ISSN 1679-1614

DOIS CRITÉRIOS PARA AVALIAÇÃO DA EFICIÊNCIA TÉCNICA[1]

Daniel Pacífico Homem de Souza[2]
Eliseu Alves[3]

Resumo - Dois critérios foram utilizados para testar três métodos que estimam a eficiência técnica: o método Varian, a fronteira estocástica e a fronteira não-paramétrica, na literatura inglesa conhecida por "data envelopment analysis-DEA" e, em português, por encapsulamento de dados. A orientação pode ser para produto ou insumo, com vistas em mudar a combinação usada pelo produtor para outra, de acordo com algum critério de escolha. O trabalho se fixou em ambas as orientações. Dados de uma amostra de produtores de leite foram usados para avaliação dos métodos. O menor distúrbio na proposta de combinação de produtos ou insumos, em relação àquela usada pelos produtores, e o maior acréscimo da renda líquida constituem os critérios. Quando houver conflitos entre eles, encontrar-se-á na área de indecisão. Convém lembrar que, quanto menos diferente a combinação proposta for da original, maiores chances terá ela, como recomendação, de ser aceita pelo produtor. O método de Varian teve melhor desempenho que o DEA, em relação aos distúrbios, visto que aquele produziu, em relação a este método, menor acréscimo na renda líquida. Os distúrbios do método Varian foram maiores que os da fronteira estocástica, mas, em compensação, o incremento na renda líquida foi bem maior. Quando o ajustamento da fronteira estocástica for muito bom, certamente produzirá menores distúrbios no produto, no entanto, o mesmo pode não ocorrer em relação aos insumos, como é o caso deste trabalho. Os dados não rejeitaram a hipótese de c-racionalização; portanto, quanto aos custos, os produtores fizeram escolhas corretas, no sentido de que o agricultor A, que produziu mais que o agricultor B, não gastou menos que este. Os testes recém-desenvolvidos para o DEA, orientação produto, foram aplicados. Foi possível testar a hipótese de retorno constante versus variável. Os dados não rejeitaram a hipótese de retorno constante. A especificação meio normal para o termo do erro também não foi rejeitada, e intervalos de confiança foram estimados para os agricultores de eficiência técnica igual a 1 e para a média.

Palavras-chave: eficiência técnica, DEA, fronteira estocástica, método Varian, critérios de avaliação.

[1] O artigo decorre da tese, em nível de doutorado, do primeiro autor, submetida ao Departamento de Economia, Administração e Sociologia da ESALQ/USP.

[2] Daniel P. H. de Souza é estudante de doutorado no Departamento de Economia, Administração e Sociologia da ESALQ/USP.

[3] Eliseu Alves é pesquisador da Embrapa.

Recebido em 08/09/2003 Aceito em 30/09/2003

1. Introdução

O objetivo deste trabalho foi discutir dois critérios para avaliar três métodos que têm sido utilizados no estudo da eficiência técnica. Dois deles, o encapsulamento de dados (DEA) e a fronteira estocástica têm tido ampla aplicação e desenvolvimento teórico. A literatura é vastíssima, e não será revisada. Dois exemplos de aplicação recente estão em Richetti e Reis (2003) e em Pereira Filho e Ferreira Filho (2003). Quanto ao encapsulamento de dados, a teoria foi ampliada para permitir o teste de hipótese, quando a orientação for produto e um único produto for produzido. Mostra-se também como a função de produção é gerada (Souza, 2003).

O terceiro método foi desenvolvido por Varian em vários trabalhos, citados em Varian (1985), o qual foi estendido para incorporar insumos fixos (Ray e Bhadra, 1993) e para suportar a presença de risco (Chavas e Cox, 1993). Os trabalhos desses autores estão discutidos, em detalhes, em Alves (2000). Como é menos conhecido, o método de Varian é apresentado, com mais detalhes, para justificar a aplicação que se fará.

Na avaliação dos métodos foi utilizada uma amostra de 143 produtores de leite, cuja distribuição, por estado, está na Tabela 1.

Tabela 1 - Distribuição das observações por estado, 2002

Estado	Número de observações	(%)
Goiás	25	17,5
Minas Gerais	31	21,7
Paraná	19	13,3
Rio Grande do Sul	34	23,8
Santa Catarina	17	11,9
São Paulo	17	11,9
Total	143	100

Fonte: CEPEA-ESALQ/USP.

A coleta dos dados foi feita pela aplicação de questionários. Foram registrados dados quantitativos, como também informações sobre o processo de produção de leite. Os critérios de seleção dos produtores foram produção média diária, acima de 150 litros/dia, e comercialização do produto, em laticínios. Todos os entrevistados receberam assistência técnica, disponibilizada pelo laticínio que compra a produção.

Na Tabela 2, a seguir, encontra-se a distribuição diária de leite e verifica-se que a amostra cobre amplo número de produtores, o que é interessante para esse tipo de estudo. A menor produção correspondeu à produção diária de 191 litros e a máxima, a 5.954 litros. Cerca de 25% da amostra produzia mais de 1.020 litros por dia; e cerca de 25% da amostra produzia, diariamente, menos de 362 litros. Portanto, 50% dos amostrados estavam no intervalo (362, 1.020) de produção diária

Tabela 2 – Produção de leite, em litros por dia, 2002

Itens	Brasil
Número de observações	143
Média	858
Coeficiente de Variação (%)	108,62
Mediana	559
Máximo	5954
Mínimo	191
Separatrizes dos quartis	[191, 362)
	[362, 559)
	[559, 1020)
	[1020, 5954]

Fonte dos dados: CEPEA-ESALQ/USP.

Quanto ao encapsulamento de dados, foi usado o *software* DEAP v 2.1 e, quanto à fronteira estocástica, o software front v 4.1-xp (Coelli et al., 1998). Os métodos determinam um "ótimo" para insumos ou produtos, o que, raramente, coincide com o valor observado. A soma dos quadrados das diferenças dos "ótimos", em relação aos valores observados, é uma medida do distúrbio, também denominada de perturbação. O termo perturbação é de Varian (1985), e a idéia tem origem nos mínimos quadrados, portanto, é velha. Como será explicado, o método Varian exige a solução de um problema de programação quadrática, em que o procedimento Congra, do SAS, foi utilizado.

Os insumos foram agregados nas seguintes variáveis:

X_1 = 3 % do valor da terra e 6 % do valor dos animais utilizados na produção e do valor da depreciação de máquinas, equipamentos e benfeitorias.

X_2 = Dispêndio com mão-de-obra familiar e contratada.

X_3 = *Dispêndio com concentrados, minerais, volumosos, inseminação* artificial e medicamentos.

X_4 = Dispêndio com energia, combustíveis, manutenção de máquinas, equipamentos e benfeitorias e demais gastos.

A renda bruta equivale à venda de leite, animais e esterco, designada por y. A renda líquida, rl ., é igual a $rl = y - x_1 - x_2 - x_3 - x_4$, resíduo que remunera o empreendedor pelo risco que corre por empreender. Obviamente, é a renda líquida de longo prazo. Entre os 143 amostrados, havia 68 produtores com renda líquida, cerca de 47,6%. Trata-se de um número elevado, mas instrumental para mostrar a capacidade dos três métodos de reverterem a renda líquida de negativa para positiva e ainda fazê-la crescer. No caso da regressão, adicionou-se uma variável binária, com valor um para os produtores de renda líquida não-negativa e zero, para os demais.

Como o objetivo do trabalho é discutir critérios de avaliação, não se deterá na análise dos índices de eficiência e nem nas limitações de cada método. No caso da encapsulamento de dados e da fronteira estocástica, as limitações têm sido amplamente discutidas, por exemplo, em Lovell e Schmidt (1988), Coelli et al. (1998) e Souza (2003). Varian (1985) examinou as principais críticas ao seu método e respondeu às objeções.

2. Metodologia

Uma discussão um pouco mais detalhada é dada para o método de Varian. Quanto aos outros dois, apenas alguns aspectos são realçados.

2.1. Método Varian

Quando se conhecem as observações, (w^i, x^i, y^i) $i = 1, 2,..., n$, de n produtores, em que w é o vetor preço de insumos; x, vetor de insumos; e y, produto, é possível saber se elas foram geradas por empreendedores que minimizam custo. Afriat (1967) desenvolveu grande parte da teoria, que foi aperfeiçoada por Varian numa série de artigos, citados em Varian (1985). Uma exposição dos artigos de Varian está em Alves (2000).

Se o agricultor **A** escolher uma combinação de insumos x para produzir o produto y, então, ela não poderá custar mais que qualquer outra combinação de outro produtor, quando o dispêndio for avaliado pelos preços do agricultor **A**. Ora, essa definição somente faz sentido se os produtores tiverem acesso à mesma tecnologia, representada pelo conjunto de produção $V(y) = \{x : x \text{ produz } y\}$.

Pode-se relaxar a hipótese de um único conjunto de produção, específico a cada produtor, que é muito forte, o que exige que ele seja um conjunto encadeado. Ora, se cada produtor, particularizado para **A**, tiver o conjunto de produção que lhe é específico, segue-se que a racionalização será trivialmente obedecida, porque qualquer outra combinação que produzisse y ou não pertencesse ao conjunto de produção do agricultor **A**, ou se pertencesse e custasse menos, teria sido escolhida.

Dois conjuntos de produção dizem-se encadeados se $y \geq z \rightarrow V(z) \supseteq V(y)$. Se a produção y for maior ou igual à produção z, e se a combinação x de insumos produzir y, então, ela também produzirá z. Admite-se um único produto. Outra propriedade importante, seja **a** e **b** duas combinações de insumos. Seja $\mathbf{a} \geq \mathbf{b}$ [4]. Se **b** produzir y, então **a** também produzirá y.

[4] a ≥ b significa que cada componente do vetor a é maior ou igual à correspondente do vetor b, e pelo menos uma é diferente.

Permite-se, assim, o desperdício de recursos, sem se impor nenhum custo. Admite-se que $V(y)$ seja um conjunto encadeado, pressuposição 1, e que permita o desperdício, pressuposição 2. Sem a pressuposição 2, é impossível falar-se de ineficiência técnica.

Formaliza-se a definição de racionalização, como se trata de custo, por *c-racionalização*.

Definição. Existirá *c-racionalização* se a seguinte condição for verificada para produzir y: quando x for a combinação de insumos observada pelo agricultor i e z, qualquer outra combinação também será observada por outros agricultores. Admite-se alcance de $\boldsymbol{i}$, mas o custo dela será avaliado pelo vetor preço do agricultor i, então, $w^i * x^i \leq w^i * z^j \;\; j = 1, 2, \ldots, n$.

C-racionalização significaria, em linguagem corrente, que o agricultor, dentro das opções possíveis para se produzir y, porque houve agricultores que as escolheram, escolherá a que lhe custar menos, em função do vetor preço de insumos que lhe estará disponível.

Com essa preparação, Varian demonstrou o teorema, um dos fundamentos do artigo, que pode ser encontrado em Alves (2000). O teorema diz que as seguintes condições são equivalentes:

Teorema

(I) Existe uma família encadeada de conjuntos de produção, $V(y)$, em que *c-racionaliza* as observações;

(II) Se $y^j \geq y^i, \rightarrow w^i * x^j \geq w^i * x^i$, quem produzir mais ou igual não gastará menos do que quem produzir a mesma quantidade ou menos, dispêndio avaliado pelos preços de quem produzir menos;

(III) Existe uma família encadeada de conjuntos de produção que são convexos, fechados e satisfazem às pressuposições 1 e 2.

A condição (ii) é o fundamento empírico. Os dados, dificilmente, não conterão algumas violações à condição (ii). É importante descobrir qual será a menor perturbação necessária, com vistas em modificar a combinação de insumos para que se satisfaça (ii), e depois testar a hipótese nula, segundo a qual os agricultores *c-racionalizam* os dados.

2.2 Fronteira estocástica

Battese (1992) representou a fronteira probabilística por

$$y_i = f(x_i, \beta)e^{(V_i - U_i)} \quad i = 1,2, \ldots, n, \tag{1}$$

em que V_i é o erro aleatório, com média zero, associado aos fatores que estão fora do controle da firma. Pelo modelo, o limite superior de produção possível é dado por $f(x_i, b)e^{(V_i)}$, portanto, estocástico. V_i e U_i são variáveis aleatórias independentes, e U_i, variável aleatória não-negativa. Admite-se que seja distribuída de acordo com uma distribuição normal truncada, e $f(x, \beta)$ seja representada por uma função Cobb-Douglas. A medida de eficiência técnica será dada por

$$ET_i = \frac{y_i}{y_i^*} = \frac{f(x_i;\beta)e^{(V_i - U_i)}}{f(x_i;\beta)e^{(V_i)}} = e^{(-U_i)}.$$

2.3. Encapsulamento de dados – DEA

Constróem-se duas matrizes. A matriz N contém entradas positivas ou nulas, e a coluna i refere-se ao agricultor i. Há uma linha para cada insumo. As entradas da coluna i são os valores dos insumos usados pelo agricultor i. Com quatro insumos e cinqüenta produtores, haverá quatro linhas e cinqüenta colunas, e N será uma matriz 4x50. Em cada linha, haverá, pelo menos, uma célula positiva, ou seja, pelo menos um agricultor terá usado o insumo. Caso contrário, por que listá-lo? Em cada coluna, existirá também uma célula positiva. A matriz M será semelhante a N, exceto cada linha que contém um produto, e obedecerá às mesmas propriedades, quanto a linhas e colunas.

O que se procura fazer é, radialmente, obter a máxima redução do vetor de insumos do agricultor i, de modo que ele tenha, pelo menos, produção igual à que obteve, que é tão somente comparado com agricultores que produziram tanto ou mais e nunca gastaram mais. Assim, trata-se de uma comparação "justa". Se não houver restrições irremovíveis, o agricultor i poderia igualar-se ao grupo, em termos da fronteira gerada. A expressão matemática dessas idéias é a seguinte:

$$Min\{t\}$$

Restrito a:

$$M * Z \geq y^{i} \tag{2}$$

$$N * Z - t * x^{i} \leq 0 \tag{3}$$

$$Z \geq 0 \tag{4}$$

Observa-se que, quanto menor for t, tanto maior será a redução radial de x. Nota-se, em (3), que t é um número real positivo. A primeira desigualdade exige que a produção não seja menor que a do agricultor i. A segunda desigualdade diz que o vetor insumo "ótimo" tenha todas as componentes iguais ou menores que as do vetor do agricultor i. Se t=1, não haverá redução de x. A função de produção gerada tem retornos constantes à escala, e é em relação a ela que se calculam os índices de eficiência técnica. A fronteira gerada é, assim, determinística. Para cada agricultor, resolve-se o problema de programação linear acima.

Adicionando-se a igualdade $\sum_{1}^{n} z_i = 1$ às três restrições descritas, gera-se uma função de produção com retornos variáveis. Há outras opções que fogem ao escopo do trabalho (Alves, 1998). Nota-se que a orientação é insumos.

Quando a orientação for produto, será preciso fazer algumas mudanças,

$Max\{t\}$

Restrito a:

$$M * Z \geq t^{*} y^{i} \qquad (5)$$

$$N * Z - x^{i} \leq 0 \qquad (6)$$

$$Z \geq 0 \qquad (7)$$

ou seja, deseja-se a máxima expansão radial do produto, desde que não se gaste mais. A função de produção gerada tem retorno constante à escala. Obtém-se retorno variável à escala, adicionando-se às restrições acima a igualdade $\sum_{1}^{n} z_i = 1.$

3. Teste de hipóteses

As hipóteses são testadas em relação à amostra de 143 produtores de leite, descrita na introdução. Os insumos serão agregados em quatro categorias, e o produto corresponde à venda de leite, animais e esterco, equivalente, portanto, à renda bruta.

3.1. Método Varian

3.1.1. Orientação insumo

Erros de medição podem levar as firmas a violarem a regra, que diz que, quem produz mais, não pode gastar menos do que quem produz menos. Varian (1985) admitiu que os desvios observados, em relação à regra descrita, sejam devidos a erros de mensuração dos insumos. A produção e os preços são conhecidos.

A regra acima, denominada de c-racionalização, é expressa, formalmente, por

$y_i \leq y_j$, que implica que $w_{i*} x_i \leq w_{i*} x_j$, $\forall i,\ i, j \in I, i \neq j$.

No caso, a firma produz um único produto. Ordenam-se as firmas segundo a ordem crescente de y. O custo da primeira unidade é comparado com o de todas as seguintes. Pela regra, tem que ser menor ou igual. Quando isto não ocorrer, haverá violação. O mesmo é feito com a segunda unidade, a terceira e, assim, sucessivamente. Contam-se as violações. Pela regra acima, o número de violações deveria ser zero. O número de comparações é dado por $\dfrac{n(n-1)}{2}$.

Na prática, o número de violações é maior que zero, mas podem-se modificar os valores dos insumos, para que a regra seja obedecida. Por um critério, que será explicado a seguir, determinam-se os valores dos insumos, de modo que obedeçam à regra de minimização de custo. A questão é saber se a diferença entre o valor observado e o calculado é, estatisticamente, pequena.

Admitindo-se a ocorrência de erros de medição, supõe-se que a demanda observada pelo fator k, na observação i, x_{ik}, esteja relacionada com a demanda "verdadeira", z_{ik}, da seguinte forma:

$x_{ik} = z_{ik} + e_{ik}$, para $i = 1,2,3,\dots,n$, $k = 1, 2, 3, \dots, m$,

em que e_{ik} é o erro, que é independente e normalmente distribuído, com média zero e variância dada, constante e igual a σ^2. A hipótese a ser considerada é que as observações sejam *c-racionalizáveis*, ou seja,

H_0 = os dados (w_i, z_i, y_i) satisfazem à regra de minimização de custo. Se fosse possível observar a demanda "verdadeira" (z_{ik}) e a variância (σ^2) de ε_{ik}, a seguinte estatística poderia ser obtida por

$$T = \frac{\sum_{i=1}^{n}\sum_{k=1}^{m}(z_{ik} - x_{ik})^2}{\sigma^2} \quad (8)$$

que tem distribuição qui-quadrado (χ^2), com $m * n$ graus de liberdade. Para um nível de significância a, se *T* for maior que o valor crítico $\chi^2_{\alpha, m*n}$, rejeitar-se-á a hipótese nula.

No entanto, a estatística *T* não é observável, já que se desconhecem a demanda "verdadeira" dos fatores (z_{ik}) e a variância (σ^2). Pode-se, mesmo assim, testar a hipótese de minimização de custos pela solução do seguinte problema de programação quadrática:

$$R = Min\sum_{i=1}^{n}\sum_{k=1}^{m}(t_{ik} - x_{ik})^2 ,$$

sujeito a:

$$\sum_{k=1}^{m} w_{ik} t_{ik} \leq \sum_{k=1}^{m} w_{ik} t_{ij} , \forall j = 1,2,\mathrm{K} ,n, \ y_i \leq y_j , \quad (9)$$

em que *R* é igual ao mínimo da soma dos quadrados dos desvios relativos às quantidades observadas, sendo observadas as restrições acima. Para completar, w_{ik} é o preço do insumo *k* para a *i*-ésima firma; e t_{ik} representa o conjunto de variáveis para as quais se busca a solução.

Considerando-se que a demanda de fatores seja medida em diferentes unidades, Varian (1985) sugeriu que a relação entre a demanda "verdadeira" e a observada tenha a especificação abaixo:

$Z_{ik} = x_{ik}(1 + e_{ik})$. Logo, $\varepsilon_{ik} = \frac{Z_{ik}}{x_{ik}} - 1$ é independentemente distribuído, com distribuição normal, média zero e variância σ^2. Os termos $\frac{Z_{ik}}{x_{ik}}$ têm, portanto, média igual a 1, são independentes e normalmente distribuídos e admite-se que tenham a mesma variância. Defini-se $\sigma = \sigma$/mn e sobre esta medida decide-se a rejeição da hipótese. Desse modo, a função objetiva do modelo de programação quadrática, toma a seguinte forma:

$$R = Min \sum_{i=1}^{n} \sum_{k=1}^{m} \left[\left(\frac{t_{ik}}{x_{ik}} \right) - 1 \right]^2 .$$

Sabe-se que T / σ^2 tem distribuição qui-quadrado. Pela forma como *R* foi calculado, $R / \sigma^2 \leq T / \sigma^2$. Logo, se o teste for realizado com *R* no lugar de *T*, aumenta-se a possibilidade de rejeição da hipótese nula.

Contudo, σ^2 é desconhecido. Seja k_α o valor crítico da tabela de qui-quadrado com $m*n$ graus de liberdade, então, $R / \sigma^2 = k_\alpha$ e R é conhecido. Logo, σ^2=R/ k_α. Se o valor for relativamente pequeno[5], a hipótese nula não deverá ser rejeitada.

Segundo ainda Varian (1985), *R/m*n* é uma estimativa de máxima verossimilhança de σ^2 , mas, infelizmente, desconhece-se a distribuição de *R/m*n*.

[5] Salienta-se que "pequeno" para o valor de é algo que irá depender da percepção do pesquisador.

Antes de fornecer os resultados do teste, é importante que sejam contadas as violações à regra de minimização de custo. O procedimento é o descrito a seguir.

O custo total é igual à soma dos valores dos quatro insumos, representado por para a observação *i*. As observações foram ordenadas do menor valor da produção para o maior. Em seguida, comparou-se o custo da observação *i* com o da observação *j*, sendo . Assim, a primeira observação foi comparada com todas as outras; a segunda, com (n - 1) observações seguintes; e, assim, sucessivamente. Pela regra acima, $C_i \leq C_j$, se $i \leq j$. Quando isto ocorrer, verificar-se-á um acerto e será atribuído o valor 1 ao acerto e 0 (zero) ao erro, ou seja, $C_i > C_j$, se $i \leq j$. Como se trata de uma progressão aritmética de n termos e razão igual a 1, o número de comparações equivalerá a n*(n-1)/2, que é o número máximo de acertos. Com 143 produtores, o número máximo de acertos será dado por 10.153 (143*142/2). O número de acertos observados equivaleu a 8.884, cerca de 87,1%, bastante elevado, considerando-se a dispersão regional da amostra e sua variabilidade, em termos do volume de leite produzido.

A seguir, na Tabela 3, apresentam-se os resultados do teste proposto por Varian. O algoritmo usado na programação quadrática foi o *Congra*, do SAS.

Tabela 3 - Teste de Varian: orientação insumo

Itens	Resultados
Número de observações	143
Programação quadrática (R)	27,28
Graus de liberdade (mn)	572
χ^2_	639,68
Desvio-padrão (R/mn)	0,2184
Desvio-padrão - programação quadrática	0,2063
$\bar{\sigma}$	0,00036

Fonte: Resultados da pesquisa.

O σ pode ser entendido como um coeficiente de variação, porque a média é 1 .O coeficiente de variação é igual a 0,036%, considerado muito pequeno. Assim, não se rejeita a hipótese de *c-racionalização*, o que era esperado, em razão do elevado número de acertos na contagem feita acima.

3.1.2. Orientação produto

O tema foi, indiretamente, tratado por Varian (1985), mas é uma extensão trivial de seu método. Os dados foram ordenados do menor dispêndio para o maior. Pela regra de racionalização, quem produziu mais não poderia ter gasto menos do que quem produziu menos. Assim, $C_i \leq C_j$ não pode implicar que $y_i > y_j$. Logo, os *y*'s seguem a mesma ordenação dos *C*'s. O processo de contagem de acertos foi igual ao descrito, com *y* no lugar de *C*; obviamente, os resultados foram iguais.

Pelo processo de contagem, o número de acertos foi muito elevado, cerca de 87,1%; espera-se que a hipótese de c-racionalização não seja rejeitada pelos dados. Encontrou-se, na Tabela 4, um coeficiente de variação de 1,91%, o que indica serem os erros de medição pequenos; assim, não se rejeita a hipótese de c-racionalização, tendo-se em conta a orientação produto.

Novamente, os agricultores foram racionais nas suas decisões, quando não requereram nem minimização de custos nem maximização da renda líquida, obedecendo-se a critérios de programação matemática, como no DEA e em procedimentos que se baseiam na maximização da renda líquida.

Tabela 4 - Resultados da programação quadrática para as amostras em estudo. Orientação produto

Itens	Brasil
Número de observações	143
Programação quadrática (R)	9,22
Graus de liberdade (mn)	143
χ^2	177,51
Desvio-padrão (R/mn)	0,25
Desvio-padrão – programação quadrática	0,228
σ	0,0016

Fonte: Resultados da pesquisa.

3.2. Fronteira estocástica

Como o método é muito conhecido, apresentam-se apenas os resultados de estimação da função de produção. A variável dependente é a renda bruta, e há cinco variáveis independentes: quatro insumos já descritos e uma variável binária, que é igual a 1, quando a renda líquida é não-negativa, e a zero, quando a renda líquida for negativa. A função de produção é Cobb-Douglas.

Tabela 5 - Estimativas dos parâmetros para a função de produção fronteira estocástica, tipo Cobb Douglas, método máxima verossimilhança, com especificação *meio-normal*

Variáveis	Coeficientes Estimados	Erro-padrão	H: coef=0	Probabilidade
β_0	1,2233	0,2822	4,34	0,0001
β_1	0,2697	0,0339	7,94	0,0001
β_2	0,0852	0,0380	2,24	0,0265
β_3	0,4810	0,0306	15,71	0,0001
β_4	0,1626	0,0325	5,01	0,0001
Dummy intercepto	0,4458	0,0354	12,61	0,0001
σ^2	0,0541	0,0202	2,68	0,0083
λ	0,3573	0,4213	0,85	0,3979
LFV[1]	24,2239			

Fonte: Resultados da pesquisa. *Software* 4.1xp.

[1] LFV é o logaritmo da função de verossimilhança (*log-likelihood function*). Diferente de zero, nível de 10 % de probabilidade, com 1 grau de liberdade, pelo teste qui-quadrado misto.

Os resultados estatísticos são muito satisfatórios. A variância é estatisticamente diferente de zero; portanto, as evidências não rejeitaram a hipótese da fronteira estocástica.

3.3. DEA orientação produto

Souza (2003) propôs um teste de hipótese. Seja a produção da observação i e θ_i, o respectivo resultado do DEA, orientação produto. O desvio d_i é dado por $d_i = y_i - \theta_i * y_i$. A hipótese é que tenham a mesma distribuição. Assim, foi possível testar as hipóteses de distribuição *meio-normal* ou *exponencial*.

Se os d_i forem iid (independente e identicamente distribuídos), com densidade comum *exponencial*, então, $2\sum_{1}^{n} d_i / s$, em que n será número de observações da amostra, e s, desvio-padrão dos d_i, que terá aproximadamente distribuição qui-quadrado, com $2n$ graus de liberdade. Se os d_i forem iid, com densidade comum *meio-normal*, então, $(1-2/\pi)\sum_{i}^{n} d_i^2 / s^2$ terá, aproximadamente, distribuição qui-quadrado, com n graus de liberdade.

A hipótese de distribuição *exponencial* foi rejeitada a 5% de probabilidade, e as evidências não foram suficientes para rejeitar a *meio-normal*. Embora os resultados não sejam reproduzidos, a conclusão foi semelhante para retorno variável. Com esses resultados, foi possível construir o intervalo de confiança para os índices de eficiência.

Tabela 6 - Ajustamento das distribuições *exponencial* e *meio-normal* (truncada em zero), retornos constantes à escala

Distribuição	Brasil		
	Graus de liberdade	Qui-quadrado	Probabilidade.
Exponencial	228	323,93	0,06
Meio-normal	114	118,26	0,93

Fonte de dados: CEPEA-ESALQ/USP.

Foi possível testar a hipótese de retornos constantes versus retornos variáveis (Souza, 2003). Então, $\sum_1^n (d_i^c)^2 / \sum_1^n (d_i^v)^2$ tem distribuição F, com n graus de liberdade para o numerador e n graus de liberdade para o denominador, no caso da distribuição *meio-normal*. Encontraram-se $F_{(143,143)} = 1{,}20$ e p = 0,14. Não houve, assim, evidências para se rejeitar a hipótese de retornos constantes[6].

Souza, com base em Souza (2003), sugeriu um método para construir os intervalos dos índices de eficiência técnica, cujos detalhes não serão discutidos para não alongar o texto.

A eficiência técnica média igualou-se a 0,41, sendo o correspondente intervalo de confiança igual a 0,24 - 0,41. Há 38 observações neste intervalo, cerca de 26,6 % da amostra.

Haverá interesse em conhecer os intervalos de confiança, quando a eficiência técnica tiver valor 1. Verifica-se enorme variabilidade da amplitude dos intervalos de confiança, como se verá a seguir.

Para a observação 27 de índice um, o intervalo igualou-se a 0,17 - 1,00 e contém 138 observações, cerca de 96,5 % das observações, que não divergiram, estatisticamente, de um. O menor intervalo correspondeu a 0,82 - 1,00 e abrigou 22 observações, cerca de 15,4% da amostra. Detalhes dos intervalos de confiança, quando a eficiência técnica se iguala a 1, estão na Tabela 7.

[6] Quando o expoente é c, trata-se de retorno constante, se v, retorno variável.

Tabela 7 - Número da observação, limite inferior do intervalo de confiança das observações de índices de eficiência técnica igual a um e número de observações que pertencem ao intervalo de confiança, que tem limite superior igual a um. Amostra de 143 observações, ordenada da menor para maior renda bruta

Número da Observação	Limite inferior	Observações do intervalo
27	0,17	138
89	0,35	104
117	0,47	78
123	0,48	77
129	0,53	67
131	0,54	66
132	0,54	66
136	0,62	44
138	0,69	35
139	0,70	33
143	0,82	22

Fonte de dados: CEPEA-ESALQ/USP.

4. Avaliação dos métodos

Teve-se o cuidado, em cada caso, de fazer a soma de quadrado compatível com o procedimento de Varian: $(x_0 / x_{obs} - 1)^2$, em que x_0 é o valor calculado pelo programa e x_{obs}, valor observado.

Em relação ao procedimento de Varian, orientação insumos, introduziu-se a restrição pela qual o custo da solução tem que ser menor ou igual ao custo observado. Na orientação produto, a nova restrição exigiu que a solução fosse maior ou igual à renda bruta observada.

Como regra geral, a menor perturbação trouxe o menor incremento da renda líquida e menor redução do número de produtores com renda líquida negativa. O método Varian, quanto à menor perturbação, somente perdeu para a fronteira estocástica, com orientação produto, e teve boa performance, quanto à renda líquida.

A finalidade do DEA, orientação insumo, é reduzir, radialmente, o consumo de insumos, pelo menos para manter a produção, no grupo de comparação de cada produtor. Assim, é natural que perturbe, com maior intensidade, os insumos, com vistas em reduzir custos. Por isto, seu efeito sobre a renda líquida foi maior. Quando a orientação é produto, expande-se, radialmente, a produção para a fronteira, obedecendo-se à restrição de insumos, qual seja, não gastar mais do que o dispêndio observado. Por isto, perturba-se a produção, no sentido de incrementá-la, sem aumentar os dispêndios. Desse modo, tanto na orientação de insumos quanto na de produtos, o DEA tendeu a incrementar a soma de quadrados e a renda líquida e a reduzir o número de agricultores com renda líquida negativa. A Tabela 8 mostra que isto realmente aconteceu.

A fronteira estocástica procura ajustar uma curva que minimiza a soma dos quadrados das diferenças entre os valores preditos e observados, quando o método é o de mínimos quadrados ordinários, e aproxima-se deste objetivo, quando o método é o de máxima verossimilhança. Mas não aduz nenhuma restrição aos insumos, com o objetivo de reduzir ou incrementar o dispêndio. Se o ajuste da regressão for muito bom, espera-se pequena perturbação. Nenhuma previsão pode ser feita sobre a renda líquida.

Quando a orientação é insumo, o método da fronteira estocástica nada tem, diretamente, a dizer, pois não se modificam os insumos. Apenas a produção prevista pelo modelo costuma ser diferente da observada. No caso da função de produção Cobb-Douglas, pode-se estimar, para cada insumo, a função de demanda condicionada ao nível de produção, sob a hipótese de que os produtores maximizem a renda líquida (Nerlove, 1965). Designando-se o valor predito da regressão por x_p e o observado por x_{obs}, a soma de quadrados, compatível com Varian, foi dada por $(x_p / x_{obs} - 1)^2$, quando x variar nos insumos e nos produtores. Verifica-se que a perturbação dos insumos foi bem maior do que no método Varian.

O método Varian teve boa performance, visto que perturbou menos os insumos, para que se obedecesse à regra de racionalização dos dados, e ainda contribuiu para diminuir o número de observações de renda líquida negativa e aumentá-la expressivamente. Suas recomendações são factíveis. No caso de um agricultor produzir menos e gastar mais, recomenda-se que, pelo menos, ajuste o gasto ao nível do que gastou menos. O método prescinde do conceito de uma fronteira e tem como base a escolha de uma ordenação dos custos, de modo que se obedeça à regra de racionalização dos dados. Assim, é um método muito conveniente para orientar grupo de produtores, porque está muito próximo do que fazem, no sentido de imitar uns aos outros.

Tabela 8 - Comparação de três métodos pelos critérios da soma de quadrados dos desvios (calculado menos observado) e da renda líquida. Renda líquida não-negativa observada, 75 observações em 143. DEA com retornos constantes (RC) e com retornos variáveis (RV)

Métodos	Orientação insumo			Orientação produto		
	Soma do Quadrado desvio	Incremento Renda líquida (%)	Produtores Renda líquida não-negativa (número)	Soma do Quadrado desvio	Incremento Renda líquida (%)	Produtores Renda líquida não-negativa (número)
Varian	51,19	256,38	142	104,9	404,9	143
DEA	RC = 158,0 RV = 88,5	RC = 325,5 RV = 225,0	RC = 141 RV = 126	RC = 586,8 RV = 509,6	RC = 794,9 RV = 703,8	RC = 141 RV = 131
Fronteira Estocástica	160,01	113,74	143	6,21	69,4	75

Fonte de dados: CEPEA-ESALQ/USP.

De forma geral, os índices de eficiência obtidos pelo modelo de fronteira estocástica foram maiores do que os obtidos pelo método de Varian, e estes foram superiores aos do modelo DEA (retornos constantes a escala), tanto para orientação insumo como para orientação produto.

No modelo estocástico, a forma funcional Cobb-Douglas ajustou-se muito bem aos dados, ou seja, ela é uma curva. O modelo DEA gera, implicitamente, uma função de produção (Souza, 2003), que é composta de segmentos lineares. É, portanto, uma aproximação mais pobre da verdadeira superfície.

O coeficiente de correlação entre os resultados do modelo de fronteira estocástica e do modelo DEA (retornos constantes à escala), ambos com orientação produto, foi igual a 0,50, estatisticamente significante a 1% de probabilidade.

Na comparação entre os resultados do modelo DEA (retornos constantes à escala) e do modelo Varian, ambos com orientação produto, o coeficiente de correlação entre os índices de eficiência foi de 0,60, significante a 1% de probabilidade. O coeficiente de correlação entre os índices de eficiência do modelo fronteira estocástica e do método Varian, orientação produto, igualou-se 0,42, estatisticamente significante a 1 % de probabilidade. O coeficiente de correlação entre os índices de eficiência, DEA (retornos constante à escala), com orientação insumo, e os índices do método Varian, também com orientação insumo, foi igual a 0,53, estatisticamente significante a 1% de probabilidade.

5. Conclusões

Dois critérios foram usados na avaliação dos métodos – perturbação dos insumos ou do produto e incremento da renda líquida. Tanto melhor quanto menor for a perturbação dos insumos e maior o acréscimo da renda líquida.

O método de Varian foi o que produziu menos distúrbios nos insumos ou no produto, pois incrementou mais a renda líquida que a fronteira estocástica e menos que o DEA. Mas sua solução fica mais próxima daquilo que é factível para cada produtor fazer. Assim, é um método conveniente de gestão e ainda útil para testar a hipótese de c-racionalização. Os dados não ofereceram evidências de sua rejeição.

A fronteira estocástica não objetiva nem aumentar a renda líquida nem reduzir os custos. Seu efeito sobre a renda líquida foi negativo e produziu maior distúrbio nos insumos e foi muito pequeno no produto, o que sempre ocorrerá, quando a função de produção ajustar-se bem aos dados.

O DEA, implicitamente, gera uma função de produção e associa-se o erro estocástico a ela. Os dados não rejeitaram a hipótese de distribuição meio normal e rejeitaram a distribuição exponencial. No confronto de retorno constante à escala com retorno variável à escala, os dados rejeitaram a hipótese de retorno variável e não rejeitaram a de retorno constante. Os intervalos de confiança dos índices de eficiência são muito sensíveis ao valor da renda bruta. O trabalho documenta a variação do intervalo de confiança, a 5% de probabilidade, para as observações de índice de eficiência igual a 1. Por exemplo, o intervalo de confiança da observação 27 continha 138 produtores, cerca de 96,5% da amostra de 143 produtores, enquanto no intervalo de confiança da unidade de decisão número 143 havia 22 produtores, cerca de 15,4% da amostra, e as observações estavam ordenadas da menor para a maior renda bruta.

Uma complicação do método Varian é o algoritmo d, programação quadrática. Há vários deles e, raramente, produzem os mesmos resultados.

Referências Bibliográficas

AFRIAT, S. N. The construction of a utility function from expenditure data. **International Economics Review**, v.8, p.67-77. 1967.

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

ALVES, E. **Teoria da produção: métodos não-paramétricos.** Brasília: Embrapa, 2000, 65 p.

BATTESE, G.E. Frontier production functions and technical efficiency: a survey of empirical applications in agricultural economics. **Agricultural Economics**, v.7, n.3/4, p.185-208, oct. 1992.

CHAVAS, J.; COX, T. L. On generalized Revealed Preference Analyses. **The Quarterly Journal of Economics.** vol. 108, p. 493-506, 1993.

COELLI, T.J.; RAO, D.S.P.; BATTESE, G.E. **An introduction to efficiency and productivity analysis.** Norwell: Kluner Academic, 1998.

275 p.

LOVELL, K. C. A.; SCHMIDT, P. A comparison of alternatives approaches to the measurement of productive efficiency. In **Applications of modern production theory.** Boston, Kluwer Academic Publishers, p. 3-32, 1988.

NERLOVE, M. **Estimation and identification of Cobb-Douglas production functions**. Rand McNally and Company, 1965.

PEREIRA FILHO, C. A.; FERREIRA FILHO, J. B. de S. Fontes de Ineficiência da pequena produção familiar agrícola na região Recôncavo do Estado da Bahia, **Revista de Economia e Sociologia Rural**. vol. 41, n. 1, p. 63-78, jan./mar., 2003.

RAY, S. C.; BHADRA, D. Nonparametric Tests of Cost Minimizing Behavior: a study of Indian Farms. **American Journal of Agricultural Economics**, vol. 75, n. 4, p.990-999, 1993.

RICHETTI, A.; REIS, R. P. Fronteira de produção e eficiência econômica da cultura de soja no Mato Grosso do Sul. vol. 41, n. 1, p. 63-78, jan./mar., 2003.

SOUZA, G.S. Funções de produção: uma abordagem estatística com o uso de modelos de encapsulamento de dados. **Texto para discussão 17**, Brasília: Embrapa, Informação Tecnológica, 2003. 49p.

VARIAN, H.R. Non-Parametric Analysis of Optimizing Behavior with Measurement Error. **Journal of Econometrics.** v.30, p.445-458, 1985.