ANÁLISE ENVOLTÓRIA DE DADOS: CONCEITOS E MODELOS BÁSICOS

Adriano Provezano Gomes¹ Antonio José Medina dos Santos Baptista²

1. Introdução

Questões relacionadas à eficiência são frequentemente abordadas por pesquisadores e tomadores de decisão, principalmente em se tratando de ambientes competitivos e dinâmicos. Identificar o verdadeiro potencial da expansão da produção e as mudanças na eficiência, no progresso tecnológico e na produtividade ao longo do tempo é condição necessária para formulação de políticas econômicas coerentes com as reais necessidades dos setores analisados. Por isso, o uso de medidas de eficiência tem crescido significativamente nas últimas décadas, tornando-se, hoje em dia, um dos principais tópicos de estudo dos economistas.

O estudo das medidas de eficiência, com base em técnicas não-paramétricas, teve início com Farrel (1957), que propôs um modelo empírico em que cada unidade de produção é avaliada em relação às outras unidades de um conjunto homogêneo e representativo. Dessa maneira, a medida de eficiência é relativa e o respectivo valor para uma unidade de produção corresponde ao desvio observado em relação àquelas unidades consideradas eficientes.

A idéia básica consiste em separar a eficiência em dois componentes: eficiência técnica, que reflete a habilidade da firma em obter máximo produto, dado um conjunto de insumos; e eficiência alocativa, que reflete a habilidade da firma em utilizar os insumos em proporções ótimas, dados seus preços relativos. Essas duas medidas são combinadas para se obter uma medida de eficiência econômica total.

As avaliações das medidas de eficiência podem, ainda, ser precedidas de duas orientações – aquela que se fundamenta na redução de insumos, denominada orientação insumo; e aquela que coloca ênfase no aumento do produto, denominada orientação

¹ Professor do Departamento de Economia da UFV, e-mail: apgomes@ufv.br.

² Doutorando em Economia Aplicada da UFV (bolsista CAPES/PEC-PG) e técnico do Instituto Nacional de Desenvolvimento das Pescas de Cabo Verde, e-mail: tozecv@yahoo.com.

produto. O gráfico (a) apresentado na Figura 1 ilustra as medidas de eficiência com orientação insumo. Já no gráfico (b) estão as medidas com orientação produto.

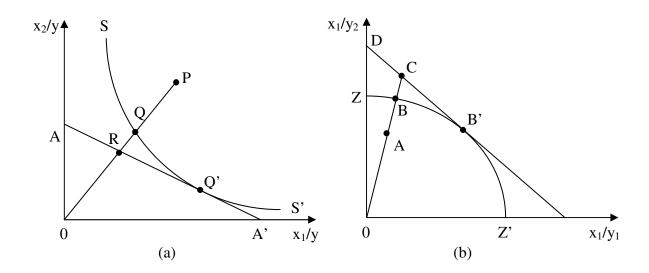


Figura 1 - Medidas de eficiência com orientação insumo (a) e orientação produto (b).

Iniciando com a orientação insumo, considere uma firma que usa dois insumos (x₁ e x₂) para produzir um único produto (y). Tendo em vista a pressuposição de retornos constantes à escala, SS' representa uma isoquanta unitária de uma firma totalmente eficiente. Se outra firma usa uma quantidade de insumos, definida pelo ponto P, para produzir uma unidade de produto, sua ineficiência técnica poderia ser representada pela distância QP, que indica a quantidade pela qual todos os insumos podem ser reduzidos sem reduzir a produção. Em relação à porcentagem, tem-se a razão QP/0P. A eficiência técnica (ET) dessa firma seria dada por:

$$ET = \frac{0Q}{0P} = 1 - \frac{QP}{0P}.$$

Como $0 < ET \le 1$, a medida encontrada fornece o grau de ineficiência técnica da firma. Se ET = 1, a firma é tecnicamente eficiente, situando-se sobre a isoquanta eficiente, como é o caso do ponto Q.

Quando se conhece a razão entre os preços dos insumos, representada pela isocusto AA', pode-se calcular a eficiência alocativa (EA). Considerando uma firma que opera em P, tem-se:

$$EA = \frac{0R}{0Q}.$$

A distância RQ representa a redução nos custos de produção que poderia acontecer, caso a produção ocorresse em um ponto de eficiência alocativa, como é o caso de Q', em vez do ponto Q, que é tecnicamente eficiente, mas alocativamente ineficiente.

Assim, a ineficiência técnica é resultante do uso excessivo de insumos, para dado nível de produção. Já a ineficiência alocativa decorre do emprego desses insumos em proporções inadequadas, dados seus respectivos preços, ou seja, quando a taxa marginal de substituição entre os insumos não for igual à razão dos seus preços. Em ambos os casos, o custo não será minimizado.

A eficiência econômica total (EE) seria dada pelo produto das eficiências técnica e alocativa:

$$EE = \frac{0Q}{0P} \times \frac{0R}{0Q} = \frac{0R}{0P}.$$

Segundo Coelli (1998), as medidas de eficiência técnica insumo orientadas procuram responder à seguinte pergunta: "Qual a quantidade de insumos que pode ser proporcionalmente reduzida, sem alterar a quantidade de produto que está sendo produzido?". Entretanto, outra questão poderia surgir: "Qual a quantidade de produto que poderia ser proporcionalmente expandida, sem alterar as quantidades de insumos utilizados?". Nesse caso, torna-se necessária a análise de medidas de orientação produto. O gráfico (b) da Figura 1 ilustra uma situação que envolve a produção de dois produtos (y_1 e y_2) e o uso de um único insumo (x_1).

Se os retornos à escala são considerados constantes, então pode-se representar a tecnologia por uma curva de possibilidades de produção unitária, descrita pela linha ZZ'. O ponto A representa uma firma ineficiente, situando-se abaixo da curva de possibilidades de produção. A distância AB representa sua ineficiência técnica, ou seja,

as quantidades de produtos que poderiam ser aumentadas sem necessidade de insumos adicionais. Nesse caso, a medida de eficiência técnica seria dada por:

$$ET = \frac{0A}{0B}$$
.

Por meio de informações sobre os preços dos produtos, pode-se traçar uma linha de "isoreceita" (DD') e, então, definir a eficiência alocativa como:

$$EA = \frac{0B}{0C}.$$

A ineficiência alocativa da firma que opera em A seria dada pela distância BC.

O cálculo da eficiência econômica total é semelhante ao realizado nas medidas de orientação insumo, isto é:

$$EE = \frac{0A}{0B} \times \frac{0B}{0C} = \frac{0A}{0C}.$$

Todas as medidas apresentadas são radiais, isto é, são medidas ao longo de um raio que sai da origem até o ponto de produção observado. Segundo Färe et al. (1994), a vantagem desse tipo de medida radial é que ela é invariante em unidade. Isso significa que as unidades de medida de insumos e produtos podem variar, pois não alteram o valor da medida de eficiência. Por exemplo, tanto faz medir o insumo mão-de-obra em dias trabalhados ou em horas trabalhadas, que a medida de eficiência radial não se altera.

2. Análise envoltória de dados (DEA)

As medidas de eficiência podem ser facilmente obtidas de problemas simples, que abrangem poucos insumos e produtos. Entretanto, em situações em que várias unidades utilizam múltiplos insumos e produzem vários produtos, o cálculo da eficiência relativa de cada unidade torna-se mais complexo. Para solucionar essas dificuldades, podem-se obter, a partir de uma amostra de dados, fronteiras eficientes, as quais servirão como referencial para as comparações entre as unidades.

Fronteiras podem ser estimadas por diferentes métodos, paramétricos ou nãoparamétricos. As fronteiras estocásticas consistem em abordagens paramétricas, sendo estimadas por métodos econométricos, enquanto a análise envoltória de dados (DEA) é uma abordagem não-paramétrica, que envolve programação matemática em sua estimação.

Na abordagem paramétrica, estima-se uma função fronteira de produção, utilizada para caracterizar uma transformação eficiente de insumos em produtos. A medida de eficiência relativa de uma firma é determinada pela comparação da produção observada nessa firma, dado um conjunto de insumos, com a produção "ideal", com os mesmos níveis de insumos. Essa produção ideal, na abordagem paramétrica, é calculada pela função de produção teórica estimada. A dificuldade reside no fato de que a função de produção teórica requer que se explicite a formulação da relação funcional entre os insumos e os produtos. Essa dificuldade aumenta quando se busca uma forma funcional teórica para processos de produção mais complexos, ou seja, processos que envolvem múltiplos insumos e produtos.

Como alternativa aos métodos estatísticos convencionais, surgiu a abordagem não-paramétrica de análise de eficiência. De acordo com Forsund e Sarafoglou (2000), embora a análise da função de produção e eficiência pela abordagem de programação matemática já tenha começado bem antes dos anos 70, foi o trabalho de Charnes et al. (1978) que introduziu na literatura o termo *Data Envelopment Analysis* (DEA), técnica não-paramétrica que utiliza a programação matemática para analisar a eficiência relativa das unidades de produção.

Para Marinho (2001), as características mais interessantes da DEA são:

- a) Caracteriza cada unidade como eficiente ou ineficiente através de uma única medida resumo de eficiência.
- b) Não faz julgamentos *a priori* sobre os valores das ponderações de insumos e produtos que levariam as unidades ao melhor nível de eficiência possível.
- c) Pode prescindir (mas não rejeita) de sistemas de preços.
- d) Dispensa (mas pode acatar) pré-especificações de funções de produção subjacentes.
- e) Pode considerar sistemas de preferências de avaliadores e de gestores.
- f) Baseia-se em observações individuais e não em valores médios.
- g) Permite a incorporação, na análise, de insumos e de produtos avaliados em unidades de medidas diferentes.

- h) Possibilita a verificação de valores ótimos de produção e de consumo respeitando restrições de factibilidade.
- i) Permite a observação de unidades eficientes de referência para aquelas que forem assinaladas como ineficientes.
- j) Produz resultados alocativos eficientes no sentido de Pareto.

Os modelos DEA são baseados em uma amostra de dados observados para diferentes unidades produtoras, também conhecidas como DMUs³. O objetivo é construir um conjunto de referências a partir dos próprios dados das DMUs e, então, classificá-las em eficientes ou ineficientes, tendo como referencial essa superfície formada.

Uma pressuposição fundamental na técnica DEA é de que, se uma DMU A é capaz de produzir Y(A) unidades de produto utilizando X(A) unidades de insumos, outras DMUs poderiam também fazer o mesmo, caso elas estejam operando eficientemente. De forma similar, se uma DMU B é capaz de produzir Y(B) unidades de produto utilizando-se X(B) de insumos, então outras DMUs poderiam ser capazes de realizar o mesmo esquema de produção. Caso as DMUs A e B sejam eficientes, elas poderiam ser combinadas para formar uma DMU composta, que utiliza uma combinação de insumos para produzir uma combinação de produtos. Desde que essa DMU composta não necessariamente exista, ela é denominada DMU virtual.

A idéia central dessa técnica é encontrar a melhor DMU virtual para cada DMU real. Se a DMU virtual, que pode ser uma combinação convexa de outras DMUs reais, conseguir produzir maiores quantidades de produtos utilizando a mesma ou menor quantidade de insumos, então a DMU real será ineficiente. As unidades eficientes que, quando combinadas, fornecem a DMU virtual para a unidade ineficiente são conhecidas como pares ou *benchmarks* daquela DMU.

Um problema apresentado por Alves e Gomes (1998) é que, em virtude das diferenças qualitativas nos insumos, uma DMU virtual gerada pela combinação de outras DMUs pode utilizar uma combinação de insumos que não existe na realidade. Segundo estes autores, tomando como exemplo dados do insumo maquinaria utilizado em fazendas, pode-se ter um produtor que utilize trator, enquanto outro pode utilizar serviços de tração animal. Um produtor virtual que combine esses dois insumos pode

.

³ Na literatura relacionada com modelos DEA, uma unidade produtora é tratada como DMU (*decision making unit*), uma vez que desses modelos provém uma medida para avaliar a eficiência relativa de unidades tomadoras de decisão. Por unidade produtora entende-se qualquer sistema produtivo que transforme insumos em produtos.

aparecer nos resultados, embora seja pouco provável que exista na prática. Para contornar dificuldades como esta, é necessário impor algumas restrições ao modelo, as quais serão abordadas no próximo capítulo.

A seguir, serão apresentados os modelos utilizados na DEA. Inicialmente, será discutido o modelo proposto por Charnes et al. (1978), o qual ficou conhecido, na literatura, como modelo CCR, em razão das iniciais dos nomes dos referidos autores. Esse modelo será descrito mais detalhadamente, a fim de que sirva de base para os demais.

2.1. Modelos com retornos constantes

Segundo Charnes et al. (1994), para estimar e analisar a eficiência relativa das DMUs, a DEA utiliza a definição de ótimo de Pareto, segundo o qual nenhum produto pode ter sua produção aumentada sem que sejam aumentados os seus insumos ou diminuída a produção de outro produto, e, de forma alternativa, quando nenhum insumo pode ser diminuído sem ter que diminuir a produção de algum produto.

Assim, para executar os modelos DEA são necessários dados referentes aos insumos e produtos de cada DMU a ser analisada. Cooper et al. (2000) identificam alguns pontos para auxiliar na escolha dos dados:

- os dados de insumos e produtos devem assumir valores não-negativos para todas as DMUs;
- os insumos, produtos e escolhas das DMUs devem refletir o interesse do pesquisador nos componentes que entrarão na avaliação da eficiência relativa das DMUs;
- a unidade de medida dos insumos e produtos não necessita ser a mesma, ou seja, pode-se ter, por exemplo, um insumo medido em dias de serviço, outros em litros, quilos, gastos monetários, etc.; e
- preferencialmente, o número de variáveis (soma de insumos e produtos) não deve superar em um terço o número de observações.

Considere que existam k insumos e m produtos para cada n DMUs. São construídas duas matrizes – a matriz X de insumos, de dimensões (k x n), e a matriz Y de produtos, de dimensões (m x n), representando os dados de todas as n DMUs:

$$\mathbf{X} = \begin{pmatrix} \mathbf{x}_{11} & \mathbf{x}_{12} & \cdots & \mathbf{x}_{1n} \\ \mathbf{x}_{21} & \mathbf{x}_{22} & \cdots & \mathbf{x}_{2n} \\ \vdots & & & & \\ \mathbf{x}_{k1} & \mathbf{x}_{k2} & \cdots & \mathbf{x}_{kn} \end{pmatrix}, \qquad \mathbf{Y} = \begin{pmatrix} \mathbf{y}_{11} & \mathbf{y}_{12} & \cdots & \mathbf{y}_{1n} \\ \mathbf{y}_{21} & \mathbf{y}_{22} & \cdots & \mathbf{y}_{2n} \\ \vdots & & & & \\ \mathbf{y}_{m1} & \mathbf{y}_{m2} & \cdots & \mathbf{y}_{mn} \end{pmatrix}.$$

Na matriz X, cada linha representa um insumo e cada coluna representa uma DMU. Já na matriz Y, cada linha representa um produto e cada coluna uma DMU. Para a matriz X, é necessário que os coeficientes sejam não-negativos e que cada linha e cada coluna contenha, pelo menos, um coeficiente positivo, isto é, cada DMU consome ao menos um insumo e uma DMU, pelo menos, consome o insumo que está em cada linha. O mesmo raciocínio se aplica para a matriz Y, ou seja, os coeficientes são não-negativos, cada produto é produzido por uma DMU, pelo menos, e cada DMU produz pelo menos um produto.

Para a i-ésima DMU, são representados os vetores x_i e y_i , respectivamente, para insumos e produtos. Pode-se, então, obter uma medida de eficiência para cada DMU da amostra, que é a razão entre todos os produtos e todos os insumos. Para a i-ésima DMU, tem-se:

$$\label{eq:efficiencia} \text{Eficiência da DMU i} = \frac{\mu \check{\ y}_i}{\nu \check{\ x}_i} = \frac{\mu_1 y_{1i} + \mu_2 y_{2i} + \dots + \mu_m y_{mi}}{\nu_1 x_{1i} + \nu_2 x_{2i} + \dots \nu_k x_{ki}},$$

em que μ é um vetor (m x 1) de pesos nos produtos e v é um vetor (k x 1) de pesos nos insumos. Verifica-se que a medida de eficiência será uma escalar, devido às ordens dos vetores que a compõem.

O cálculo da eficiência dessa forma pressupõe a aplicação de um conjunto comum de pesos em todas as DMUs. No entanto, existe certa dificuldade em obter um conjunto comum de pesos para determinar a eficiência relativa de cada DMU, pois as DMUs podem estabelecer diferentes valores para os insumos e produtos e adotarem diferentes pesos. É necessário, então, estabelecer um problema que permita que cada DMU possa adotar o conjunto de pesos que for mais favorável, comparativamente com as outras unidades. Para selecionar os pesos ótimos para cada DMU, especifica-se um problema de programação matemática. O modelo adaptado de Charnes et al. (1978) para calcular a eficiência da i-ésima DMU possui a seguinte forma:

$$\begin{aligned} \text{MAX}_{u,v} & \left(\mu ` y_i / v ` x_i \right), \\ \text{sujeito a:} & \\ & \mu' y_j / v' x_j \leq 1, \quad j = 1, 2, ..., n, \\ & \mu, v \geq 0. \end{aligned} \tag{1}$$

Essa formulação envolve a obtenção de valores para μ e ν , de tal forma que a medida de eficiência para a i-ésima DMU seja maximizada, sujeita à restrição de que as medidas de eficiência de todas as DMUs sejam menores ou iguais a um. A característica principal desse modelo é que os pesos μ e ν são tratados como incógnitas, sendo escolhidos de maneira que a eficiência da i-ésima DMU seja maximizada. Caso a eficiência obtida para a DMU que está sendo testada seja igual a um, ela será eficiente em relação às demais; caso contrário, será ineficiente, pois existem DMUs que combinam, de forma mais eficiente, os seus insumos e produtos.

Contudo, a restrição de não-negatividade imposta aos vetores de pesos não é matematicamente suficiente para obter um único valor definido. Isso porque, segundo Charnes et al. (1978), o modelo nessa forma é uma formulação de programação não-linear estendida de um problema de programação fracionário. Torna-se necessário, portanto, transformar o problema de programação fracionária, com infinitas soluções possíveis, em um problema de programação linear, com uma única solução. Essa transformação permitirá a utilização de um grande número de observações (DMUs) e um menor número de insumos e produtos, característica de muitas aplicações econômicas⁴. A formulação linearizada é a seguinte:

$$MAX_{u,v}$$
 $(\mu' y_i)$,
 $sujeito \ a:$
 $\mu' y_j - v' x_j \le 0, \quad j = 1,2,...,n,$
 $vx_i = 1,$
 $\mu, v \ge 0.$ (2)

Para Bowlin (1998), o modelo apresentado nessa forma linearizada objetiva maximizar a produção virtual, mantendo a condição de que esta produção virtual não pode exceder os insumos virtuais em nenhuma DMU. Nesse sentido, o modelo condiz

⁴ Para maiores detalhes sobre a teoria de programação fracionária, veja o artigo pioneiro de Charnes e Cooper (1962).

com as condições de ótimo de Pareto, uma vez que aumentos no valor máximo somente são possíveis aumentando o valor de ao menos um insumo ou reduzindo a produção de algum produto.

O problema anterior consiste em uma maximização do produto. Entretanto, por meio da dualidade em programação linear, pode-se derivar uma forma envoltória do problema anterior, a qual terá orientação insumo, da seguinte forma⁵:

$$MIN_{\theta,\lambda} \quad \theta,$$
 $sujeito \ a:$

$$-y_i + Y\lambda \ge 0,$$

$$\theta x_i - X\lambda \ge 0,$$

$$\lambda \ge 0,$$
(3)

em que θ é uma escalar, cujo valor será a medida de eficiência da i-ésima DMU. Caso o valor de θ seja igual a um, a DMU será eficiente; caso contrário, será menor que um. O parâmetro λ é um vetor (n x 1), cujos valores são calculados de forma a obter a solução ótima. Para uma DMU eficiente, todos os valores de λ serão zero; para uma DMU ineficiente, os valores de λ serão os pesos utilizados na combinação linear de outras DMUs eficientes, que influenciam a projeção da DMU ineficiente sobre a fronteira calculada. Isso significa que, para uma unidade ineficiente, existe pelo menos uma unidade eficiente, cujos pesos calculados fornecerão a DMU virtual da unidade ineficiente, mediante combinação linear. Essas unidades eficientes serão os *benchmarks* daquela DMU.

Segundo Coelli et al. (1998), o problema apresentado nessa forma envoltória (dual) envolve menor número de restrições do que o primal. Como no primal existem (k+m) variáveis, o dual terá (k + m) restrições, que é menor que as (n + 1) restrições do primal, uma vez que o número de DMUs é superior à soma do número de insumos mais o número de produtos. Nesse sentido, o dual é preferível ao primal, pois consome menos tempo para ser resolvido. É importante destacar que o problema deve ser resolvido n vezes, isto é, uma vez para cada DMU na amostra.

_

⁵ Tratamentos mais rigorosos dos modelos, bem como as provas matemáticas das transformações, podem ser encontrados em Seiford e Thrall (1990).

Um exemplo com dados fictícios ajudará na compreensão. Considere a Tabela 1, em que são analisadas seis DMUs que utilizam dois insumos $(x_1 e x_2)$ para produzir um único produto (y).

Tabela 1 - Produto e insumos de seis DMUs

DMU	Y	\mathbf{x}_1	x ₂	x ₁ /y	x ₂ /y
1	2	6	14	3	7
2	4	8	16	2	4
3	3	12	12	4	4
4	2	6	4	3	2
5	1	6	3	6	3
6	3	15	3	5	1

Fonte: Dados fictícios.

Em um problema simples como este, envolvendo dois insumos e um produto, uma forma de avaliar a eficiência das DMUs é construir uma isoquanta unitária e comparar, graficamente, a posição de cada unidade. Para isso, os dados apresentados nas duas últimas colunas da Tabela 1 referem-se às relações insumo/produto, as quais permitem a construção da isoquanta unitária eficiente (Figura 2).

Observando a Figura 2, verifica-se que a isoquanta eficiente é composta pelas DMUs 2, 4 e 6. Nenhuma DMU encontra-se abaixo desta isoquanta; as DMUs 1, 3 e 5 encontram-se acima dela, ou seja, gastam proporcionalmente mais, em pelo menos um insumo, do que aquelas que estão na isoquanta.

Identificada a fronteira eficiente, pode-se, utilizando movimentos radiais, projetar as DMUs ineficientes para a fronteira. Esses movimentos permitem identificar quais DMUs eficientes foram responsáveis por outra DMU ter sido considerada ineficiente, ou seja, quais são os *benchmarks* da DMU ineficiente. Assim, a projeção da DMU 1 para a fronteira eficiente permite dizer que apenas a DMU 2 foi responsável pela sua ineficiência. Já a DMU 3 tem como *benchmarks* as DMUs 2 e 4, enquanto os *benchmarks* da DMU 5 são as unidades 4 e 6.

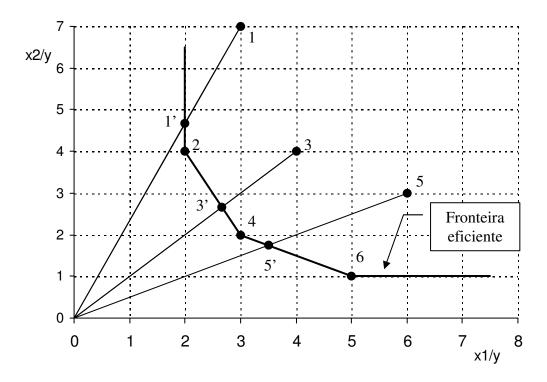


Figura 2 - Isoquanta eficiente no modelo com retornos constantes.

A obtenção da fronteira eficiente foi resultado da aplicação de seis problemas de programação linear, um para cada DMU. Por exemplo, para a DMU 5, o problema é escrito da seguinte forma:

MIN
$$\theta$$
,
sujeito a :

$$-y_5 + (y_1\lambda_1 + y_2\lambda_2 + y_3\lambda_3 + y_4\lambda_4 + y_5\lambda_5 + y_6\lambda_6) \ge 0,$$

$$\theta x_{15} - (x_{11}\lambda_1 + x_{12}\lambda_2 + x_{13}\lambda_3 + x_{14}\lambda_4 + x_{15}\lambda_5 + x_{16}\lambda_6) \ge 0,$$

$$\theta x_{25} - (x_{21}\lambda_1 + x_{22}\lambda_2 + x_{23}\lambda_3 + x_{24}\lambda_4 + x_{25}\lambda_5 + x_{26}\lambda_6) \ge 0,$$

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) \ge 0$$

Substituindo os dados da Tabela 1, o problema para obtenção da medida de eficiência para a DMU 5 fica:

MIN
$$\theta$$
,
sujeito a :
 $-1+(2\lambda_1+4\lambda_2+3\lambda_3+2\lambda_4+1\lambda_5+3\lambda_6) \ge 0$,
 $6\theta-(6\lambda_1+8\lambda_2+12\lambda_3+6\lambda_4+6\lambda_5+15\lambda_6) \ge 0$,
 $3\theta-(14\lambda_1+16\lambda_2+12\lambda_3+4\lambda_4+3\lambda_5+3\lambda_6) \ge 0$,
 $(\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5,\lambda_6) \ge 0$

Analogamente, podem-se construir os problemas para as demais DMUs. A Tabela 2 apresenta os valores dos θ s e λ s obtidos com a aplicação da DEA nas seis unidades analisadas.

Tabela 2 - Resultados fornecidos pela DEA com orientação insumo e retornos constantes à escala

DMU	θ	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
1	0,6667	_	0,5000	_	-	_	-
2	1,0000	-	1,0000	-	-	-	-
3	0,6667	-	0,2500	-	1,0000	-	-
4	1,0000	_	_	-	1,0000	-	-
5	0,5833	_	_	_	0,3750	_	0,0833
6	1,0000	_	_	_	_	_	1,0000

Fonte: Resultados da pesquisa.

Conforme mencionado, das seis DMUs analisadas, apenas as DMUs 2, 4 e 6 são eficientes. As demais estão utilizando insumos em excesso, os quais poderiam ser reduzidos sem ocasionar redução na produção. Por exemplo, a DMU 5 possui eficiência técnica igual a 0,5833, isto é, ela pode reduzir o uso dos insumos em 41,67%, sem reduzir o nível de produção. Isso implicaria uma projeção do ponto 5 para o ponto 5', na Figura 2.

Considerando a medida de eficiência obtida para a DMU 5 de 0,5833, pode-se calcular a quantidade de insumos x_1 e x_2 que esta DMU deveria utilizar:

$$x_{15}^* = \theta x_{15} = 0,5833 \ x \ 6 = 3,5$$

 $x_{25}^* = \theta x_{25} = 0,5833 \ x \ 3 = 1,75$

De modo semelhante, podem-se obter as quantidades de insumos que as outras DMUs ineficientes deveriam utilizar. Para isso, basta multiplicar o valor da eficiência da DMU pela quantidade utilizada de cada insumo.

Observando a Figura 2, nota-se que o ponto 5' localiza-se na fronteira eficiente na faixa delimitada pelas DMUs 4 e 6. Com isso, essas DMUs são os *benchmarks* da DMU 5. O ponto 5' é uma combinação linear dos pontos 4 e 6, em que os pesos utilizados são os valores encontrados para λ_4 e λ_6 referentes ao problema de programação aplicado à DMU 5. Podem-se, então, obter as quantidades dos insumos x_1 e x_2 que a DMU 5 deveria utilizar mediante a combinação linear dos *benchmarks*:

$$x_{15}^* = (x_{14}\lambda_4) + (x_{16}\lambda_6) = (6 \times 0.3750) + (15 \times 0.0833) = 3.5$$

 $x_{25}^* = (x_{24}\lambda_4) + (x_{26}\lambda_6) = (4 \times 0.3750) + (3 \times 0.0833) = 1.75$

Novamente, podem-se utilizar combinações lineares dos *benchmarks* das outras DMUs ineficientes para calcular as quantidades de insumos que estas DMUs deveriam utilizar. De posse desses valores, obtêm-se as coordenadas das DMUs ineficientes projetadas na fronteira eficiente. Para a DMU 5, as coordenadas do ponto 5' são:

$$\frac{x_{15}^*}{y_5} = \frac{3.5}{1} = 3.5$$
$$\frac{x_{25}^*}{y_5} = \frac{1.75}{1} = 1.75$$

O mesmo raciocínio se aplica às demais DMUs. Os valores projetados na fronteira das DMUs ineficientes estão descritos na Tabela 3.

Tabela 3 - Valores projetados para as DMUs ineficientes

DMUs	\mathbf{x}_1	x ₂	x ₁ /y	x ₂ /y
1 → 1'	4,00	9,33	2,00	4,66
3 → 3'	8,00	8,00	2,67	2,67
5 → 5'	3,50	1,75	3,50	1,75

Fonte: Resultados da pesquisa.

A forma linear da fronteira não-paramétrica, obtida pela DEA, pode causar algumas dificuldades na mensuração da eficiência. Esse problema ocorre devido a algumas "faixas" da fronteira linear que são paralelas aos eixos. Quando uma DMU apresentar combinação ótima (DMU virtual) projetada na fronteira paralela aos eixos, caberá questionar se o ponto projetado é realmente ótimo.

Por exemplo, a medida de eficiência para a DMU 1 seria igual a 01'/01, ou seja, a razão entre as coordenadas do ponto 1', localizado na parte da fronteira que é paralela ao eixo x_2 , e do ponto 1. Entretanto, é questionável se o ponto 1' é eficiente, já que se pode reduzir a quantidade do insumo x_2 e, ainda, produzir a mesma quantidade de produto. Observando o ponto 2, percebe-se que este é mais eficiente do que o ponto 1', uma vez que utiliza menor quantidade do insumo x_2 para produzir o mesmo produto. Assim, mesmo que o movimento radial tenha projetado a DMU 1 para a fronteira eficiente, o ponto projetado ainda apresenta alguma ineficiência. Pode-se dizer que existe uma folga no insumo x_2 (*input slack*).

Para a i-ésima DMU da formulação anterior, as folgas de insumo somente serão zero, quando θx_i - $X\lambda = 0$, para valores ótimos de θ e λ . Assim, como a DMU 2 é o único *benchmark* da DMU 1, as folgas no uso dos insumos (I_S) para a DMU 1 são:

$$I_{S11} = \theta x_{11} - x_{12} \lambda_2 = (0,6667 \ x \ 6) - (8 \ x \ 0,5) = 0$$

$$I_{S21} = \theta x_{21} - x_{22} \lambda_2 = (0,6667 \ x \ 14) - (16 \ x \ 0,5) = 1,3338$$

Percebe-se que a folga na utilização do insumo x_2 é de cerca de 1,33, que deve ser reduzida das 9,33 unidades obtidas anteriormente para o ponto 1'. Com isso, as quantidades de insumos que a DMU 1 deveria utilizar para se tornar eficiente, já

descontados os movimentos radial e de folga, são $x_1 = 4$ e $x_2 = 8$. O ponto da DMU 1, na fronteira eficiente, seria dado por:

$$\frac{x_{11}^*}{y_1} = \frac{4}{2} = 2$$
$$\frac{x_{21}^*}{y_1} = \frac{8}{2} = 4$$

$$\frac{x_{21}^*}{y_1} = \frac{8}{2} = 4$$

o qual coincide com as coordenadas da DMU 2.

Analogamente, caso se considere uma situação que envolva mais insumos e, ou, múltiplos produtos, pode-se obter a ocorrência de um conceito relacionado, conhecido como folga de produto (output slack), que é a possibilidade de aumentar a produção de um produto sem que seja necessário aumentar a utilização de um ou mais insumos. Nesse caso, existem outras DMUs que, ao utilizar a mesma combinação de insumos, conseguem produzir maiores quantidades de produtos do que a DMU que está sendo analisada. Assim, para a i-ésima DMU, as folgas de produto serão zero, somente se Yλ $y_i = 0$.

No exemplo anterior, a folga no insumo x₂ associada a DMU 1 foi fácil de ser identificada. Entretanto, para casos em que existam mais insumos e produtos, a identificação de um ponto mais próximo na fronteira eficiente, tal como a DMU 2, não é uma questão tão trivial. Todavia, de acordo com Lins e Meza (2000), é necessário considerar todas as variáveis de folgas que ainda possam ter valores positivos na solução inicial do problema, para que sejam obtidos coeficientes de eficiência que satisfaçam as condições de Pareto-Koopmans (escore de eficiência = 1 e variáveis de folga = 0).

Autores como Ali e Seiford (1993) e Cooper et al. (2000) utilizam o método de dois estágios como alternativa para resolver o problema de folgas na análise de eficiência. De modo geral, no primeiro estágio são computados os escores de eficiência para as DMUs – na Figura 2, isso significa a projeção da DMU 1 para o ponto 1'. No segundo estágio, obtém-se a projeção do ponto ineficiente 1' para um ponto eficiente na fronteira, tal como 2, considerando-se as possíveis folgas existentes.

Conforme mencionado anteriormente, as folgas nos insumos (I_s) e nos produtos (O_S) para a i-ésima DMU são dadas por:

$$I_S = \theta x_i - X\lambda$$
$$O_S = Y\lambda - y_i$$

O segundo estágio utiliza os valores para as medidas de eficiência (θ) , obtidos no primeiro estágio, e resolve um problema de programação linear utilizando λ , I_S e O_S como variáveis. Esse problema é dado por:

$$MIN_{\lambda,O_{S},I_{S}} - (M_{1}O_{S} + K_{1}I_{S}),$$
sujeito a:
$$-y_{i} + Y\lambda - O_{S} = 0,$$

$$\theta x_{i} - X\lambda - I_{S} = 0,$$

$$\lambda \geq 0, O_{S} \geq 0, I_{S} \geq 0,$$

$$(4)$$

em que O_S é um vetor (m x 1) de folgas de produtos; I_S é um vetor (k x 1) de folgas de insumos; e M_1 e K_1 são vetores unitários (1,...,1), com dimensões (m x 1) e (k x 1), respectivamente. O valor da medida de eficiência θ é fornecido pelo primeiro estágio do problema de programação linear. Note que o segundo estágio também deve ser aplicado a cada DMU da amostra. Por exemplo, para a DMU 1 do exemplo anterior, o segundo estágio do problema é dado por:

$$\begin{split} MIN & - \left(O_{S1} + I_{S11} + I_{S21}\right), \\ sujeito \ a: \\ & - y_1 + \left(y_1\lambda_1 + y_2\lambda_2 + y_3\lambda_3 + y_4\lambda_4 + y_5\lambda_5 + y_6\lambda_6\right) - O_{S1} = 0, \\ & \theta x_{11} - \left(x_{11}\lambda_1 + x_{12}\lambda_2 + x_{13}\lambda_3 + x_{14}\lambda_4 + x_{15}\lambda_5 + x_{16}\lambda_6\right) - I_{S11} = 0, \\ & \theta x_{21} - \left(x_{11}\lambda_1 + x_{12}\lambda_2 + x_{13}\lambda_3 + x_{14}\lambda_4 + x_{15}\lambda_5 + x_{16}\lambda_6\right) - I_{S21} = 0, \\ & \left(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_5, \lambda_6, O_{S1}, I_{S11}, I_{S21}\right) \geq 0 \end{split}$$

Substituindo os valores da Tabela 1, tem-se:

MIN
$$-(O_{S1} + I_{S11} + I_{S21})$$
,
sujeito a:
 $-2 + (2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 2\lambda_4 + 1\lambda_5 + 3\lambda_6) - O_{S1} = 0$,
 $(0,6667 \times 6) - (6\lambda_1 + 8\lambda_2 + 12\lambda_3 + 6\lambda_4 + 6\lambda_5 + 15\lambda_6) - I_{S11} = 0$,
 $(0,6667 \times 14) - (6\lambda_1 + 8\lambda_2 + 12\lambda_3 + 6\lambda_4 + 6\lambda_5 + 15\lambda_6) - I_{S21} = 0$,
 $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_5, \lambda_6, O_{S1}, I_{S11}, I_{S21}) \ge 0$

Uma vez obtidas as folgas, pode-se projetar o ponto ineficiente da i-ésima DMU para a fronteira eficiente, da seguinte forma:

$$y_i^* = y_i + O_S, x_i^* = \theta x_i - I_S,$$

em que o asterisco indica o ponto projetado.

Ao utilizar o modelo com dois estágios, a solução de interesse no primeiro estágio é apenas a obtenção de θ ótimo, que indica a redução proporcional de insumos máxima ou a eficiência técnica. No segundo estágio, seriam calculados o vetor λ e as folgas, possibilitando-se a projeção do ponto para a fronteira eficiente.

Segundo Coelli (1996), um problema associado à programação de dois estágios é que as unidades de medida não são invariantes, uma vez que, no segundo estágio, obtém-se uma medida para todas as folgas em conjunto. Assim, alterações nas unidades de medida de insumos e produtos poderiam resultar na identificação de diferentes faixas-limite de eficiência e, então, em diferentes valores para as folgas.

Para contornar esse problema, Coelli (1998) propôs a utilização de múltiplos estágios de programação, para obtenção do ponto exato das DMUs ineficientes, na fronteira eficiente. O método consiste na estimação da eficiência em seis passos, cujo ponto final projetado (chamado de "ponto mais representativo da eficiência") será invariante para as unidades de medidas escolhidas. A idéia básica do modelo é realizar uma série de movimentos radiais para obter o ponto eficiente, em vez de apenas um. Para maiores detalhes sobre o método, consulte os comentários feitos por Cherchye e Puyenbroeck (2001).

Uma vez que as folgas ocorrem pela própria forma como a fronteira é obtida, isto é, o uso de amostras finitas sempre irá gerar faixas na fronteira que são paralelas aos eixos, seu tratamento varia entre os pesquisadores. Autores como Ferrier e Lovell (1990) consideram a existência de folgas como sendo uma ineficiência alocativa, visto

que reflete uma combinação de insumos inadequada. Färe et al. (1994) tratam as folgas como um tipo separado de ineficiência, distinto das ineficiências técnica e alocativa, ambas obtidas por movimentos radiais. Para uma discussão mais detalhada sobre a questão das folgas, veja, por exemplo, Cooper et al. (2001), Sengupta (1998) e Cooper e Tone (1997).

Independentemente de qual seja a forma como serão tratadas as folgas, o importante é considerá-las na análise. Vale ressaltar que esse tratamento varia entre os diversos *softwares* utilizados para computar as medidas de eficiência.

2.2. Modelo com retornos variáveis

O modelo de retornos constantes à escala pode ser reformulado, com o objetivo de possibilitar retornos variáveis às DMUs analisadas. Essa proposta foi inicialmente feita por Banker et al. (1984), cujo modelo ficou conhecido como BCC, devido às iniciais dos nomes dos autores. A idéia é introduzir uma restrição de convexidade ao modelo CCR apresentado em (3). O modelo BCC, que pressupõe retornos variáveis à escala, pode ser representado pela seguinte notação algébrica:

$$MIN_{\theta,\lambda} \quad \theta,$$
 $sujeito \ a:$

$$-y_i + Y\lambda \ge 0,$$

$$\theta x_i - X\lambda \ge 0,$$

$$N_1 \lambda = 1,$$

$$\lambda \ge 0,$$
(5)

em que N₁ é um vetor (n x 1) de algarismos unitários (1,...,1). Essa abordagem forma uma superfície convexa de planos em interseção, a qual envolve os dados de forma mais compacta do que a superfície formada pelo modelo com retornos constantes. Com isso, os valores obtidos para eficiência técnica, com a pressuposição de retornos variáveis, são maiores do que aqueles obtidos com retornos constantes. Se uma DMU é eficiente no modelo CCR, então ela também é eficiente no modelo BCC. Isso ocorre porque a medida de eficiência técnica, obtida no modelo com retornos constantes, é composta pela medida de eficiência técnica no modelo com retornos variáveis, também chamada de pura eficiência técnica, pela medida de eficiência de escala e pela medida de

congestão (Seiford e Zhu, 1999; Färe e Grosskopf, 2000). Essa decomposição da medida de eficiência será apresentada a seguir.

2.2.1. Eficiência de escala

O uso da especificação de retornos constantes, quando nem todas as DMUs estão operando em escala ótima, resultará em medidas de eficiência técnica que podem ser confundidas com eficiência de escala. Para separar essas medidas, realiza-se o procedimento, conduzindo ambos, retornos constantes e variáveis, ao mesmo conjunto de dados. Se existir uma diferença nos valores de eficiência técnica para uma DMU qualquer, isso indica que esta DMU tem ineficiência de escala, que pode ser calculada pela diferença entre os valores das eficiências técnicas com retornos variáveis e com retornos constantes.

A Figura 3 ilustra uma situação que envolve um insumo e um produto. Podem-se traçar as fronteiras eficientes calculadas pela DEA, isto é, a fronteira obtida com retornos constantes (RC) e a obtida com retornos variáveis (RV), sendo esta última descrita pela linha pontilhada. Note que a fronteira com retornos variáveis é composta por uma parte de retornos crescentes (iniciando no ponto B até o ponto C, sendo este último o único ponto de retornos constantes da curva) e outra parte de retornos decrescentes.

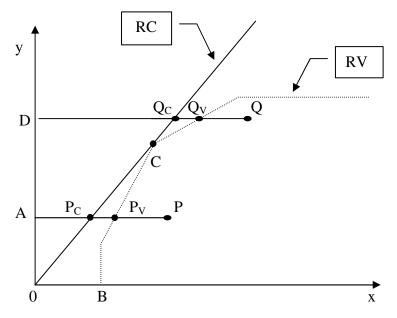


Figura 3 - Eficiência técnica e eficiência de escala.

Considere o ponto P na Figura 3. Sob a pressuposição de retornos constantes, a ineficiência técnica do ponto P é dada pela distância PP_C, enquanto a ineficiência técnica é dada pela distância PP_V, sob a pressuposição de retornos variáveis. A diferença entre essas duas, P_CP_V, fornece a ineficiência de escala. As medidas de eficiência do ponto P, em termos de razão, isto é, limitadas entre zero e um, são dadas por:

$$ET_{I,RC} = AP_C/AP,$$

$$ET_{I,RV} = AP_V/AP,$$

$$EE_I = AP_C/AP_V,$$

em que ET é a medida de eficiência técnica; EE, a medida de eficiência de escala; o subscrito I indica modelos com orientação insumo; RC, retornos constantes; e RV, retornos variáveis.

Como $AP_C/AP = (AP_V/AP)x(AP_C/AP_V)$, então $ET_{I,RC} = ET_{I,RV}$ x EE_I , isto é, a medida de eficiência técnica com retornos constantes à escala é composta pela eficiência técnica pura e pela eficiência de escala.

Considere, agora, o ponto Q. De forma semelhante, as medidas de eficiência são dadas por:

$$ET_{I,RC} = DQ_C/DQ,$$

 $ET_{I,RV} = DQ_V/DQ,$
 $EE_I = DQ_C/DQ_V,$

uma vez que $ET_{I,RC} = ET_{I,RV} \times EE_{I}$.

Apesar de a forma como são calculadas as medidas de eficiência ser semelhante, percebe-se que o ponto P está na faixa de retornos crescentes, enquanto o ponto Q está na faixa de retornos decrescentes. O modelo na forma como é apresentado em (5) não permite essa distinção, ou seja, sabe-se apenas que, se a medida de eficiência de escala for igual a um, a firma estará operando com retornos constantes à escala. No entanto, se for menor que um, poderão ocorrer retornos crescentes ou decrescentes. Em outras palavras, mesmo que seja evidenciada a existência de ineficiência de escala, ainda não se sabe qual a natureza dessa ineficiência, isto é, se ela se deve a retornos crescentes ou a retornos decrescentes à escala.

Para contornar essa situação, é necessário formular outro problema da programação, impondo a pressuposição de retornos não-crescentes ou não-decrescentes. Considerando o caso de retornos não-crescentes, a formulação consiste em substituir a restrição $N_1\lambda=1$, em (5), pela restrição $N_1\lambda\leq 1$, obtendo-se o seguinte modelo:

$$MIN_{\theta,\lambda} \quad \theta,$$
 $sujeito \ a:$

$$-y_i + Y\lambda \ge 0,$$

$$\theta x_i - X\lambda \ge 0,$$

$$N_1 \lambda \le 1,$$

$$\lambda \ge 0.$$
(6)

A fronteira obtida para o modelo com retornos não-crescentes pode ser observada na Figura 3. Nota-se que ela é composta, inicialmente, por uma faixa da fronteira com retornos constantes, com origem em 0 até o ponto C, e depois por uma faixa da fronteira de retornos variáveis (decrescentes). Para determinar a natureza da escala de uma DMU qualquer, basta verificar se o coeficiente de eficiência técnica no modelo com retornos não-crescentes (RNC) é igual ao do modelo com retornos variáveis (RV). Se forem diferentes, então a DMU terá retornos crescentes à escala. Se forem iguais, ocorrerá uma situação de retornos decrescentes, isto é:

Se
$$ET_{RNC} = ET_{RV} \Rightarrow$$
 retornos decrescentes,
Se $ET_{RNC} \neq ET_{RV} \Rightarrow$ retornos crescentes.

De forma alternativa, pode-se formular um problema de programação impondo a pressuposição de retornos não-decrescentes à escala. Para isso, basta substituir a restrição $N_1\lambda \leq 1$, no modelo com retornos não-crescentes, pela restrição $N_1\lambda \geq 1$. O modelo com retornos não-decrescentes à escala seria dado por:

$$MIN_{\theta,\lambda} \quad \theta,$$
 $sujeito \ a:$

$$-y_i + Y\lambda \ge 0,$$

$$\theta x_i - X\lambda \ge 0,$$

$$N_1 \lambda \ge 1,$$

$$\lambda \ge 0.$$
(7)

Na Figura 3, essa fronteira inicia-se com uma faixa de retornos crescentes (segmento BC) e continua com uma faixa de retornos constantes. Para identificar se a firma está operando com retornos crescentes ou decrescentes, basta comparar o resultado encontrado para eficiência técnica, no modelo com retornos variáveis (RV), com aquele encontrado no modelo com retornos não-decrescentes (RND), ou seja:

Se
$$ET_{RND} = ET_{RV} \rightarrow$$
 retornos crescentes,
Se $ET_{RND} \neq ET_{RV} \rightarrow$ retornos decrescentes.

Vale ressaltar, de novo, que esses modelos devem ser aplicados n vezes, ou seja, um modelo para cada DMU da amostra. Considere o exemplo anterior, cujos dados são apresentados na Tabela 1. Para a DMU 5, o modelo com retornos variáveis seria dado por:

MIN
$$\theta$$
,
sujeito a :

$$-1 + (2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 2\lambda_4 + 1\lambda_5 + 3\lambda_6) \ge 0,$$

$$6\theta - (6\lambda_1 + 8\lambda_2 + 12\lambda_3 + 6\lambda_4 + 6\lambda_5 + 15\lambda_6) \ge 0,$$

$$3\theta - (14\lambda_1 + 16\lambda_2 + 12\lambda_3 + 4\lambda_4 + 3\lambda_5 + 3\lambda_6) \ge 0,$$

$$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 = 1$$

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) \ge 0$$

Para modelar os outros modelos, basta alterar a terceira restrição, da seguinte forma:

Retornos não-crescentes
$$\Rightarrow$$
 substituir por $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 \le 1$; e Retornos não-decrescentes \Rightarrow substituir por $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 \ge 1$.

Calculando os modelos para todas as DMUs da amostra, obtiveram-se os seguintes resultados, apresentados na Tabela 4.

Tabela 4 - Resultados dos modelos com retornos constantes, variáveis, não-crescentes e não-decrescentes

DMU	Med	ida de eficiê	Eficiência	Tipo de		
DMU	RC	RV	RNC	RND	de escala	retorno
1	0,6667	1,0000	0,6667	1,0000	0,6667	Crescente
2	1,0000	1,0000	1,0000	1,0000	1,0000	Constante
3	0,6667	0,7167	0,7167	0,6667	0,9302	Decrescente
4	1,0000	1,0000	1,0000	1,0000	1,0000	Constante
5	0,5833	1,0000	0,5833	1,0000	0,5833	Crescente
6	1,0000	1,0000	1,0000	1,0000	1,0000	Constante

Fonte: Resultados da pesquisa.

RC = retornos constantes; RV = retornos variáveis; RNC = retornos não-crescentes; e RND = retornos não-decrescentes.

Conforme visto anteriormente, as DMUs 2, 4 e 6 apresentaram máxima eficiência no modelo com retornos constantes à escala. Com isso, suas medidas de eficiência, obtidas no modelo com retornos variáveis, são também máximas e, conseqüentemente, estão operando sem problema de escala. O mesmo não ocorre com as DMUs 1, 3 e 5.

Analisando as unidades ineficientes, percebe-se que a ineficiência da DMU 1 é exclusivamente devido à escala incorreta, uma vez que sua medida de eficiência com retornos variáveis é máxima. Além disso, para esta DMU, $ET_{RV} = ET_{RND} \neq ET_{RNC}$, o que permite dizer que ela está operando na faixa de retornos crescentes. Situação semelhante ocorre com a DMU 5. Já a DMU 3 apresentou resultados diferentes das demais ineficientes. Note que, para esta DMU, $ET_{RV} = ET_{RNC} \neq ET_{RND}$, o que implica retornos decrescentes à escala.

2.2.2. Congestão de insumos

Pode-se, ainda, definir uma medida de eficiência que consiga captar a existência de congestão em algum insumo, isto é, uma DMU pode utilizar determinada quantidade de insumos que faça com que a produção decresça. Está-se, então, no terceiro estágio da

função de produção, com retornos negativos. A Figura 4 ilustra essa situação, considerando-se a utilização de um insumo e um produto. A linha pontilhada descreve a fronteira obtida com retornos variáveis (RV), e a linha contínua descreve a fronteira com retornos constantes. O ponto R localiza-se na faixa de retornos negativos da curva, ou seja, há certa congestão no uso do insumo x.

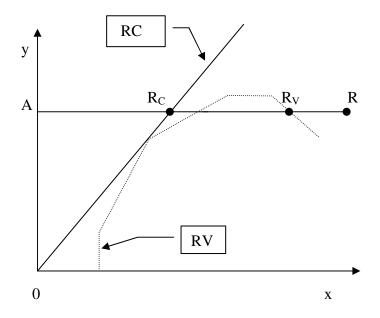


Figura 4 - Eficiência técnica e congestão de insumo.

Sob a pressuposição de retornos constantes, a eficiência técnica do ponto R é dada pela distância RR_C . Essa distância é composta, por uma parte, devido à pura ineficiência técnica ou à ineficiência com retornos variáveis (R_VR) e, por outra parte, devido à ineficiência de escala (R_CR_V) .

Entretanto, o ponto R localiza-se na faixa de retornos negativos à escala, e não na faixa de retornos decrescentes. Para determinar se há congestão em algum insumo, é necessário comparar dois modelos: um modelo sob disponibilidade forte (ou plena) e outro sob disponibilidade fraca. Segundo Souza et al. (1997), a disponibilidade forte ocorre quando se usam mais insumos e consegue-se produzir, pelo menos, a mesma quantidade de produto que se tinha anteriormente. Por outro lado, a disponibilidade fraca permite que o nível de produto diminua com o aumento no uso dos insumos. Note que em todos os modelos vistos até aqui foi assumida a hipótese de disponibilidade forte.

A medida de eficiência técnica para a i-ésima DMU, pressupondo-se retornos constantes e disponibilidade fraca, é dada por:

$$MIN_{\theta_{G},\lambda}$$
 θ_{G} ,

 $sujeito\ a$:
$$-y_{i} + Y\lambda \geq 0,$$

$$\theta_{G}x_{i} - X\lambda = 0,$$

$$\lambda \geq 0,$$
(8)

em que θ_G é a medida de eficiência no modelo com disponibilidade fraca. A única diferença deste modelo para aquele apresentado em (3) encontra-se na igualdade verificada na segunda restrição. Esse problema também deve ser executado n vezes, ou seja, uma vez para cada DMU da amostra.

A existência de congestão ocorrerá quando a medida de eficiência encontrada neste modelo for diferente daquela obtida no modelo com disponibilidade forte, ou seja, a medida de congestão nos insumos é dada por:

Congestão =
$$\frac{\theta}{\theta_G}$$

Para maiores detalhes sobre congestão de insumos na DEA, consulte, por exemplo, Färe e Grosskopf (1998), Färe e Grosskopf (2001), Cooper et al. (2000), Cooper et al. (2001), McDonald (1996) e Sueyoshi (2003).

Aplicando esses modelos aos dados apresentados na Tabela 1, chega-se aos resultados apresentados na Tabela 5.

Das unidades ineficientes, apenas a DMU 1 está localizada na fronteira que admite disponibilidade fraca. As DMUs 3 e 5, apesar de serem ineficientes, não estão operando com congestão de insumos.

Tabela 5 - Análise da congestão utilizando-se modelos com disponibilidades forte e fraça

DMU	ET_{RC}	$ET_{RC,F}$	Congestão
1	0,6667	1,0000	0,6667
2	1,0000	1,0000	1,0000
3	0,6667	0,6667	1,0000
4	1,0000	1,0000	1,0000
5	0,5833	0,5833	1,0000
6	1,0000	1,0000	1,0000

Fonte: Resultados da pesquisa.

O subscrito F indica o modelo com disponibilidade fraca.

2.3. Eficiência econômica e eficiência alocativa

Se houver informações sobre preços dos insumos, pode-se, após obter a eficiência técnica, calcular a eficiência econômica da DMU, também conhecida como eficiência custo. A idéia consiste em obter quantidades ótimas de insumos que minimizem o custo de produção. O problema de minimização de custos para a i-ésima DMU, pressupondo-se retornos constantes à escala, é dado por:

$$MIN_{\lambda, x_{i}^{E}} \quad w_{i} x_{i}^{E},$$

$$sujeito \ a:$$

$$-y_{i} + Y\lambda \geq 0,$$

$$x_{i}^{E} - X\lambda \geq 0,$$

$$\lambda \geq 0,$$
(9)

em que w_i é o vetor de preço dos insumos para a i-ésima DMU e x_i^E é o vetor de insumos que minimiza os custos. Embora o problema de minimização de custos seja semelhante àquele utilizado para calcular a medida de eficiência técnica, existem duas diferenças significativas: a primeira é relacionada à função objetivo, ou seja, no modelo de minimização de custos não se está interessado em calcular a medida de eficiência, mas sim identificar o menor custo possível dos insumos. A segunda diferença ocorre na restrição dos insumos. Enquanto no modelo utilizado para medir a eficiência os insumos são tomados como fixos, na minimização de custos as quantidades dos insumos podem

variar, isto é, representam variáveis a serem obtidas pelo modelo. Isso significa que a composição de insumos utilizados pela DMU pode variar, dependendo da relação de preços dos insumos.

Obtida a composição ideal de insumos para a DMU e conhecendo os preços destes insumos, obtém-se o custo mínimo de produção. A medida de eficiência econômica (EE), para a i-ésima DMU, é dada pela razão entre o custo mínimo e o custo observado, ou seja:

$$EE = \frac{\overrightarrow{w_i} x_i^E}{\overrightarrow{w_i} x_i}.$$

Uma vez identificada a eficiência econômica (EE) e conhecendo-se a eficiência técnica (ET), pode-se calcular, residualmente, a eficiência alocativa (EA) da seguinte forma: EA = EE/ET.

Considere o exemplo apresentado na Tabela 1. Para facilitar, suponha que ambos os preços dos insumos sejam 1 ($P_{x1} = 1$ e $P_{x2} = 1$) para todas as DMUs. Pode-se, então, calcular o custo total da produção, obtendo os dados apresentados na Tabela 6. Vale ressaltar que o modelo admite que os preços possam variar de DMU para DMU.

Tabela 6 - Produto, insumos e custo total de seis DMUs

DMU	Y	\mathbf{x}_1	\mathbf{x}_2	P_{x1}	P_{x2}	Custo total
1	2	6	14	1	1	20
2	4	8	16	1	1	24
3	3	12	12	1	1	24
4	2	6	4	1	1	10
5	1	6	3	1	1	9
6	3	15	3	1	1	18

Fonte: Dados fictícios.

Uma vez que a relação de preços é de um para um, pode-se traçar uma isocusto, como a representada pela linha tracejada na Figura 5.

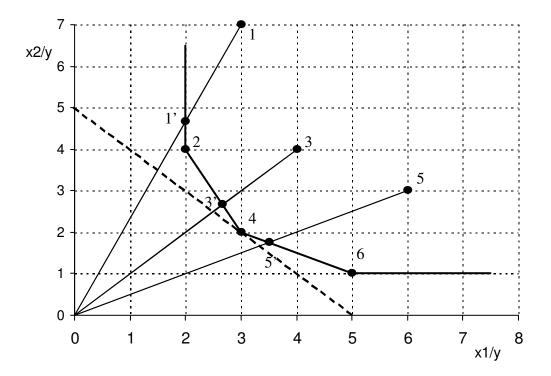


Figura 5 - Isoquanta eficiente e isocusto.

Percebe-se que, apesar de as DMUs 2, 4 e 6 serem tecnicamente eficientes, apenas a DMU 4 possui eficiência econômica. Isso significa que as DMUs 2 e 6 possuem certa ineficiência alocativa. A idéia do modelo consiste em obter valores ótimos de x_1 e x_2 para as DMUs, de forma a minimizar o custo. Por exemplo, para a DMU 5, o modelo assume a seguinte forma:

MIN
$$P_{x_1}x_{15}^E + P_{x_2}x_{25}^E$$
,
sujeito a :
 $-y_5 + (y_1\lambda_1 + y_2\lambda_2 + y_3\lambda_3 + y_4\lambda_4 + y_5\lambda_5 + y_6\lambda_6) \ge 0$,
 $x_{15}^E - (x_{11}\lambda_1 + x_{12}\lambda_2 + x_{13}\lambda_3 + x_{14}\lambda_4 + x_{15}\lambda_5 + x_{16}\lambda_6) \ge 0$,
 $x_{25}^E - (x_{21}\lambda_1 + x_{22}\lambda_2 + x_{23}\lambda_3 + x_{24}\lambda_4 + x_{25}\lambda_5 + x_{26}\lambda_6) \ge 0$,
 $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) \ge 0$

Substituindo os dados da Tabela 6, o problema para obtenção da medida de eficiência para a DMU 5 fica:

MIN
$$1x_{15}^{E} + 1x_{25}^{E}$$
,
sujeito a :
 $-1 + (2\lambda_{1} + 4\lambda_{2} + 3\lambda_{3} + 2\lambda_{4} + 1\lambda_{5} + 3\lambda_{6}) \ge 0$,
 $x_{15}^{E} - (6\lambda_{1} + 8\lambda_{2} + 12\lambda_{3} + 6\lambda_{4} + 6\lambda_{5} + 15\lambda_{6}) \ge 0$,
 $x_{25}^{E} - (14\lambda_{1} + 16\lambda_{2} + 12\lambda_{3} + 4\lambda_{4} + 3\lambda_{5} + 3\lambda_{6}) \ge 0$,
 $(\lambda_{1}, \lambda_{2}, \lambda_{3}, \lambda_{4}, \lambda_{5}, \lambda_{6}) \ge 0$

Executando o problema de minimização de custos para todas as DMUs, obtêmse os valores ótimos para os insumos apresentados na Tabela 7. Note que, como a DMU 4 é a única que possui máxima eficiência econômica, as quantidades dos insumos calculadas para as demais DMUs devem ter a mesma proporção da verificada para aquela DMU eficiente. Isso significa que, como a DMU 4 utiliza seis unidades de x₁ e quatro unidades de x₂, ou seja, uma proporção de 3 para 2, após estimar as quantidades ótimas de insumo que as DMUs ineficientes deveriam utilizar, essas quantidades devem seguir essa mesma proporção.

Tabela 7 - Resultados do modelo de minimização de custos e medidas de eficiência

Eficiência técnica	Eficiência alocativa
0.6667	0,7500
· ·	,
	0,8333
0,6667	0,9375
1,0000	1,0000
0,5833	0,9525
1,0000	0,8333
	0,6667 1,0000 0,6667 1,0000 0,5833

Fonte: Resultados da pesquisa.

Conhecendo os valores ótimos para os insumos, calcula-se o custo mínimo de produção, permitindo, assim, identificar a eficiência econômica. A eficiência alocativa é obtida pela razão entre as eficiências econômica e técnica, esta última previamente determinada. Para a DMU 5, por exemplo, a sequência é a seguinte:

Custo mínimo =
$$(1 \times 3) + (1 \times 2) = 5$$

Ef. econômica = $\frac{Custo mínimo}{Custo observado} = \frac{5}{9} = 0,5556$
Ef. alocativa = $\frac{Ef. econômica}{Ef. técnica} = \frac{0,5556}{0,5833} = 0,9525$

2.4. Modelos com orientação produto

Os modelos discutidos anteriormente buscam identificar a ineficiência técnica das DMUs, mediante redução proporcional na utilização dos insumos, isto é, são modelos com orientação insumo. Entretanto, podem-se também obter medidas de eficiência com o aumento proporcional na produção, conhecidas como medidas de eficiência com orientação produto.

Tais medidas permitem verificar qual a quantidade máxima que pode ser produzida, dada uma quantidade fixa de insumos. Assim, ao contrário dos modelos insumo orientados, em que a idéia é reduzir insumos, na orientação produto os insumos são fixos e o produto pode ser expandido.

Apesar da diferença de princípio de orientação, os modelos possuem grande similaridade. De forma semelhante aos modelos com orientação insumo, as medidas de eficiência nos modelos produto orientados podem ser decompostas em técnica, alocativa e econômica, sendo nesta última requeridas informações sobre os preços dos produtos.

Na maioria das vezes, a escolha da orientação não exerce grande influência na magnitude do valor de eficiência técnica. Além disso, o critério de escolha dependerá da finalidade do trabalho, isto é, se a decisão primária for regularizar a utilização dos insumos ou possibilitar o aumento da produção.

O problema com orientação produto, pressupondo-se retornos constantes e disponibilidade forte, pode ser escrito da seguinte forma:

$$MAX_{\phi,\lambda} \quad \phi,$$
 $sujeito \ a:$

$$-\phi y_i + Y\lambda \ge 0,$$

$$x_i - X\lambda \ge 0,$$

$$\lambda \ge 0,$$
(10)

em que $1 \le \phi < \infty$ e ϕ - 1 é o aumento proporcional nos produtos que poderiam ser obtidos pela i-ésima DMU, mantendo-se constante a utilização de insumos. A medida de eficiência técnica seria dada por $1/\phi$, que varia de zero a um. Adicionando a restrição de convexidade $N_1\lambda=1$, obtém-se um modelo de orientação produto com retornos variáveis. Conseqüentemente, alterando essa restrição para $N_1\lambda \le 1$, tem-se o modelo com retornos não-crescentes e, alterando para $N_1\lambda \ge 1$, tem-se o modelo com retornos não-decrescentes.

Os modelos com orientação produto serão mostrados a seguir, na mesma seqüência em que foram apresentados os modelos com orientação insumo. Para facilitar o entendimento, os modelos serão exemplificados utilizando-se os dados da DMU 5, contidos na Tabela 1.

Retornos constantes	Para a DMU 5:
$MAX_{\phi,\lambda}$ ϕ ,	$MAX \phi,$
sujeito a :	sujeito a :
$-\phi y_i + Y\lambda \ge 0,$	$-1\phi + (2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 2\lambda_4 + 1\lambda_5 + 3\lambda_6) \ge 0,$
$x_i - X\lambda \ge 0,$	$6 - (6\lambda_1 + 8\lambda_2 + 12\lambda_3 + 6\lambda_4 + 6\lambda_5 + 15\lambda_6) \ge 0,$
$\lambda \ge 0$	$3 - (14\lambda_1 + 16\lambda_2 + 12\lambda_3 + 4\lambda_4 + 3\lambda_5 + 3\lambda_6) \ge 0,$
	$(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6)' \ge 0$

Retornos variáveis	Para a DMU 5:
$MAX_{\phi,\lambda}$ ϕ ,	$MAX \phi$
sujeito a:	sujeito a:
$-\phi y_i + Y\lambda \ge 0,$	$-1\phi + (2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 2\lambda_4 + 1\lambda_5 + 3\lambda_6) \ge 0,$
$x_i - X\lambda \ge 0$,	$6 - (6\lambda_1 + 8\lambda_2 + 12\lambda_3 + 6\lambda_4 + 6\lambda_5 + 15\lambda_6) \ge 0,$
$N_1 \lambda = 1,$	$3 - (14\lambda_1 + 16\lambda_2 + 12\lambda_3 + 4\lambda_4 + 3\lambda_5 + 3\lambda_6) \ge 0,$
$\lambda \geq 0$	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 = 1$
	$(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) \ge 0$

Retornos não-crescentes $MAX_{\phi,\lambda} \quad \phi,$ $sujeito \ a:$ $-\phi y_i + Y\lambda \ge 0,$ $x_i - X\lambda \ge 0,$ $N_1 \ \lambda \le 1,$ $\lambda \ge 0$

Para a DMU 5:
MAX
$$\phi$$
,
sujeito a :
 $-1\phi + (2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 2\lambda_4 + 1\lambda_5 + 3\lambda_6) \ge 0$,
 $6 - (6\lambda_1 + 8\lambda_2 + 12\lambda_3 + 6\lambda_4 + 6\lambda_5 + 15\lambda_6) \ge 0$,
 $3 - (14\lambda_1 + 16\lambda_2 + 12\lambda_3 + 4\lambda_4 + 3\lambda_5 + 3\lambda_6) \ge 0$,
 $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 \le 1$
 $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) \ge 0$

Retornos não-decrescentes
$$MAX_{\phi,\lambda}$$
 ϕ , sujeito a :
$$-\phi y_i + Y\lambda \ge 0,$$

$$x_i - X\lambda \ge 0,$$

$$N_1 \lambda \ge 1,$$

$$\lambda \ge 0$$

Para a DMU 5:
MAX
$$\phi$$
,
sujeito a:
 $-1\phi + (2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 2\lambda_4 + 1\lambda_5 + 3\lambda_6) \ge 0$,
 $6 - (6\lambda_1 + 8\lambda_2 + 12\lambda_3 + 6\lambda_4 + 6\lambda_5 + 15\lambda_6) \ge 0$,
 $3 - (14\lambda_1 + 16\lambda_2 + 12\lambda_3 + 4\lambda_4 + 3\lambda_5 + 3\lambda_6) \ge 0$,
 $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 \ge 1$
 $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) \ge 0$

Retornos constantes
Disponibilidade fraca

$$MAX_{\phi_G,\lambda} \quad \phi_G,$$
 $sujeito \ a:$

$$-\phi_G y_i + Y\lambda = 0,$$

$$x_i - X\lambda \ge 0,$$

$$\lambda \ge 0$$

Para a DMU 5:

$$MAX \quad \phi_{G}$$
,
 $sujeito \ a$:
 $-1\phi_{G} + (2\lambda_{1} + 4\lambda_{2} + 3\lambda_{3} + 2\lambda_{4} + 1\lambda_{5} + 3\lambda_{6}) = 0$,
 $6 - (6\lambda_{1} + 8\lambda_{2} + 12\lambda_{3} + 6\lambda_{4} + 6\lambda_{5} + 15\lambda_{6}) \ge 0$,
 $3 - (14\lambda_{1} + 16\lambda_{2} + 12\lambda_{3} + 4\lambda_{4} + 3\lambda_{5} + 3\lambda_{6}) \ge 0$,
 $(\lambda_{1}, \lambda_{2}, \lambda_{3}, \lambda_{4}, \lambda_{5}, \lambda_{6}) \ge 0$

Maximização da receita
Retornos constantes MAX_{λ,y_i^E} $r_i^{'}y_i^E$, $sujeito\ a$: $-y_i^E + Y\lambda \ge 0$, $x_i - X\lambda \ge 0$, $\lambda \ge 0$ em que r_i é o vetor de preço dos produtos para a i-ésima

DMU; e y_i^E é o vetor de produtos que maximiza a receita.

Considerando $P_{y} = 2$, para a DMU 5: $MAX = 2y_{5}^{E}$, $sujeito \ a$: $-y_{5}^{E} + (2\lambda_{1} + 4\lambda_{2} + 3\lambda_{3} + 2\lambda_{4} + 1\lambda_{5} + 3\lambda_{6}) \ge 0,$ $6 - (6\lambda_{1} + 8\lambda_{2} + 12\lambda_{3} + 6\lambda_{4} + 6\lambda_{5} + 15\lambda_{6}) \ge 0,$ $3 - (14\lambda_{1} + 16\lambda_{2} + 12\lambda_{3} + 4\lambda_{4} + 3\lambda_{5} + 3\lambda_{6}) \ge 0,$ $(\lambda_{1}, \lambda_{2}, \lambda_{3}, \lambda_{4}, \lambda_{5}, \lambda_{6})' \ge 0$

3. Considerações finais

Com o interesse crescente da análise de eficiência nos mais diversos setores da economia, cresceu também o uso da técnica não-paramétrica conhecida como análise envoltória de dados. Entretanto, como qualquer instrumental analítico, a DEA apresenta algumas limitações de uso que devem ser observadas. Por se tratar de uma técnica determinística, e não estatística, possui elevada sensibilidade a observações extremas (outliers). A existência de apenas uma observação discrepante na amostra influenciará todas as outras medidas de eficiência. É necessário que o pesquisador analise previamente os dados que serão utilizados nos cálculos das medidas de eficiência. Caso contrário, certamente obterá resultados não-satisfatórios, os quais podem passar despercebidos.

Outra limitação encontra-se no fato de que a DEA é um instrumental matemático que analisa a eficiência relativa das DMUs, mas não a eficiência absoluta. Isso significa que a medida de eficiência encontrada é relacionada às outras unidades da amostra, não podendo fazer inferências a um máximo teórico nem comparar resultados entre estudos distintos. Por fim, cabe destacar que, sendo a DEA uma técnica não-paramétrica, o uso de testes de hipóteses estatísticas é extremamente difícil.

A despeito das limitações, a DEA é uma ferramenta muito útil para analisar a eficiência de unidades. É preciso que se tenha consciência de suas restrições, mas também é importante conhecer suas vantagens em comparação com outros métodos. Vários autores reconhecem que a DEA é uma estrutura que oferece muitas vantagens em relação às medidas de eficiência tradicionais.

Dentre as vantagens, além daquelas citadas anteriormente, destaca-se a possibilidade de decompor a eficiência econômica em técnica e alocativa, além dos efeitos da escala. A DEA permite, ainda, encontrar quais foram as DMUs eficientes que foram as responsáveis por determinada DMU ter sido considerada ineficiente, indicando em quais variáveis devem ser realizadas modificações. Assim, a DEA não mede somente eficiência, mas também provém um guia para as DMUs eliminarem ineficiências, ou seja, a DMU ineficiente pode ter como referência seus pares, para tentar aumentar a eficiência na produção.

Outras características, como o fato de não ser necessária a especificação de nenhuma forma funcional entre insumos e produtos, além da facilidade de trabalhar com múltiplas variáveis, fizeram com que a DEA se tornasse o instrumental de análise de eficiência mais utilizado nos últimos anos.

4. Referências bibliográficas

ALI, A.A., SEIFORD, L.M. The mathematical approach to efficiency analysis. In: FRIED, H., LOVELL, C.A.K., SCHMIDT, S. (Eds.). **The measurement of productive efficiency: techniques and applications**. Oxford: Oxford University, 1993.

ALVES, E., GOMES, A.P. Medidas de eficiência na produção de leite. **Revista Brasileira de Economia**, v. 52, n. 1, p. 145-167, 1998.

BANKER, R.D., CHARNES, H., COOPER, W.W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. **Management Science**, v. 30, n. 9, p. 1078-1092, 1984.

BOWLIN, W. F., Measuring performance: an introduction to data envelopment analysis (DEA). **Journal of Cost Analysis & Management**, Fall, p. 3-27, 1998.

CHARNES, A., COOPER, W.W., LEWIN, A.Y., SEIFORD, L.M. **Data envelopment analysis: theory, methodology, and application**. Dordrecht: Kluwer Academic, 1994.

CHARNES, A., COOPER, W.W. Programming with linear fractional functions. **Naval Research Logistics**, v. 9, p. 181-185, 1962.

CHARNES, A., COOPER, W.W., RHODES, E. Measuring the efficiency of decision making units. **European Journal of Operational Research**, v. 2, n. 6, p. 429-444, 1978.

CHERCHYE, L., PUYENBROECK, T.V. A comment on multi-stage DEA methodology. **Operations Research Letters**, v. 28, Issue 2, p. 93-98, 2001.

COELLI, T.J. A guide to DEAP version 2.1: a data envelopment analysis program. Armidale, Austrália: University of New England, 1996. 49 p. (CEPA Working Papers 8).

COELLI, T.J. Multi-stage methodology for the solution of orientated dea models. **Operations Research Letters**, v. 23, Issues 3-5, p. 143-149, 1998.

COELLI, T.J., RAO, D.S.P., BATTESE, G.E. An introduction to efficiency and productivity analysis. Norwell: Kluwer Academic Publishers, 1998. 275 p.

COOPER, W.W., SEIFORD, L.M ZHU, J. A unified additive model approach for evaluating inefficiency and congestion with associated measures in DEA. **Socio-Economic Planning Sciences**, v. 34, n. 1, p. 1-25, 2000.

COOPER, W.W., SEIFORD, L.M ZHU, J. Slacks and congestion: response to a comment by R.F. are and S. Grosskopf. **Socio-Economic Planning Sciences**, v. 35, p. 205-215, 2001.

COOPER, W.W., SEIFORD, L.M., TONE, K. Data envelopment analysis: a comprehensive text with models, applications, references and DEA-Solver software. Norwell: Kluwer Academic Publishers, 2000. 318 p.

COOPER, W.W., TONE, K. Measures of inefficiency in data envelopment analysis and stochastic frontier estimation. **European Journal of Operational Research**, v. 99, p. 72-88, 1997.

FÄRE, R., GROSSKOPF, S. When can slacks be used to identify congestion? An answer to W.W. Cooper, L. Seiford and J. Zhu. **Socio-Economic Planning Sciences**, v. 35, n. 3, p. 217-221, 2001.

FÄRE, R., GROSSKOPF, S. Congestion: a note. **Socio-Economic Planning Sciences**, v. 32, n. 1, p. 21-23, 1998.

FÄRE, R., GROSSKOPF, S. **Reference guide to OnFront**. Sweden: Economic Measurement and Quality Corporation, 2000. 45 p.

FÄRE, R., GROSSKOPF, S., LOVELL, C.A.K. **Production frontiers**. New York: Cambridge University Press, 1994. 296 p.

FARREL, M.J. The measurement of productive efficiency. **Journal of the Royal Statistical Society**, Series A, part III, p. 253-290, 1957.

FORSUND, R.F., SARAFOGLOU, N. **On the origins of data envelopment analysis**. Oslo, Norway: Department of Economics, University of Oslo, 2000. (Memorandum, 24).

LINS, M.P.E., MEZA, L.A. Análise envoltória de dados e perspectivas de integração no ambiente de apoio à tomada de decisão. Rio de Janeiro: COPPE/UFRJ, 2000.

MARINHO, A. Estudo de eficiência em alguns hospitais públicos e privados com a geração de rankings. Rio de Janeiro: IPEA, 2001. 17 p. (Texto para discussão, 794).

McDONALD, J, A problem with the decomposition of technical inefficiency into scale and congestion components. **Management Science**, v. 42, n. 3, p. 473-474, 1996.

SEIFORD, L.M., THRALL, R.M. Recent developments in DEA: the mathematical approach to frontier analysis. **Journal of Econometrics**, v. 46, n. 1, p. 7-38, 1990.

SEIFORD, L.M., ZHU, J. An Investigation of returns to scale in data envelopment analysis. **Omega – The Journal of Management Science**, v. 27, n. 1, p. 1-11, 1999.

SENGUPTA J.K. New efficiency theory: extensions and new applications of data envelopment analysis. **International Journal of Systems Science**, v. 29, Issue 3, p. 255-265, 1998.

SOUZA, G.S., ALVES, E., ÁVILA, A.F.D., CRUZ, E.R. Produtividade e eficiência relativa de produção em sistemas de produção de pesquisa agropecuária. **Revista Brasileira de Economia**, v. 51, n. 3, p. 281-307, 1997.

SUEYOSHI, T. DEA implications of congestion. **Asia Pacific Management Review**, v. 8, n. 1, p. 59-70, 2003.