.4	8.1	0.4	6.9	8.8
FRANÇA	15.3	6.9	0.8	7.6	7.5
GRECIA	18.3	9.3	0.7	8.3	10.3
HOLANDA	13.7	8.4	1.0	4.3	6.2
IRLANDA	18.2	10.0	0.9	7.3	10.3
ISLANDIA	34.8	30.2	1.3	3.3	27.4
ITALIA	17.6	10.5	0.5	6.6	10.5
NORUEGA	14.5	8.4	0.6	5.5	7.1
PORTUGAL*	27.2	19.6	1.1	6.5	20.8
REINO UNIDO	13.1	8.3	0.3	4.5	8.7
SUECIA	13.7	6.0	0.5	7.2	7.3
SUIÇA	12.1	6.7	0.9	4.5	5.0
MEDIA**	16.0	9.6	0.7	5.7	9.0

* Os dados portugueses referem-se ao período 1974-84

** Não ponderada e sem incluir Portugal

FONTE: OECD - "Health Care Systems in Transition - The Search for Efficiency"-Paris 1990

O crescimento das despesas correntes em saúde, na generalidade dos países da OCDE, vem explicado em parte pelas variações dos preços dos serviços sanitários, mas também pelo importante crescimento do consumo de serviços sanitários per capita, quer

dizer, pela intensidade da utilização. A construção do quadro 3 está baseada na decomposição da taxa de crescimento da despesa sanitária, nas taxas de crescimento dos preços, da população e das quantidades físicas per capita, quer dizer:

$$DSPSS = PREÇO \times POPUL \times \frac{Q}{POPUL}$$

onde:

- . PREÇO é o incremento no nível geral de preços dos serviços sanitários
- . POPUL é o incremento da população, e
- . Q é o incremento da quantidade física oferecida de serviços de saúde

Pode-se ainda observar no quadro 3 que o crescimento dos preços no sector sanitário foi, na generalidade dos países, superior ao crescimento do deflactor do PIB.

Na maior parte dos países da OCDE a despesa sanitária foi sobretudo financiada pelo Estado. No quadro 4 pode-se observar a distribuição da despesa sanitária pública pelas suas principais componentes, podendo-se comprovar que o seu principal destino é o sector hospitalar.

QUADRO 4 - DECOMPOSIÇÃO DA DESPESA PÚBLICA COM A SAÚDE

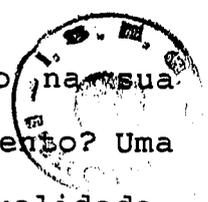
PAÍS E ANO	HOSPITAL.	AMBULAT.	FARMACEU.	OTROS
AUSTRIA, 1983	25.3	20.3	9.9	44.7
BELGICA, 1981	21.0	37.7	11.8	29.5
DINAMARCA, 1984	73.9	22.0	4.8	-
FINLANDIA, 1983	55.2	28.1	5.9	10.7
FRANÇA, 1984	59.5	22.9	13.1	7.9
ALEMANHA, 1983	43.0	25.5	19.2	12.3
GRECIA, 1982	49.5	13.4	14.8	22.3
IRLANDA, 1983	73.4	11.5	7.0	9.7
ITALIA, 1984	55.3	27.8	13.0	4.6
HOLANDA, 1984	69.3	23.2	7.2	3.7
NORUEGA, 1981	69.9	15.3	7.2	7.6
PORTUGAL, 1983	46.3	20.7	20.3	12.7
ESPANHA, 1981	42.5	16.7	15.8	25.7
SUECIA, 1983	72.6	10.2	4.9	12.3
REINO UNIDO, 1979	59.7	11.2	10.3	20.0
MEDIA*	54.4	20.4	11.0	14.9

* Não ponderada

FONTE: OECD - "Health Care Systems in Transition - The Search for Efficiency"-Paris 1990

Alguns autores tentaram explicar o comportamento da despesa sanitária pública. É o caso de Buchanan (1965) e Leu (1986). Para Buchanan, que se baseou na análise do SNS britânico, cada indivíduo enfrenta decisões contraditórias na sua dupla condição de consumidor e contribuinte. Na sua condição de consumidor de um bem (cuidados de saúde) do qual obtém utilidade e pelo qual não paga nada directamente, aumentará a sua procura indefinidamente, enquanto a sua utilidade marginal for positiva. Contudo, na sua condição de contribuinte, reconhece que cada escudo de imposto compete com outras utilizações (educação, segurança social, etc) das quais também obtém utilidade.

O resultado desta dupla posição é o racionamento, na sua qualidade de consumidor. Mas como se realiza este racionamento? Uma possibilidade é que o sistema ofereça serviços de baixa qualidade, ou estabeleça filas de espera produzidas pela saturação dos serviços, como forma de inibir a procura.



Uma análise alternativa é a de Leu (1986). Para Leu existem três tipos de agentes que determinam o nível de despesa sanitária: os doentes, quem oferece os cuidados de saúde e quem os financia. Estes últimos (o Estado, os seguros privados ou a beneficência) contribuem para o aumento da despesa sanitária ao baixar o preço ao qual o consumidor paga os cuidados de saúde, independentemente da remuneração recebida pela entidade que oferece os serviços de saúde.

Deve-se esperar que, em geral, a despesa sanitária aumente com a percentagem de financiamento por parte do Estado. Note-se que a argumentação anterior refere-se ao financiamento da despesa sanitária por um terceiro pagador. Este terceiro agente pode não ser quem forneça os cuidados de saúde. Contudo, argumenta Leu, se quem fornece é o Estado devem-se esperar custos maiores para cada quantidade oferecida. A falta de concorrência no sector público, e a conseqüente falta de incentivos para que os administradores minimizem os custos, além da sua conduta maximizadora de orçamento, faz com que seja de esperar um excesso de oferta com inflação de custos.

Para Leu, o nível de despesa sanitária cresce com o aumento da percentagem da oferta pública na oferta total. Leu considera um modelo explicativo do nível da despesa sanitária per capita nos diferentes países da OCDE, trazendo evidência empírica às suas teses.

Contudo, o argumento de que a oferta de cuidados de saúde pelo sector público é relativamente ineficiente é contraditória com resultados de outros autores. Assim, Bays (1977) com dados dos Estados Unidos, e Buttler, no caso australiano, recolhem evidência empírica que os hospitais com fins lucrativos especializaram-se em case-mix menos complexos, concentrando-se, por exemplo, em cirurgia de menor complexidade. Por seu lado, Stoddart e Labelle (1985) passaram em revista o tema e concluíram que a evidência empírica não confirma a existência de custos de produção mais baixos no sector com fins lucrativos. Uma apresentação mais detalhada de este tema encontra-se mais adiante no nosso trabalho.

Apesar de a longo prazo o crescimento da despesa sanitária ser mais rápida que o crescimento do PIB, nos últimos anos pode assinalar-se uma tendência para a estabilização do peso das despesas em saúde no PIB, nos países da OCDE (figura 2.2).

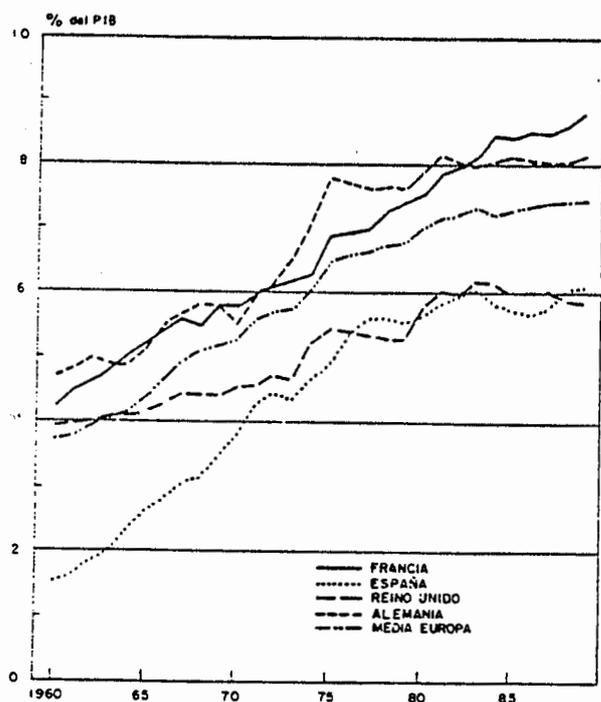


Figura 2.2

QUADRO 5 - INFLAÇÃO ESPECÍFICA DOS PREÇOS DOS SERVIÇOS DE SAÚDE (1960/87)(*) - TAXAS ANUAIS DE CRESCIMENTO

PAISES	60/87	60/75	75/87	60/70	70/80	80/87
ALEMANHA	0.8	1.2	0.4	1.5	0.5	0.5
AUSTRIA	2.5	3.3	1.6	4.0	2.8	1.2
BÉLGICA	0.9	1.0	0.7	1.5	0.3	0.8
DINAMARCA	-0.1	0.6	-0.9	0.6	-0.8	0.0
ESPAÑA	0.7	1.0	0.4	2.6	0.0	1.1
FINLANDIA	-0.6	-1.9	0.7	-2.3	-0.5	1.3
FRANÇA	-0.6	-0.3	-1.1	0.3	-1.1	-1.2
GRECIA	-0.4	-0.5	-0.3	-2.3	0.5	-1.3
HOLANDA	1.8	2.6	0.7	1.2	3.6	0.0
IRLANDA	-0.1	-1.2	1.3	-0.8	-0.1	0.7
ISLANDIA	1.5	2.3	0.7	1.3	3.2	-0.4
ITALIA	0.6	0.6	0.1	-0.7	-0.3	0.2
NORUEGA	1.6	2.2	0.9	0.0	0.8	1.5
REINO UNIDO	-0.2	-1.1	0.9	1.5	-0.5	1.2
SUECIA	0.4	0.1	1.0	1.2	1.4	0.7
SUIÇA	1.6	2.0	1.0	-0.4	2.3	0.8
MEDIA**	0.7	0.8	0.5	0.6	0.8	0.4

* Aumento de preços dos serviços sanitários acima do índice geral de preços

** Não ponderada

FONTE: OECD - "Health Care Systems in Transition - The Search for Efficiency"-Paris 1990

A contenção da despesa sanitária, em termos reais, depende da contenção das suas duas componentes: por um lado a inflação específica dos preços dos serviços sanitários, posto que é sabido que os preços dos serviços sanitários evoluíram a um ritmo superior aos preços da generalidade dos bens e serviços, e, por outro lado, do crescimento real do consumo de serviços sanitários. Os quadros 5 e 6 mostram a desaceleração de ambas as componentes nos diferentes países da OCDE.

QUADRO 6 - CRESCIMENTO REAL DO CONSUMO SANITÁRIO PER CAPITA (1960/87) - TAXAS ANUAIS DE CRESCIMENTO

PAISES	60/87	60/75	75/87	60/70	70/80	80/87
ALEMANHA	4.0	5.3	2.1	3.8	5.8	1.3
AUSTRIA	2.9	3.8	1.8	1.8	4.5	1.0
BÉLGICA	5.0	6.7	3.1	4.6	7.9	1.8
DINAMARCA	4.6	6.4	2.4	8.5	3.8	0.3
ESPAÑA	6.7	10.2	2.4	5.0	6.3	0.9
FINLÂNDIA	6.7	9.6	3.2	10.1	5.0	3.2
FRANÇA	6.4	7.6	4.7	7.1	6.7	4.2
GRÉCIA	6.5	8.5	4.0	13.2	4.0	4.4
HOLANDA	3.5	5.3	1.3	7.0	1.7	1.1
IRLÂNDIA	5.5	9.5	1.4	8.3	7.7	-0.6
ISLÂNDIA	6.7	7.7	5.0	5.6	5.9	5.9
ITALIA	5.4	7.2	4.2	9.7	6.5	2.6
NORUEGA	4.5	6.2	3.4	3.6	6.3	3.0
REINO UNIDO	3.9	5.5	1.9	6.4	4.9	1.6
SUECIA	4.5	6.9	1.5	11.3	3.1	0.3
SUIÇA	3.5	5.3	1.2	8.9	2.1	1.3
MEDIA*	5.0	7.2	2.7	7.2	5.2	2.0

* Não ponderada

FONTE: OECD - "Health Care Systems in Transition - The Search for Efficiency"-Paris 1990

Este facto mostra o relativo êxito das políticas restritivas que foram impostas na última década no sector sanitário, conjuntamente com a busca de incentivos de ordem empresarial, com o objectivo de conter a despesa sanitária e aumentar a eficiência. Um dos sectores onde se manifestou esta busca de eficiência foi o sector hospitalar. As políticas adoptadas tiveram como consequência a descida generalizada da duração média.

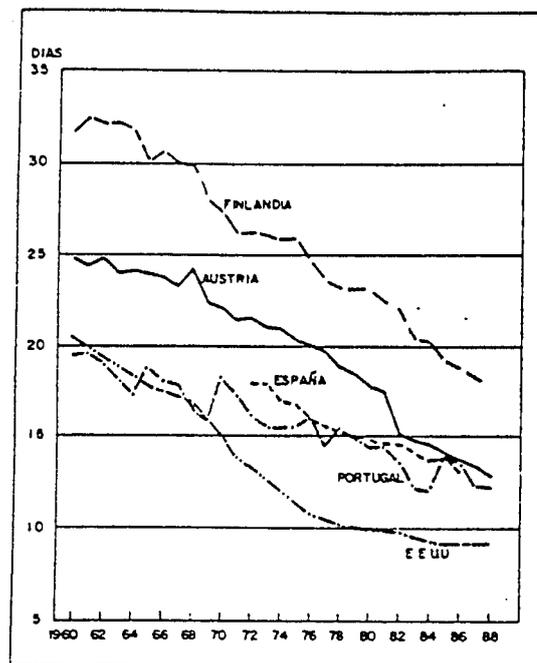


Figura 2.3

Uma das origens da grande variabilidade da duração média entre países é o sistema de financiamento hospitalar. Os sistemas de financiamento que associam o seu volume ao número de dias da estadia no hospital tendem a estar associados a maiores estadias. Por outro lado, as formas de pagamento prospectivo, que associam o financiamento a uma determinada quantidade e composição do output,

tenderão a reduzir a duração média e o custo por caso.

Uma das formas de pagamento prospectivo é a ligada à implantação dos DRG's (Grupos de Diagnóstico Homogéneo) como medida do output hospitalar. A introdução do sistema de financiamento baseado nos DRG's procura a diminuição dos custos por dia de estadia e mais acentuadamente do custo por caso.

Uma reforma do sistema hospitalar que parece estar directamente associada à introdução de incentivos de ordem empresarial é a recente proposta de financiamento dos hospitais do Reino Unido (United Kingdom Department of Health (1989)). Esta reforma está baseada na ideia de que quem financia (e é responsável por que a assistência sanitária esteja disponível) pode não coincidir com quem fornece a assistência sanitária. Esta característica manifesta-se pela existência de hospitais com administração autónoma (constituída por um corpo de directores) que concorrem com hospitais sob a direcção das autoridades sanitárias, pelos doentes utentes do SNS. Os contratos entre as autoridades sanitárias e os hospitais com administração autónoma especificam a quantidade e qualidade do output. Espera-se que os incentivos de tipo empresarial sejam transmitidos ao sector hospitalar sob a direcção das autoridades sanitárias, que também têm fixados objectivos de quantidade e qualidade como os outros centros.

CAPÍTULO 3

FUNÇÕES FRONTEIRA E MEDIDA DA EFICIÊNCIA

O objectivo deste capítulo é a abordagem dos problemas das unidades produtivas em termos de eficiência. Esta abordagem tem consequências na especificação de funções de produção e de custo e nos métodos de estimação resultantes. O estudo da função de produção e da função de custo tem um interesse concreto na avaliação da actividade das empresas e em particular de serviços públicos gerando-se orientações para a melhoria da utilização dos recursos disponíveis.

As funções de produção, quando se pretende medir a eficiência das unidades produtivas, são especificadas como funções de fronteira desde Farrell (1957), isto é , há aproximadamente três décadas. Embora o conceito de função de fronteira tenha sido desenvolvido no contexto das funções de produção, podem, com adaptações simples, as funções de custo e de lucro também ser especificadas em termos de funções de fronteira.

Este capítulo trata o problema geral das funções de fronteira quer do ponto de vista da sua especificação econométrica como da estimação dos seus parâmetros. Na primeira secção definem-se as

funções fronteira assim como a forma de medição da eficiência com a sua utilização. Também se realiza um estudo comparativo com outros instrumentos de medida do desempenho das empresas correntemente utilizados.

Na segunda secção expõem-se os problemas suscitados pela estimação de fronteiras com dados "cross-section" e numa terceira secção os relativos à sua estimação quando os dados são do tipo painel.

3.1. - FUNÇÕES DE FRONTEIRA E EFICIÊNCIA

A função de produção define-se tradicionalmente como a fronteira do conjunto de produção, ou, de uma forma equivalente, o output máximo que se pode produzir com uma determinada dotação de inputs. Trata-se de um conceito limite que tem implícita uma conduta otimizadora que implica a noção de eficiência produtiva. Esta situação pode-se ilustrar, no caso mais simples de um só input e um só output, pela figura seguinte:

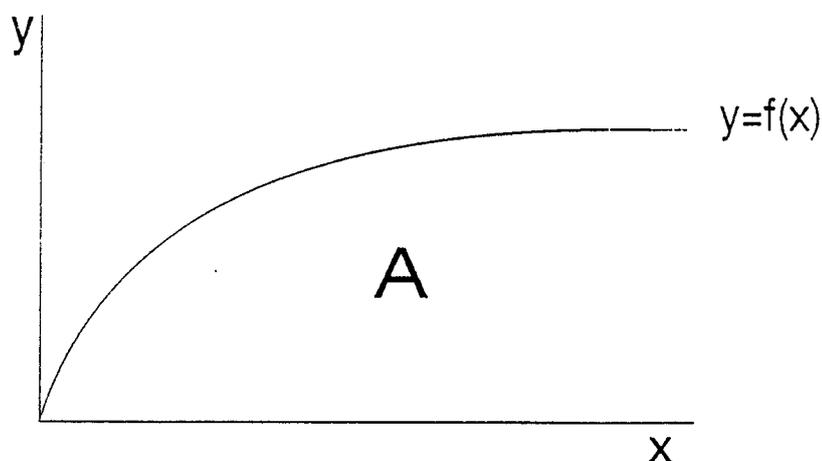


Figura 3.1.

Se x representa o input e y o output, a superfície A , como lugar geométrico de todos os outputs que se podem produzir com uma tecnologia dada, é o conjunto de produção. A fronteira deste conjunto, expressa por $y=f(x)$, é o máximo output que se pode produzir com um nível dado de input, ou dito de outra forma, é a função de produção. Então, uma empresa concreta estará sobre a fronteira do conjunto de produção ou no seu interior. Na teoria micro-económica aparecem igualmente outros exemplos da utilização de este conceito limite.

Outras funções, como a função lucro ou a função de custo definem-se também em termos de fronteira; desta forma resulta que:

- a função de custo é o custo mínimo com o qual é possível é possível produzir um nível determinado de output, dados os preços dos inputs.

. a função lucro é o lucro máximo que pode ser obtido, dados os preços do output e dos inputs.

Os métodos de estimação habituais destas funções são tais que algumas observações (empresas) têm resíduo positivo e outras resíduo negativo , o que significa que se situam acima e por debaixo da função de produção (custo ou lucro), violando, portanto, o carácter de óptimo associado àquela função.

O conceito de função fronteira repõe o carácter óptimo daquelas funções. Um dos objectivos da estimação de funções de fronteira é a medida da ineficiência, sendo esta definida como a razão entre a produção observada efectivamente, e o seu valor máximo, ou seja o seu valor de fronteira. Análogamente, a medida da ineficiência com funções de custo (ou lucro), resulta do rácio entre o custo (ou lucro) observados e o seu valor de fronteira, ou seja o custo mínimo (ou lucro máximo).

Antecipando os problemas econométricos que surgem na estimação de fronteiras, diremos que os valores observados da produção e lucro (ou custo) estão limitados superiormente (ou inferiormente) pelo máximo (ou mínimo) que se possa obter ou, dito de outra forma, pela função fronteira. Este problema resolve-se com a utilização de perturbações aleatórias que são só negativas (ou positivas). Os problemas econométricos suscitados pelo facto das hipóteses de normalidade e média nula das perturbações se deixarem de verificar, afastam a estimação de fronteiras da situação clássica do Modelo de

Regressão Linear Múltipla.

3.1.1. - A EFICIÊNCIA

Consideremos uma empresa produzindo o output y com um vector de inputs de dimensão n : $x=(x_1, \dots, x_n)$, cujos preços são fixos: p , o do output, e $w=(w_1, \dots, w_n)$, os dos inputs.

Então definem-se:

- . a função de produção como a fronteira do conjunto de produção:
 $y=f(x)$
- . a função de custo: $c(y,w)=\min_x \{ w'x \mid f(x) \geq y, x \geq 0 \}$
- . a função de lucro: $\pi(p,w)=\max_{y,x} \{ py - w'x \mid f(x) \geq y, x \geq 0, y \geq 0 \}$.

Associados a estes conceitos definem-se três tipos de medidas de eficiência (ineficiência). Dizemos que um plano de produção (y_0, x_0) é técnicamente eficiente se

$$y_0 = f(x_0)$$

y técnicamente ineficiente se

$$y_0 < f(x_0).$$

A ineficiência técnica mede-se pelo rácio $0 \leq y_0 / f(x_0) \leq 1$. De este modo, um rácio inferior à unidade significa que o plano de produção se situa no interior do conjunto de produção, ou seja, que se pode obter mais output com aquela dotação de inputs, ou

alternativamente, que o mesmo output pode ser produzido com menos inputs. Também, em situações de ineficiência técnica, custos e lucros diferem dos seus valores óptimos:

$$w'x_0 > c(y_0, w)$$

e

$$py_0 - w'x_0 < \pi(p, w).$$

Além da eficiência técnica, a empresa procura a eficiência de afectação. Supondo que a função de produção é diferenciável pelo menos até à primeira ordem, um plano de produção (y_0, x_0) tem eficiência de afectação, se

$$f_i(x_0)/f_j(x_0) = w_i/w_j,$$

e é ineficiente do ponto de vista da afectação de recursos, se

$$f_i(x_0)/f_j(x_0) \neq w_i/w_j,$$

onde f_i e f_j representam as derivadas parciais da função de produção em relação aos inputs x_i e x_j . No caso da ineficiência de afectação de recursos os custos e lucros diferem dos seus valores óptimos devido ao facto de as proporções de inputs que se utilizam serem diferentes das óptimas. Nesta situação não nos situaremos sobre a fronteira de custo (ou de lucro), podendo-se diminuir os custos (aumentar os lucros) por simples modificação das proporções em que os inputs são utilizados.

Um terceiro tipo de eficiência que as empresas procuram é a eficiência de escala. Diz-se que numa empresa há eficiência de escala, se

$$p = c_y(y_0, w),$$

e ineficiência de escala se

$$p \neq c_y(y_0, w)$$

(onde c_y é a função de custo marginal). Numa situação de ineficiência de escala é possível aumentar os lucros aumentando só a escala da produção.

Em síntese, podemos afirmar o seguinte:

- . para que $y_0 = f(x_0)$, basta que a empresa seja tecnicamente eficiente. Pelo contrário, uma empresa é tecnicamente ineficiente quando o output dado pela função de produção é diferente do observado.
- . para que $w'x_0 = c(y_0, w)$, é necessário que a empresa, além de ser tecnicamente eficiente, tenha também eficiência de afectação de recursos; pelo contrário, se o custo mínimo difere do efectivamente observado, operando a empresa com ineficiência de custo, tal sucede quer seja devido ao custo da ineficiência técnica, quer seja devido ao custo da ineficiência de afectação, quer devido a ambas. Ou dito de outra forma, a empresa está em condições de reduzir os custos aumentando a eficiência técnica ou por via da reafectação de factores produtivos.
- . para que $py_0 - w'x_0 = \pi(p, w)$, é necessário que a empresa seja tecnicamente eficiente, tenha eficiência de afectação de recursos, e além disso, tenha também eficiência de escala.

3.1.2. - PRODUTIVIDADE E EFICIÊNCIA

Além da eficiência, outro conceito ao qual se recorre com frequência quando se pretende medir o desempenho de uma empresa, é o de produtividade. A noção corrente de produtividade média de um factor é a do rácio entre a quantidade produzida e a quantidade empregue de aquele factor. Na utilização deste conceito deve-se ter em conta as suas limitações. Vamos seguidamente referir-nos a duas importantes.

A primeira delas é que o conceito de produtividade não distingue entre deslocações da fronteira do conjunto de produção devidas a mudanças tecnológicas, e movimentos sobre a fronteira do conjunto de produção, mantendo-se aquela fronteira fixa. Esta situação ilustra-se na figura 3.2.

Um deslocamento da fronteira do conjunto de produção de t_0 para t_1 , e um aumento da produção de Y_c para Y_d (sobre a fronteira do conjunto de produção) deixa a produtividade inalterada, apesar da evolução tecnológica. Quer dizer, os efeitos da renovação do equipamento ou da melhoria dos métodos de gestão poderão ser captados através de uma função de produção incluindo um termo independente do tempo, mas não pelo conceito de produtividade.

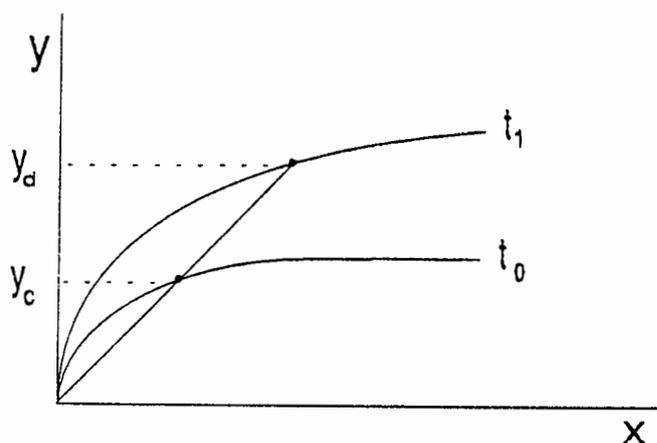


Figura 3.2.

Um segundo argumento que mostra as limitações da produtividade como indicador do desempenho das empresas, é que se torna impossível distinguir entre variações da produtividade que implicam variações de eficiência e as que não implicam. A insuficiência do conceito de produtividade, por não referir-se ao conceito fundamental de função de produção pode-se apreciar (figura 3.3) no seguinte aparente paradoxo: o plano de produção B, apesar de ser ineficiente, tem produtividade superior ao plano de produção A, que está sobre a fronteira do conjunto de produção.

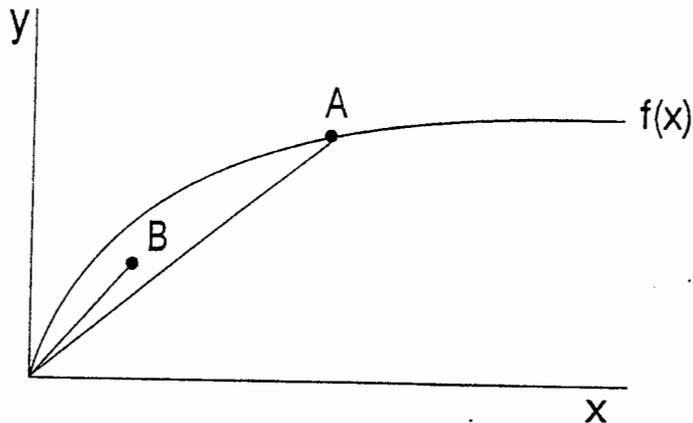


Figura 3.3.

3.1.3. - A EVOLUÇÃO TECNOLÓGICA

Outra forma de encarar o desempenho de uma unidade produtiva, também muito usada em estudos empíricos, é a medição do deslocamento da função de produção no tempo, quer dizer, a taxa de evolução tecnológica. Como se verá este conceito está directamente relacionado com o conceito de produtividade. Consideremos agora uma função de produção $Y = f(x, t)$, onde x é o vector dos inputs e t é o tempo. Por diferenciação total em relação ao tempo desta função obtemos:

$$\frac{dY}{dt} = \sum_j \frac{\delta f}{\delta x_j} \cdot \frac{dx_j}{dt} + \frac{\delta f}{\delta t}$$

e, dividindo ambos os membros por Y , teremos:

$$\frac{d \log Y}{d t} = \sum_j e_j \frac{d \log x_j}{d t} + T(x, t)$$

sendo:

$$e_j = \frac{\partial f}{\partial x_j} \cdot \frac{x_j}{Y}$$

e $T(x, t)$ é a taxa de evolução tecnológica ou a medida do deslocamento da função de produção no tempo.

Sob a hipótese de maximização do lucro, a elasticidade e_j é igual à participação do factor j no rendimento total:

$$e_j = \frac{w_j x_j}{p Y}$$

então teremos:

$$T(x, t) = \frac{d \log Y}{d t} - \sum_j \frac{w_j x_j}{p Y} \cdot \frac{d \log x_j}{d t}$$

o que significa que se a evolução do output e dos inputs no tempo é conhecida, assim como os seus preços e níveis de utilização e produção, então o deslocamento da função de produção no tempo torna-se conhecido sem ser necessário estimar aquela função.

O somatório da expressão anterior designa-se como índice de inputs Divisia (exposto por Divisia (1926)). Se temos observações discretas $T(x, t)$ pode ser aproximado por:

$$T(x, t) = \log Y_t - \log Y_{t-1} - \sum_{i=1}^M [V_{it} + V_{i,t-1}] [\log x_{it} - \log x_{i,t-1}]$$

onde V_{it} é o rácio entre o custo do input i e o rendimento no período t .

O somatório desta aproximação é conhecido como aproximação de Tornqvist ao índice Divisia (exposto pela primeira vez por Tornqvist (1936)). Esta expressão para $T(x,t)$ é uma medida exacta da evolução tecnológica, se a função de produção é uma translog com rendimentos de escala constantes (ver Diewert (1976) e Caves et al. (1982)). Se se supõe um comportamento do produtor minimizador dos custos, então a expressão para $T(x,t)$ compreende a elasticidade de escala e as participações dos factores nos custos; aquela elasticidade só se pode conhecer com a estimação da função de produção ou da função de custo; contudo esta estimação dá imediatamente a taxa de evolução tecnológica.

Uma relação entre o conceito de produtividade e o de deslocamento da função de produção é a que vem dada pelo conceito de produtividade total (TFP) que se define pela expressão $TFP=Y/X$, onde X é um índice dos inputs. Tomando logaritmos e diferenciando ambos os membros em relação ao tempo, teremos:

$$TFP = \dot{Y} - \dot{X}$$

Se, como taxa de crescimento do índice dos inputs, tomamos a média ponderada das taxas de crescimento dos inputs pelas participações de cada factor no rendimento total (ou as participações de cada factor no custo total), quer dizer:

ou

$$\dot{X} = \sum_j \frac{w_j X_j}{PY} \dot{x}_j$$

$$\dot{X} = \sum_j \frac{w_j X_j}{C} \dot{x}_j$$

teremos o índice de inputs Divisia (ou a sua expressão para um comportamento minimizador dos custos e uma tecnologia com rendimentos de escala constantes). Deste modo pode afirmar-se que a taxa de evolução tecnológica sob a hipótese de maximização do lucro ou de minimização dos custos e rendimentos de escala constantes é igual à taxa de crescimento da produtividade total, se se considera um índice de inputs tal que a sua taxa de crescimento iguala o índice de inputs de Divisia.

Admitiu-se que a unidade de produção actuava na fronteira do conjunto de produção. Em caso de tal não suceder, é possível que que ocorram aumentos da produção sem que se produza um deslocamento da função de produção e com a utilização dos mesmos inputs pelo que, neste caso, não estaríamos medindo a taxa de evolução tecnológica, mas um aumento da eficiência.

3.2. - ESPECIFICAÇÃO E ESTIMAÇÃO DE FUNÇÕES FRONTEIRA COM DADOS SECCIONAIS

Pode-se admitir que seja possível chegar ao conhecimento do máximo output alcançável com quantidades de inputs dadas, com o conhecimento que os engenheiros têm da tecnologia. Contudo, a forma de estimação de fronteiras que se utiliza correntemente é utilizando observações quantificadas das variáveis incluídas na função de produção. Vamo-nos centrar aqui na obtenção de funções de produção fronteira (a estimação de funções de custo e de lucro de fronteira é uma adaptação simples deste último caso).

Os diferentes modelos e os métodos de estimação dos seus parâmetros podem classificar-se pela forma como se especifica a fronteira de produção (Forsund et al.(1980)). Deste modo podemos distinguir entre:

- . Métodos paramétricos e métodos não paramétricos conforme se especifique ou não uma forma funcional dependendo de parâmetros constantes.
- . Uma segunda distinção é entre métodos estatísticos e não estatísticos, segundo se explicita ou não um modelo estatístico para o processo gerador dos dados.
- . Finalmente, a fronteira pode ser determinística ou estocástica, conforme se assuma que as perturbações aleatórias do modelo igualam exactamente a ineficiência das empresas, ou que as perturbações incluem não só ineficiência, mas também

ruído estatístico, representado num termo adicional.

Seguidamente desenvolvemos brevemente as características de alguns destes métodos.

3.2.1. - MODELOS COM FRONTEIRAS NÃO PARAMÉTRICAS

O ponto de partida de este tipo de método é o trabalho de Farrell (1957) que é o primeiro autor a sugerir um método para medir a eficiência técnica baseado numa fronteira de produção para um único output e com rendimentos de escala constantes. Muitos outros trabalhos de extensão e aplicação da sua metodologia foram publicados desde então.

Num destes trabalhos, Deprins et al. (1984) propuseram um método não paramétrico de determinação da fronteira de produção baseando-se só na hipótese de disponibilidade dos inputs. Neste método uma observação é declarada ineficiente se é possível encontrar outra (ou outras) observação (observações) na amostra na(s) qual (quais) a quantidade de output produzida é superior e a(s) quantidade(s) de input(s) utilizada(s) são inferior(es). Nesta situação a primeira observação diz-se que está dominada pela segunda (ou pelas outras). As observações pertencentes à amostra

que não são dominadas por outras observações são declaradas eficientes. Esta situação está representada na figura seguinte tomada do artigo daqueles autores (figura 3.4):

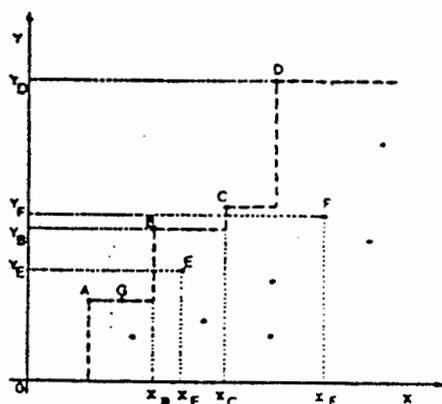


Figura 3.4

A empresa E está estritamente dominada (e portanto é ineficiente) pela empresa B visto que esta produz mais output com menos input. A fronteira do conjunto de produção tem o aspecto de linha em escada da figura 3.4. Define-se além disso que:

- . Uma observação é dominada em output se existe pelo menos outra observação cujo output é estritamente superior e cujo input é igual ou inferior.
- . Uma observação é dominada em input se há pelo menos outra observação cujo output é igual ou superior e o input é estritamente inferior ao da primeira observação.

Para este tipo de modelos a medida do grau de eficiência realiza-se da seguinte forma:

- a) Os pontos declarados eficientes têm uma medida de eficiência igual a 1
- b) Os pontos ineficientes têm uma medida de eficiência inferior a 1 determinada como segue:
 - . o grau de eficiência em output é o rácio entre o output da observação que se pretende caracterizar e o maior dos outputs produzidos pelas empresas que a dominam em output, e
 - . o grau de eficiência em input é o rácio entre o menor dos inputs utilizados pelas empresas que a dominam em input e o input da observação que está sendo caracterizada

Este método é facilmente generalizável para o caso de vários inputs e outputs. Contudo, esta generalização tem o problema de que as condições que declaram uma observação ineficiente são cada vez mais difíceis de cumprir, e portanto, o número de empresas ineficientes é reduzido¹.

¹ Uma aplicação de uma fronteira não paramétrica ao sector hospitalar pode ser encontrada em Valor(1989)

3.2.2. - MODELOS NÃO ESTATÍSTICOS DE FRONTEIRA PARAMÉTRICA E DETERMINÍSTICA

Supondo, sem perda de generalidade, que a fronteira se pode representar por:

$$\log Y = \log f(X;\alpha)$$

os métodos de estimação não estatística recorrem à resolução do problema:

$$\min_{\alpha} \sum_{i=1}^N [\log f(x_i ; \alpha) - \log y_i]$$

$$\text{s.a.} \quad \log f(x_i ; \alpha) - \log y_i \geq 0 \quad (i=1, \dots, N)$$

Trata-se de um problema de programação linear no caso da função $\log f(x_i ; \alpha)$ ser linear. Também se pode recorrer à resolução do problema de programação quadrática que consiste em tomar como função objectivo a soma dos quadrados dos desvios entre a produção observada para cada unidade e a fronteira.

Observe-se que x_i é um vector que contém as quantidades dos k inputs utilizados na empresa i e que se impõe a negatividade dos "resíduos". A exponencial dos desvios entre os valores observados e a fronteira permite uma avaliação da ineficiência. Este método foi desenvolvido por Aigner e Chu (1968), no contexto de uma função de produção Cobb-Douglas, mas é imediatamente generalizável a outras formas funcionais.

Apesar do método anterior ter a vantagem de conduzir à

"estimação" de uma função contínua (dependente de um número reduzido de parâmetros) para representar a fronteira do conjunto de produção, ele tem alguns inconvenientes:

- . a primeira limitação é o facto de o número de observações que são consideradas eficientes depender do número de parâmetros que se "estimam" (porque o número de restrições saturadas depende da dimensão do vector α),
- . o método é muito sensível à presença de "outliers". Para evitar este problema existe a possibilidade de suprimir algumas observações, procedendo-se desta forma até que os resultados estabilizem.
- . um último inconveniente deste método é que as "estimativas" não têm as propriedades estatísticas habituais por não se tratar de facto de inferência estatística.

3.2.3. - MODELOS ESTATÍSTICOS DE UMA FRONTEIRA PARAMÉTRICA E DETERMINÍSTICA

Contrariamente ao método exposto na secção anterior supõe-se agora uma distribuição estatística para as perturbações (desvios entre os valores efectivamente observados e a fronteira) do modelo. Para a estimação estatística de fronteiras, devem ser assumidas

algumas hipóteses sobre o modelo que agora representamos na forma:

$$\log Y = \log [f(x;\alpha)] - u$$

A primeira hipótese é que $u \geq 0$; além disso assume-se que u é um vector de variáveis aleatórias independentes e identicamente distribuídas e que x não está correlacionado com u . A estimação dos parâmetros deste modelo pode efectuar-se seguindo diferentes métodos que seguidamente referimos brevemente:

a) Estimação por máxima verosimilhança: para tal devemos propôr uma distribuição concreta para u . Foram propostas muitas distribuições tais como a Exponencial, a Normal Truncada, a distribuição Gama, Weibull, Log-Logística e a distribuição Log-Normal. Um problema da estimação por máxima verosimilhança é que o intervalo de variação da variável dependente é função dos parâmetros a estimar (supõe-se que $\log y \leq \log [f(x;\alpha)]$), o que viola as condições de regularidade invocadas na demonstração da consistência e eficiência do estimador de máxima verosimilhança. São necessárias hipóteses adicionais no que respeita ao comportamento das perturbações aleatórias para garantir as propriedades desejáveis dos estimadores de máxima verosimilhança.

A este respeito devem-se destacar os resultados de Green (1980) que mostrou que se se restringe a classe das distribuições possíveis para u às que possuam as seguintes propriedades:

. a densidade de u é zero quando $u = 0$, e

. a derivada da densidade de u em relação aos seus parâmetros tende para zero quando u tende para zero.

então os estimadores de máxima verosimilhança possuem as propriedades habituais (e em particular são consistentes e eficientes). Como nota Green a função de densidade Gama satisfaz este critério; mas é um pouco forçado considerar que as propriedades estatísticas dos estimadores dependam da distribuição que se adopte para as perturbações aleatórias, sobretudo quando não há a priori motivos para preferir uma distribuição concreta.

Outro resultado com interesse é que a resolução do problema de programação linear apresentado na secção anterior é equivalente à estimação por máxima verosimilhança com uma distribuição Exponencial, e o problema de programação quadrática é por sua vez equivalente à estimação por máxima verosimilhança utilizando a Normal Truncada. Contudo, estas duas últimas distribuições violam as hipóteses de Green, pelo que as propriedades estatísticas dos estimadores são desconhecidas.

b) Outro método de estimação para os modelos estatísticos de fronteira determinística é o método dos mínimos quadrados corrigidos. O ponto de partida deste método é a constatação de que a não normalidade das perturbações não impede a estimação por mínimos quadrados ordinários; contudo, o facto de que as perturbações não terem média nula, e a importância de ter boas estimativas do termo independente para o cálculo da ineficiência das empresas, impõe cautelas no uso de MQO.

O método dos mínimos quadrados corrigidos (COLS) procede da seguinte forma (admite-se sem perda de generalidade uma forma funcional Cobb-Douglas): em primeiro lugar corrigem-se as perturbações para que tenham média nula, ou seja, estima-se por MQO o seguinte modelo:

$$\log y = (\alpha_0 - \mu) + \sum_{i=1}^k \alpha_i \log x_i - (u - \mu)$$

ao qual, como se pode ver se subtrai às perturbações e ao termo independente a média da variável aleatória residual μ . Deste modo as estimativas de $(\alpha_0 - \mu)$ y α_i são evidentemente BLUE.

O segundo passo consiste em encontrar uma estimativa para μ , para estimar o termo independente da relação inicial, ou seja α_0 . O procedimento consiste em supor uma distribuição específica para $(u - \mu)$, de tal forma que os parâmetros dessa distribuição possam ser estimados pelo método dos momentos. Como μ é função daqueles parâmetros, também pode ser estimado consistentemente e essa estimação pode ser utilizada para "corrigir" o termo independente.

O método dos mínimos quadrados corrigidos tem duas importantes desvantagens: por um lado, alguns resíduos podem ter o "sinal trocado", mesmo depois da correcção do termo independente. Ou seja algumas observações podem estar acima da fronteira de eficiência pelo que este método é pouco fiável na medição da eficiência. Um segundo problema é que a correcção do termo independente é função da distribuição assumida para as perturbações aleatórias, variando

com a distribuição adoptada.

c) O terceiro método que vamos comentar distingue-se do anterior pela forma de corrigir a estimação do termo independente. Essa correcção obtem-se ajustando o termo independente da função de produção média, por forma a que todas as observações estejam abaixo da fronteira e pelo menos uma sobre a fronteira. Este método devido a Green (1980), é chamado método dos Mínimos Quadrados Ordinários Deslocados (DOLS). É um método muito usado nas aplicações empíricas, já que as estimativas não dependem da distribuição das perturbações. A estimação do termo independente apesar de consistente é enviesada e pouco eficiente.

3.2.4. - FRONTEIRAS ESTOCÁSTICAS

Os métodos anteriores estimam fronteiras determinísticas: todas as empresas têm uma fronteira do conjunto de produção que é comum, e todos os desvios em relação àquela fronteira são considerados como correspondendo a ineficiência. Contudo, há dois tipos de críticas a esta ideia:

. em primeiro lugar, críticas com conteúdo económico: o comportamento das empresas pode estar afectado por factores totalmente fora do seu controle, tais como avarias, mau tempo, rupturas de abastecimento de inputs, ou noutro sentido, a presença

de um gestor mais hábil que a média, etc. Estes factores estão incluídos nos resíduos estimados pelos métodos anteriores e portanto na ineficiência estimada. Contudo seria mais razoável incluir, além da variável aleatória que expressa a ineficiência, outra variável aleatória simétrica e de média nula que pudesse traduzir a existência de choques exógenos num sentido ou noutro.

. em segundo lugar, existem também críticas com conteúdo estatístico, que se referem à existência de ruído que afecta qualquer relação empírica. Este ruído pode provir de erros de medição na variável dependente ou de variáveis explicativas de pouca importância omitidas na relação (a consideração de erros de medida nas variáveis independentes conduziria a problemas estatísticos mais graves).

Estes argumentos levam à construção de modelos de fronteiras estocásticas. A ideia básica é considerar que a variável aleatória residual está composta por duas partes: uma simétrica que capta o ruído estatístico e os choques exógenos, e uma componente unilateral e assimétrica traduzindo a ineficiência. Teremos então:

$$\log Y = \log f(x;\alpha) + (v - u)$$

onde $f(x;\alpha) \cdot \exp(v)$ representa a fronteira estocástica com v com uma distribuição simétrica. A fronteira variará então de empresa para empresa. A ineficiência técnica está representada por u , com uma distribuição tal que $u \geq 0$, pelo que todas as empresas estão, por

construção, abaixo da fronteira estocástica. O problema das fronteiras estocásticas é que aparentemente não se podem decompor os resíduos em ineficiência e choque aleatório. Pode-se contudo medir a ineficiência média das empresas que constituem a amostra.

O método de estimação mais utilizado é o da máxima verosimilhança, assumindo-se distribuições para u e v . Este estimador tem as habituais propriedades estatísticas visto que a variável dependente não tem uma variação limitada dependente dos parâmetros a estimar, devido à presença do termo simétrico.

O cálculo da ineficiência técnica quando se estima uma função de produção paramétrica e estocástica, quer dizer, quando as perturbações são do tipo:

$$\phi_i = v_i - u_i$$

onde v_i é ruído branco e u_i é um termo não negativo, pode fazer-se como segue. O método, é baseado nos resultados de Jondrow (1982), e obtém a esperança matemática de u_i condicionada aos valores de ϕ_i . Esta esperança matemática obtém-se através da distribuição de u_i condicionada por ϕ_i . Supondo uma distribuição semi-normal para u_i , a esperança matemática toma a forma:

$$E(u_i | \phi_i) = \sigma_x \left\{ \frac{f(\phi_i \delta / \sigma_\phi)}{1 - F(\phi_i \delta / \sigma_\phi)} - (\phi_i \delta / \sigma_\phi) \right\}$$

onde $\delta = \sigma_u / \sigma_v$ $\sigma_x = \sigma_u \sigma_v / \sigma_\phi$.

A estimação da função de produção fronteira estocástica pelo método da máxima verosimilhança proporciona estimativas de σ_u^2 .

O modelo também pode ser estimado por Mínimos Quadrados Corrigidos (COLS) ajustando para tal o termo independente com $E(u)$, e estimando-o pelo método dos momentos aplicados aos resíduos da estimação MQO.

3.2.5. - FACTORES-X E FACTORES-Z

Muitas vezes, na estimação de funções de produção fronteira, reconhecem-se factores que influem sobre o nível de eficiência técnica encontrado: são factores ambientais tais como o congestionamento do tráfego num estudo de transportes urbanos, ou a idade da população utente numa função de produção hospitalar. Estes factores não costumam ser explicitados na especificação de funções de produção, mas se se deseja obter uma medida de eficiência líquida daqueles efeitos, então é necessário tê-los em conta.

Para tal deve-se distinguir entre factores produtivos (factores-x) e factores exógenos de ineficiência (factores-z). A obtenção de uma medida de eficiência líquida da influência dos

factores-z poderá ser conseguida fazendo a regressão dos resíduos de estimação de uma função de produção fronteira, onde só se tenha considerado os factores-x, sobre os factores-z e observar seguidamente os resíduos da segunda estimação.

Alternativamente, podem-se considerar aqueles factores directamente no modelo inicial. Os factores-z podem-se introduzir na regressão da seguinte forma:

Seja a função de produção fronteira expressa da seguinte forma:

$$\log Y_i = \sum_j \alpha_j \log x_{ji} - u_i \quad u_i \geq 0$$

onde supomos, além disso, que as perturbações são explicadas pelos factores-z no seguinte modo:

$$u_i = \exp(c'z_i) - v_i$$

A forma exponencial destina-se a garantir, juntamente com a condição $v_i \leq \exp(c'z_i)$, o sinal positivo de u_i . A variável aleatória v_i tem média zero e representa a parte da produção não explicada pelos outros dois termos.

Substituindo a segunda equação na primeira, teremos:

$$\log Y_i = \sum_j \alpha_j \log x_{ji} - \exp(c'z_i) + v_i,$$

com:

$$v_i \leq \exp(c'z_i)$$

A medida de eficiência, tendo em conta os factores exógenos de ineficiência, vem dada por:

$$EF_i = Y_i / \hat{Y}_i = \exp(\hat{v}_i),$$

com:

$$\hat{v}_i = \log Y_i - \sum_j \hat{\alpha}_j \log x_{ji} + \exp(c'z_i)$$

ou

$$EF_i = \exp(\log Y_i - \sum_j \hat{\alpha}_j \log x_{ji}) \cdot \exp[\exp(c'z_i)]$$

O primeiro factor que é, -por construção, igual ou menor que a unidade, é a medida da eficiência sem ter em conta os factores exógenos de ineficiência e o segundo factor, que é por construção maior que a unidade, pode ser interpretado como a correcção devida aos factores ambientais. Como o termo EF_i pode ser maior que a unidade, uma correcção que facilita a comparação entre empresas, obtem-se simplesmente por divisão de cada elemento EF_i pelo valor máximo encontrado ($\max(EF_i)$).

O método que consiste em estimar a função de produção fronteira sem ter em conta os factores-z e depois fazer a regressão dos resíduos daquela estimação sobre os factores-z, só proporciona estimadores não enviesados e consistentes no caso de os factores-x e os factores-z serem ortogonais.

O modelo pode-se estimar por Mínimos Quadrados (mínimos quadrados não lineares) ou por Máxima Verosimilhança. O método dos Mínimos Quadrados não lineares é a minimização em relação aos parâmetros a estimar da função:

$$\sum_{i=1}^N v_i^2 = \sum_{i=1}^N (\log Y_i - \sum_{j=1}^K \alpha_j \log x_{ji} + \exp(c'z_i))^2$$

Deve-se notar que se não há factores exógenos de ineficiência, ou seja, se $z_i=1$ ($i=1, \dots, N$), então o termo independente da função de produção não está identificado (as condições de identificação do modelo podem ser vistas em Deprins e Simar (1989 (b))). Quando ocorre este caso deve-se, evidentemente, utilizar os métodos de estimação habituais de funções de produção fronteira determinísticas.

Também se pode estimar este modelo por Máxima Verosimilhança. Neste caso uma forma de ter em conta a restrições dadas por:

$$u_i = \exp(c'z_i) - v_i \geq 0$$

consiste em especificar uma distribuição unilateral para u_i , parametrizada por forma a que $E(u_i) = \exp(c'z_i)$. Em Deprins e Simar (1989 (b)) pode-se ver qual é a forma da função de verosimilhança a maximizar para as várias especificações da distribuição das perturbações aleatórias e também a forma da matriz de variâncias e covariâncias daqueles estimadores.

As dificuldades computacionais do método aqui exposto, devido

a Deprins e Simar (1989), e a simplicidade do método em duas etapas, talvez justifiquem que a maioria dos trabalhos empíricos utilizem o procedimento em duas etapas exposto anteriormente.

3.3. - FRONTEIRAS ESTOCÁSTICAS E DADOS DE PAINEL

A medição da eficiência das empresas vem, na sua precisão, consideravelmente melhorada quando são utilizados dados de tipo painel. Este tipo de informação refere-se à actividade de diferentes unidades de produção (ou de consumo) em sucessivos momentos do tempo. Este tipo de informação, que constitui um avanço em relação aos dados puramente seccionais ou cronológicos e implicou um desenvolvimento específico da metodologia econométrica, teve também, no que se refere às funções fronteira, consequências na sua formulação. Nesta secção fazemos referência em primeiro lugar aos problemas da estimação com dados de tipo painel e seguidamente às questões relativas à estimação de fronteiras com este tipo de informação.

3.3.1. - ESTIMAÇÃO COM DADOS DE PAINEL

A estimação de modelos onde as variáveis (dependente e explicativas) são observadas sobre diferentes unidades econômicas e em diferentes períodos de tempo é uma situação cada vez mais frequente em econometria e põe problemas específicos. O modelo geral para este tipo de situações pode-se escrever:

$$y_{it} = \alpha_{it} + \sum_{k=2}^K \beta_{kit} x_{kit} + u_{it}$$

(i=1, ..., N; t=1, ..., T)

Pode-se postular, como nesta especificação do modelo, que os coeficientes das variáveis explicativas variam com os indivíduos e com o tempo; que variam só com os indivíduos, ou só com o tempo; ou que, finalmente são os mesmos para os diferentes indivíduos e os diferentes períodos de tempo. Na análise seguinte da função de produção e função de custo, admitiremos que os coeficientes dos regressores são estáveis entre indivíduos e no tempo, variando só os efeitos diferenciadores concentrados no termo independente.

Supondo, por agora, que os efeitos diferenciadores só distinguem os indivíduos, sendo constantes no tempo, poderemos escrever o modelo na forma:

$$y_{it} = \alpha_1 + \mu_i + \sum_{k=2}^K \beta_k x_{kit} + u_{it}$$

Neste modelo, os efeitos diferenciadores μ_i podem ser considerados como fixos ou aleatórios. Uma modelização óbvia no primeiro caso, denominado modelo de efeitos fixos, é através da utilização de variáveis artificiais (tantas quantas o número de indivíduos, menos um, para evitar a colinealidade perfeita) . O método de estimação aconselhado (assumindo que as perturbações são independentes e identicamente distribuídas) é o dos Mínimos Quadrados Ordinários, sendo os estimadores resultantes BLUE. Contudo, MQO aplicados ao modelo anterior implicam a inversão de uma matriz de ordem $N+K$. Uma formulação alternativa, e que demonstra ser equivalente à anterior, consiste em exprimir as variáveis na forma de desvios em relação à média das observações de cada unidade económica no tempo. Nesta situação, tanto o termo independente como as variáveis artificiais desaparecem. Poderemos, então, escrever o modelo na forma:

$$y_{it} - \bar{y}_i = \sum_{k=2}^K \beta_k (x_{kit} - \bar{x}_{ki.}) + u_{it} - \bar{u}_i.$$

Do mesmo modo a matriz Q_1 , que se define da seguinte forma:

$$Q_1 = I_T - \frac{1}{T} e_T e_T' \quad (e_T = (1 \ 1 \ \dots \ 1)' \text{ com dimensão } T),$$

é uma matriz idempotente tal que operando sobre as observações em T períodos do indivíduo i , em qualquer das variáveis, reproduz as observações dessa variável em forma de desvios em relação à média do indivíduo; a matriz

$$(I_N \otimes Q_1)$$

operando sobre as variáveis originais, transforma-as em variáveis expressas em termos de desvios em relação à média de cada indivíduo. A aplicação de MQO ao modelo com as variáveis expressas em termos de desvios em relação às médias individuais permite obter o estimador dos coeficientes dos regressores do modelo:

$$\beta^w = [X'_s (I_N \otimes Q_1) X_s]^{-1} X'_s (I_N \otimes Q_1) y$$

onde X_s designa a matriz das observações das variáveis exógenas (excluindo o termo independente e as variáveis artificiais dos efeitos individuais). Este estimador é o chamado estimador "within".

Estimativas dos termos independentes $\alpha_{1i} = (\alpha_1 + \mu_i)$ podem obter-se da seguinte forma:

$$\hat{\alpha}_{1i} = \bar{y}_i - \sum_{k=2}^K \beta_k^w \bar{x}_{ki}$$

enquanto que as estimativas dos efeitos individuais vêm dadas por:

$$\hat{\mu}_i = \hat{\alpha}_{1i} - \sum_{j=1}^N \frac{\hat{\alpha}_{1j}}{N}$$

Poderemos alternativamente supôr que os efeitos individuais são aleatórios. Admite-se então que:

$$E(\mu_i) = 0, \quad E(\mu_i^2) = \sigma_\mu^2, \quad E(\mu_i \cdot \mu_j) = 0 \text{ com } i \neq j$$

supõe-se além disso que μ_i e u_{it} não estão correlacionados. Então o modelo pode-se escrever:

$$y = \alpha_1 + X_s \beta_s + [(\mu \otimes e_T) + u]$$

onde $\mu = (\mu_1, \dots, \mu_N)'$. A matriz de variâncias e covariâncias do termo residual é diagonal por blocos e vem dada por:

$$\phi = E[(\mu \otimes e_T + u) (\mu \otimes e_T + u)']$$

O estimador obvio para este modelo será o de Mínimos Quadrados Generalizados (MQG). Demonstra-se que o estimador MQG para este modelo é uma média ponderada do estimador "within" e de outro estimador chamado de "between". Este último estimador para os coeficientes dos regressores vem dado por:

$$\beta_s^b = (X_s' Q_2 X_s)^{-1} X_s' Q_2 y$$

onde

$$Q_2 = I_N \otimes \frac{e_T e_T'}{T} - \frac{e_{NT} e_{NT}'}{NT}$$

é uma matriz idempotente que ao ser pós-multiplicada pela matriz X_s conduz ao produto de Kronecker de uma matriz com elemento genérico $\bar{x}_{ki} - \bar{x}_{k..}$, pelo vector e_T . Em consequência, este estimador só considera a variabilidade dos indivíduos em relação à média global. O estimador dos MQG dos coeficientes dos regressores pode-se então escrever:

$$\beta_s^{MQG} = \left[\frac{X_s' Q_2 X_s}{\sigma_1^2} + \frac{X_s' Q_1 X_s}{\sigma_u^2} \right]^{-1} \left[\frac{X_s' Q_2 X_s}{\sigma_1^2} \beta_s^b + \frac{X_s' Q_1 X_s}{\sigma_u^2} \beta_s^w \right]$$

onde

$$\sigma_1^2 = T\sigma_\mu^2 + \sigma_u^2;$$

e β_s^w é o estimador "within" dos coeficientes dos regressores.

O termo independente pode ser estimado por:

$$\hat{\alpha}_1^{MOG} = \bar{y}_{..} - \sum_{k=2}^K \beta_k^{MOG} \bar{X}_{k..}$$

enquanto que os efeitos individuais se obtêm a partir de:

$$\hat{\mu}_i = \left(\frac{T\sigma_\mu^2}{\sigma_1^2} \right) (\bar{y}_{i.} - \hat{\alpha}_1^{MOG} - \sum_{k=2}^K \beta_k^{MOG} \bar{X}_{ki.})$$

e, finalmente, os estimadores das variâncias são:

$$\hat{\sigma}_1^2 = \frac{\hat{u}^b \hat{u}^b}{N-K}$$

$$\hat{\sigma}_u^2 = \frac{\hat{u}^w \hat{u}^w}{N(T-1) - (K-1)}$$

onde

$$\hat{u}^b = Q_2 y - Q_2 X_s \beta_s^b$$

$$\hat{u}^w = Q_1 y - Q_1 X_s \beta_s^w;$$

sendo, por seu lado:

$$\hat{\sigma}_\mu^2 = \frac{\hat{\sigma}_1^2 - \hat{\sigma}_u^2}{T}$$

Aquele último estimador tem a desvantagem de poder ser negativo. Se tal ocorrer Maddala (1971) sugere que se pode tomar zero como estimativa de σ_μ^2 , e atribui este facto a uma incorrecta omissão dos efeitos temporais.

É interessante, num contexto de modelização das unidades produtivas, a formulação de Mundlak (1978). O seu problema é que os factores produtivos (variáveis exógenas) e a ineficiência técnica (efeitos individuais) poderão estar correlacionados, fazendo com que os estimadores anteriores (MQG) sejam inconsistentes. Intuitivamente poderemos admitir que a dimensão das empresas, medida pelo stock de capital, possa estar correlacionada com o termo correspondente à ineficiência.

Mundlak sugere que aquela associação seja modelizada por:

$$\mu_i = x_i' a + w_i \text{ com } w_i \text{ IID}$$

com $E(w_i) = 0$, e $E(w_i^2) = \sigma_w^2$. Com esta informação chega-se aos estimadores MQG:

$$a_1^{MQG} = \bar{y}_{..} - \sum_{k=2}^K \beta_k^b \bar{x}_{k..}$$

$$\beta_B^{MQG} = \beta_B^w$$

$$a^{MQG} = \beta_B^b - \beta_B^w$$

Um teste para a pertinência da formulação de Mundlak, quer dizer um ensaio para a hipótese:

$$H_0: a = 0 \text{ vs } H_1: a \neq 0$$

pode ser encontrado em Hsiao (1986), pg. 48.

O modelo para a estimação com dados painel onde existem

efeitos individuais e efeitos temporais não estocásticos, pode ser escrito como:

$$y_{it} = \alpha_1 + \mu_i + \pi_t + \sum_{k=2}^K \beta_k x_{kit} + u_{it}$$

Este modelo pode ser estimado por MQO depois de uma reparametrização para ter em conta a redundância de um dos μ_i e um dos π_t . Alternativamente, podem-se estimar separadamente tanto o vector β_s , como o termo independente e os efeitos individuais e temporais, como no caso do estimador só com efeitos individuais:

$$\beta_s^1 = (X_s' Q_3 X_s)^{-1} X_s' Q_3 y$$

onde

$$Q_3 = I_{NT} - I_N \otimes \frac{e_T e_T'}{T} - \frac{e_N e_N'}{N} \otimes I_T + \frac{e_{NT} e_{NT}'}{NT}$$

é uma matriz simétrica e idempotente que, quando se pós-multiplica por X_s produz uma matriz com elemento genérico:

$$x_{kit} - \bar{x}_{ki.} - \bar{x}_{k.t} + \bar{x}_{k..}$$

Os restantes restantes parâmetros do modelo são estimados por:

$$\hat{\mu}_i = (\bar{y}_{i.} - \bar{y}_{..}) - \sum_{k=2}^K (\bar{x}_{ki.} - \bar{x}_{k..}) \beta_k$$

$$\hat{\pi}_t = (\bar{y}_{.t} - \bar{y}_{..}) - \sum_{k=2}^K (\bar{x}_{k.t} - \bar{x}_{k..}) \beta_k$$

$$\hat{\alpha}_1 = \bar{y}_{..} - \sum_{k=2}^K \bar{x}_{k..} \beta_k$$

Se μ_i e π_t são considerados aleatórios, com $E(\mu_i) = E(\pi_t) = 0$, e $E(\mu_i^2) = \sigma_\mu^2$, $E(\pi_t^2) = \sigma_\pi^2$, $E(\mu_i \mu_j) = E(\pi_t \pi_s) = 0$, para $i \neq j$ e $t \neq s$, e μ_i , π_t , u_{it} não estão correlacionados; então o modelo pode-se escrever, de uma forma compacta:

$$y = \alpha_1 + X_s \beta_s + (\mu \otimes e_T) + (e_N \otimes I_T) \pi + u$$

Se ϕ é a matriz de variâncias e covariâncias da parte aleatoria, então o estimador BLUE para α_1 e β_s será:

$$\begin{pmatrix} \hat{\alpha}_1 \\ \hat{\beta}_s \end{pmatrix} = (X' \phi^{-1} X)^{-1} X' \phi^{-1} y$$

Contudo, pode-se provar que:

$$\phi^{-1} = \frac{Q_3}{\sigma_u^2} + \frac{Q_2}{\sigma_1^2} + \frac{Q_4}{\sigma_2^2} + \frac{Q_5}{\sigma_3^2}$$

onde Q_2 e Q_3 foram definidas anteriormente e:

$$Q_4 = \frac{e_N e_N'}{N} \otimes I_T - \frac{e_{NT} e_{NT}'}{NT}$$

$$Q_5 = \frac{e_{NT} e_{NT}'}{NT}$$

$$\sigma_1^2 = \sigma_u^2 + T \sigma_\mu^2$$

$$\sigma_2^2 = \sigma_u^2 + N \sigma_\pi^2$$

$$\sigma_3^2 = \sigma_u^2 + T \sigma_\mu^2 + N \sigma_\pi^2$$

Então o estimador MQG para β_s é:

$$\beta_s^{MQG} = \left[\frac{X'_s Q_2 X_s}{\sigma_1^2} + \frac{X'_s Q_4 X_s}{\sigma_2^2} + \frac{X'_s Q_3 X_s}{\sigma_u^2} \right]^{-1} \times$$

$$\left[\frac{X'_s Q_2 X_s}{\sigma_1^2} \beta_s^* + \frac{X'_s Q_4 X_s}{\sigma_2^2} \beta_s^0 + \frac{X'_s Q_3 X_s}{\sigma_u^2} \beta_s^1 \right]$$

e

$$\hat{\alpha}_1 = \bar{y}_{..} - \sum_{k=2}^K \beta_k^{MQG} \bar{X}_{k..}$$

onde:

$$\beta_s^* = (X'_s Q_2 X_s)^{-1} X'_s Q_2 y$$

$$\beta_s^0 = (X'_s Q_4 X_s)^{-1} X'_s Q_4 y$$

e β_s^1 , é o estimador de variáveis artificiais. Os estimadores BLUE para μ_i e π_t são:

$$\hat{\mu}_i = \left(\frac{T\sigma_u^2}{\sigma_1^2} \right) (\bar{y}_{i.} - \hat{\alpha}_1 - \sum_{k=2}^K \beta_k^{MQG} \bar{X}_{ki.})$$

$$\hat{\pi}_t = \left(\frac{N\sigma_\pi^2}{\sigma_2^2} \right) (\bar{y}_{.t} - \hat{\alpha}_1 - \sum_{k=2}^K \beta_k^{MQG} \bar{X}_{k.t})$$

Estimadores não enviesados para σ_1^2 , σ_2^2 e σ_u^2 serão:

$$\hat{\sigma}_1^2 = \frac{\hat{u}^{*'} \hat{u}^*}{N-K}$$

$$\hat{\sigma}_2^2 = \frac{\hat{u}^{0'} \hat{u}^0}{T-K}$$

$$\hat{\sigma}_u^2 = \frac{\hat{u}' \hat{u}}{(N-1)(T-1) - (K-1)}$$

onde:

$$\hat{u}^* = Q_2 y - Q_2 X_B \beta_B^*$$

$$\hat{u}^0 = Q_4 y - Q_4 X_B \beta_B^0$$

$$\hat{u} = Q_3 y - Q_3 X_B \beta_B^1$$

e poderemos estimar:

$$\hat{\sigma}_\mu^2 = \frac{\hat{\sigma}_1^2 - \hat{\sigma}_u^2}{T}$$

$$\hat{\sigma}_\pi^2 = \frac{\hat{\sigma}_2^2 - \hat{\sigma}_u^2}{N}$$

Vamos fazer uma observação sobre as propriedades estatísticas dos estimadores MQG quando na matriz de variâncias e covariâncias, que depende de parâmetros desconhecidos, se substituem esses parâmetros pelas suas estimativas obtidas da forma acima indicada (MQGE). No caso dos efeitos individuais serem invariantes no tempo,

Swamy (1971) e Fuller e Battese (1973) examinaram as condições sob as quais as propriedades assintóticas do estimador MQGE são idênticas às do estimador MQG. No que respeita às propriedades em amostras finitas, Taylor (1980) prova que o estimador MQGE é mais eficiente que o estimador "within" quando os graus de liberdade excedem um determinado valor suficientemente pequeno. No caso do modelo com efeitos individuais e temporais, Fuller e Battese (1974) obtiveram as condições sob as quais o estimador MQGE é não enviesado e assintoticamente equivalente ao de MQG.

Como no modelo de efeitos fixos não são necessárias hipóteses sobre a distribuição dos efeitos individuais (e/ou temporais), esta especificação pode ser utilizada num maior número de situações. Contudo, se se cumprem as hipóteses do modelo de efeitos aleatórios o estimador MQG será mais eficiente que o estimador do modelo de variáveis artificiais. Concretamente se as variáveis incluídas como regressores estão correlacionadas com os efeitos individuais, o estimador MQG para o modelo de efeitos aleatórios não será consistente. A estimação com variáveis artificiais será neste caso BLUE, mas a inferência estará condicionada aos μ_i na amostra.

Como "a priori" não existem motivos para escolher o modelo de efeitos fixos ou o modelo de efeitos aleatórios, quando se trabalha num estudo concreto, a selecção de um deles deve-se basear nas suas propriedades estatísticas. Um teste de hipóteses de efeitos individuais aleatórios (H_0) versus efeitos individuais fixos, é o de Hausman y Taylor (1981):

$$m = (\beta_S^{(1)} - \beta_S^{(2)})' (M_1 - M_0)^{-1} (\beta_S^{(1)} - \beta_S^{(2)}) \sim \chi_{k'}^2$$

onde:

$$\beta_S^{(1)}, \beta_S^{(2)}$$

são os estimadores do modelo com efeitos fixos e aleatórios respectivamente.

$$M_1 = \sigma_u^2 [X_S' (I_N \otimes Q_1) X_S]^{-1}$$

é a matriz de variâncias e covariâncias do estimador "within", e

$$M_0 = \left[\frac{1}{\sigma_1^2} X_S' Q_2 X_S + \frac{1}{\sigma_u^2} (X_S' (I_N \otimes Q_1) X_S) \right]^{-1}$$

é a matriz de variâncias e covariâncias do estimador MQG. Se a hipótese nula é verdadeira, então os dois estimadores diferem só devido a flutuações amostrais. Se a alternativa é verdadeira, e concretamente se μ_i e X_s estão correlacionados, a diferença entre os dois estimadores pode ser mais ampla, esperando-se que este facto conduza a uma estatística de teste suficientemente grande para rejeitar a hipótese nula. Nesta situação optaremos, então pelo estimador "within".

3.3.2. - ESTIMAÇÃO COM DADOS DE PAINEL E FUNÇÕES DE FRONTEIRA

Seguidamente passamos a expôr o modelo para funções de fronteira, utilizando painel de dados. Consideremos o modelo:

$$y_{it} = \alpha + X'_{it}\beta - u_i + v_{it} \quad (i=1, \dots, N ; t=1, \dots, T)$$

onde representamos por y_{it} o output (ou qualquer função do output) da unidade i no período t ; X_{it} o vector dos inputs (ou qualquer função de esses inputs); v_{it} é uma variável aleatória simétrica e u_i representa o termo de ineficiência que, por definição, é sempre positivo. Supõe-se que u_i é IID com média τ y variância σ_u^2 ; e supõe-se além disso, que u_i distribui-se independentemente de v_{it} .

Estamos, então, num contexto de funções de fronteira estocástica, onde o termo que representa a ineficiência supõe-se invariante no tempo. A analogia desta especificação com a dos modelos de painel de dados é imediata: o termo de ineficiência representa os efeitos individuais e v_{it} as perturbações aleatórias. A única diferença em relação aos modelos de painel de dados é que a média dos efeitos individuais, é $\tau > 0$. Contudo, se redefinirmos $\alpha^* = \alpha - \tau$, y $u_i^* = u_i - \tau$, então o modelo com as variáveis assim transformadas é idêntico ao modelo com painel de dados com os

parâmetros dos regressores constantes e sem efeitos temporais (o deslocamento da fronteira no tempo pode ser introduzido com variáveis artificiais).

A correção do termo independente e dos efeitos individuais para obter os termos de ineficiência sempre positivos ou nulos, faz-se com o pressuposto que a unidade mais eficiente é 100% eficiente; quer dizer $u_i = 0$ para a unidade mais eficiente. Isto corresponde, no caso da função de produção, a fazer:

$$u_i = \max(\mu_i) - \mu_i; \alpha = \alpha^* + \max(\mu_i)$$

Voltemos de novo ao problema da escolha entre efeitos fixos e efeitos aleatórios. A principal vantagem de considerar os efeitos como fixos é, como já dissemos, o facto de a consistência do estimador não depender da não correlação entre os efeitos individuais e as variáveis explicativas. O seu principal inconveniente consiste no facto de não ser possível a inclusão de regressores que sejam constantes no tempo, mas variem entre indivíduos. Os efeitos individuais estimados incluem, neste caso, os efeitos de todas estas variáveis que podem não ter qualquer relação com a ineficiência das empresas.

Em ambos os casos, ou seja tanto no modelo de efeitos fixos como no de efeitos aleatórios o estimador dos parâmetros dos regressores é consistente quando $N \rightarrow \infty$, ou $T \rightarrow \infty$, enquanto que o estimador dos efeitos individuais é consistente quando $T \rightarrow \infty$. O facto de em muitas ocasiões T ser pequeno não é preocupante na estimação dos efeitos individuais, por ser a estimação com dados de

painel já uma considerável melhoria em relação à estimação com dados seccionais, onde se estimam os factores de ineficiência só com uma observação por elemento. Finalmente, notemos que o facto de não ter considerado os efeitos temporais na especificação do modelo pode ser modificado quando se consideram as variáveis artificiais correspondentes a cada período de tempo.

3.4. - FUNÇÕES DE CUSTO FRONTEIRA

Uma tecnologia é univocamente definida pela função de produção e pela função de custo. A decisão sobre qual das duas funções devemos estimar depende dos objectivos do estudo e dos dados disponíveis. Sem dados sobre os preços dos factores, a função de produção dará só informação sobre o grau de ineficiência técnica. Uma função de custo fronteira dará informação sobre o custo da ineficiência técnica e de afectação conjuntamente, mas não sobre as duas em separado. Obviamente as funções de custo fronteira poderão ser determinísticas ou estocásticas, como as funções de produção.

A separação da ineficiência de custo em ineficiência técnica e ineficiência de afectação pode realizar-se seguindo o procedimento sugerido por Schmidt e Lovell (1979). Partindo de uma especificação de uma função de custo, fronteira e estocástica, do tipo Cobb-Douglas, quer dizer, com:

$$\log C = C_0 + \beta_0 \log y + \sum_{j=1}^n \beta_j \log w_j - (v-u) , u \geq 0$$

onde:

$$\beta_0 = \frac{1}{r} , r = \sum_{j=1}^n \alpha_j , \beta_j = \frac{\alpha_j}{r} , v \sim N(0, \sigma_v)$$

sendo α_j a elasticidade do output em relação ao input X_j ; e u um vector de variáveis aleatórias com média positiva e que representam a ineficiência de custo.

O custo efectivamente observado é igual ao custo mínimo associado a um nível dado de output e de remuneração dos factores, multiplicado por e^{-v} , que é um factor que representa os choques aleatórios fora do controle da empresa e com o qual se obtém a fronteira estocástica, uma vez tida em conta a presença do elemento e^u , que representa a ineficiência de custo. A ineficiência de custo inclui um factor derivado da ineficiência técnica e um factor derivado da ineficiência de afectação. Contudo, sem mais informação não poderemos distinguir os dois factores e só obteremos, com a estimação da função de custo, as duas ineficiências num único termo representativo da ineficiência de custo.

O mérito de Schmidt e Lovell foi a obtenção separada dos dois factores de ineficiência - técnica e de afectação - sendo a

ineficiência de custo o produto de ambos. Concretamente demonstraram que:

$$\log C = K + \frac{1}{r} \log y + \sum_{j=1}^n \frac{\alpha_j}{r} \log w_j - \frac{1}{r} (\epsilon - \gamma) + (E - \log r)$$

onde k é uma função dos parâmetros da função de produção Cobb-Douglas especificada como:

$$\log y = A + \sum_{i=1}^n \alpha_i \log X_i + (\epsilon - \gamma), \quad \gamma_i \geq 0, \quad \epsilon_i \sim N(0, \sigma_\epsilon)$$

na qual γ_i é uma variável aleatória com média não nula e que toma só valores positivos, representando a ineficiência técnica. Deste modo, $e^{\frac{1}{r}\gamma_i}$ é o factor de ineficiência técnica e $e^{(E - \log r)}$ é o factor de ineficiência de afectação.

No que respeita ao elemento E , devemos notar que podemos definir a_i , tal que:

$$\log X_1 - \log X_i = B_i + a_i, \quad (i=2, \dots, n)$$

onde:

$$B_i = \log \left(\frac{w_i \alpha_1}{w_1 \alpha_i} \right)$$

A primeira expressão é equivalente a:

$$\frac{\frac{\partial f}{\partial X_i}}{\frac{\partial f}{\partial X_1}} = \frac{w_i}{w_1} e^{a_i}$$

ou seja, e^{a_j} é um factor de ineficiência de afectação. Mas:

$$E = \sum_{j=2}^n \frac{\alpha_j}{r} a_j + \log [\alpha_1 + \sum_{j=2}^n \alpha_j e^{-a_j}]$$

e pode-se demonstrar que para $a_2 = a_3 = \dots = a_n = 0$, E alcança o seu valor mínimo, sendo então E igual a $\log(r)$.

O método de estimação proposto é o de máxima verosimilhança, assumindo-se uma distribuição Normal Multidimensional para a_i , assumindo-se além disso que u_i são independentes e indenticamente distribuídos com uma Normal Truncada, e os v_i são também independentes e indenticamente distribuídos com distribuição Normal. Obtem-se a distribuição conjunta dos a_i e das perturbações aleatórias da função de produção, que como função dos parâmetros a estimar é a função de verosimilhança a maximizar.

CAPÍTULO 4

A EFICIÊNCIA NO SISTEMA HOSPITALAR

Neste capítulo apresentamos alguns tópicos relativos a problemas que podem surgir na estimação de fronteiras no contexto do sistema hospitalar.

Assim, na primeira secção introduzimos o problema da presença de ineficiência estrutural no sector. Os hospitais, com o seu estatuto de propriedade mais corrente, não procuram em geral a maximização do lucro, e conseqüentemente pode-se prevêr que não se cumprirão os equilíbrios competitivos que a teoria da produção pressupõe para a empresa que maximiza o lucro e actua num contexto de concorrência perfeita, que é suposta ser inteiramente eficiente.

Na segunda secção efectuamos uma apreciação crítica de algumas das contribuições mais relevantes no que respeita à especificação e estimação de funções de custo no sector hospitalar. Na terceira secção estudamos um problema decisivo na estimação de funções de produção e de custo hospitalares: o da medição do output hospitalar. A unidade de output mais usada é o caso tratado; contudo, é consensual na literatura considerar a extrema heterogeneidade daquela unidade. As tentativas de melhoria da quantificação do output e a sua apreciação crítica são o conteúdo desta terceira secção.

4.1. - A UNIDADE HOSPITALAR NUM CONTEXTO NÃO LUCRATIVO

Como referimos no capítulo 2, assistiu-se nas últimas décadas a aumentos crescentes do custo por caso tratado no sistema hospitalar, com a conseqüente pressão sobre o nível das despesas públicas. Na impossibilidade de financiar indefinidamente níveis crescentes de despesa sanitária, assistiram-se a apelos crescentes a uma maior "eficiência" do sistema na utilização dos recursos disponíveis. Este apelo deixa antever a possibilidade de ineficiência estrutural na combinação dos recursos disponíveis, ou talvez de outros objectivos e comportamentos diferentes dos da empresa que maximiza o lucro e actua num contexto de concorrência perfeita.

Seguidamente analisaremos o comportamento da unidade hospitalar como agente em que o estatuto de propriedade é diferente do da empresa privada, pelo que os objectivos que se põem à sua gestão são distintos dos daquele tipo de empresas.

O comportamento económico dos gestores hospitalares e dos médicos no contexto de unidades hospitalares não lucrativas, foi objecto de sucessivas teorizações, em especial nos EUA onde as organizações não lucrativas absorviam a grande maioria das despesas do sector. O primeiro destes modelos é o proposto por Newhouse

(1970). Para Newhouse as principais decisões são tomadas pelos administradores e seu "staff" e referem-se à qualidade e quantidade do output produzido. A administração prefere mais quantidade produzida a menos, visto que tal decisão significa uma maior esfera de actuação. Prefere também mais qualidade a menos, visto que com uma maior qualidade do output consegue aumentar o seu prestígio.

O gestor enfrenta, contudo, restrições na sua actuação. Em primeiro lugar a curva da procura: Newhouse admite portanto, a existência de certo poder monopolístico da unidade hospitalar. Em segundo lugar, a tecnologia, que limita o conjunto dos planos de produção admissíveis. Finalmente as restrições resultantes do estatuto de propriedade: o hospital não pode ter lucros, pelo que deverá igualar o preço ao custo médio, no caso de não obter recursos adicionais para o financiamento das despesas correntes. A primeira conclusão do modelo é que o hospital produzirá qualquer combinação quantidade-qualidade com um custo mínimo. A segunda conclusão é a existência de um enviesamento em direcção à qualidade.

Vejamos como Newhouse desenvolve este argumento. Como se disse a unidade hospitalar não lucrativa deve igualar o preço ao custo médio. Mas para cada quantidade produzida existe uma família de curvas de custo médio correspondentes aos diferentes níveis de qualidade. O equilíbrio da unidade hospitalar produz-se no ponto de intersecção da curva da procura - que se supõe ter uma inclinação negativa e deslocar-se para a direita, quando a qualidade aumenta -

com a curva de custo médio. Representemos esta situação graficamente (figura 4.1):

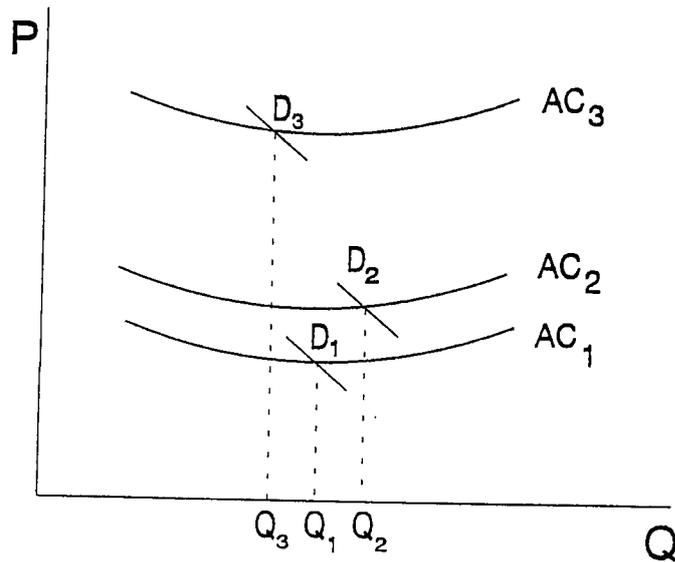


Figura 4.1.

É possível que haja uma zona onde os incrementos de qualidade venham acompanhados de um incremento da quantidade de equilíbrio

$$(AC_1 \rightarrow AC_2)$$

e que, por seu lado, esta esteja seguida de uma zona na qual os incrementos de qualidade venham acompanhados de uma diminuição da quantidade de equilíbrio.

A figura 4.2 representa então a fronteira quantidade/qualidade do hospital e as curvas de indiferença do administrador.

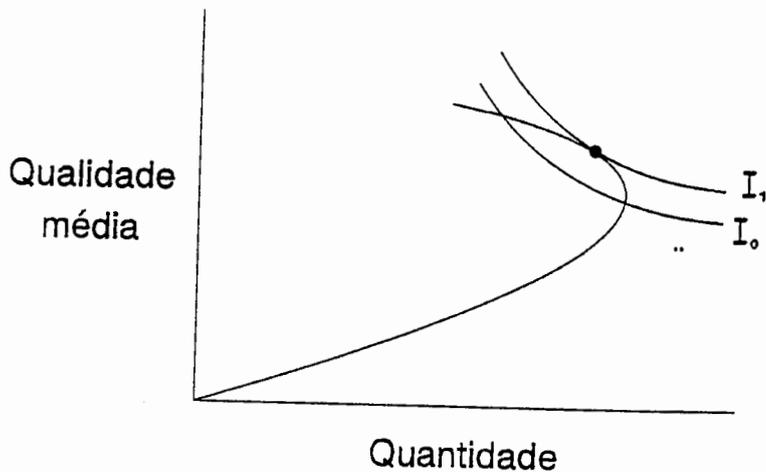


Figura 4.2.

A fronteira quantidade/qualidade com a qual se enfrenta o administrador, apresenta, como admitimos uma primeira zona onde quantidade e qualidade aumentam e uma segunda zona onde a expansão da qualidade vem acompanhada de uma diminuição da quantidade. As curvas de indiferença representam as combinações de quantidade e qualidade que produzem idêntica satisfação para a administração do hospital. A maximização da utilidade, requer que o ponto escolhido seja aquele onde a fronteira é tangente à curva de indiferença mais afastada.

Mas reparemos que uma eficiente utilização dos recursos implica que qualquer quantidade seja oferecida até que a disponibilidade do consumidor para pagar uma unidade adicional de output iguale o custo marginal de produzir esta unidade. Admitamos que a unidade hospitalar está em equilíbrio (do ponto de vista do

administrador) em Q_1 e que há consumidores dispostos a pagar ao custo marginal a quantidade Q_0 . O administrador não achará compensatório (do ponto de vista da sua função de utilidade) a redução da qualidade para aumentar o output. Existe, conseqüentemente um enviesamento em direcção à qualidade.

Uma manifestação desse enviesamento é o sobre-equipamento dos hospitais: a presença de determinados equipamentos, independentemente do seu grau de utilização, está associada a uma maior qualidade e prestígio da instituição, apesar de este efeito vir reforçado pelo facto de os donativos tornarem o factor capital relativamente mais barato a este tipo de instituições. O enviesamento em direcção à qualidade e a ineficiência inerente, permite prognosticar um espaço para a concorrência e para o papel regulador do mercado. A concorrência não poderia vir de outras organizações não lucrativas, porque não é a perspectiva de lucro que as atrai ao mercado e é pouco credível que seja este tipo de instituições cuja concorrência faça mais eficientes as restantes organizações não lucrativas.

Também a filantropia e a situação mais favorável das instituições não lucrativas do ponto de vista fiscal, faz com que as organizações que têm como objectivo o lucro não entrem no mercado. Como afirma Newhouse (1970):

"Even though the model implies least cost production, there are two reasons why it does not lead to an optimal outcome. (...) The two reasons are a bias against producing lower

quality products and barriers to entry resulting from non-profit status"

Um factor adicional de ineficiência para Newhouse, é o reembolso das despesas hospitalares por parte de uma terceira entidade. Neste contexto o administrador pode aumentar a quantidade e qualidade produzida até que a sua utilidade marginal seja nula. A única restrição que enfrenta é que um aumento dos preços pode-se repercutir no preço dos seguros e portanto diminuir a procura. Há, então, na forma de financiamento do sistema um factor de ineficiência adicional.

Um segundo modelo do comportamento do hospital não lucrativo deve-se a Pauly e Redisch (1973). Estes autores argumentam que é o "staff" médico que detem de facto o controle das principais decisões no hospital e que actua com a intenção de maximizar o seu rendimento líquido. Admitamos que o hospital enfrenta uma curva de procura com inclinação negativa e compra os seus inputs (além do "staff" médico) a preços dados. Os médicos dividem entre eles o excedente gerado depois de pagar os restantes inputs. Dado o volume do "staff" médico, cada input é comprado até que o valor do produto marginal iguale o seu preço (caso contrário o "staff" médico poderia aumentar o seu rendimento). Admitamos que ao aumentar o número de médicos, o rendimento médio deste input aumente numa primeira fase devido a economias de escala, alcançando um máximo para logo decrescer numa segunda fase.

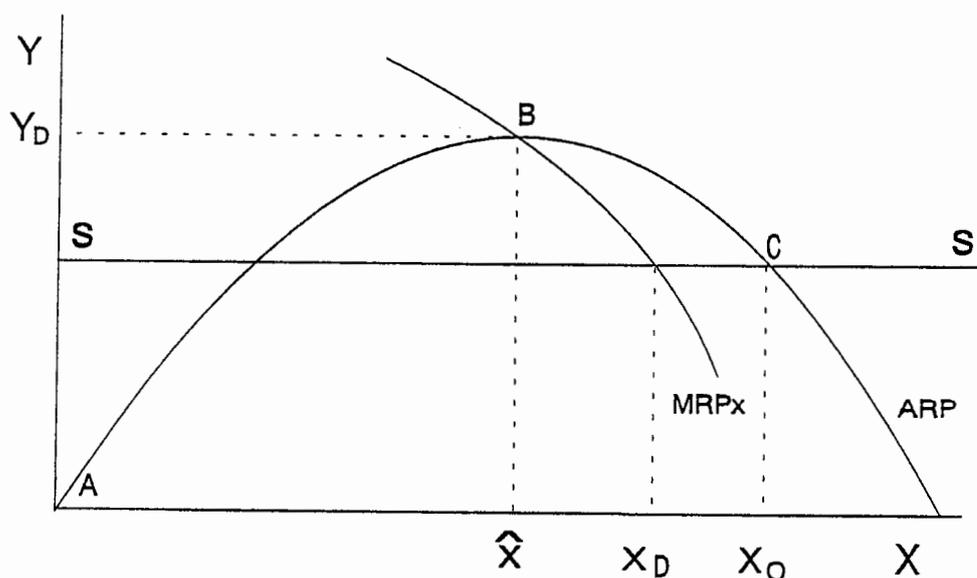


Figura 4.3.

Admitindo que o "staff" médico controla a sua dimensão e maximiza o seu rendimento médio, então esta dimensão vem representada por \hat{X} , o que justifica a afirmação: "Physicians will be willing to add members to the staff as long as it causes each member's net income to rise" (Pauly e Redisch (1973)).

Então o rendimento total do hospital pode ser aumentado com a expansão do "staff" médico. Com efeito, o rendimento marginal iguala o rendimento médio no seu máximo e este por seu lado, excede o custo de oportunidade do factor. O output será também menor que o que se obteria com outra política de admissão de pessoal médico, como veremos mais adiante.

Além desta política de "closed staff", que consiste em o

"staff" médico controlar a sua dimensão e no facto dos novos médicos admitidos terem o mesmo estatuto dos restantes membros do "staff", podemos admitir outras situações. Por um lado, pode-se admitir que o "staff" médico que controla a instituição possa contratar outros médicos a um preço igual ao seu custo de oportunidade. Então a dimensão do "staff" médico aumentará até X_p , onde o valor do seu produto marginal iguala o salário S . Os médicos que controlam a instituição apropriam-se do excedente entre o valor do produto marginal de cada médico contratado e o seu salário. Há, contudo, discriminação na remuneração entre o pessoal médico que controla a instituição e o pessoal médico assalariado, o que pode produzir uma situação instável entre os profissionais médicos. Uma terceira situação é a de "open staff", que se produz quando a dimensão do "staff médico" não é restringida e todos os médicos admitidos têm os mesmos privilégios. Numa situação destas o rendimento médio por médico será igual ao salário e a dimensão do "staff" médico será X_0 . Numa situação de "closed staff" a dimensão do "staff" médico e portanto do output é menor que em qualquer das outras situações.

Este é o equilíbrio de curto prazo, no longo prazo, enquanto houver restrições à admissão de médicos ou os admitidos o forem com um estatuto assalariado, quer dizer, enquanto haja médicos remunerados de forma diferenciada, haverá tendência a que os médicos com remuneração menor mudem de hospital ou formem outra instituição hospitalar, até que o rendimento médio iguale o rendimento marginal e o salário dos médicos.

Para Pauly e Redisch, qualquer que seja a política do "staff" médico, não é o seu controle sobre a produção o que causa a ineficiência, a longo prazo. Esta ineficiência pode resultar tanto da cooperação imperfeita entre o "staff" médico, o que sucede quando um médico ou um grupo de médicos controlam a produção e pretendem aumentar a qualidade do seu output final com a utilização de mais inputs que os outros médicos, assim como no caso em que uma terceira parte pague as despesas hospitalares dos consumidores de tal modo que os médicos possam aumentar os inputs hospitalares em relação ao seu próprio trabalho, numa tentativa de aumentar a sua produtividade marginal.

Como afirmam Pauly e Redisch (1973): "the main thrust of the model we have suggested here, (...) is the use of maximization of physicians' income as the characteristic function". Este modelo é, definitivamente, uma tentativa de explicação de factos tão diferentes como o aumento dos custos reais unitários do sector ou o enviesamento em direcção à qualidade, e neste sentido podemos afirmar que se trata de uma teoria com grande poder explicativo.

O terceiro modelo do hospital que comentamos é o que se atribui a Harris (1977). Ao contrário dos dois anteriores, não evoca na construção do modelo, o estatuto de propriedade da unidade hospitalar. Este modelo é um compromisso entre o modelo do hospital dirigido pelo administrador de Newhouse e o modelo do hospital dirigido pelo "staff" médico de Pauly e Redisch. Para Harris o hospital é constituído por duas empresas actuando numa só. Por um

lado existe a empresa do administrador e por outro a empresa do "staff" médico. Estas duas empresas estão em luta permanente entre si. Como diz este autor: "(...) for all intents and purposes, the typical hospital is two firms loosely connected by a complex set of non-market relations".

Harris, como Pauly e Redisch, realça o papel do "staff" médico na combinação de recursos que a produção do output hospitalar implica. De facto, este processo produtivo resulta de um complicado processo de decisão que consiste no diagnóstico e tratamento de doenças. Neste processo as decisões de utilização de recursos são decisões de curto prazo, consistindo na requisição por parte do "staff" médico de testes e tratamentos para cada doente.

Sobre estas decisões o poder do administrador é reduzido, sendo o médico o único agente qualificado no hospital para este processo de decisão. Então a produção no hospital deve estar organizada como se cada ordem do "staff" médico fosse uma necessidade absoluta, devendo-se garantir que os inputs estão sempre disponíveis. Deste modo, o funcionamento do hospital pode encarar-se como um processo onde, do lado da oferta actua o administrador garantindo a disponibilidade dos inputs, enquanto que do lado da procura actua o médico, requisitando inputs na sua qualidade de agente do doente. O processo de tratamento pode ser encarado como um conjunto de respostas ou transacções, entre a oferta e a procura, quer dizer, entre o médico (na sua qualidade de agente do doente) e a administração do hospital (na sua qualidade

de organizadora dos serviços que fornecem os inputs).

O médico deverá manter-se independente da administração hospitalar para poder actuar em representação do interesse do doente. Por outro lado, a melhor solução para o doente é ter contratos independentes com ambos agentes (médico e administração). A dupla estrutura de poderes actua do seguinte modo: o médico quer garantir que os inputs estão sempre disponíveis para o tratamento dos seus doentes e o administrador pretende garantir que a capacidade produtiva está sendo plenamente utilizada.

Mas quando se caminha em direcção a uma maior intensidade de utilização da capacidade produtiva incrementa-se também a pressão do "staff" médico para a ampliação daquela capacidade, como consequência de que possa haver situações em que alguns médicos não possam ver satisfeitas imediatamente as suas necessidades de inputs. Há, então, uma constante pressão para o aumento da capacidade produtiva e uma possibilidade real de sobre-dimensionamento.

Para finalizar esta secção desenvolveremos a contribuição de Spicer (1982). Este autor baseia-se na análise do SNS britânico, que, ao contrário das organizações não lucrativas ou privadas, é uma organização estatal. Por isso, para Spicer, o SNS é sobretudo uma organização burocrática. Por organização burocrática entende ser uma organização com as seguintes três características:

. O output não é pago directamente pelos utentes, mas por um

terceiro pagador.

- . Os trabalhadores da organização não podem apropriar-se das eventuais poupanças geradas por iniciativas suas.
- . A organização é um monopólio no que respeita ao serviço que produz.

Estas três características estão claramente presentes num SNS. Numa actuação com estas características, as principais causas de ineficiência da organização, podem resumir-se nos seguintes argumentos:

- . Como a burocracia não recebe directamente pelos serviços que produz, é difícil a medição do output da organização. Este facto reduz a capacidade da entidade financiadora de medir a ineficiência dos serviços.
- . O estatuto monopolístico reduz a pressão exterior para uma maior eficiência. E, finalmente:
- . Os incentivos à eficiência, internos à organização são reduzidos, visto que os funcionários não beneficiam das eventuais poupanças resultantes da sua acção.

Estes três factores fazem com que a X-eficiência (excesso dos custos observados sobre o custo mínimo alcançável) seja demasiado grande.

Outra origem da ineficiência reside no poder descricionário que têm os médicos neste sistema. Os procedimentos clínicos variam muito entre hospitais, sem que haja uma avaliação dos procedimentos

utilizados, e sem que os médicos tenham em conta os custos resultantes dos seus actos. Das consequências do carácter burocrático da organização hospitalar sobre a eficiência, Spicer afirma:

"The absence of effective measures of performance, combined with weak external and internal pressures for efficiency, means that those working with the bureaucracy can enjoy greater discretion concerning the activity they perform, the rate at which they perform those activities, and the quality of their performance."

Uma última origem da ineficiência é o estatuto monopolístico e o facto de o consumidor não pagar os serviços do SNS, o que faz que este apareça como indiferente às preferências dos consumidores. Se esse comportamento do SNS lhe permite reduzir os seus custos, também faz com que os utentes suportem outros custos adicionais reduzindo o valor dos serviços produzidos.

Introduzimos até aqui o tema da possibilidade do funcionamento do sistema hospitalar com ineficiência estrutural sistemática. McGuire (1987) associa a inexistência de ineficiência estrutural a um sistema hospitalar com custos de transacção nulos e direitos de propriedade plenos e privados. Este autor argumenta que o movimento em direcção a uma maior eficiência global do sistema resulta mais de acontecimentos exógenos tais como a organização institucional e profissional do hospital. Veremos mais adiante quais são as implicações da existência de ineficiência estrutural no sistema

hospitalar para o nosso estudo.

4.2. - FUNÇÕES DE CUSTO HOSPITALARES

A tradição de especificação e estimação de funções de custo para o sistema hospitalar remonta a Feldstein (1967). Apesar de inicialmente esta prática estar relacionada com o objectivo da determinação da dimensão óptima da unidade hospitalar (Feldstein, 1967; Evans e Walker, 1972; Jenkins, 1980), adquiriu, posteriormente, outros objectivos tais como a determinação dos custos unitários dos serviços prestados (Jenkins, 1980), a medição da ineficiência específica de cada unidade hospitalar (Wagstaff, 1989) ou a caracterização da tecnologia (Vita, 1990).

Do ponto de vista da especificação existem alguns traços comuns às abordagens anteriores. Exceptuando o trabalho de Vita, que estima uma função de custo do tipo "translog", os restantes autores especificam funções que a literatura chama de "behavioural" ou "ad hoc", quer dizer funções de custo que não são duais de uma função de produção. A especificação de funções do tipo "ad hoc" não tem, pois, um fundamento claro na teoria da produção.

As funções de custo especificadas incluem, em geral, como

variáveis explicativas quer o número de camas e o fluxo de casos, quer o número de camas, o índice de ocupação e a duração média, sob uma forma aditiva. A estas variáveis junta-se habitualmente e sem uma justificação clara, potências de diversa ordem dos próprios regressores. Aparentemente a origem desta tradição situa-se em Feldstein (1967) que ao pretender determinar a dimensão óptima das unidades hospitalares inglesas pretende encontrar a curva de custo médio por regressão do custo por caso no número de camas e seu quadrado. Trata-se de obter a curva em U do custo médio como função da variável dimensão número de camas e seu quadrado. Apesar da qualidade do ajustamento obtido ser reduzida, Feldstein observa que se se inclui o fluxo de casos e o seu quadrado como variáveis explicativas adicionais, os resultados melhoram notavelmente.

Feldstein só se decide por esta especificação depois de várias outras tentativas de menor êxito. Trata-se, em resumo, de uma função de custo que resulta de um processo aproximativo sem fundamento na teoria da produção, excepto no que se refere à forma de curva em U que postula para a curva de custo médio. A função de custo adoptada por Feldstein é a seguinte:

$$[C/y] = \beta_1 + \beta_2 B + \beta_3 B^2 + \beta_4 F + \beta_5 F^2 + p\gamma + u$$

onde: C/y é o custo por caso

B é o número de camas

F é o fluxo de casos

p é o vector de case-mix, quer dizer, o quociente entre o número de casos de cada especialidade

médica e o número total de casos tratados pelo hospital

u é um termo de perturbação aleatório cumprindo as hipóteses básicas do modelo de regressão linear múltipla.

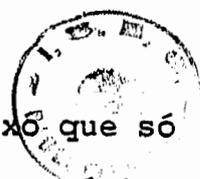
As tentativas de especificação da função de custo feitas por Feldstein diferenciam-se pelas variáveis incluídas como regressores. Apesar disso, este autor justifica a inclusão do fluxo de casos na função de custo pela sua contribuição à medição da intensidade de utilização da capacidade produtiva. O número de camas é considerado por Feldstein como a variável de dimensão mais apropriada. A inclusão do vector de case-mix como regressor é obviamente uma forma de ter em conta o carácter heterogéneo do output. Feldstein encontra um mínimo do custo médio para instituições hospitalares com dimensão entre 250 e 300 camas, argumentando que acima deste limite há deseconomias de escala.

Com esta aproximação Feldstein logra fixar um padrão para a função de custo que terá posteriormente grande popularidade. A proposta de Feldstein traduz-se na forma aditiva da função, na não linearidade nas variáveis, capaz de produzir uma curva em U quando os custos por caso se simulam como função do número de camas, e uma tentativa implícita de medir o output através de uma "proxy" da capacidade produtiva, de uma "proxy" da intensidade de utilização da capacidade produtiva e uma forma de ter em conta a heterogeneidade do output através do vector do "case-mix".

Certamente Evans e Walker (1972) tiveram uma contribuição importante para o aperfeiçoamento da especificação de funções de custo. A sua originalidade reside na forma como têm em conta a intensidade e complexidade dos casos médicos, contribuindo desta forma para uma medição mais precisa do output. A especificação da função de custo de Evans e Walker, para além da variável de complexidade, é muito semelhante à proposta por Feldstein. A função inclui a variável camas e o seu quadrado, o índice de ocupação e a duração média ou, alternativamente a estas duas variáveis, o fluxo de casos. Estas variáveis são incluídas sob uma forma aditiva, como já o fazia Feldstein. Também como este autor, a utilização que fazem Evans e Walker da função de custo relaciona-se com o estudo da dimensão óptima das unidades hospitalares.

Vejamos com um pouco mais de detalhe em que consiste a variável de complexidade introduzida por Evans e Walker. O seu raciocínio é que os casos mais complexos de cada especialidade médica tendem a ser tratados em poucos hospitais onde se concentram os equipamentos e o pessoal especializado, enquanto que os casos mais simples tendem a distribuir-se de uma forma homogénea por todas as unidades do sistema hospitalar.

Evans e Walker obtêm a sua variável de complexidade e especialização medindo o grau de concentração de cada hospital em determinada especialidade médica e atribuindo um valor maior à variável de complexidade quanto maior for a concentração numa (ou em várias) especialidades médicas. Como casos limites Evans e



Walker falam de situações onde o diagnóstico é tão complexo que só pode ser tratado num único hospital do sistema ou, pelo contrário, um diagnóstico é tão simples que pode ser tratado em qualquer hospital. Voltaremos mais adiante e com mais pormenor ao estudo deste índice de complexidade e especialização.

Sobre a abordagem de Jenkins (1980) da função de custo, diremos simplesmente que se trata duma abordagem mista no que se refere à medição do output: mede-se através da inclusão de variáveis tais como o número de estadias ou o número de casos tratados (medida do tipo "case-type") ou através de variáveis relativas aos diversos outputs intermédios (raios-X, análises ou tratamentos de fisioterapia, etc) produzidos pelo hospital (medida do tipo "service-mix"). Só se distingue das abordagens anteriores pela forma de incluir as variáveis dimensão, mantendo a característica de considerar uma forma funcional aditiva.

A multiplicidade de variáveis incluídas permite a Jenkins ter uma ideia detalhada da estrutura dos custos do hospital. Contudo, a forte colinearidade põe o problema da precisão das estimativas obtidas. Também Jenkins utiliza a sua função de custo na análise das economias de escala. Uma característica comum a Evans e Walker e a Jenkins (os dois estudos trabalham com dados canadianos) é que encontram dimensões óptimas menores que a de Feldstein (100-200 camas no estudo de Evans e Walker; 100-300 camas no de Jenkins).

Apesar de, pelas variáveis que inclui na função de custo, o

trabalho de Wagstaff (1989) não se distinguir da especificação da função de custo de Feldstein, a sua especificação estocástica e a utilização que dela faz representam uma inovação em relação a trabalhos anteriores. As principais novidades da abordagem de Wagstaff são, em primeiro lugar, a especificação da função de custo em termos de função de fronteira, o que lhe permite, como já escrevíamos no capítulo anterior, medir a ineficiência em cada unidade hospitalar e, em segundo lugar, a utilização de dados de tipo painel o que supõe uma melhoria em relação a trabalhos anteriores que só usam dados seccionais.

Torna-se possível a utilização de uma função de custo deste tipo em política sanitária e, em especial, na elaboração de indicadores de eficiência da gestão das unidades hospitalares. Contudo, vamos seguidamente formular algumas críticas a este trabalho, que, todavia, não invalidam o seu carácter inovador nem de referência obrigatória quando se aborda o problema da eficiência hospitalar.

Como se viu, a especificação estocástica e o método de estimação utilizados por Wagstaff podem conduzir a inferências inconsistentes. Mundlak e outros autores argumentam que normalmente os factores de ineficiência - ou seja, os efeitos individuais, que Wagstaff considera aleatórios - estão correlacionados com a dimensão dos centros, quando se modeliza o comportamento do produtor. A validação deste modelo depende de um ensaio de hipóteses que afaste aquele tipo de situação. Na ausência deste

teste, a literatura econométrica - Mundlak, Hsiao e Judge, entre outros - é explícita sobre os riscos de inconsistência do método, utilizado naquele contexto.

Por outro lado, e também numa perspectiva econométrica, o objectivo de Wagstaff é a comparação de diversos estimadores alternativos com diferentes especificações estocásticas. Pretende deste modo verificar a robustez dos resultados em relação a diferentes métodos de estimação de fronteiras. Encontra Wagstaff uma grande instabilidade das estimativas, passando a concluir que talvez esteja na forma de estimação de fronteiras a origem da instabilidade. Contudo devemos contrapor a possível inconsistência da especificação de efeitos individuais aleatórios e o elevado grau de colinearidade envolvida por alguns dos métodos de estimação. Com efeito, o vector do case-mix apresenta, no caso espanhol, uma elevada multicolinearidade pelo que, quando se introduzem estas variáveis como regressores é fácil provocar a instabilidade das estimativas.

Uma terceira crítica constitui o conceito de eficiência que utiliza Wagstaff. De uma forma geral diremos que, quando se utilizam funções de fronteira, a ineficiência mede-se pelo rácio entre o valor do custo (ou output) efectivamente observado e a "estimativa" do custo (ou output) obtida com a parte principal do modelo. Esta parte principal do modelo estima o custo mínimo (ou o output máximo) que é possível alcançar. Como a função de custo utilizada por Wagstaff não é dual de uma função de produção, ou

seja, não resulta da minimização dos custos sujeitos a uma função de produção, a estimação efectuada pode não coincidir com o custo mínimo.

A tentativa de Wagstaff, apesar das imprecisões aqui apontadas, é inovadora e é uma consequência da utilização dos resultados de uma década de avanços na teoria das funções de fronteira, assim como da contribuição de Schmidt e Sickles (1984) sobre a estimação de funções de fronteira com painel de dados. A aplicação destes métodos econométricos à economia sanitária, num período onde se faz um crescente apelo à eficiência do sector, é certamente de grande oportunidade.

Outro trabalho sugestivo é o que apresenta Vita (1990). A função de custo ajustada, abandona os pressupostos das funções de custo "behavioural" ou "ad hoc", para ser, mais de acordo com a teoria microeconómica, uma função dual de uma função de produção; mais concretamente dual da função de produção "translog". Como é conhecido, esta função de produção é uma aproximação pontual exacta à verdadeira função de produção até à segunda ordem de diferenciação. Também a função de custo "translog" é uma aproximação pontual exacta da verdadeira função de custo até à segunda ordem de diferenciação.

Se tivermos em conta os teoremas da dualidade entre a função de custo e a função de produção, estamos em condições de poder caracterizar a tecnologia empregue pelo sistema hospitalar também

em termos da função de custo. Por isso, Vita obtém para o sistema hospitalar os diferentes tipos de economias de escala e de alcance ("scope economies").

A abordagem de Vita tem, contudo, alguns problemas. A estimação da função de custo "translog" não é directa. Vita utiliza a via indirecta com a estimação das equações das percentagens factoriais ("factor cost shares equations") que se obtém sob a hipótese de que as unidades hospitalares minimizam os custos. Contudo, e tal como vimos anteriormente, o comportamento do sector não lucrativo do sistema hospitalar, dominante nos EUA, é compatível com a existência de ineficiência de custo. Vita deveria ter tido em conta a ineficiência de afectação deste subsector de forma explícita no processo de estimação da sua função de custo. Essa ineficiência sistemática, como se viu, traduz-se na contratação de menos médicos que o equilíbrio competitivo exige (Pauly e Redisch) ou, alternativamente, no sobreequipamento das instalações hospitalares (Harris). Este facto tem como consequência que as correspondentes equações de percentagens factoriais devem especificar-se por forma a que as perturbações aleatórias tenham média não nula nas observações correspondentes aos centros não lucrativos (em correspondência com a ineficiência de afectação sistemática). Na forma que Vita estima as equações de percentagens factoriais, existe enviesamento nos termos independentes daquelas equações e portanto um enviesamento na estimação dos parâmetros correspondentes àquelas variáveis na função de custo.

4.3. - A MEDIÇÃO DO OUTPUT DOS HOSPITAIS

Na análise da eficiência dos centros hospitalares fizemos referência às funções de custo e de produção. Contudo, um problema decisivo na especificação destas funções é o da medição do output, que é variável dependente na função de produção e um dos regressores na função de custo.

A correcta medição do output hospitalar é um problema complexo, e aparentemente de difícil solução. Isto resulta do produto da actuação hospitalar - a melhoria do estado de saúde dos doentes - ser, enquanto tal, de difícil medição. Não obstante, é sobre este resultado final da acção do sistema hospitalar, quer dizer, sobre a melhoria do estado de saúde de uma determinada população, que alguns autores pretendem medir o output dos hospitais.

Contudo, existe uma multiplicidade de factores, que estão para além da acção do sistema hospitalar, e que influem no estado de saúde dos indivíduos. Estes factores podem ser de ordem genética, biológica, cognitiva, etc. Além disso, o próprio meio físico, social e económico (factores ambientais, o grupo social ao qual pertence o indivíduo, o estado da economia, etc.), fazem com que seja difícil isolar o resultado de um tratamento hospitalar episódico sobre o estado de saúde de um determinado indivíduo.

A noção de que se deve medir o resultado de cada acção hospitalar em termos de QALY's ("Quality Adjusted Life Years") resulta, apesar da sua dificuldade de concretização, certamente paradigmática. Assim um episódio de tratamento tem como output um determinado número de anos de vida, ajustados pela sua qualidade, além dos que teria o doente na ausência de tratamento hospitalar (Culier et al. (1971); Williams (1985)).

Apesar do seu atractivo, a medição do output hospitalar através de QALY's, tem uma implantação prática complexa pela dificuldade de isolamento dos factores que influem no estado de saúde de uma população determinada, e também pela sua própria forma de medição. De momento, a economia da saúde preferiu tomar como objecto de medição não o resultado final da acção hospitalar, mas os outputs intermédios desta actividade.

A tentativa de medir o resultado da actividade hospitalar através da quantidade de serviços prestados, resulta até certo ponto compreensível no contexto do sistema sanitário norte-americano, onde o doente contrata separadamente os serviços de um hospital e os cuidados de um médico. A quantidade de serviços prestados refere-se tanto a dias de internamento, como a meios auxiliares de diagnóstico, tais como radiografias, análises laboratoriais, como aos tratamentos de fisioterapia, etc. É o processo de medição do output através do "service-mix".

Cohen (1967 e 1970) desenvolve uma forma de agregação destes

outputs intermédios heterogêneos. Para isso, expressa o output dos diferentes serviços do hospital em unidades físicas e adopta como medida do output do hospital a soma ponderada desses outputs, usando como ponderadores os rácios entre os custos médios de cada serviço e o custo médio de um dia de estadia. Esta abordagem é criticável pelo facto de parecer ignorar as diferenças no output provenientes da sua diferente composição em termos das patologias tratadas. Também ignora as diferenças de qualidade e eficiência dos serviços prestados pelos hospitais.

Outra abordagem da medição do output hospitalar é a que utiliza como medida o número de casos tratados. Evidentemente uma dificuldade relativa a este procedimento é a extrema heterogeneidade daquilo que se está medindo. É a chamada abordagem do case-mix.

Na literatura especializada encontram-se diversas formas de tratar o problema da heterogeneidade dos casos. Por um lado, pretende-se chegar a grupos ou categorias homogêneas de patologias, e, por outro, propõem-se medidas agregadas de output tendo em conta essa heterogeneidade.

A classificação das patologias baseou-se frequentemente na Classificação Internacional das Doenças (ICD) que na sua forma mais detalhada considera mais de 1000 tipos de patologias distintas. Contudo, muitas vezes, não há informação sobre os diagnósticos e a

única referência em relação a cada caso tratado é a especialidade médica em que se enquadra. Esta é a situação da qual parte Feldstein (1967). Este autor só dispunha do número de casos em cada especialidade médica, agrupadas em 8 categorias mutuamente exclusivas.

É interessante observar a utilização que Feldstein faz desta informação. Em primeiro lugar, calcula as percentagens de casos em cada uma das 8 especialidades médicas, depois, sempre que pretende incluir o output hospitalar como variável numa regressão, representa-o pelo número de casos, sendo a sua heterogeneidade tida em conta com a inclusão, como regressor adicional, do vector das percentagens das 8 especialidades médicas de cada hospital.

Outra forma usada por Feldstein, para ter em conta a heterogeneidade do número de casos como medida do output hospitalar, foi a agregação dos casos de diferentes especialidades médicas numa única medida resultante da diferente ponderação dos casos das 8 distintas especialidades médicas. O ponderador utilizado foi o custo médio no conjunto do sistema hospitalar de um caso da especialidade médica em questão.

Pretende-se chegar assim ao custo total esperado de produzir um determinado case-mix. O argumento é que o custo médio esperado (o custo médio no conjunto do sistema hospitalar) de um caso de um tipo determinado é proporcional ao custo social médio de um caso desse tipo. Então aquela medida agregada deve ser proporcional ao

valor social produzido num hospital dado. Apesar desta última argumentação ser demasiado forte nos seus pressupostos, permaneceu a ideia na literatura da economia da saúde, de aproximar o custo marginal do tratamento de um caso de um tipo determinado no conjunto do sistema hospitalar pelo custo médio de um caso desse tipo, também no conjunto do sistema.

Este pressuposto está também presente na medida agregada que propõem Jensen y Morrisey (1986):

$$y_h^* = I_h \cdot N_h$$

com

$$I_h = \frac{\sum_{i=1}^N w_i p_{hi}}{\frac{1}{H} \sum_{h=1}^H \sum_{i=1}^N w_i p_{hi}}$$

A medida de Jensen e Morrisey está baseada na correcção do número de casos tratados (N_h) num hospital dado por um valor específico desse hospital (I_h) que pretende traduzir (em relação ao conjunto do sistema hospitalar) a relativa complexidade do case-mix tratado.

Este valor (I_h) resulta da comparação entre o "valor" do case-mix no hospital h e o "valor" médio dos case-mix do conjunto do sistema hospitalar. Na expressão anterior, w_i representa o custo médio de um caso do tipo i no conjunto do sistema e p_{hi} representa a percentagem de casos do tipo i no hospital h.

Evidentemente que o valor destas medidas de output baseadas nas percentagens de casos em cada especialidade médica depende da homogeneidade dos casos em cada categoria. De facto, não se deveria trabalhar com os casos classificados em especialidades médicas, onde existe uma grande heterogeneidade intra-categoria, mas sim adoptando uma classificação dos casos em patologias de tratamento homogéneo.

A classificação mediante os DRG's (Grupos de Diagnóstico Homogéneo) teve um grande desenvolvimento nos Estados Unidos desde meados da década de 70. O seu objectivo é o de agrupar na mesma categoria casos homogéneos do ponto de vista do consumo de recursos e da prática clínica (Fetter et al. (1980)). Os 383 DRG's (baseados na ICD) estabelecidos por Fetter, pretendem ser "a manageable number of patient classes that are exhaustive and mutually exclusive with respect to the types of patients seen in an acute care setting. Further, the DRG's provide patient classes that are clinically consistent and that have similar patterns of output utilization as measured by length of stay"².

O objectivo desta classificação é não só chegar a categorias de case-mix homogéneas, mas também à normalização de tratamentos e durações médias. É possível também a sua utilização no cálculo do financiamento do sistema hospitalar tendo em conta a composição e a complexidade dos casos tratados. Diversas críticas foram feitas

² - Fetter, R. B. et al. "Case-mix definition by Diagnostic Related Groups" citado por Thatchel (1984).

ao DRG's: primeiro, a sua possível não adequação a regiões para as quais não foram delineados; segundo, o insuficiente poder de discriminação dos casos tratados (casos atribuídos ao mesmo DRG podem ser tratados de uma forma diferente em diferentes hospitais); terceiro, a importância do diagnóstico na atribuição do DRG e o espaço para a subjectividade na classificação pode levar a uma utilização enviesada no financiamento dos hospitais; e, por último, a grande incidência de erros de medida e de apreciação nos diagnósticos.

Desde meados da década de 80 que o Ministério da Saúde faz esforços para implantar um sistema de financiamento dos hospitais baseado nos DRG's. Este sistema começou a ser utilizado no ano de 1990 e estará totalmente implantado em meados da presente década. Mas quer se utilize especialidades médicas, quer se utilize os DRG's, ou outro agrupamento das categorias do ICD, será importante sintetizar tanto quanto possível a informação, quando se pretende medir o output.

Uma das técnicas de condensação da informação é a proposta pelo índice de especialização e complexidade desenvolvido por Evans e Walker (1972). Esta medida está baseada na ideia de que os casos mais complexos deverão ser tratados em poucos hospitais onde existem os equipamentos e os meios humanos suficientemente sofisticados para tratar esses casos, enquanto que os casos mais simples deverão distribuir-se uniformemente por todo o sistema hospitalar.

Vejamos como se calcula o índice de Evans e Walker. Seja N_{ij} , o número de casos do tipo j tratados no hospital i ; $N_{i.}$, o número de casos de todos os tipos tratados no hospital i ; e seja $N_{.j}$, o número de casos do tipo j tratados no sistema hospitalar, então:

$$p_{ij} = \frac{N_{ij}}{N_{i.}}$$

é a percentagem dos casos tratados pelo hospital i , que pertencem à categoria j , e

$$q_{ij} = \frac{N_{ij}}{N_{.j}}$$

é a percentagem dos casos de tipo j que são tratados no hospital i .

A complexidade de cada categoria define-se como

$$H_j^1 = \sum_i q_{ij} \log(H \cdot q_{ij})$$

no caso de $q_{ij}=0$, $q_{ij} \log(H \cdot q_{ij})$ não está definido, mas toma-se então $q_{ij} \log(H \cdot q_{ij})=0$, visto que $\lim_{x \rightarrow 0} (x \log(x))=0$. Por outro lado,

$$p_{i.} = \frac{N_{i.}}{N}$$

é a percentagem de casos de todos os tipos tratados no hospital i , e

$$q_{.j} = \frac{N_{.j}}{N}$$

é a percentagem de casos de tipo j tratados pelo sistema. Normalizando H_j^1 teremos:

$$\bar{H}_j^1 = \frac{H_j^1}{\sum_j H_j^1 q_{.j}}$$

Esta normalização faz com que a soma ponderada do número de casos em cada patologia pela sua complexidade seja igual ao número total de casos, quer dizer:

$$\sum_j N_{.j} \bar{H}_j^1 = N$$

O índice de complexidade de cada hospital será então:

$$\phi_i = \sum_j \bar{H}_j^1 p_{ij}$$

que não é outra coisa senão a média ponderada dos indicadores de complexidade das patologias (ou especialidades), sendo os pesos as proporções do case-mix. Esta medida de complexidade revelou um grande poder explicativo das variações dos custos entre hospitais, tal como mostra um estudo empírico realizado por Evans e Walker. Contudo, existem algumas limitações na utilização desta metodologia. Em primeiro lugar, quando o índice se aplica a um grande número de categorias de diagnóstico, a raridade do diagnóstico pode confundir-se com a sua complexidade e, em segundo lugar, ignoram-se as variações de complexidade dentro de cada categoria do case-mix visto que se trabalha como se estas fossem categorias homogêneas.

4.4. - A MEDIÇÃO DO OUTPUT E A ESTIMAÇÃO DE FUNÇÕES DE CUSTO E DE PRODUÇÃO.

Uma das formas mais populares de medição do output nas funções de custo e de produção consiste em considerar o carácter de múltiplo output da produção hospitalar. Esta consideração já se podia encontrar implícita na especificação de Feldstein. A razão que o faz considerar o vector de case-mix na função de produção é a sua especificação como argumento de uma actividade produzindo múltiplos outputs a partir de inputs múltiplos, quer dizer:

$$\phi(X_{1i}, \dots, X_{si}, C_{1i}, \dots, C_{mi}, u_i) = 0$$

Feldstein inclui na sua forma implícita s inputs e m outputs que se identificam com o número de casos nas diferentes especialidades médicas. Uma forma alternativa de apresentar aquela função é:

$$\phi(X_{1i}, \dots, X_{si}, Y_i, P_{1i}, \dots, P_{mi}, u_i) = 0$$

onde:

$$y_i = \sum_{j=1}^m C_{ji} , P_i = \frac{C_i}{\sum_j C_{ji}} .$$

Explicitando esta função em relação a y_i , teremos:

$$y_i = \phi(X_{1i}, \dots, X_{si}, P_{1i}, \dots, P_{mi}, u_i)$$

A forma funcional considerada por Feldstein para a função de produção é a Cobb-Douglas, mas outras formas funcionais são, evidentemente, compatíveis com esta especificação. Lopez e Wagstaff (1988) adoptam a especificação de Feldstein com a mesma forma funcional e a mesma forma de medição do output que este autor para dados referentes aos hospitais espanhóis.

Uma forma alternativa de considerar o caráter de output múltiplo da produção hospitalar é a de Conrad et al. (1983) e Vita (1990). Estes autores consideram uma função de custo da seguinte forma:

$$C_i = \phi(w_{1i}, \dots, w_{si}, y_{1i}, \dots, y_{mi}, u_i)$$

onde C_i são os custos do hospital i , w_{ji} é o preço do input j no hospital i e y_{ji} é o número de casos da especialidade j no hospital i .

Para estes autores o número de categorias de case-mix adoptado

é de quatro ($m=4$) pelo que parece ser uma hipótese demasiado forte a da homogeneidade dos quatro outputs e insatisfatória a forma de captar a variabilidade do case-mix pela especificação adoptada. Talvez seja por esta razão que Vita inclui além dos quatro tipos de produto, uma variável a que chama índice de case-mix, cuja natureza não explicita e que se destina a introduzir a variabilidade do case-mix não incluída nas quatro variáveis de output.

Apesar do atractivo deste tipo de análise, problemas de ordem prática surgem quando se pretende utilizar o vector de case-mix como regressor. De facto Lopez e Wagstaff e Wagstaff, trabalhando uma amostra de hospitais espanhóis, encontraram sérios problemas de colinearidade, cuja origem foi possível diagnosticar (usando o "condition number") como residindo em grande parte naquele vector. Também no caso português foi possível verificar um elevado grau de colinearidade no vector do case-mix. Voltaremos mais adiante a este tema.

Este facto aconselha a utilização de medidas agregadas do output, baseadas no case-mix como, por exemplo, o índice de complexidade e especialização de Evans e Walker e a normalização dos casos tratados pelo índice de Jensen e Morrisey. A comparação das duas medidas agregadas, a priori, e sem ter em conta a sua aplicação empírica, é difícil.

A medida de Jensen e Morrisey tem o atractivo da medição do output com uma medida agregada que tem uma interpretação física

evidente: casos, normalizados por meio de um índice que tem em conta a composição do output nas diferentes especialidades médicas e os custos médios no conjunto do sistema hospitalar. A justificação para a utilização dos custos no conjunto do sistema é, na opinião de Feldstein: "if hospital i produces an additional case of type j , one less case of that type can be produced elsewhere without changing the amount of hospital care in the system as a whole". A interpretação de Feldstein do custo médio (do sistema) como o valor social marginal de um caso desse tipo, já foi criticada por Lopez e a ela não voltaremos. Contudo, a utilização dos custos médios do sistema na ponderação do case-mix está suficientemente divulgada e utilizá-la-emos sem mais justificações.

CAPÍTULO 5

A EFICIÊNCIA NO SISTEMA HOSPITALAR PORTUGUÊS

A função de produção, a função de custo e a função de lucro, além de caracterizarem uma tecnologia dada, permitem a medição das diferentes formas de eficiência económica, quando especificadas como funções de fronteira. É a estimação de fronteiras para os hospitais portugueses que constitui o conteúdo essencial deste capítulo.

Começamos o capítulo com uma breve descrição da amostra utilizada e com a comparação de algumas medidas de output hospitalar. Numa segunda secção obtemos a ineficiência de custo de cada centro ao especificarmos a função de custo como função fronteira. Numa terceira secção, e com a estimação da função de produção, pudémos separar a ineficiência de custo em ineficiência técnica e de afectação.

Com a estimação da eficiência técnica pudémos constatar a sua grande variabilidade entre centros. Como origem dessa variabilidade conseguimos estabelecer que eram factores relacionados com a dimensão dos centros e factores associados à duração média, que não tinham sido tidos em conta na função de produção estimada. Fizemos a regressão da eficiência técnica sobre as variáveis camas, duração

média e tipo de centro, tendo obtido assim a ineficiência técnica líquida daqueles factores. Esta medida de eficiência tem já muito menor variabilidade do que a medida anterior.

Na segunda secção estimámos uma função de custo que é um compromisso entre as funções de custo "behavioural" habituais e as funções de custo duais de funções de produção. Também a ineficiência de custo tinha grande variabilidade entre centros. Os factores que encontrámos como explicativos desta variação foram a dimensão e o tipo de centro. Este resultado permitiu-nos, por um lado a simulação do custo médio como função do número de camas, tendo-se obtido uma curva em U com um mínimo cerca das 400 camas, e por outro obter a eficiência de custo líquida daqueles factores.

Uma aproximação ao estudo das elasticidades de substituição entre factores no sistema hospitalar português é o que nos propusémos fazer na quarta secção quando estimámos uma função de custo do tipo "translog".

5.1. - OS DADOS

Um problema prévio que devemos resolver - quer se trabalhe com a função de produção, quer se trabalhe com a função de custo - é o

da medição do output hospitalar. Das diferentes formas de medição que apresentámos no capítulo 4, escolhemos uma com base no seu poder de síntese da informação nas suas possibilidades de utilização prática. Apresentamos na segunda sub-secção desta secção uma comparação, usando a amostra dos hospitais portugueses, entre o índice de Evans e Walker e o índice de Jensen e Morrisey, com a finalidade de verificar o seu comportamento em aplicações empíricas.

Outro problema é o de encontrar uma "proxy" para o stock de capital porque não é possível medir, mesmo de uma forma aproximada, aquela quantidade. Na literatura consultada é manifesto que se toma tradicionalmente o número de camas como "proxy" do stock de capital, ainda que em certas ocasiões seja pouco precisa esta aproximação (pode-se, por exemplo, afirmar que para igual número de camas, centros com especialidades médicas diferentes, necessitam de stocks de capital diferentes), é a melhor entre as disponíveis.

Um terceiro problema no trabalho empírico desenvolvido foi o de encontrar um preço para o factor capital. Optámos, como Lopez e Wagstaff (1988), pelo custo de aluguer de uma cama de hospital. Adoptando como custo de oportunidade do capital a taxa de juro real à qual o Estado é financiado e um período de vida útil de 20 anos para o investimento, encontrámos, partindo de um custo de investimento por cama correspondente ao praticado pelo Ministério da Saúde nos últimos anos, o custo de aluguer de uma cama de hospital. É calculado da seguinte forma:

$$I = \sum_{t=1}^{20} \frac{A}{(1+i)^t}$$

onde: I é o custo de construção e equipamento por cama de um hospital, i é o custo de oportunidade do capital e A é o custo de aluguer da cama de hospital.

As restantes variáveis (custos, casos, estadias, quantidades físicas dos inputs e seus preços) foram cedidos pelo Ministério da Saúde. Nas secções seguintes apresentamos uma breve descrição da amostra (restringimos a análise ao corte transversal de 1988) e, como referimos, um estudo empírico do problema da medição do output.

5.1.1. - DESCRIÇÃO DOS DADOS DA AMOSTRA

Começamos por uma descrição geral dos dados da amostra. O quadro 1 descreve as variáveis incluídas nos modelos que se estimam mais adiante. Os dados referem-se ao ano de 1988 e incluem as variáveis camas, casos, custo por caso, duração média e índice de ocupação; incluem além destas as variáveis relativas ao pessoal empregue: médicos, enfermeiras e pessoal auxiliar, assim como os respectivos salários médios.

Pode-se comprovar a grande dispersão das variáveis relacionadas com a dimensão dos centros: camas, casos, pessoal;

enquanto as variáveis relativas aos salários apresentam muito menor dispersão. Outras variáveis como a estadia média, o índice de ocupação e o custo por caso, têm uma dispersão intermédia em termos absolutos, mas significando provavelmente condições de exploração muito diferentes entre hospitais.

QUADRO 1

VARIÁVEL	max	min	média	desv padr
camas	1494.000	57.000	363.256	323.82
custo/caso	230.960	29.399	110.703	44.39
estad. med	18.258	7.304	11.128	2.66
índice ocu	0.957	0.542	0.751	0.08
casos	31637.431	1476.125	8842.022	7115.71
médicos	1066.000	19.000	161.436	231.63
enfermeir.	1223.000	48.000	264.026	283.79
auxiliares	2117.000	127.000	491.000	516.33
sal med	3711.053	1229.044	2547.814	458.04
sal enfer	1959.542	1038.424	1345.278	175.00
sal auxil	886.167	643.993	728.367	53.75

Uma primeira aproximação ao problema das economias de escala, pode ser obtida da observação do seguinte quadro:

QUADRO 2

	camas	cust/caso	est.med.	índice ocup	índice
4 quartil	820.5556	146.7020	12.9316	0.8067	1.2804
3 quartil	337.3000	99.9864	11.2108	0.7341	1.2579
2 quartil	213.6000	94.2549	10.0242	0.7106	1.1984
1 quartil	127.3000	105.4668	10.5256	0.7587	1.1930

Dividiu-se a amostra dos centros em quartis de acordo com a dimensão (camas). Para cada um dos quartis calculou-se a média amostral para as seguintes variáveis: camas, custo por caso, duração média índice de ocupação e índice de complexidade (Jensen e Morrisey). Pode-se constatar que o custo por caso, a duração média e o índice de ocupação têm um mínimo no 2º quartil, que tem como dimensão média 214 camas; além disso, o índice de complexidade está positivamente correlacionado com a dimensão dos centros.

5.1.2. - ALGUMAS MEDIDAS DE OUTPUT NO CASO PORTUGUÊS

Apresentamos seguidamente os resultados do cálculo de algumas medidas de output, para o caso português. A primeira das medidas de output estudadas foi a inclusão do vector de case-mix na regressão com a utilização do número de casos como medida do output (Feldstein - 1967). O vector de case-mix introduz, neste caso, uma forte colinearidade na regressão, sendo o "condition number":

$$CN = \left(\frac{\lambda_1}{\lambda_k} \right)^{1/2}$$

(ou seja: a raiz quadrada do quociente entre o maior e o menor dos valores próprios da matriz $X'X$, depois de normalizadas as variá-

veis), $CN=100.28$,³ quando se inclui na matriz X, além do termo independente, 13 categorias de case-mix. Considerou-se um case-mix com uma dimensão de 16 e ao qual se amputaram 3 categorias para evitar a colinearidade perfeita. As categorias de case-mix correspondem às especialidades mais frequentes nos Hospitais Distritais; as restantes especialidades existentes nos Hospitais Centrais foram agrupadas naquelas 16. A colinearidade introduzida pelo vector de case-mix numa regressão onde seja incluído, é provavelmente tão forte que os parâmetros estimados serão altamente instáveis.

Também se calculou o índice de Evans e Walker e o índice de Jensen e Morrisey. Em relação a estes índices, calcularam-se as correlações entre o custo por caso e os índices referidos, a correlação entre os dois índices, para cada um dos 5 anos da amostra de 46 hospitais.

A correlação entre os custos médios e o índice baseado na teoria da informação varia entre 0.4 e 0.51 (em 1988 e 1984 respectivamente); a correlação entre o custo médio e o índice de Jensen e Morrisey varia entre 0.38 e 0.47 (1985 e 1988 respectivamente); a correlação entre os dois índices entre 0.57 e 0.72 (1985 e 1987 respectivamente). Todos estes resultados podem ser vistos com mais detalhe no quadro seguinte:

³ O valor acima do qual se admite como demasiado fortes os efeitos da colinearidade é $CN = 30$

Correlação	1984	1985	1986	1987	1988
Custo, Teo. Inf.	0.506	0.497	0.407	0.407	0.402
Custo, Ind. Jen. Mo	0.396	0.382	0.402	0.440	0.471
Te. In., In. Je. Mo.	0.655	0.574	0.713	0.721	0.638

Os dois índices estão aparentemente medindo coisas não muito diferentes (dada a grande correlação entre ambos), e a variável da qual são "proxies" tem uma correlação algo forte com o custo por caso (dados os valores das correlações de ambos com o custo médio). Não há uma vantagem nítida de um índice sobre o outro dados os valores semelhantes das correlações. Os valores de ambos os índices para o conjunto dos hospitais são os seguintes:

ÍNDICE DE ESPECIALIZAÇÃO E COMPLEXIDADE

Centros	1984	1985	1986	1987	1988
Centro 1	-	-	-	-	-
Centro 2	1.11472	1.10638	1.09998	1.09846	1.09087
Centro 3	1.20058	1.06724	1.04926	1.05543	1.05802
Centro 4	1.17360	1.20417	1.24029	1.22117	1.27353
Centro 5	1.13022	1.15899	1.15159	1.12136	1.09032
Centro 6	1.09066	1.08940	1.04484	1.09468	1.07172
Centro 7	-	-	-	-	-
Centro 8	0.98485	0.95903	0.93278	0.93466	0.88487
Centro 9	-	-	-	-	-
Centro 10	0.68538	0.73522	0.72064	0.74566	0.76417
Centro 11	-	-	-	-	-
Centro 12	-	-	-	-	-

Centro 13	1.07018	1.05596	1.04935	1.11025	1.05879
Centro 14	-	-	-	-	-
Centro 15	-	-	-	-	-
Centro 16	-	-	-	-	-
Centro 17	0.96048	0.94299	0.96507	1.02130	1.05552
Centro 18	1.01637	0.96162	0.93940	0.93894	0.93529
Centro 19	-	-	-	-	-
Centro 20	0.73721	0.76720	0.75239	0.75114	0.73706
Centro 21	1.17035	1.24938	1.19894	1.16415	1.18671
Centro 22	-	-	-	-	-
Centro 23	-	-	-	-	-
Centro 24	-	-	-	-	-
Centro 25	0.93547	0.95952	0.90775	0.95904	0.88961
Centro 26	1.12552	1.09777	1.15680	1.16201	1.14968
Centro 27	1.03378	0.96368	0.95476	0.92632	0.91151
Centro 28	-	-	-	-	-
Centro 29	0.69682	0.72091	0.74651	0.76183	0.78674
Centro 30	0.78802	0.87417	0.79801	0.77658	1.10748
Centro 31	0.84653	0.84466	0.84870	0.79581	0.81883
Centro 32	0.92969	0.93742	0.94312	0.93674	0.90693
Centro 33	-	-	-	-	-
Centro 34	1.11090	1.09766	1.06536	1.09115	1.03893
Centro 35	0.80862	0.84757	0.81107	0.76861	0.76278
Centro 36	-	-	-	-	-
Centro 37	1.09739	1.10290	1.17474	1.29421	1.19720
Centro 38	0.80803	0.82125	0.80206	0.77822	0.78013
Centro 39	0.81308	0.91695	0.87501	0.85036	0.85994
Centro 40	0.70087	0.71347	0.69201	0.70179	0.70133
Centro 41	-	-	-	-	-
Centro 42	-	-	-	-	-
Centro 43	-	-	-	-	-
Centro 44	0.72967	0.74661	0.73384	0.73824	0.72810
Centro 45	0.89951	0.89613	0.94426	0.92114	0.94910
Centro 46	0.97768	0.99124	0.96939	0.94264	0.90584

ÍNDICE DE JENSEN E MORRISEY

Centros	1984	1985	1986	1987	1988
Centro 1	-	-	-	-	-
Centro 2	1.13718	1.20258	1.16268	1.19161	1.33919
Centro 3	1.11551	1.20227	1.15825	1.21970	1.34723
Centro 4	1.08707	1.17825	1.16440	1.22945	1.39869
Centro 5	1.12111	1.19360	1.16510	1.19728	1.35280
Centro 6	0.90108	0.96408	1.00321	1.03775	1.17487
Centro 7	-	-	-	-	-
Centro 8	0.94880	1.05028	1.03798	1.05495	1.18340
Centro 9	-	-	-	-	-
Centro 10	0.94069	1.02011	0.99479	1.04010	1.21051
Centro 11	-	-	-	-	-
Centro 12	-	-	-	-	-
Centro 13	0.97514	1.00878	1.02781	1.03664	1.14895
Centro 14	-	-	-	-	-
Centro 15	-	-	-	-	-
Centro 16	-	-	-	-	-
Centro 17	0.86441	0.94923	0.95464	1.05012	1.21500
Centro 18	1.06219	1.16305	1.16929	1.17494	1.31871
Centro 19	-	-	-	-	-
Centro 20	0.89943	0.98315	0.98407	1.01743	1.13275
Centro 21	1.12464	1.14893	1.19963	1.25417	1.39497
Centro 22	-	-	-	-	-
Centro 23	-	-	-	-	-
Centro 24	-	-	-	-	-
Centro 25	1.01652	1.08076	1.05853	1.17843	1.29945
Centro 26	0.97970	1.04228	1.05503	1.13237	1.29264
Centro 27	1.06221	1.14343	1.01858	1.08885	1.17305
Centro 28	-	-	-	-	-
Centro 29	0.85579	0.94773	0.88657	0.97073	1.12319
Centro 30	1.02477	1.11369	1.04004	1.04283	1.11270

Centro 31	0.96323	1.04086	0.99608	1.01999	1.16804
Centro 32	0.98166	1.00350	0.98958	1.05966	1.19542
Centro 33	-	-	-	-	-
Centro 34	1.10908	1.14016	1.10735	1.16114	1.34480
Centro 35	0.94917	1.09050	1.01891	1.06115	1.20754
Centro 36	-	-	-	-	-
Centro 37	1.04061	1.09367	1.11268	1.20300	1.36596
Centro 38	1.04790	1.10780	1.05494	1.06379	1.21652
Centro 39	0.99791	1.02126	1.00233	1.05154	1.18609
Centro 40	0.81788	0.92450	0.86360	0.93787	1.08593
Centro 41	-	-	-	-	-
Centro 42	-	-	-	-	-
Centro 43	-	-	-	-	-
Centro 44	0.92848	0.99362	0.99955	1.06435	1.20675
Centro 45	1.03441	1.12245	1.08389	1.10399	1.26285
Centro 46	1.01343	1.04116	1.02299	1.05383	1.19895

5.2. - OBTENÇÃO DA INEFICIÊNCIA DE CUSTO

Na Economia da Saúde tem-se procurado modelizar com funções de custo do tipo "behavioural" as diferenças de custos unitários entre centros hospitalares; pretendia-se, além disso, estudar o problema da dimensão óptima dos hospitais, e também em recentes trabalhos, medir a ineficiência dos centros.

A estimação que fizemos da função de custo para os hospitais portugueses tem as seguintes características individualizadoras:

- . Na especificação: ao contrário das especificações "ad hoc", que predominam na função de custo hospitalar, explicitou-se uma função de custo do tipo "Cobb-Douglas".

- . Medição do output: considerou-se como medida do output a medida agregada de Jensen e Morrisey, que tem a vantagem de considerar uma medida física do output (ou seja: o número de casos), corrigida por um índice que tem em conta o facto de o case-mix ser mais ou menos "caro" do que o "valor" médio dos case-mix do conjunto do sistema hospitalar.
- . Medição da ineficiência: ao contrário da medida de eficiência de Wagstaff (1989), os factores de ineficiência obtidos podem ser interpretados como a percentagem em que os custos observados de um centro excedem os custos mínimos alcançáveis por esse centro, por especificarmos uma função de custo que supõe a minimização dos custos, sujeita a uma função de produção.
- . Estimação com painel de dados: a especificação dos efeitos individuais como fixos, permite não só não considerar o preço do capital, que se supõe ser o mesmo para todos os centros em cada ano, como também evitar a inconsistência causada pela correlação entre as variáveis independentes e os efeitos individuais. (A estatística de Hausman e Taylor permite-nos rejeitar que as diferenças entre o estimador próprio do modelo de efeitos fixos e o estimador MQG são só devidas a flutuações amostrais).
- . A forma de ter em conta o factor dimensão dos centros: por tratar-se de uma função de custo dual de uma função de produção, a dimensão dos centros é originariamente medida pelo output. Contudo, pode-se decompôr o output no produto de outras três variáveis: camas, índice de ocupação e duração

média; isto permite-nos aproximar a nossa especificação das especificações "ad hoc" e além disso fazer a simulação do custo médio como função do número de camas. Voltemos a alguns destes pontos.

Em primeiro lugar, abordemos a especificação da função de custo. Vejamos em que consiste a especificação intermédia entre a função de custo "Cobb-Douglas" e as especificações "behavioural". A função de custo Cobb-Douglas é a que se obtém como solução do problema:

$$\min_{K,L} w_L L + w_K K$$

$$s.a. \quad y \leq AL^\alpha K^\beta$$

$$L \geq 0, K \geq 0$$

A solução deste problema tem a forma:

$$C = C_0 y^{\frac{1}{\alpha+\beta}} w_L^{\frac{\alpha}{\alpha+\beta}} w_K^{\frac{\beta}{\alpha+\beta}}$$

onde C_0 é uma função dos parâmetros da função de produção. Tomando logaritmos:

$$\log C = \beta_0 + \beta_1 \log y + \beta_2 \log w_K + \beta_3 \log w_L$$

ou

$$\log C/y = \beta_0 + \beta'_1 \log y + \beta_2 \log w_K + \beta_3 \log w_L$$

onde

$$\beta_0 = \log C_0, \quad \beta_1 = \frac{1}{\alpha + \beta}, \quad \beta_2 = \frac{\alpha}{\alpha + \beta}, \quad \beta_3 = \frac{\beta}{\alpha + \beta}$$

$$\beta'_1 = \beta_1 - 1$$

Voltemos agora à função de custo de Feldstein:

$$C/y = \beta_0 + \beta_1 B + \beta_2 B^2 + \beta_3 F + \beta_4 F^2 + pY$$

Se, em vez de uma forma funcional linear nas variáveis e seus quadrados, tomarmos uma forma funcional linear nos logaritmos das variáveis, e se tomarmos como medida do output o número de casos corrigido pelo índice de Jensen e Morrisey, em vez do vector de case-mix, teremos:

$$\log C/y^* = \beta_0 + \beta_1 \log B + \beta_2 \log F^*$$

onde:

$$y^* = B \cdot \frac{Y^*}{B} = B \cdot F^*$$

Tomando na função de custo Cobb-Douglas o output de cada hospital medido pelo número de casos corrigidos pelo índice de Jensen e Morrisey, tendo em consideração a expressão anterior e afastando a hipótese de que as variáveis camas e fluxo de casos têm o mesmo parâmetro, para passar a admitir que estão afectadas de

parâmetros não necessariamente iguais, teremos uma função de custo Cobb-Douglas transformada, com o seguinte aspecto:

$$\log C/y^* = \beta_0 + \beta_1 \log B + \beta_2 \log F^* + \beta_3 \log w_K + \beta_4 \log w_L$$

Quer dizer, a diferença entre a função de custo Cobb-Douglas transformada e a função de custo "ad hoc" de Feldstein é, além da sua forma funcional e forma de medição do output, a inclusão na função de custo Cobb-Douglas do vector de preços dos inputs.

A amostra que utilizámos na estimação desta função de custo cobre 29 hospitais (6 Hospitais Centrais e 23 Hospitais Distritais) no período de 1984-1988. Porque tínhamos uma amostra com dados de tipo painel, e pretendíamos proceder à medição da eficiência, isto é, pretendíamos estimar uma função de fronteira, o método de estimação apropriado (e a especificação estocástica correspondente) é um dos propostos por Schmidt e Sickles (1984), ao qual nos referimos anteriormente. Considerámos em primeiro lugar o modelo de efeitos individuais e temporais fixos (não aleatórios), mas testámos também aquela especificação versus efeitos aleatórios (teste de Hausman e Taylor (1981)).

A especificação dos efeitos individuais e temporais como fixos pareceu-nos aconselhável por três tipos de argumentos. Primeiro, porque ao incluir-se como regressores variáveis de dimensão esperávamos que estivessem correlacionadas com os factores de ineficiência que estão incluídos nas perturbações do modelo de efeitos aleatórios, o que faz o estimador MQG inconsistente. Em

segundo lugar, porque os efeitos temporais incluíam dois tipos de efeitos: a deslocação da função de custo no tempo, que é habitualmente modelizado de uma forma determinística, e o preço do factor capital, considerado, em cada ano, idêntico para todos os centros. A razão da inclusão desta última parcela, está, por um lado no facto de tratarem-se só de hospitais do sector público, pelo que o preço do capital deveria ser idêntico para todos, e por outro lado, no facto de, ao incluir-se uma parcela que não varia entre indivíduos, só variando no tempo, esta não se poder distinguir dos factores temporais. Uma terceira razão é o facto da amostra quase coincidir com o universo dos hospitais (Centrais e Distritais) e não poder ser encarada como uma amostra aleatória de uma população mais vasta (Judge et al. (1988))

A especificação final que adoptámos para a nossa função de custo, parte também da identidade seguinte:

$$y^* = 365 \cdot B \cdot \frac{\text{estadias}}{365 \cdot B} \cdot \frac{1}{\frac{\text{estadias}}{y^*}} =$$

$$= 365 \cdot \text{camas} \cdot \text{índice de ocupação} \cdot \frac{1}{\text{Duração média}^*}$$

a incorporação desta equação na função de custo Cobb-Douglas, faz com que tenhamos a seguinte versão final da função de custo:

$$\log C = \beta_0 + \beta_1 \log B + \beta_2 \log INDOC - \beta_3 \log ESTMED + \beta_4 \log w_L$$

A estimativa obtida para a função de custo, foi:

VARIÁVEL	EST. VARI. ARTIF.	MQG
CONSTANTE	7.0638	6.6785
CAMAS	0.8345 (6.46)	1.1374 (34.94)
ÍNDICE OCUPAÇÃO	0.711 (4.27)	1.0067 (11.54)
DURAÇÃO MEDIA	- 0.2136 (-1.26)	- 0.2487 (-5.94)
SALÁRIO	0.2735 (2.18)	0.1057 (1.55)
R2	0.9927	0.9451
ESTAT. MULT. LAGR.		177.5
ESTAT. HAUS.E TAYLOR		47.5
N = 29		
T = 5		

Deve-se rejeitar a hipótese de que o estimador MQG é consistente, por a estatística de Hausman e Taylor assumir um valor que se situa claramente na região crítica. Situar-nos-emos no que se segue no contexto do estimador de variáveis artificiais. As estimativas têm os sinais e magnitudes esperados e só a variável duração média não é estatisticamente significativa. Este último facto pode ser a consequência de a duração média, no período da amostra, ter sido um dos indicadores de gestão associados ao financiamento dos centros, pelo que poderá ter baixado artificialmente o seu valor no período.

Com estas estimativas obtiveram-se os factores individuais, que depois de corrigidos, assumindo que a unidade mais eficiente é 100% eficiente, nos dão os factores de ineficiência seguintes:

INEFICIÊNCIA DE CUSTO

CENTRO	FACTOR
centro 1	-
centro 2	4.133
centro 3	2.836
centro 4	3.217
centro 5	2.927
centro 6	3.092
centro 7	-
centro 8	2.198
centro 9	-
centro 10	1.470
centro 11	-
centro 12	-
centro 13	1.319
centro 14	1.148
centro 15	-
centro 16	-
centro 17	1.671
centro 18	1.213
centro 19	-
centro 20	1.590
centro 21	1.335
centro 22	-
centro 23	-
centro 24	-
centro 25	1.356
centro 26	1.621
centro 27	1.225
centro 28	-
centro 29	1.517
centro 30	1.000
centro 31	1.305
centro 32	1.649
centro 33	-
centro 34	1.428
centro 35	1.104
centro 36	-
centro 37	1.706
centro 38	1.279

centro 39	1.274
centro 40	1.378
centro 41	-
centro 42	-
centro 43	-
centro 44	1.473
centro 45	1.093
centro 46	1.306

Média	1.719406
Desv. padrão	0.756038
Valor mínimo	1
Valor máximo	4.13298

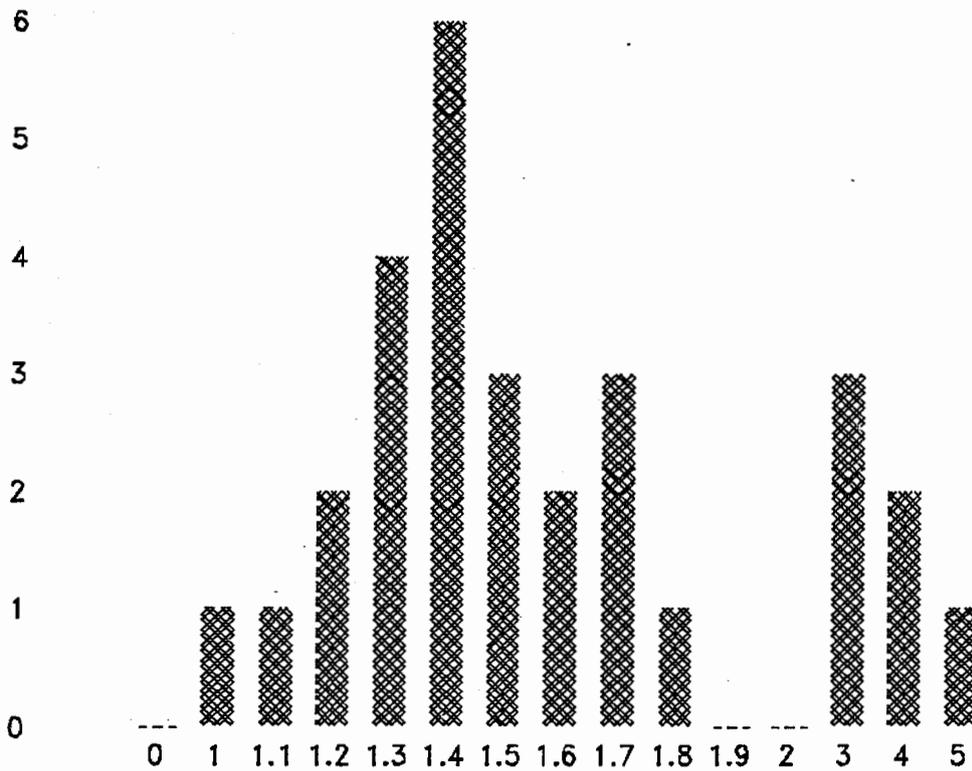


Figura 5.1.

Estes factores de ineficiencia ($\exp(\mu_i - \min(\mu_i))$) indicam que o centro mais eficiente é o hospital 30; os valores dos factores de ineficiência são excepcionalmente elevados para os

Hospitais Centrais (os primeiros 9 centros) e apresentam, além disso, grandes variações para os Hospitais Distritais. Estes factores de ineficiência que são factores que multiplicam a fronteira de custo para se obter os custos efectivamente observados, devem ser interpretados como a percentagem em que os custos de cada centro excedem os custos mínimos alcançáveis, dado um nível de output e um vector de preços dos inputs.

O deslocamento da função de custo média até à função de custo fronteira faz-se pressupondo que a unidade mais eficiente é 100% eficiente. Sobre este pressuposto, deveremos notar que, como vimos, pode haver no sistema hospitalar ineficiência estrutural; deste modo é possível que a unidade mais eficiente da amostra não seja 100% eficiente. Contudo, se associarmos à unidade mais eficiente o maior grau de eficiência alcançável num quadro institucional dado, admitindo que para além desta fronteira observável, há outra fronteira de custo absoluta e inobservável onde não há ineficiência, teremos justificada a utilização do conceito de função fronteira nesta situação,

Para a validação da especificação apresentada fizeram-se os seguintes testes:

1. Teste do multiplicador de Lagrange, de Breusch e Pagan (1980). A sua finalidade é testar:

$$H_0: \sigma_\mu^2=0 \text{ vs } H_1: \sigma_\mu^2 \neq 0$$

e a estatística de teste é:

$$LM = \frac{NT}{2(T-1)} \left[\frac{\tilde{u}'(I_N \otimes e_T e_T') \tilde{u}}{\tilde{u}'\tilde{u}} - 1 \right]^2$$

onde e_T é o vector coluna de dimensão T com todos os elementos iguais a 1. Esta estatística só inclui os resíduos dos mínimos quadrados da regressão da variável dependente sobre as variáveis independentes, com termo constante (\tilde{u}), e que tem uma distribuição assintótica do qui-quadrado com um grau de liberdade. O valor alcançado por esta estatística é de 177.5 pelo que se rejeita a hipótese nula em favor da alternativa de os efeitos individuais serem estatisticamente significativos.

2. A estatística de teste de Hausman e Taylor:

$$(\beta_B^w - \beta_B^{MQG})' (M_1 - M_0)^{-1} (\beta_B^w - \beta_B^{MQG})$$

alcança o valor de 47.5, (onde M_1 e M_0 são, respectivamente, as matrizes de variâncias e covariâncias dos estimadores de variáveis artificiais e MQG) pelo que rejeitaremos a hipótese nula de o estimador MQG ser diferente do estimador de variáveis artificiais apenas por flutuações amostrais (a estatística de teste tem, sob a hipótese nula, uma distribuição do qui-quadrado com tantos graus de liberdade quantos os regressores). Assim, o estimador MQG não é consistente.

Contudo, pode-se observar que os centros maiores e os Hospitais Centrais têm factores de ineficiência maiores. Isto

resulta dos factores que permanecem constantes no tempo, mas variam com o centro, como os factores relacionados com a dimensão, estarem incluídos nos factores individuais. Fez-se, portanto, a regressão dos efeitos individuais sobre a variável camas e uma variável artificial representando o tipo de centro.⁴

As estimativas obtidas com esta última regressão foram:

$$\mu = - 0.2574 + 0.00036435 B + 0.51808 DUM$$

(5.25) (2.39) (3.82)

$$N=29 , R^2=0.8516 , \sigma_e^2=0.0212 , \bar{B}=305.82$$

Obtiveram-se os resíduos de estimação, que foram normalizados com o mesmo critério que os efeitos individuais, para obtermos os factores de ineficiência líquidos dos efeitos dimensão e tipo de centro:

⁴ A docência nos centros hospitalares, além do tipo de centro (central ou distrital) não resultou estatisticamente significativa na regressão. Os hospitais parecem diferenciar-se em Hospitais Centrais e Distritais, mas os Hospitais Centrais não se diferenciam pelo seu estatuto de docência ou não docência.

INEFICIÊNCIA DE CUSTO LÍQUIDA

CENTRO	FACTOR
Centro 1	-
Centro 2	1.622
Centro 3	1.111
Centro 4	1.282
Centro 5	1.368
Centro 6	1.558
Centro 7	-
Centro 8	1.111
Centro 9	-
Centro 10	1.489
Centro 11	-
Centro 12	-
Centro 13	1.243
Centro 14	1.146
Centro 15	-
Centro 16	-
Centro 17	1.493
Centro 18	1.181
Centro 19	-
Centro 20	1.580
Centro 21	1.261
Centro 22	-
Centro 23	-
Centro 24	-
Centro 25	1.240
Centro 26	1.505
Centro 27	1.228
Centro 28	-
Centro 29	1.486
Centro 30	1.000
Centro 31	1.313
Centro 32	1.691
Centro 33	-
Centro 34	1.423
Centro 35	1.114
Centro 36	-
Centro 37	1.635
Centro 38	1.326
Centro 39	1.304
Centro 40	1.414
Centro 41	-
Centro 42	-
Centro 43	-
Centro 44	1.454
Centro 45	1.077
Centro 46	1.173

Media	1.338872
Desv. padrão	0.184857
Mínimo	1
Máximo	1.691

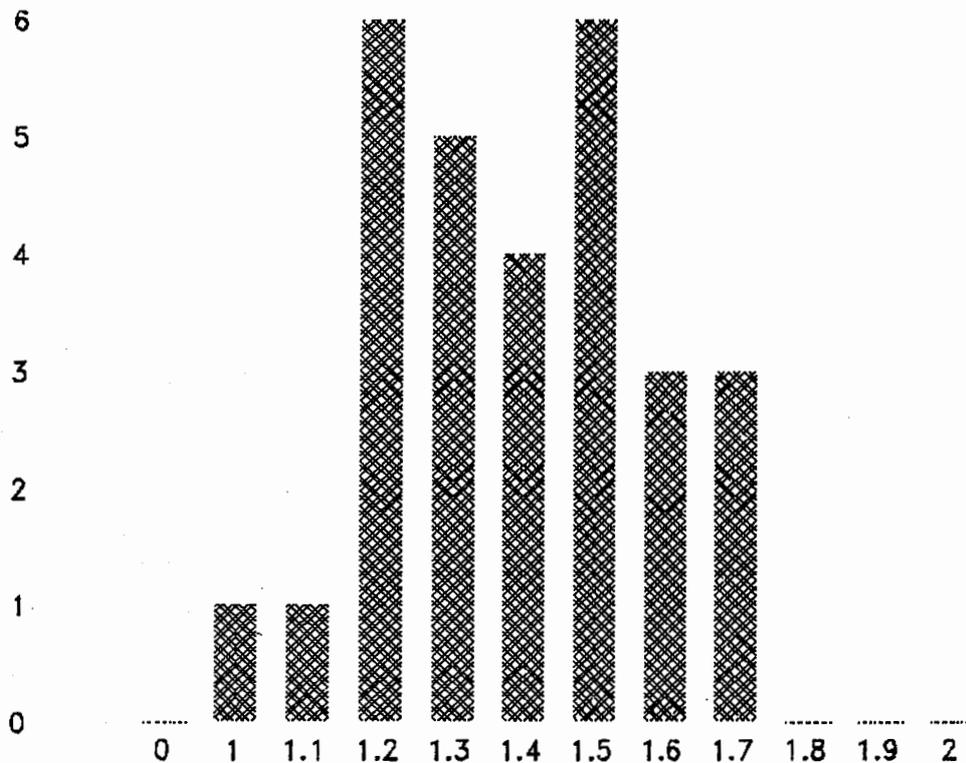


Figura 5.2.

Podemos observar que a maior parte das observações fica no intervalo $(1.1;1.5]$, pelo que os valores anómalos e a grande dispersão da ineficiência dos centros, desapareceram; permanece, no entanto, ainda uma elevada ineficiência de custo para a maior parte dos centros. A observação dos factores de ineficiência permite-nos concluir que o centro mais eficiente é o hospital 30 (com um factor de ineficiência igual à unidade) e o centro mais ineficiente é o centro 32, com um factor de ineficiência de 1.691.

Substituindo a expressão desta última regressão na função de custo, pode-se fazer a simulação dos custos médios, como função do número de camas, fixando no seu nível médio as restantes variáveis. Obtem-se uma curva de custo médio em U, com um mínimo cerca das 400 camas, um pouco acima da média amostral.

Um último resultado interessante que se pode ressaltar é que as despesas de pessoal representam 0.796 das despesas totais, enquanto que a elasticidade dos custos em relação ao salário é de 0.234. Estes valores, que pelo lema de Shephard se deveriam igualar, parecem indicar um excessivo peso dos salários no custo total.

5.3. - FUNÇÃO DE PRODUÇÃO HOSPITALAR E SEPARAÇÃO DA INEFICIÊNCIA DE CUSTO EM INEFICIÊNCIA TÉCNICA E DE AFECTAÇÃO

O estimador de Schmidt e Lovell (1979) (ver capítulo 3) da função de custo e o método de separação da ineficiência de custo em ineficiência técnica e ineficiência de afectação, baseia-se na estimação por máxima verosimilhança dos parâmetros da função de produção. Os dados utilizados são seccionais e a função de produção é do tipo Cobb-Douglas. Assume-se que se obtém a distribuição conjunta das perturbações da função de produção:

$$\xi_t = \log y_t - A - \sum_{i=1}^R \alpha_i \log x_{it}$$

e dos a_{it} :

$$a_{it} = \begin{pmatrix} \log x_{1t} - \log x_{2t} - B_{2t} \\ \dots\dots\dots \\ \log x_{1t} - \log x_{nt} - B_{nt} \end{pmatrix}$$

que, como função dos parâmetros a estimar, é a função de verosimilhança a maximizar.

Na estimação que fizemos, porque dispúnhamos de dados de tipo painel para as variáveis incluídas na função de produção e dados de tipo seccional das remunerações dos factores (variáveis incluídas nos B_{it}), optámos por estimar separadamente as elasticidades da função de produção e os a_{it} .

A estimação das elasticidades com painel de dados beneficia da inclusão de informação adicional (mais 4 anos por cada uma das observações de corte transversal), enquanto que a estimação dos a_{it} realiza-se com as elasticidades obtidas anteriormente e com dados de tipo seccional das remunerações e quantidades físicas dos factores.

A função de produção que estimámos é uma função de tipo Cobb-Douglas, tal como é especificada por Schmidt e Lovell, onde incorporámos como factores de produção o número de camas, como "proxy" do stock de capital, o número de médicos, o número de enfermeiras e o volume do pessoal auxiliar. O output é medido pelo

número de casos, normalizado pelo índice de Jensen e Morrisey.

O método de estimação utilizado foi o estimador de efeitos individuais e temporais fixos. Contudo, fizemos também o teste de Hausman e Taylor para a validação da opção efeitos individuais fixos versus efeitos aleatórios.

Obtivemos as seguintes estimativas da função de produção:

VARIÁVEIS	EST. VARI. ARTIF.	MQG
CONSTANTE	4.0268	4.5625
CAMAS	0.5059 (9.10)	0.6119 (13.68)
MEDICOS	0.0701 (1.71)	0.0600 (1.81)
EFERMEIRAS	0.1407 (2.53)	0.1466 (2.95)
PESSOAL AUXILIAR	0.1902 (2.48)	0.0272 (0.44)
R2	0.9944	0.9171
EST. HAUSMAN Y TAYLOR		51.88
N = 46		
T = 5		

A estatística de Hausman e Taylor alcança o valor de 51.88 pelo que rejeitaremos a hipótese nula de igualdade (além de flutuações amostrais) do estimador de variáveis artificiais e do estimador MQG; quer dizer, não se rejeita que o estimador MQG seja inconsistente.

O coeficiente de determinação é muito elevado: 0.9944, devido à presença de $(N-1)+(T-1)$ variáveis artificiais, correspondentes aos efeitos individuais e temporais. Os parâmetros têm a magnitude e sinal esperados e só a elasticidade da variável médicos não é estatisticamente significativa.

Obteve-se também, operando a transformação dos efeitos individuais, $EF_i = \lambda_i - \max \lambda_i$, a ineficiência técnica e^{EF_i} .

INEFICIENCIA TÉCNICA

CENTRO	FACTOR
Centro 1	0.5962
Centro 2	0.5684
Centro 3	0.6612
Centro 4	0.6054
Centro 5	0.5826
Centro 6	0.5862
Centro 7	0.2989
Centro 8	0.4913
Centro 9	0.3667
Centro 10	0.5287
Centro 11	0.4500
Centro 12	0.8533
Centro 13	0.7439
Centro 14	0.6111
Centro 15	0.5137
Centro 16	0.7664
Centro 17	0.6122
Centro 18	0.6985
Centro 19	0.7379
Centro 20	0.5243
Centro 21	0.7270
Centro 22	0.5462
Centro 23	0.8225
Centro 24	0.7169
Centro 25	0.6041
Centro 26	0.6052
Centro 27	0.8142
Centro 28	0.6838
Centro 29	0.8674
Centro 30	0.8499
Centro 31	0.8649

Centro	32	0.5976
Centro	33	1.0000
Centro	34	0.7221
Centro	35	0.6289
Centro	36	0.5440
Centro	37	0.5947
Centro	38	0.5798
Centro	39	0.5607
Centro	40	0.5696
Centro	41	0.9053
Centro	42	0.6342
Centro	43	0.4798
Centro	44	0.8443
Centro	45	0.8893
Centro	46	0.9003

Media 0.6576
 desv. padr. 0.147991
 Mínimo 0.2989
 Máximo 0.147991

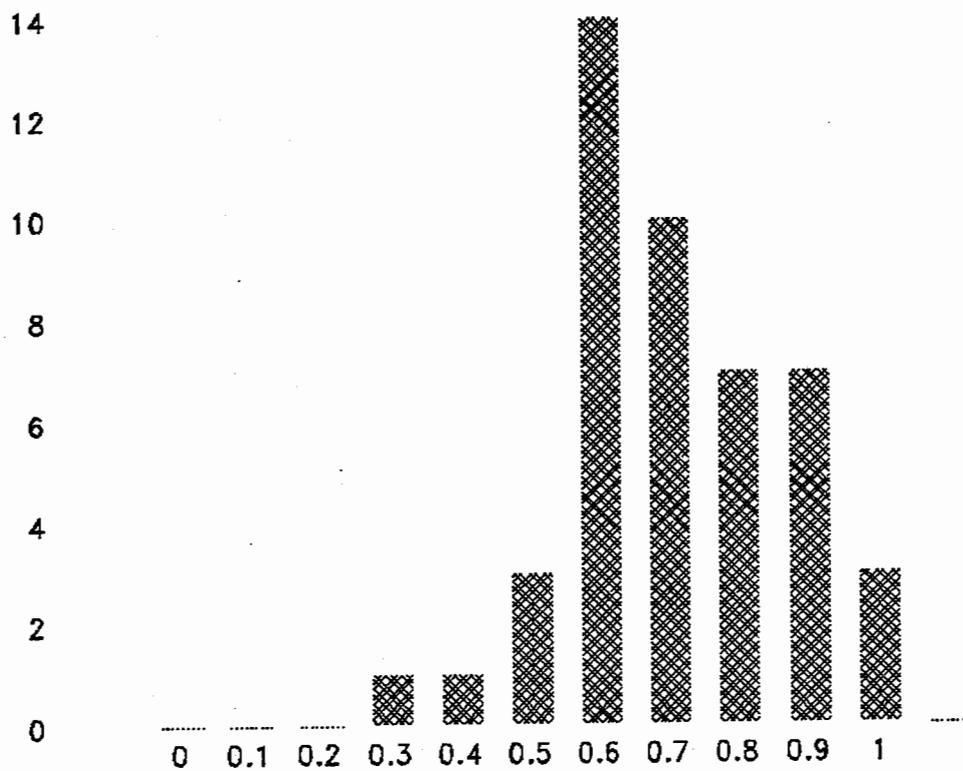


Figura 5.3.



Uma grande parte dos factores de ineficiência técnica fica no intervalo $[0.5;0.7)$, a que corresponde um grau de ineficiência técnica elevado.

Para a separação da ineficiência de custo em ineficiência técnica e ineficiência de afectação, é necessário o conhecimento das remunerações dos factores produtivos.

Enquanto que as remunerações correspondentes aos factores de trabalho foi possível serem obtidas através do Ministério da Saúde, sobre o factor capital o desconhecimento era considerável, visto que o próprio stock de capital tinha sido aproximado pelo número de camas.

Contudo, partindo da informação disponível de que:

- . o custo médio de um hospital (custo de construção e equipamento) é de aproximadamente 17 500 contos por cama nos hospitais construídos em 1991. (A fonte foi o Ministério da Saúde),
- . é possível tomar o período de vida útil média de 20 anos para este investimento.

Tomando uma taxa de actualização de 8%⁵, como custo de oportunidade do capital, foi possível chegar a um custo de aluguer de uma cama de hospital a preços de 1991. Este valor foi deflacionado para chegar a um custo a preços de 1988.

Com o conhecimento dos preços dos factores e das elasticidades da função de produção Cobb-Douglas obtivemos os rácios óptimos das despesas com cada um dos factores produtivos, que foram comparados com os rácios efectivamente observados. Estes rácios apresentam-se seguidamente no quadro:

RÁCIO	REAL: $\frac{p_i x_i}{p_j x_j}$	ÓPTIMO: $\frac{\alpha_i}{\alpha_j}$
Medico / cama	0.7365	0.1386
Medico / enfermeira	1.0658	0.4982
Medico / rest. pess.	1.1014	0.3686
Enferm./ cama	0.6911	0.2781
Enferm./ rest. pess.	1.0334	0.7398
Cama / rest. pess.	1.4953	2.6598

⁵ Ao assumirmos um período de vida útil e uma taxa de actualização idênticos aos que Lopez e Wagstaff (1988) tomam para os hospitais espanhóis, admitimos que as condições que prevalecem em Espanha não são muito diferentes das de Portugal.

Daqui podemos extrair as seguintes conclusões:

- . Os rácios das despesas médico/enfermeira e médico/restante pessoal são superiores ao óptimo, quer dizer: está-se a gastar excessivamente em médicos em relação às despesas com enfermeiras e restante pessoal.

- . Idênticamente o rácio enfermeiras/restante pessoal é superior ao óptimo, o que significa que o peso nas despesas de pessoal das despesas com restante pessoal deveria ser maior e inferior o peso das despesas com médicos. No que respeita a enfermeiras o sentido de variação não está definido.

- . Pelo contrário, parece que as despesas de pessoal são demasiado elevadas em relação às despesas de capital. Este resultado está de acordo com os obtidos com a estimação da função de custo.

- . Estes resultados concordam com os de Lopez e Wagstaff (1988), que trabalham dados espanhóis, no que respeita a reafectação implícita das despesas de pessoal e de capital. No que respeita ao rácio médicos/enfermeiros as conclusões vêm ao contrário, mas a diferença de base de dados e o elevado grau de colinearidade com que trabalham Lopez e Wagstaff na estimação da função de produção, faz com que as suas estimativas sejam pouco precisas.

Procedeu-se à separação da ineficiência de custo em ineficiência de afectação e técnica, com um programa baseado na metodologia de Schmidt e Lovell (1979); conseguiram-se os factores de custo da ineficiência de afectação e ineficiência técnica, que apresentamos seguidamente:

FACTORES DE CUSTO DA INEFICIÊNCIA DE AFECTAÇÃO

CENTRO	FACTOR
Centro 1	-
Centro 2	1.33162
Centro 3	1.25240
Centro 4	1.16052
Centro 5	1.29692
Centro 6	1.27139
Centro 7	-
Centro 8	1.17586
Centro 9	-
Centro 10	1.13056
Centro 11	1.24173
Centro 12	-
Centro 13	1.18993
Centro 14	1.03753
Centro 15	-
Centro 16	1.11612
Centro 17	1.15366
Centro 18	1.02094
Centro 19	1.27916
Centro 20	1.22984
Centro 21	1.07110
Centro 22	1.12487
Centro 23	1.08034
Centro 24	-
Centro 25	1.07348
Centro 26	1.17652
Centro 27	1.16935
Centro 28	1.01411
Centro 29	1.07588
Centro 30	1.01356
Centro 31	1.13062
Centro 32	1.19177
Centro 33	1.04678

Centro	34	1.09896
Centro	35	1.14748
Centro	36	1.17193
Centro	37	1.16296
Centro	38	1.12986
Centro	39	1.11382
Centro	40	1.16835
Centro	41	-
Centro	42	1.08753
Centro	43	1.30966
Centro	44	1.10511
Centro	45	1.07540
Centro	46	1.06665

Media	1.1460
Desv. Padr.	0.0829
Minimo	1.0136
Maximo	1.3316

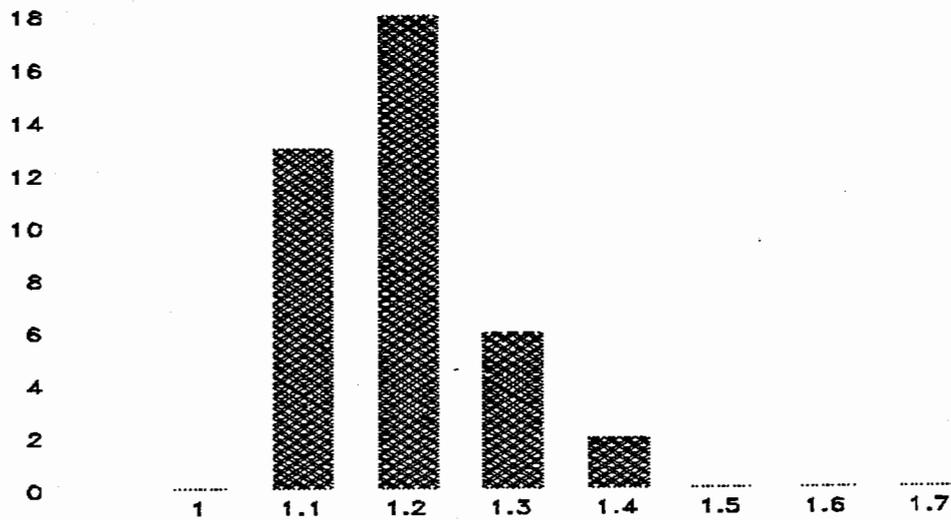


Figura 5.4.

Pode-se dizer que a ineficiência de afectação com um valor médio muito baixo e também uma baixa dispersão, não é a principal origem do facto dos custos observados excederem os custos mínimos alcançáveis.

FACTORES DE CUSTO DA INEFICIÊNCIA TÉCNICA

CENTRO	FACTOR
Centro 1	1.76862
Centro 2	1.86417
Centro 3	1.57800
Centro 4	1.73897
Centro 5	1.81425
Centro 6	1.80187
Centro 7	3.78741
Centro 8	2.18929
Centro 9	3.02300
Centro 10	2.01930
Centro 11	2.41215
Centro 12	1.19108
Centro 13	1.38571
Centro 14	1.72131
Centro 15	2.08434
Centro 16	1.34099
Centro 17	1.71786
Centro 18	1.48544
Centro 19	1.39816
Centro 20	2.03784
Centro 21	1.42118
Centro 22	1.94796
Centro 23	1.24039
Centro 24	1.44332
Centro 25	1.74315
Centro 26	1.73968
Centro 27	1.25434
Centro 28	1.52059
Centro 29	1.16984
Centro 30	1.19640
Centro 31	1.17357
Centro 32	1.76423
Centro 33	0.99999
Centro 34	1.43180
Centro 35	1.66744
Centro 36	1.95675

Centro	37	1.77362
Centro	38	1.82403
Centro	39	1.89249
Centro	40	1.85986
Centro	41	1.11592
Centro	42	1.65220
Centro	43	2.24748
Centro	44	1.20521
Centro	45	1.29813
Centro	46	1.12281

Media 1.696133
 Desv Padr. 0.499453
 Mínimo 0.999999
 Máximo 3.78741

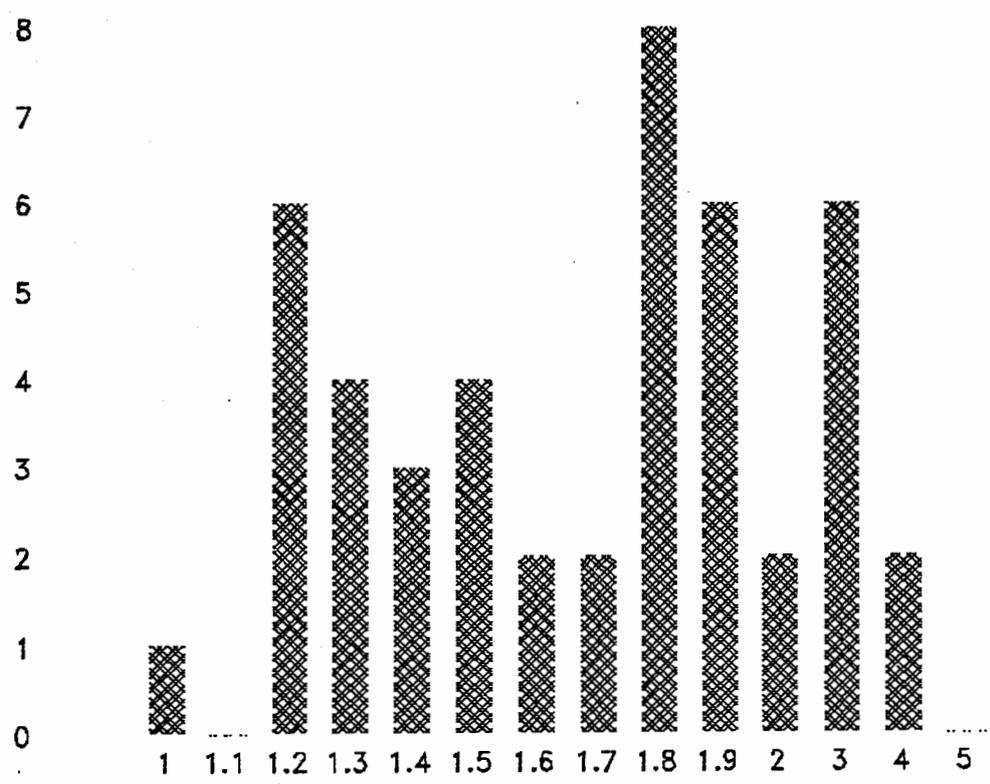


Figura 5.5.

Uma primeira característica é a muito maior variabilidade do factor de custo da ineficiência técnica do que o factor de custo da ineficiência de afectação. Isto significa que a origem dos custos efectivamente observados excederem os custos mínimos alcançáveis é principalmente devido à ineficiência técnica. Quer dizer, a ineficiência de afectação é pequena quando comparada com outras origens da ineficiência.

Esta situação levou-nos , como no caso da função de custo da secção anterior, a procurar os factores que explicam as diferenças de eficiência técnica entre centros. Consideraram-se dois factores: a dimensão dos centros medida pelo número de camas (x1), e duração média (x2). Considerou-se também uma variável artificial representando o tipo de centro (Hospital Central ou Distrital). A regressão obtida foi:

$$\mu = 0.54825 + 0.00050385 B - 0.071577 \text{ ESTMED} - 0.31621 \text{ DUM}$$

(3.89)	(3.34)	(-4.20)	(-2.41)
--------	--------	---------	---------

$$N=39 , R^2=0.4178 , \sigma_e^2=0.1537$$

os factores de ineficiência técnica líquidos de aqueles efeitos, foram:

FACTORES DE CUSTO DA INEFICIÊNCIA TÉCNICA LÍQUIDA

CENTRO	FACTOR
Centro 1	-
Centro 2	1.59485
Centro 3	1.34081
Centro 4	1.70860
Centro 5	1.45372
Centro 6	1.43003
Centro 7	-
Centro 8	1.24581
Centro 9	-
Centro 10	1.64550
Centro 11	1.78875
Centro 12	-
Centro 13	1.51710
Centro 14	1.54537
Centro 15	-
Centro 16	1.34884
Centro 17	1.79390
Centro 18	1.36561
Centro 19	1.64105
Centro 20	1.88214
Centro 21	1.29517
Centro 22	1.75482
Centro 23	1.07757
Centro 24	-
Centro 25	1.54958
Centro 26	1.41014
Centro 27	1.31539
Centro 28	1.13147
Centro 29	1.19056
Centro 30	1.22856
Centro 31	1.15862
Centro 32	1.74390
Centro 33	1.00001
Centro 34	1.50772
Centro 35	1.35777
Centro 36	1.78396
Centro 37	1.58539
Centro 38	1.46664
Centro 39	1.60763
Centro 40	1.79464
Centro 41	-
Centro 42	1.42012
Centro 43	1.90447
Centro 44	1.31110
Centro 45	1.32305
Centro 46	1.22470

Media	1.47295
Desv padr.	0.232201
Minimo	1.00001
Maximo	1.90447

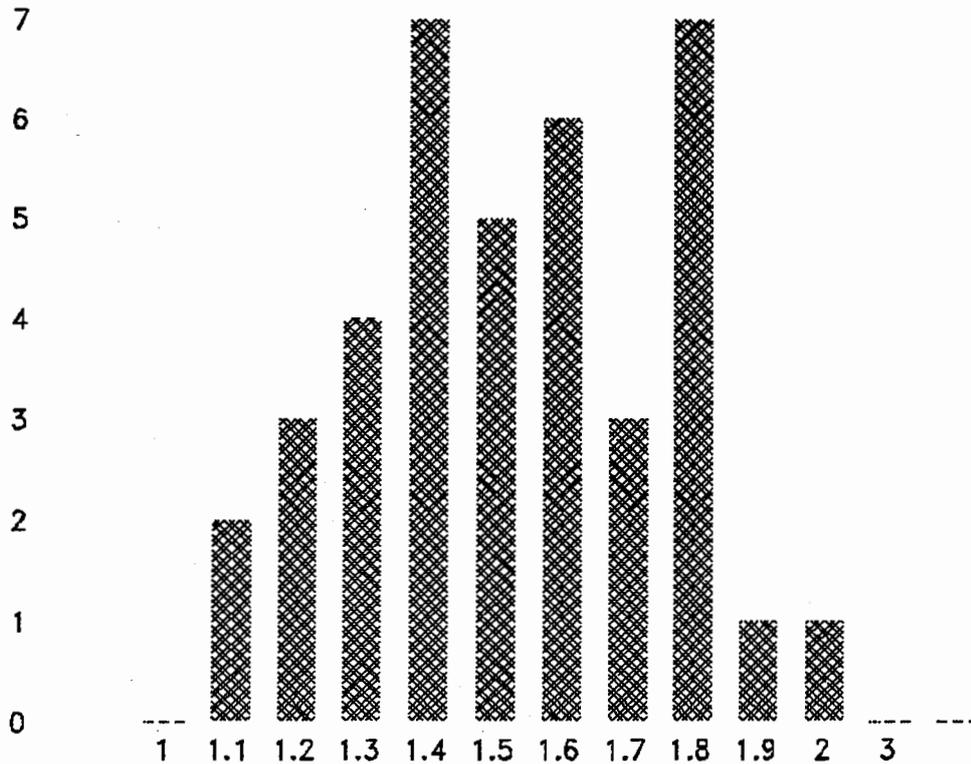


Figura 5.6.

Os factores de ineficiência técnica líquidos da influência dos factores dimensão e duração média, têm muito menor variabilidade que o factor de ineficiência técnica incluindo aquelas influências, mas apesar disso a ineficiência técnica líquida é, em termos absolutos, elevada, podendo-se afirmar que os custos efectivamente observados excedem os mínimos alcançáveis, em média, cerca de 47%,

devido à ineficiência técnica.

Sobre os resultados desta regressão devemos fazer as seguintes considerações:

- . As observações da variável dependente são os factores individuais obtidos aquando da estimação da função de produção. A eficiência técnica é: $EFT_i = \exp [RES_i - \max (RES_j)]$ com $\max (RES_j) = 0.4461$. $0 < EFT_i \leq 1$ é um valor tanto maior quanto maior for a eficiência técnica e portanto os efeitos individuais são tanto maiores quanto maior for eficiência técnica. Assim:
- . devemos explicar o sinal positivo da variável número de camas pelas possíveis economias de escala nos efeitos associados à dimensão dos centros,
- . o sinal negativo da variável duração média é explicável por uma maior complexidade dos casos associados a uma maior duração média, e
- . a maior complexidade dos casos associada aos Hospitais Centrais pode explicar também o sinal negativo da variável artificial correspondente.
- . o baixo R^2 associado a esta regressão poderá ser explicado pelo elevado nível de ineficiência técnica líquida subsistente

5.4. - ESTIMAÇÃO DE UMA FORMA FUNCIONAL FLEXÍVEL: A TRANSLOG

É conhecido que a função de produção translog é uma aproximação pontual exacta até à segunda ordem de diferenciação da verdadeira função de produção. Era, então, importante na caracterização da tecnologia da unidade hospitalar média (admite-se que é no ponto correspondente à média amostral que a translog aproxima exactamente a verdadeira função de produção) a estimação da translog.

Contudo, há alguns problemas na estimação da translog:

1. Na forma original da translog estão envolvidos os logaritmos das quantidades de inputs consumidas, dos seus quadrados e os produtos cruzados daquelas variáveis logaritmizadas:

$$\log y = \gamma_0 + \sum_i \gamma_i \log X_i + \frac{1}{2} \sum_i \sum_j \gamma_{ij} \log X_i \log X_j$$

Esta forma funcional origina habitualmente uma elevada colinearidade entre os regressores, quando se tenta a sua estimação directa.

A forma habitual de estimação da translog é através das equações de participação factorial ("factor cost shares"): derivando a translog em relação a $\log X_i$ e sob a hipótese de maximização do lucro, teremos o primeiro membro da equação:

$$\frac{\partial \log y}{\partial \log X_i} = \frac{P_i X_i}{C_i}$$

que é a participação do factor X_i na despesa total.

Se em vez da função de produção tivéssemos a função de custo:

$$\log C = \beta_0 + \sum_i \beta_i \log w_i + \frac{1}{2} \sum_i \sum_j \beta_{ij} \log w_i \log w_j$$

(pode-se provar que a função de custo, como a função de produção, é uma aproximação pontual exacta até à segunda ordem de diferenciação da verdadeira função de custo), e se derivamos esta função em relação ao logaritmo do preço do factor X_i , o primeiro membro da equação é, sob a hipótese de minimização dos custos e devido ao lema de Shephard $\frac{P_i X_i}{C_i}$, ou a participação do factor X_i na despesa total.

Estimámos, seguidamente, a função de custo ou mais concretamente as suas equações de participação factorial.

2. Se a utilização de cada factor se afasta da utilização óptima, ou seja, se existe ineficiência de afectação sistemática, este facto tem como consequência a existência de perturbações aleatórias de média não nula nas equações de participação factorial. Vimos anteriormente, que apesar do seu valor em termos de custo não ser muito importante, existia no sector hospitalar ineficiência de afectação sistemática.

A consequência da estimação das equações de participação factorial sob a hipótese de perturbações de média nula, quando, de facto, as perturbações não têm média nula, é que em vez de estimarmos os termos independentes das equações de participação factorial, estimamos esses termos mais a correspondente média das perturbações, ou seja: estes termos são estimados com enviesamento.

Como é impossível estimar a função de custo na sua forma original devido aos problemas de colinearidade, estimámos as equações de participação factorial, obtidas por derivação da função de custo. Para tal, tivémos em conta:

a) que há enviesamento na estimação dos termos independentes das equações de participação factorial.

b) as restrições sob as quais estimámos a função de custo, ou seja, rendimentos de escala constantes e simetria das segundas derivadas parciais do logaritmo da função de custo.

Os factores produtivos considerados foram como na especificação da função de produção Cobb-Douglas, médicos, enfermeiras, restante pessoal e capital, aproximado pelo número de camas; então a função de custo translog vem dada por:

$$\begin{aligned}
\log C = & \beta_0 + \beta_k \log w_k + \beta_M \log w_M + \beta_N \log w_N + \beta_{RP} \log w_{RP} + \\
& + \frac{1}{2} \beta_{kk} \log^2 w_k + \beta_{kM} \log w_k \log w_M + \beta_{kN} \log w_k \log w_N + \\
& + \beta_{kRP} \log w_k \log w_{RP} + \frac{1}{2} \beta_{MM} \log^2 w_M + \beta_{MN} \log w_M \log w_N + \\
& + \beta_{MRP} \log w_M \log w_{RP} + \frac{1}{2} \beta_{NN} \log^2 w_N + \beta_{NRP} \log w_N \log w_{RP} + \\
& + \frac{1}{2} \beta_{RPRP} \log^2 w_{RP}
\end{aligned}$$

Diferenciando em relação aos logaritmos dos preços dos factores, obteremos:

$$S_k = \frac{p_k X_k}{C} = \frac{\partial \log C}{\partial \log w_k} = \beta_k + \beta_{kk} \log w_k + \beta_{kM} \log w_M + \beta_{kN} \log w_N + \beta_{kRP} \log w_{RP}$$

$$S_M = \beta_M + \beta_{kM} \log w_k + \beta_{MM} \log w_M + \beta_{MN} \log w_N + \beta_{MRP} \log w_{RP}$$

$$S_N = \beta_N + \beta_{kN} \log w_k + \beta_{MN} \log w_M + \beta_{NN} \log w_N + \beta_{NRP} \log w_{RP}$$

$$S_{RP} = \beta_{RP} + \beta_{kRP} \log w_k + \beta_{MRP} \log w_M + \beta_{NRP} \log w_N + \beta_{RPRP} \log w_{RP}$$

A hipótese de rendimentos de escala constantes implica que $\beta_k + \beta_M + \beta_N + \beta_{RP} = 1$ e que $\sum_i \beta_{ij} = \sum_j \beta_{ij} = 0$, então as três últimas equações podem-se escrever:

$$S_M = \beta_M + \beta_{MM} \log(w_M/w_k) + \beta_{MN} \log(w_N/w_k) + \beta_{MRP} \log(w_{RP}/w_k)$$

$$S_N = \beta_N + \beta_{MN} \log(w_M/w_k) + \beta_{NN} \log(w_N/w_k) + \beta_{NRP} \log(w_{RP}/w_k)$$

$$S_{RP} = \beta_{RP} + \beta_{MRP} \log(w_M/w_k) + \beta_{NRP} \log(w_N/w_k) + \beta_{RPRP} \log(w_{RP}/w_k)$$

Como temos restrições (de simetria e de rendimentos de escala constantes) sobre os parâmetros, então só 9 dos coeficientes podem ser estimados independentemente.

Como as quatro participações factoriais somam a unidade é de esperar que as perturbações das equações de participação factorial estejam contemporaneamente correlacionadas.

O modelo apropriado neste caso é o SURE (Seemingly Unrelated Regression Equations) que se baseia na aplicação de mínimos quadrados generalizados àquelas equações. Como a matriz de variâncias e covariâncias contemporâneas das perturbações das equações (Σ) não é conhecida, utiliza-se MQGE onde esta matriz é estimada com os resíduos da estimação com MQO.

Pode-se além disso utilizar o estimador que consiste na iteração entre estimativas de β e Σ que no caso de perturbações com distribuição normal multivariada coincide com o estimador de máxima verosimilhança (Judge et al. (1984)).

As estimativas obtidas foram:

PARÂMETRO	VALOR	ESTATÍSTICA t
β_M	0.15322	5.3453
β_{MM}	- 0.03783	- 1.1712
β_{MN}	- 0.03963	- 2.2956
β_{MRP}	- 0.06152	- 1.9571
β_N	0.18667	7.0714
β_{NN}	0.06912	2.4490
β_{NRP}	- 0.03201	- 0.9033
β_{RP}	0.30054	4.7269
β_{RPRP}	0.11022	1.2456

N=39

Sobre as estimativas obtidas devemos fazer as seguintes considerações:

- . foi usada para a estimação uma rotina específica do TSP
- . todos os termos independentes são estatisticamente significativos.
- . três dos seis coeficientes dos regressores não são estatisticamente significativos e um está muito próximo do valor crítico de 1.96. Os restantes parâmetros são estatisticamente significativos

Obtiveram-se com os coeficientes dos regressores das equações de participação factorial as elasticidades de substituição de Allen, Morishima e sombra. Notemos em primeiro lugar que as elasticidades de substituição dependem só dos coeficientes dos regressores e não dos termos independentes. Como vimos anteriormente as estimativas dos termos independentes podem ser enviesadas; os coeficientes dos regressores, ainda que exista ineficiência de afectação sistemática, não temos razão para supor a sua estimação como sendo enviesada.

Definamos agora alguns conceitos utilizados em seguida:

. Os inputs i e j dizem-se substitutos no sentido de Allen se um aumento no preço do input i conduz a um aumento na utilização do input j . Pelo contrário, diz-se que os inputs i e j são complementares no sentido de Allen, se um aumento do preço de i conduz a um decréscimo na utilização de j .

. Os inputs i e j dizem-se substitutos no sentido de Morishima se um crescimento no preço do input j provoca um aumento no rácio x_i/x_j . Serão complementares no sentido de Morishima os inputs onde aquela variação for em sentido inverso.

. As elasticidades de substituição de Allen e Morishima podem ser expressas em termos das elasticidades cruzadas procura-preço:

$$\epsilon_{ij} = \frac{\partial x_i}{\partial w_j} \frac{w_j}{x_i}$$

Assim, a elasticidade de substituição de Allen é: $\sigma_{ij} = e_{ij} / S_j$.
 Pode-se provar, além disso, que:

$$\sigma_{ij} = \frac{\beta_{ij} + S_i S_j}{S_i S_j}$$

$$\sigma_{ii} = \frac{\beta_{ii} + S_i^2 - S_i}{S_i^2}$$

Onde S_i é a participação do factor i na despesa total e β_{ij} é o coeficiente do produto cruzado dos logaritmos de w_i e w_j na função de custo translog.

A elasticidade de substituição de Morishima é:

$$\sigma_{ij}^M = e_{ij} - e_{jj} = \frac{\partial \log(x_i/x_j)}{\partial \log w_j}$$

A elasticidade de substituição sombra é uma média ponderada das elasticidades σ_{ij}^M e σ_{ji}^M . É:

$$\sigma_{ij}^s = \frac{S_i}{S_i + S_j} \sigma_{ij}^M + \frac{S_j}{S_i + S_j} \sigma_{ji}^M$$

Com estas hipóteses estimámos:

ELASTICIDADES DE SUBSTITUIÇÃO DE ALLEN

-5.15247	-0.12002	-0.79698	2.60037
-0.12002	-2.40647	0.00361	1.03093
-0.79698	0.00361	-1.12601	0.78831
2.60037	1.03093	0.78831	-1.86025

ELASTICIDADES DE SUBSTITUIÇÃO DE MORISHIMA

0.00000	0.41659	0.05801	1.99479
0.97730	0.00000	0.19915	1.29294
0.84584	0.43912	0.00000	1.18444
1.50560	0.62629	0.33749	0.00000

ELASTICIDADES DE SUBSTITUIÇÃO SOMBRA

0.00000	0.68801	0.43289	1.65372
0.68801	0.00000	0.31716	0.81927
0.43289	0.31716	0.00000	0.57697
1.65372	0.81927	0.57697	0.00000

Algumas observações sobre as estimativas obtidas:

. O factor médicos é complementar no sentido de Allen dos factores enfermeiras e restante pessoal.

. A elasticidade de substituição de Allen entre enfermeiras e restante pessoal é positiva, se bem que, provavelmente não seja estatisticamente diferente de zero.

. As elasticidades de Morishima e sombra são todas positivas.

Os resultados de Salkever e outros, citados por Lopez e Wagstaff segundo os quais algumas tarefas executadas por médicos poderiam sê-lo por enfermeiras e a hipótese de algumas tarefas executadas por enfermeiras também poderiam sê-lo por pessoal com menor qualificação, não são suficientemente confirmados pela

evidência empírica recolhida aqui, por serem estatisticamente não significativas algumas elasticidades de substituição, ou por terem sinais opostos os diferentes conceitos de elasticidade de substituição, para o mesmo par de inputs.

CAPÍTULO 6

CONCLUSÕES

Poderemos resumir as principais conclusões deste trabalho nos seguintes pontos:

1. As especificações da função de custo como dual de uma função de produção e como função de custo "ad hoc", têm algumas semelhanças. Assim, a função de custo "ad hoc" (tomamos para exemplo a função de custo de Feldstein, mas uma argumentação semelhante poderia ser usada com outras funções de custo do mesmo tipo) distingue-se de uma função de custo Cobb-Douglas onde a variável dimensão número de casos, é decomposta nas variáveis de dimensão, camas e fluxo de casos (assumindo coeficientes não necessariamente iguais para estas duas variáveis), pelas seguintes características:

. forma funcional - a função de custo transformada inclui sob uma forma aditiva os logaritmos das variáveis camas e fluxo de casos, enquanto a função de custo de Feldstein inclui sob uma forma aditiva as variáveis camas, fluxos de casos e respectivos quadrados.

. a função de custo Cobb-Douglas transformada inclui o vector dos logaritmos dos preços dos factores produtivos, vector que não é incluído na especificação "ad hoc". A função de custo de Feldstein é uma função de custo que tem como explicativas só variáveis de dimensão (as mesmas que a função de custo de Cobb-

Douglas transformada, mas com uma forma funcional diferente) mas não considera os preços dos factores produtivos.

2. Com a função de custo Cobb-Douglas transformada obtivemos a eficiência de custo dos diferentes centros. Além de um elevado valor médio da ineficiência (1.72), encontramos uma grande dispersão daqueles factores. Um dos motivos deste facto é o elevado grau de ineficiência dos Hospitais Centrais. Como variáveis explicativas destas variações, encontrou-se um factor de dimensão, o número de camas e uma variável artificial representando o tipo de centro (Hospital Central ou Distrital). Obtiveram-se os factores de ineficiência líquidos da influência daquelas variáveis. Estes factores líquidos têm um valor médio claramente inferior ao anterior (1.34 versus 1.72), mas em termos absolutos ainda elevado: em média o custo observado excede em 34% o custo mínimo alcançável, dado o nível de output e os preços dos factores.

3. Estimamos um valor mínimo para o custo por caso no quartil cuja dimensão média é de 214 camas. Este valor está associado também a um mínimo da variável duração média e a um valor próximo do mínimo do índice de Jensen e Morrisey. Deve-se esperar então que, quando corrigimos aquela relação tendo em conta a variação das outras variáveis que influem sobre o comportamento do custo médio (duração média e índice de Jensen e Morrisey, por exemplo), teremos um mínimo do custo médio para dimensões superiores àquelas.

4. Juntando à função de custo Cobb-Douglas transformada a

relação entre os efeitos individuais e as variáveis camas e a variável artificial representando o tipo de centro, obtemos o custo médio como função do número de camas e seu logaritmo. Simulando o custo médio como função do número de camas, obtemos uma curva em U com um mínimo um pouco acima da média amostral de 363 camas. Contudo, se temos em conta que alguns dos últimos hospitais construídos em Portugal têm uma dimensão muito superior à referida média amostral, podemos afirmar que se está seguindo uma política de sobredimensionamento das unidades hospitalares.

5. A ineficiência de custo calculada inclui, além da ineficiência técnica, a ineficiência de afectação. Foi nossa intenção separar a ineficiência de custo naqueles dois factores. Esta separação recorre à dualidade entre a função de custo e a função de produção e exige que se estime esta última função. Com esta estimação foi possível comparar as relações entre as despesas com cada um dos factores produtivos, com as relações entre as despesas óptimas com aqueles factores. Assim foi possível estabelecer que se está a gastar excessivamente com cada uma das categorias de pessoal em relação às despesas com o factor capital (camas). Sendo plausível esta estimativa, as despesas de pessoal deveriam reduzir-se (tomando como fixos custo de aluguer de uma cama de hospital e, no curto prazo o número de camas). Os rácios de despesas médico/enfermeiro e médico/restante pessoal indicam que há um peso excessivo das despesas com pessoal médico no conjunto das despesas de pessoal. Pelo contrário a categoria restante pessoal parece ter uma participação nas despesas de pessoal inferior ao

aconselhável. O sentido da variação do peso das despesas com enfermeiros no conjunto das despesas de pessoal não é conclusivo, por ser aconselhável que diminuam os rácios das despesas médicos/enfermeiras e enfermeiras/restante pessoal.

6. Obteve-se a separação da ineficiência de custo em ineficiência técnica e ineficiência de afectação. Os factores de custo da ineficiência de afectação (com uma média de 1.15) parecem ser responsáveis em muito menor grau que os factores de custo da ineficiência técnica (com uma média de 1.70) pelo facto de os custos excederem os mínimos alcançáveis. Para explicar a grande variabilidade da eficiência técnica entre centros fizemos a regressão dos factores de ineficiência técnica sobre um conjunto de variáveis, tendo resultado significativas: o factor dimensão (camas), duração média, e uma variável artificial representando o tipo de centro. Quando estes factores são tidos em consideração, a média dos factores de custo da ineficiência técnica reduz-se a 1.47. Este valor da média dos factores de ineficiência técnica líquidos dos factores dimensão (camas), duração média e tipo de centro é ainda muito elevado, pelo que é de admitir que persista no sistema hospitalar português um elevado grau de ineficiência técnica.

7. Uma possível reafectação dos factores produtivos, no sentido de uma maior eficiência, deverá ter em conta o sinal e a magnitude das elasticidades de substituição entre factores produtivos. Estas elasticidades foram estimadas com uma função de

custo translog. As elasticidades de Allen de médicos/enfermeiras e médicos/restante pessoal têm sinal negativo e são calculadas com parâmetros estatisticamente significativos. Estes factores são, assim, complementares no sentido de Allen. As restantes elasticidades de substituição entre factores são obtidas com as restrições sobre os parâmetros da translog ou não são estatisticamente significativas. As elasticidades de substituição de Morishima e sombra são positivas pelo que aqueles factores são substitutos sob aqueles critérios.

Bibliografia

- . Aigner, D.J. e Chu, S.F. (1968), "On Estimating the Industry Production Function". American Economic Review, Vol.58, 226-239
- . Arellano, M. e Bover, O. (1990), "La Econometria de Datos de Panel". Investigaciones Económicas, Vol.XIV, Nº1, 3-45
- . Bays, C.W. (1977), "Case-Mix Differencies between Non-Profit and for Profit Hospitals". Inquiry, 14(1), 17-21
- . Berndt, E.R. e Christensen, L.R. (1973), "The translog Function and the Substitution of Equipment, Structure and Labor in U.S. Manufacturing 1929-68". Journal of Econometrics, Nº1, 81-114
- . Breusch, T.S. e Pagan, A.R. (1980), " The Lagrange Multiplier Test and its Applications to the Model Specification in Econometrics". Review of Economic Studies, Vol.47, 239-253
- . Buchanan, J.M. The Inconsistencies of the National Health Service, (1965), Institute of Economic Affairs, London
- . Campos, A. (1983), Saude - O Custo de um Valor sem Preço. Editora Portuguesa de Livros Científicos, Lisboa
- . Campos, A. (1990) "A Hora da Reforma: Linhas Gerais para a Revisão do Serviço Nacional de Saúde em Portugal". Farmácia Portuguesa, Nº63, 45-49.
- . Campos, A. (1991(a)) "Estado-Providência. Perspectivas e Financiamento. O Caso da Saúde". Sociologia - Problemas e Práticas, Nº9, 9-26.
- . Campos, A. (1991(b)), "Public Needs, Private Services. The NHS and the Public-Private Mix in Portugal". Health Economics, Vol. 9, Nº1, 5-8.
- . Campos, A., Caldeira da Silva, Costa, C. (1990) "Market Elements and New Developments for Competition in the Health Care System in Portugal" in Casparie, A.F., Hermans, H., Paelink, J. Competitive Health Care in Europe - Future Prospects, Vermont, 1990
- . Caves, D.W., Christensen, L.R. e Diewert, W.E. (1982), " The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity". Econometrica, 50/6, 1393-1414

- . Chambers, R.G. (1988), Applied Production Analysis, Cambridge University Press, Cambridge
- . Cohen, H.A. (1967), "Variations in Cost among Hospitals of Different Sizes". Southern Economic Journal, Vol.33, 355-366
- . Cohen, H.A. (1970), "Hospital Cost Curves with Emphasis on Measuring Patient Care Output" in Empirical Studies in Health Economics (Editado por Klarman, H.E.), 279-294, John Hopkins Press, Baltimore
- . Conrad, R.F. e Strauss, R.P. (1983), "A Multiple-Output Multiple-Input Model of the Hospital Industry in North Carolina". Applied Economics, Vol.15, 341-352
- . Culyer, A., Lavers, R.J. e Williams, A. (1971), "Social Indicators: Health". Social Trends, Vol.2 , 21-41
- . Deprins, D. e Simar, L. (1989(a)) "Estimating Technical Ineficiências with Correction for Environmental Conditions with an Application to Railway Companies". Annals of Public and Cooperative Economics, Vol.60, Nº1, 81-102
- . Deprins, D. e Simar, L. (1989(b)), "Estimation de Frontières Déterministes avec Facteurs Exogènes d'Inefficacité". Cahiers du SMASH 8801, Facultés Universitaires Saint-Louis, Bruxelles
- . Deprins, D., Simar, L. e Tulkens, H. (1984), "Measuring Labor Inefficiency in Post Offices". In Marchand, M., Pestieu, P. e Tulkens, H. The Performance of Public Enterprises: Concepts and Measurement, Amsterdam, North-Holland
- . Diewert, W.E. (1976)" Exact and Superlative Index Numbers". Journal of Econometrics, 115-145
- . Divisia, F.(1926), L'Indice Monétaire et la Théorie de la Monnaie, Paris, Sirey
- . "DRGs in Portugal"- Relatório sobre a aplicação dos DRGs no financiamento dos hospitais portugueses.
- . Evans, R.G. e Walker, M.D. (1972), "Information Theory and the Analysis of Hospital Cost Structure". Canadian Journal of Economics, Vol.5, 398-418
- . Farrell, M.J. (1957), "The Measurement of Productive Efficiency". Journal of the Royal Statistical Society, Vol.120, 253-281
- . Feldstein, M.S.(1967), Economic Analysis for Health Service Efficiency: Econometric Studies of the British National Health Service, North-Holland, Amsterdam,

- . Fetter, R.B., Shin, Y., Freeman, J.L., Averill, R.F. e Thompson, J.D. (1980), "Case-mix Definition by Diagnostic-Related Groups". Medical Care, Vol.18, Suppl.38
- . Försund, F.R., Lovell, C.A.K. e Schmidt, P. (1980), "A Survey of Frontier Production Functions and their Relationship to Efficiency Measurement". Journal of Econometrics, Vol.13, 5-25
- . Fuller, W.A. e Battese, G.E. (1973), " Transformations for Estimation of Linear Models with Nested Error Structure". Journal of the American Statistical Association, Vol.68, 626-632
- . Fuller, W.A. e Battese, G.E. (1974), "Estimation of Linear Models with Crossed-Error Structure". Journal of Econometrics, N°2, 67-78
- . Gathon, H.J. (1989), "Indicators of Partial Productivity and Technical Efficiency in the European Urban Transit Sector". Annals of Public and Cooperative Economics, Vol.60, N°1, 43-60
- . Green, W.H. (1980), "Maximum Likelihood Estimation of Econometric Frontier Functions". Journal of Econometrics, Vol.13, 27-56
- . Harris, J.E. (1977), "The Internal Organization of Hospitals: Some Economic Implications". The Bell Journal of Economics, Vol.8, N°2, 467-482
- . Hausman, J.A. e Taylor, W.E. (1981), "Panel Data and Unobservable Individual Effects". Econometrica, Vol.49, 1377-1398
- . Hsiao, C. (1986) Analysis of Panel Data, Cambridge University Press, Cambridge
- . Jenkins, A.W. (1980), "Multiproduct Cost Analysis: Service and Case-Type Cost Equations for Ontario Hospitals". Applied Economics, Vol.12, 103-113
- . Jensen, G.A. e Morrisey, M.A. (1986(a)) "Medical Staff Speciality Mix and Hospital Production". Journal of Health Economics, Vol.5, 253-276
- . Jensen, G.A. e Morrisey, M.A. (1986(b)) "The Role of Physicians in Hospital Production". Review of Economics and Statistics, Vol. 68, 432-442
- . Johnston, J. (1987), Econometric Methods, McGraw-Hill, New York
- . Jondrow, J., Lovell, C.A.K., Materov, I. e Schmidt, P. (1982), "On the Estimation of Technical Ineficiency in the Stochastic Frontier Production Function Model". Journal of Econometrics, Vol.19, 233-238
- . Judge, G.G., Hill, R.C., Griffiths, W.E., Lütkepohl, H. e Lee, T.C. (1985), Theory and Practice of Econometrics, Wiley, New York

. Judge, G.G., Hill, R.C., Griffiths, W.E., Lütkepohl, H. e Lee, T.C. (1988), Introduction to the Theory and Practice of Econometrics, Wiley, New York

. Kalirajan, K.P. e Shand, R.T. (1989), "A Generalized Measurem of Technical Efficiency". Applied Economics, Vol.21, 25-34

. Leu, R.R. (1986), "The Public-Private Mix and International Health Care Costs". In Culyer, A.J. e Jönsson, B. (Edts) Public and Private Health Services: Complementarities and Conflicts, Basil Blackwell, Oxford, England

. Lopez Casanovas, G., Wagstaff, A. (1988), "La Combinación de los Factores Productivos en el Hospital: Una Aproximación a la Función de Producción". Investigaciones Económicas (Segunda época), Vol.XII, nº2, 305-327

. Maddala, G.S. (1971), "The Use of Variance Components in Pooling Cross Section and Time Series Data". Econometrica, Vol.39, 341-358

. Mantas, A., Costa, C., Ramos, F. (1989), "Financiamento Hospitalar: que Contributo para a Equidade? O Caso Português". Associação Portuguesa de Economia da Saúde, Documento de Trabalho Nº3/89, Lisboa.

. McGuire, A. (1987), "The measurement of Hospital Efficiency". Social Science and Medicine, Vol. 24, 719-724

. Mundlak, Y. (1978), "On the Pooling of Time Series and Cross Section Data". Econometrica, Vol.49, 1399-1416

. Newhouse, J.P. (1970), "Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital". The American Economic Review, Vol. 60, Nº1, 64-74.

. OECD (1990) Health Care Systems in Transition - The Search for Efficiency, OECD, Social Policy Studies Nº7, Paris

. Pauly, M.V. e Redisch, M. (1973), "The Not-for-Profit Hospital as a Physician's Cooperative". The American Economic Review, Vol. 63, Nº1, 87-100

. Poullier, J.P. (1990), "El Afan por la Eficiencia - Una Perspectiva Internacional". Cuadernos Económicos de I.C.E., Mayo-Junio

. Raymond, J.L. (1989), "Análisis de las Economías de Escala en el Sector Cajas de Ahorros". Documentos de Trabajo de la Fundación FIES, nº51/ 1989

. Schmidt, P. e Lovell, C.A.K. (1979), "Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Frontiers". Journal of Econometrics, Vol. 9, 343-366

- . Schmidt, P. e Sickles, R.C. (1984), "Production Frontiers and Panel Data". Journal of Business and Economic Statistics, Vol.2, Nº4, 367-374
- . Spicer, M.W. (1982), "The Economics of Burocracy and the British National Health Service". Milbank Memorial Fund Quarterly, Vol. 60, Nº4, 657-672
- . Stoddart, G.L. e Labelle, R.J.(1985) Privatisation in the Canadian Health Care System, Ministry of Supply and Services, Ottawa, Canada
- . Swamy, P.A.V.B.(1971), Statistical Inference in Random Coefficient Regression Models, Springer-Verlag, New York
- . Tatchel, M.(1983), "Measuring Hospital Output: a Review of the Service-mix and Casemix Approaches". Social Science and Medicine, Vol.17, 871-883
- . Taylor, W.E. (1980), "Small Sample Considerations in Estimation From Panel Data". Journal of Econometrics, Vol.13, 203-223
- . Thiry, B., Tulkens, H. (1989), "Productivity, Efficiency and Technical Progress". Annals of Public and Cooperative Economics, Vol.60, Nº1, 20-42
- . Tornqvist, L. (1936), "The Bank of Finland's Consumption Price Index". Bank of Finland, Monthly Bulletin, Nº10, 1-8
- . United Kingdom Department of Health "Self-Governing Hospitals". Working Paper 1, 1989, London
- . Vita, M.G.(1990) "Exploring Hospital Production Relationships with Flexible Functional Forms". Journal of Health Economics, Vol.9, 1-21
- . Wagstaff, A. (1989), "Estimating Efficiency in the Hospital Sector: A Comparison of three Statistical Cost Frontier Models". Applied Economics, Vol.21, 659-672
- . Wagstaff, A. (1989), "The Measurement of Hospital Efficiency: a Comment". Social Science and Medicine, Vol.28, Nº1, 13-17
- . Williams, A. (1985), "The Economics of Coronary Artery Bypass Grafting". British Medical Journal, Vol. 291, 326-329

A N E X O 1

Desenvolvemos vários programas para chegar à inferência sobre os modelos propostos, partindo dos dados que recolhemos. Os dados foram inicialmente introduzidos com "Lotus" e seguidamente passados a ASCII.

Sobre estes dados correram-se um conjunto programas (em FORTRAN 77) que tinham como output final as amostras utilizadas nas estimações. Este grupo era constituído pelos seguintes programas:

- . "CHICO" que faz a leitura dos ficheiros de pessoal.
- . "SILVIA" que faz a leitura dos restantes ficheiros de dados.
- . "SAO" faz um conjunto de cálculos simples sobre os dados iniciais.
- . "RUI" calcula o índice de especialização e complexidade de Evans e Walker.
- . "TONICO" que é o programa principal e que nas suas diferentes versões calcula as variáveis das regressões.

Um segundo grupo de programas (em FORTRAN 77) está desenhado para o cálculo das estimativas sobre as regressões com painel de dados. Estes programas são: "CORRE" que é o programa principal e que fornece a matriz de dados à subrutina "FREDEF" que estima os parâmetros do modelo de efeitos individuais e temporais fixos.

Um terceiro grupo de programas é constituído por programas em TSP que calculam a estatística de Hausman y Taylor o SURE, bem como regressões lineares.

Apresentamos seguidamente a título de exemplo as rotinas "FREDEF" e "RUI".

```
subroutine fredef5(beta1,beta,efi,eft,x1)
```

```
INTEGER*4 K,L,M,LM,IFAIL
```

```
REAL*8 X(230,20),Q(230,230),INT(230,230),IN(46,46),JTJT(5,5),
```

```
1 P,T,KRN1(230,230),C1(5,5),JNJN(46,46),IT(5,5),KRN2(230,230),
```

```
2 JNT(230,230),Q1(230,230),Q2(230,230),XQ(20,230),XQX(20,20),
```

```
3 XQ1(20,230),XQ1X(20,20),XQ2(20,230),XQ2X(20,20),XT(20,230),
```

```
4 ZE(1),XQXS(20,20),XQ1XS(20,20),XQ2XS(20,20),AA(20,20),BB(20,20),
```

```
5 I20(20,20),WSPC(20),X1(230,21),Y(230,1),XQY(20,1),XQ1Y(20,1),
```

```
6 XQ2Y(20,1),BETA(20,1),BETA2(20,1),QY(230,1),
```

```
7 XBETA(230,1),BETA1,XBETA2(20,1),RES(230,1),RES1(230,1),
```

```
8 Q1Y(230,1),Q2Y(230,1),RES2(230,1),SGM,SGM1,SGM2,KK,POND(20,20),
```

```
9 POND$ (20,20),QB(20,1),Q1B(20,1),Q2B(20,1),B(20,1),SMU,SLA,SGM3,
```

```
A PHI(230,230),XPHI(20,230),VAR(20,20),YIBAR(46,1),XIBAR(46,20),
```

```
B YTBAR(5,1),XTBAR(5,20),YBAR,XBAR(1,20),EFI(46,1),EFT(5,1),
```

```
C XB(230,1),REGS(230,1),YYBAR(230,1),STQ,SQE,R2,B1(46,1),
```

```
D XC(230,20),YC(230,1)
```

```
C K -NUM DE VAR
```

```
C L -NUM DE ANOS
```

```
C M -NUM DE INDIV
```

```
CALL DADOSI(X1,K,L,M)
```

```
LM=L*M
```

```
IFAIL=0
```

```
P=DFLOAT(M)
```

```
T=DFLOAT(L)
```

```
open(unit=8,file='output5.dat',status='new')
```

```
KK=DFLOAT(K)
```

```
DO 21 I=1,LM
```

```
DO 21 J=1,K
```

```
X(I,J)=X1(I,J+1)
```

```
21 CONTINUE
```

```
DO 22 I=1,LM
```

```
Y(I,1)=X1(I,1)
```

```
22 CONTINUE
```

```
YBAR=0.0D0
```

```
DO 44 I=1,LM
```

```
YBAR=YBAR+Y(I,1)
```

```
44 CONTINUE
```

```
YBAR=YBAR/(P*T)
```

```
DO 45 I=1,K
```

```
XBAR(1,I)=0.0D0
```

```
45 CONTINUE
```

```
DO 46 I=1,LM
```

```
DO 46 J=1,K
```

```
XBAR(1,J)=XBAR(1,J)+X(I,J)
```

```
46 CONTINUE
```

```
DO 47 I=1,K
```

```
XBAR(1,I)=XBAR(1,I)/(P*T)
```

```
47 CONTINUE
```

```
DO 28 I=1,M
```

```
YIBAR(I,1)=0.0D0
```

```
28 CONTINUE
```

```
DO 49 J=1,M
```

```
DO 49 I=1,L
```

```
YIBAR(J,1)=YIBAR(J,1)+Y(J+(I-1)*M,1)
```

```
49 CONTINUE
```

```
DO 99 J=1,M
```

```
YIBAR(J,1)=YIBAR(J,1)/T
```

```
99 CONTINUE
```

```
DO 29 J=1,M
```

```
DO 29 I=1,K
```

```
XIBAR(J,I)=0.0D0
```

```
29 CONTINUE
```

```
DO 51 J=1,M
```

```
DO 51 I=1,L
```

```
DO 51 II=1,K
```

```
XIBAR(J,II)=XIBAR(J,II)+X(J+(I-1)*M,II)
```

```

51 CONTINUE
   DO 52 J=1,M
   DO 52 I=1,K
   XIBAR(J,I)=XIBAR(J,I)/T
52 CONTINUE
   DO 27 I=1,L
   YTBAR(I,1)=0.0D0
27 CONTINUE
   DO 56 I=1,L
   DO 56 J=1,M
   YTBAR(I,1)=YTBAR(I,1)+Y(J+(I-1)*M,1)
56 CONTINUE
   DO 57 J=1,L
   YTBAR(J,1)=YTBAR(J,1)/P
57 CONTINUE
   DO 26 J=1,L
   DO 26 I=1,K
   XTBAR(J,I)=0.0D0
26 CONTINUE
   DO 58 J=1,M
   DO 58 I=1,L
   DO 58 II=1,K
   XTBAR(I,II)=XTBAR(I,II)+X(J+(I-1)*M,II)
58 CONTINUE
   DO 59 J=1,L
   DO 59 I=1,K
   XTBAR(J,I)=XTBAR(J,I)/P
59 CONTINUE
   DO 92 I=1,M
   DO 92 J=1,L
   DO 92 II=1,20
   XC(I+(J-1)*M,II)=X(I+(J-1)*M,II)-XIBAR(I,II)-XTBAR(J,II)+XBAR(1,II
1)
92 CONTINUE
   DO 91 I=1,M
   DO 91 J=1,L
   YC(I+(J-1)*M,1)=Y(I+(J-1)*M,1)-YIBAR(I,1)-YTBAR(J,1)+YBAR
91 CONTINUE
   DO 32 I=1,M*L
   DO 32 J=1,20
   XT(J,I)=XC(I,J)
32 CONTINUE
   DO 33 I=1,20
   DO 33 J=1,20
   I20(I,J)=0.0D0
33 CONTINUE
   DO 34 I=1,20
   I20(I,I)=1.0D0
34 CONTINUE
   CALL F01CKF(XQX,XT,X,20,20,LM,ZE,1,1,IFAIL)
   CALL F04AEF(XQX,20,I20,20,K,K,XQX$,20,WSPC,AA,20,BB,20,IFAIL)
   CALL F01CKF(XQY,XT,YC,20,1,230,ZE,1,1,IFAIL)
   CALL F01CKF(BETA,XQX$,XQY,20,1,20,ZE,1,1,IFAIL)
   CALL F01CKF(XBETA,X,BETA,LM,1,20,ZE,1,1,IFAIL)
   DO 95 J=1,M
   B1(J,1)=0.0D0
95 CONTINUE
   DO 89 J=1,M
   YIBAR(J,1)=YIBAR(J,1)-YBAR
89 CONTINUE
   DO 82 J=1,M
   DO 82 I=1,K
   XIBAR(J,I)=XIBAR(J,I)-XBAR(1,I)
82 CONTINUE
   DO 48 I=1,K
   DO 48 J=1,M

```

```

      B1(J,1)=B1(J,1)+XIBAR(J,I)*BETA(I,1)
48  CONTINUE
      DO 54 J=1,M
      EFI(J,1)=YIBAR(J,1)-B1(J,1)
54  CONTINUE
      DO 97 J=1,L
      YTBAR(J,1)=YTBAR(J,1)-YBAR
97  CONTINUE
      DO 85 J=1,L
      DO 85 I=1,K
      XTBAR(J,I)=XTBAR(J,I)-XBAR(1,I)
85  CONTINUE
      DO 60 J=1,L
      EFT(J,1)=0.0D0
60  CONTINUE
      DO 50 J=1,L
      DO 50 I=1,K
      EFT(J,1)=EFT(J,1)+XTBAR(J,I)*BETA(I,1)
50  CONTINUE
      DO 61 J=1,L
      EFT(J,1)=YTBAR(J,1)-EFT(J,1)
61  CONTINUE
      BETA1=0.0D0
      DO 80 I=1,K
      BETA1=BETA1+XBAR(1,I)*BETA(I,1)
80  CONTINUE
      BETA1=YBAR-BETA1
      DO 63 J=1,M
      DO 63 I=1,L
      REGS(J+(I-1)*M,1)=Y(J+(I-1)*M,1)-BETA1-XBETA(J+(I-1)*M,1)-
1EFT(I,1)-EFI(J,1)
63  CONTINUE
      STQ=0.0D0
      SQE=0.0D0
      DO 64 I=1,LM
      YYBAR(I,1)=Y(I,1)-YBAR
64  CONTINUE
      DO 65 J=1,LM
      STQ=STQ+YYBAR(J,1)*YYBAR(J,1)
      SQE=SQE+REGS(J,1)*REGS(J,1)
65  CONTINUE
      SGM1=SQE/((P*T)-(KK+P+T-1.0D0))
      DO 86 I=1,K
      DO 86 J=1,K
      VAR(I,J)=XQX$(I,J)*SGM1
86  CONTINUE
      R2=1.0D0-(SQE/STQ)
      WRITE(8,2222)
2222 FORMAT('COEFICIENTES DO MODELO')
      WRITE(8,2223) BETA1
2223 FORMAT('TERMO INDEPENDENTE',F12.4)
      DO 25 J=1,K
      VAR(J,J)=BETA(J,1)/(VAR(J,J)**0.50D0)
25  CONTINUE
      DO 66 I=2,K+1
      WRITE(8,2224)I,BETA(I-1,1),VAR(I-1,I-1)
2224 FORMAT('BETA',I2,',','=','F12.4,', ' RACIO T=',F10.4)
66  CONTINUE
      WRITE(8,2225) R2
2225 FORMAT('COEFIC. DE DETERMIN.',F10.4)
      WRITE(8,2226)
2226 FORMAT('EFEITOS INDIVIDUAIS')
      WRITE(8,2227)(EFI(I,1),I=1,M)
2227 FORMAT(F12.4)
      WRITE(8,2228)
2228 FORMAT('EFEITOS TEMPORAIS')

```

```
WRITE(8,2229)(EFT(I,1),I=1,L)
2229 FORMAT(F12.4)
return
END
```

```

REAL*8 CFDCHC(45,16),DS1HD(185,16),DSIJ(230,16),DS(5,1),
1 DSI(230,1),DSJ(5,16),QIJ(230,16),PI(230,1),PIJ(230,16),
2 H2IJ(230,16),H2J(5,16),HBAR(5,1),QJ(5,16),TI(230,1),
3 um(5)
CALL SAO (CFDCHC,DS1HD)
DO 10 I=1,16
DO 10 J=1,37,9
DO 10 K=1,9
DSIJ(46*((J-1)/9)+K,I)=CFDCHC(K+(J-1),I)
10 CONTINUE
DO 11 I=1,16
DO 11 J=1,149,37
DO 11 K=1,37
DSIJ(((j-1)/37)*46+k+9,I)=DS1HD(K+(J-1),I)
11 CONTINUE
DO 20 I=1,5
DS(I,1)=0.0D0
20 CONTINUE
DO 12 I=1,16
DO 12 J=1,5
DO 12 K=1,46
DS(J,1)=DS(J,1)+DSIJ((J-1)*46+K,I)
12 CONTINUE
DO 21 I=1,16
DO 21 J=1,230
DSI(J,1)=0.0D0
21 CONTINUE
DO 13 I=1,16
DO 13 J=1,230
DSI(J,1)=DSI(J,1)+DSIJ(J,I)
13 CONTINUE
DO 22 I=1,16
DO 22 J=1,5
DSJ(J,I)=0.0D0
22 CONTINUE
DO 14 I=1,16
DO 14 J=1,5
DO 14 K=1,46
DSJ(J,I)=DSJ(J,I)+DSIJ((J-1)*46+K,I)
14 CONTINUE
DO 15 I=1,46
DO 15 K=1,5
DO 15 J=1,16
QIJ(I+(K-1)*46,J)=DSIJ(I+(K-1)*46,J)/DSJ(K,J)
15 CONTINUE
DO 16 I=1,16
DO 16 J=1,5
DO 16 K=1,46
PI((J-1)*46+K,1)=DSI((J-1)*46+K,1)/DS(J,1)
16 CONTINUE
DO 17 I=1,16
DO 17 J=1,230
PIJ(J,I)=DSIJ(J,I)/DSI(J,1)
17 CONTINUE
DO 18 I=1,16
DO 18 J=1,5
do 18 k=1,46
H2IJ((J-1)*46+k,I)=qij((J-1)*46+k,I)*46.0d0
18 CONTINUE
DO 19 I=1,16
DO 19 J=1,230
IF(H2IJ(J,I).LT.0.000000001) THEN
H2IJ(J,I)=0.0D0
ELSE
H2IJ(J,I)=qij(j,i)*LOG(H2IJ(J,I))
ENDIF

```

```

19 CONTINUE
   DO 25 I=1,16
   DO 25 J=1,5
   H2J(J,I)=0.0D0
25 CONTINUE
   DO 30 I=1,16
   DO 30 J=1,5
   DO 30 K=1,46
   H2J(J,I)=H2J(J,I)+H2IJ((J-1)*46+K,I)
30 CONTINUE
   DO 26 J=1,5
   HBAR(J,1)=0.0D0
26 CONTINUE
   DO 27 I=1,5
   DO 27 J=1,16
   QJ(I,J)=DSJ(I,J)/DS(I,1)
27 CONTINUE
   DO 31 I=1,16
   DO 31 J=1,5
   HBAR(J,1)=HBAR(J,1)+H2J(J,I)*QJ(J,I)
31 CONTINUE
   DO 32 I=1,16
   DO 32 J=1,5
   H2J(J,I)=H2J(J,I)/HBAR(J,1)
   DO 28 J=1,230
   TI(J,1)=0.0D0
28 CONTINUE
   do 75 j=1,5
   um(j)=0.0d0
75 continue
   do 76 j=1,5
   do 76 i=1,16
   um(j)=um(j)+h2j(j,i)*qj(j,i)
76 continue
   DO 33 I=1,16
   DO 33 J=1,5
   DO 33 K=1,46
   TI((J-1)*46+K,1)=TI((J-1)*46+K,1)+H2J(J,I)*PIJ((J-1)*46+K,I)
33 CONTINUE
end

```

A N E X O 2



A amostra é constituída pelos seguintes centros:

HOSPITAIS CENTRAIS

Centro 1	Hospitais Civis de Lisboa
Centro 2	Hospital Santa Maria
Centro 3	Hospital S. José
Centro 4	Hospital Universitário de Coimbra
Centro 5	Hospital de Santo António
Centro 6	Centro Hospitalar de Coimbra
Centro 7	Hospital Pulido Valente
Centro 8	Centro Hospitalar de Vila Nova de Gaia
Centro 9	Hospital Egas Moniz

HOSPITAIS DISTRITAIS

Centro 10	H D Abrantes
Centro 11	H D Almada
Centro 12	C H Aveiro Sul
Centro 13	C H Aveiro Norte
Centro 14	H D Barcelos
Centro 15	H D Barreiro
Centro 16	H D Beja
Centro 17	H D Braga
Centro 18	H D Bragança
Centro 19	C H Caldas da Rainha

Centro 20	H D Cascais
Centro 21	H D Castelo Branco
Centro 22	H D Chaves
Centro 23	H D Covilhã
Centro 24	H D Elvas
Centro 25	H D Évora
Centro 26	H D Faro
Centro 27	H D Figueira da Foz
Centro 28	H D Guarda
Centro 29	H D Guimarães
Centro 30	H D Lamego
Centro 31	H D Leiria
Centro 32	H D Matosinhos
Centro 33	H D Mirandela
Centro 34	H D Portalegre
Centro 35	H D Portimão
Centro 36	H D Santarém
Centro 37	H D Setúbal
Centro 38	H D Tomar
Centro 39	H D Torres Novas
Centro 40	H D Torres Vedras
Centro 41	C H Vale do Sousa
Centro 42	H D Viana do Castelo
Centro 43	H D Vila Franca de Xira
Centro 44	H D Famalicão
Centro 45	H D Vila Real
Centro 46	H D Viseu