

4 Modelos de Fronteira Estocástica

4.1. Considerações Gerais⁷

A produção de qualquer bem ou serviço é um processo onde um conjunto de insumos (*inputs*) é transformado em um conjunto de produtos (*outputs*). Esta transformação pode ser feita de diversas maneiras, todas representadas no conjunto de possibilidades de produção (CPP), conforme ilustrado na figura 2, para o caso de x unidades de um tipo de *input* ser usado na produção de y unidades de um tipo de produto (Varian, 1999).

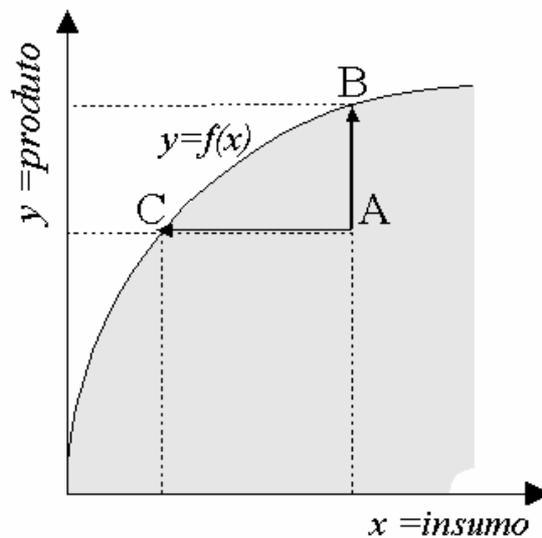


Figura 2. Conjunto de possibilidades de produção e fronteira de produção $y=f(x)$.

Na figura 2 qualquer ponto $(x,y) \in \text{CPP}$ indica uma maneira tecnologicamente viável de transformar uma quantidade x de insumo em uma quantidade y de produto, mas devido às restrições de natureza tecnológica, o CPP é limitado pela função fronteira de produção ($y=f(x)$), uma função que indica a máxima quantidade de produto que pode ser obtida a partir de uma quantidade x

⁷ Este capítulo se baseou em estudo feito por Pessanha & Souza (2003) denominado “Modelos de Fronteira Estocástica na Definição de Metas Globais de Continuidade das Distribuidoras de Energia Elétrica”.

de *inputs*. Portanto, para uma tecnologia dada, a função fronteira de produção caracteriza o mínimo conjunto de *inputs* necessários para produzir quantidades fixas de vários tipos de *outputs* (orientação segundo a ótica dos *inputs*), ou de forma análoga, a fronteira de produção caracteriza a máxima produção possível dada uma quantidade fixa de *inputs* (orientação segundo a ótica dos *outputs*).

Assim, todo produtor localizado na fronteira é classificado como tecnicamente eficiente⁸, enquanto os demais produtores do CPP são considerados tecnicamente ineficientes. Por exemplo, na figura 2, o produtor A é tecnicamente ineficiente, enquanto os produtores B e C são tecnicamente eficientes. Observe que um produtor situado no ponto A é ineficiente, pois com a mesma quantidade de insumo consumida em A pode-se obter uma produção igual a do produtor B, superior a produzida em A. De forma análoga, o nível de produção em A pode ser obtido com uma quantidade de insumo igual a do produtor C, menor que a usada em A.

Supondo que são conhecidos os preços, $w \in R_+^n$, de um vetor de *inputs* $x \in R_+^n$, um produtor racional procura produzir um conjunto de *outputs* $y \in R_+^m$ com o menor custo de produção possível. Neste caso, o padrão contra o qual a performance do produtor é avaliada desloca-se da fronteira de produção para a fronteira de custo e a eficiência passa a ter duas componentes : técnica e alocativa⁹ (Kumbhakar & Lovell, 2000; Varian, 1999).

A função fronteira de custo é uma representação alternativa da tecnologia de produção e indica o mínimo custo necessário para produzir um conjunto de *outputs*, dados o nível de produção, os preços dos *inputs* e a tecnologia existente. O custo de um produtor ineficiente é maior que o mínimo definido pela fronteira de custo, então, os produtores ineficientes estão localizados acima da fronteira de

⁸ Segundo a definição Pareto-Koopmans um processo de produção é tecnicamente eficiente se um incremento na quantidade de qualquer produto, requer um incremento na quantidade de pelo menos um insumo ou a redução de pelo menos a quantidade de outro produto. De forma análoga, um processo de produção é tecnicamente eficiente se a redução na quantidade de qualquer insumo requer o incremento na quantidade de outro insumo, para manter o mesmo nível de produção, ou a redução na quantidade de algum produto. Observe que os pontos B e C na figura 2 satisfazem a definição de Pareto-Koopmans para a eficiência técnica.

⁹ Um produtor tecnicamente eficiente não apresenta necessariamente a eficiência econômica, pois pode utilizar um conjunto de inputs que não seja o de menor custo. Para ser economicamente eficiente, o produtor deve ser tecnicamente eficiente e também utilizar os insumos de menor custo, isto é, ter eficiência alocativa.

custo, enquanto os produtores eficientes estão na fronteira, conforme mostra a figura 3.

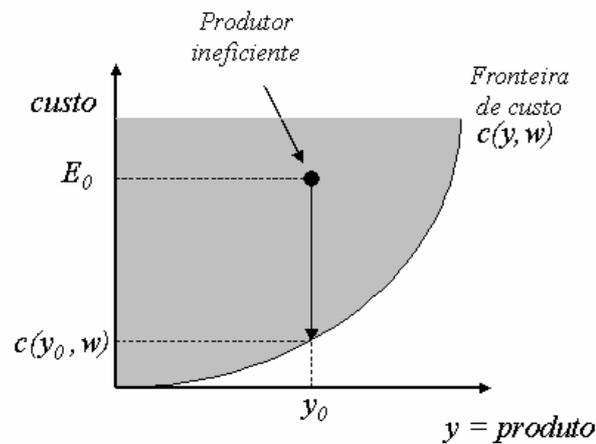


Figura 3. Fronteira de custo

A fronteira de produção e a fronteira de custo funcionam como uma referência, ou melhor, um *benchmarking* contra o qual podem ser comparados os desempenhos de diferentes produtores ou DMU's (*Decision Making Units*) que atuam no mesmo segmento da indústria. Uma simples comparação com a função fronteira permite discriminar os produtores em duas categorias: eficientes ou ineficientes.

Os desvios em relação à função fronteira refletem falhas na otimização do processo de produção. Isto sugere que o grau de eficiência relativa de um produtor pode ser avaliado por meio da distância vertical entre ele e a função fronteira, conforme a métrica radial proposta por Debreu (Kumbhakar & Lovell, 2000). Esta métrica é um número no intervalo $[0,1]$ e o produtor é considerado eficiente se a métrica assume um valor unitário, caso contrário ele é considerado ineficiente.

Segundo esta métrica radial, a eficiência técnica de um produtor que produz um vetor de *outputs* y , a partir de um vetor de *inputs* x é a razão $TE = y/f(x)$, onde $f(x) \geq y$ é a máxima produção possível, especificada pela função fronteira de produção. De forma análoga, admitindo que os preços dos *inputs* (vetor w) sejam conhecidos, o custo de produção é $E = w^T x$ e a eficiência econômica do produtor é a razão $CE = c(y,w)/E$, onde $c(y,w) \leq E$ é o mínimo custo de produção definido pela função fronteira de custo.

Na prática não se conhece plenamente o CPP, portanto, a fronteira de produção ou custo é desconhecida e a eficiência de cada produtor não pode ser mensurada diretamente. De fato, na maioria das vezes, os poucos dados disponíveis limitam-se às observações acerca das quantidades e preços de *inputs* e *outputs* de uma amostra tipo *cross-section* com I produtores que atuam no mesmo segmento da indústria. Uma das abordagens usadas para estimar o grau de eficiência de cada um dos produtores consiste em estimar a função fronteira por meio de modelos econométricos, tradicionalmente classificados em modelos de fronteira determinística e modelos de fronteira estocástica (Kumbhakar & Lovell, 2000). A seguir são apresentados alguns modelos econométricos usados na estimação das funções fronteira de produção e fronteira de custo.

4.2. Estimando a Fronteira de Produção

A seguir são apresentadas duas extensões do modelo de regressão linear usadas na estimação da fronteira de produção: mínimos quadrados corrigidos e fronteira estocástica, ambas observam a premissa microeconômica que a função fronteira não é uma linha de tendência central, mas uma função envelope que representa a máxima produção possível.

No método dos mínimos quadrados corrigidos, a equação da fronteira é um modelo de regressão linear estimado por mínimo quadrados ordinários, onde o intercepto é corrigido para atender a premissa microeconômica. Por sua vez, no modelo de fronteira estocástica, a fronteira é um modelo de regressão com erro assimétrico, cujos coeficientes são estimados por máxima verossimilhança.

4.2.1. Mínimos Quadrados Corrigidos

Considere que cada produtor, em uma amostra tipo *cross-section* com I produtores utilize n tipos de insumos representados pelo vetor de quantidades de insumos $x_i \in R_+^n, i=1, I$, na produção de uma quantidade y_i de um único tipo de produto. Como visto na seção 4.1, a medida de eficiência técnica com orientação *output* para o i -ésimo produtor (TE_i) é dada por :

$$TE_i = \frac{y_i}{f(x_i)} \quad (4.1)$$

Arranjando os termos da igualdade 4.1, tem-se que :

$$y_i = TE_i \cdot f(x_i, \beta) \quad (4.2)$$

onde $f(x, \beta)$ é a fronteira determinística de produção que define a maior quantidade possível de *output* produzida com o vetor de *inputs* x , β é o vetor de parâmetros da função de produção a ser estimada e $0 \leq TE_i \leq 1$.

Aplicando uma transformação logarítmica em 4.2, tem-se a seguinte equação :

$$Lny_i = Lnf(x_i, \beta) + LnTE_i \Rightarrow Lny_i = Lnf(x_i, \beta) - u_i \quad (4.3)$$

onde $u_i = -Ln TE_i \geq 0$ é uma medida da ineficiência técnica, pois $u_i \approx 1 - TE_i$.

Então, a eficiência técnica do i -ésimo produtor é dada por :

$$TE_i = e^{-u_i} \quad (4.4)$$

Admitindo que a função de produção seja linear nos logaritmos das variáveis *inputs*, por exemplo, uma função Cobb-Douglas¹⁰ e denotando por $\varepsilon_i = -u_i$, um termo aleatório, tem-se o seguinte modelo de regressão linear :

$$Lny_i = \beta_0 + \beta_1 Lnx_i + \dots + \beta_n Lnx_n + \varepsilon_i \quad (4.5)$$

Lembrando que o grau de eficiência técnica é expresso pelo desvio da produção observada em relação à produção ideal, definida pela função de produção, é fácil perceber que na abordagem econométrica este desvio é definido por ε_i .

No modelo de regressão linear em 4.5, o termo aleatório tem média não nula ($E(\varepsilon_i) < 0$) e admite-se que este erro tenha variância constante, não tem autocorrelação serial e é independente de todas as variáveis explicativas. Para transformar o modelo 4.5 em um modelo de regressão linear clássico, faz-se a seguinte transformação :

$$Lny_i = [\beta_0 + E(\varepsilon_i)] + \beta_1 Lnx_i + \dots + \beta_n Lnx_n + \varepsilon_i - E(\varepsilon_i) \quad (4.6)$$

Fazendo $\beta_0^* = \beta_0 + E(\varepsilon_i)$ e $\varepsilon_i^* = \varepsilon_i - E(\varepsilon_i)$, obtém-se o seguinte modelo de regressão linear :

$$Lny_i = \beta_0^* + \beta_1 Lnx_i + \dots + \beta_n Lnx_n + \varepsilon_i^* \quad (4.7)$$

¹⁰ Função Cobb-Douglas : $Y = \beta_0 X_1^{\beta_1} X_2^{\beta_2}$

No modelo de regressão transformado 4.7, o novo termo aleatório ε_i^* tem distribuição normal, pois é expresso pela diferença entre uma variável aleatória e a sua média. Assim, com exceção do intercepto β_0 , os coeficientes de regressão podem ser estimados de forma consistente por meio dos mínimos quadrados ordinários (OLS - *Ordinary Least Squares*).

Para garantir que a fronteira de produção estimada envolva todos os produtores e passe pelo produtor mais eficiente, o intercepto estimado por OLS é ajustado até que todos os resíduos, exceto um, sejam negativos, conforme mostra a figura 4. Este procedimento é conhecido como mínimos quadrados corrigidos ou COLS (*Corrected OLS*).

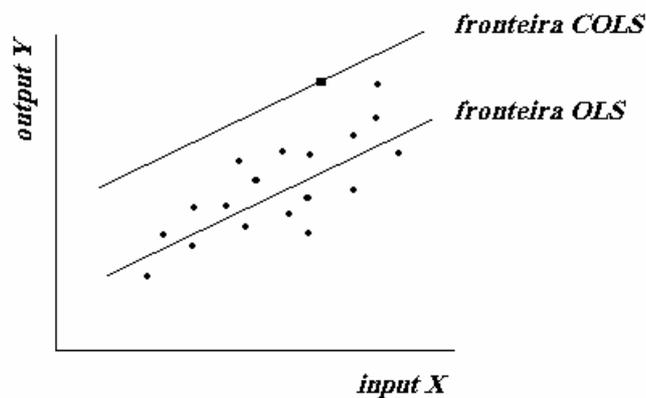


Figura 4. Fronteiras de produção estimadas por OLS e COLS

Os resíduos corrigidos e a estimativa corrigida do intercepto são dadas por :

$$\beta_0^{COLS} = \beta_0^* + \max \varepsilon_i^* \quad (4.8)$$

$$\varepsilon_i^{COLS} = \varepsilon_i^* - \max \varepsilon_i^* \quad (4.9)$$

Finalmente, a estimativa da eficiência do i -ésimo produtor é dada por :

$$TE_i = e^{\varepsilon_i^{COLS}} \quad (4.10)$$

A estimação da fronteira pelo método COLS é facilmente implementada, entretanto, como apenas o intercepto é corrigido, a fronteira de produção estimada é paralela à linha de tendência central estimada pelo OLS, uma hipótese pouco plausível e muito restrita (Kumbhakar & Lovell, 2000). Para escapar desta restrição, deve-se utilizar um modelo de fronteira estocástica.

4.2.2. Modelo de Fronteira Estocástica

Nos modelos de fronteira determinística qualquer desvio em relação à fronteira de produção é atribuído à ineficiência técnica do produtor. Tais modelos ignoram o fato de que a produção pode ser afetada por choques aleatórios fora do controle do produtor, por exemplo, greves e condições ambientais sob a produção.

Um avanço neste sentido são os modelos de fronteira estocástica introduzidos, simultaneamente, em 1977, por Aigner, Lovell & Schmidt (ALS) e Meeusen & van den Broeck (MB) (Kumbhakar & Lovell, 2000), cuja principal virtude está no fato deles reconhecerem que os desvios em relação à fronteira de produção podem ter origem na ineficiência técnica dos produtores ou são causados por choques aleatórios fora do controle dos produtores. Tais modelos são especificados como :

$$y_i = TE_i \cdot f(x_i, \beta) e^{v_i} \quad (4.11)$$

onde $f(x_i, \beta) e^{v_i}$ é a fronteira de produção estocástica e v_i é uma variável aleatória irrestrita em sinal.

A fronteira estocástica é composta de duas partes : a fronteira determinística $f(x_i, \beta)$, comum a todos os produtores, e um termo e^{v_i} , cuja finalidade é capturar o efeito de choques aleatórios que afeta especificamente o i -ésimo produtor.

Pela equação 4.4, tem-se que $TE_i = e^{-u_i}$, logo a equação 4.11 pode ser escrita como:

$$y_i = f(x_i, \beta) e^{v_i - u_i} \quad (4.12)$$

A partir de 4.11 tem-se que a medida de eficiência técnica do i -ésimo produtor é definida pela razão entre a produção observada e a máxima produção possível, especificada pela fronteira estocástica de produção, em um ambiente caracterizado por e^{v_i} , ou seja:

$$TE_i = \frac{y_i}{f(x_i, \beta) e^{v_i}} \quad (4.13)$$

Admitindo que $f(x_i, \beta)$ seja linear nos logaritmos, tem-se o seguinte modelo, após a aplicação de uma transformação logarítmica em 4.11 ou 4.12 :

$$\text{Lny}_i = \beta_0 + \beta_1 \text{Lnx}_i + \dots + \beta_n \text{Lnx}_n + \text{LnTE}_i + v_i \quad (4.14)$$

$$\text{Lny}_i = \beta_0 + \beta_1 \text{Lnx}_i + \dots + \beta_n \text{Lnx}_n + v_i - u_i \quad (4.15)$$

onde $u_i = -\text{LnTE}_i \geq 0$.

Na equação 4.15, o desvio entre o nível de produção y e a parte determinística da fronteira de produção é dado pela combinação de duas componentes:

- u_i , um erro que assume apenas valores não negativos e que captura o efeito da ineficiência técnica
- v_i , um erro simétrico que captura qualquer choque aleatório fora do controle do produtor. A hipótese de simetria da distribuição de v_i é suportada pelo fato de que condições ambientais desfavoráveis e favoráveis são igualmente prováveis (Kumbhakar & Lovell, 2000).

Fazendo $\varepsilon_i = v_i - u_i$ tem-se :

$$\text{Lny}_i = \beta_0 + \beta_1 \text{Lnx}_i + \dots + \beta_n \text{Lnx}_n + \varepsilon_i \quad (4.16)$$

Admitindo que v_i seja i.i.d. com distribuição simétrica, usualmente uma distribuição normal, e que seja independente de u_i , então, desde que u_i sempre assume um valor positivo, a distribuição do erro composto ε_i é assimétrica. Portanto, modelos de regressão com erro simetricamente distribuído e com média zero não são apropriados para estimar a função fronteira estocástica.

A estimação do modelo 4.16 por OLS produz estimativas consistentes dos coeficientes de regressão, com exceção do intercepto, desde que $E(\varepsilon_i) = -E(u_i) < 0$.

Embora os resultados obtidos por OLS não permitam estimar a eficiência técnica de cada produtor, a análise dos resíduos oriundos desta estimação provê uma forma simples de testar a presença de ineficiência técnica. Se $u_i = 0$, então o erro é simétrico ($\varepsilon_i = v_i$) e os dados não suportam a hipótese de ineficiência técnica. Entretanto, se $u_i > 0$, então, ε_i , é negativamente assimétrico e a hipótese de ineficiência técnica é aceita. Inspirados nesta idéia Schmidt e Lin propuseram a seguinte estatística de teste (Kumbhakar & Lovell, 2000):

$$\sqrt{b_1} = \frac{m_3}{(m_2)^{3/2}} \quad (4.17)$$

onde m_2 e m_3 são respectivamente os momentos de segunda e terceira ordem dos resíduos da estimação por mínimos quadrados.

Desde que v_i tem distribuição simétrica, m_3 é o momento de terceira ordem de u_i . Então $m_3 < 0$ indica que os resíduos são negativamente assimétricos e sugere a presença de ineficiência técnica. Por outro lado, quando $m_3 > 0$, os

resíduos são positivamente assimétricos, o que não faz sentido neste caso e, portanto, indica problemas de especificação do modelo.

Como a distribuição de $\sqrt{b_1}$ não é publicada, Coelli propôs a seguinte estatística de teste, assintoticamente distribuída como uma $N(0,1)$ (Kumbhakar & Lovell, 2000):

$$\frac{m_3}{\sqrt{6m_2^3 / n}} \tag{4.18}$$

Para estimar a eficiência técnica de cada produtor é necessário extrair da estimativa de ε_i as estimativas de v_i e u_i . Isto requer alguma suposição sobre as distribuições das duas componentes de erro. Estas suposições permitem a utilização do método de máxima verossimilhança na estimação dos parâmetros do modelo de regressão e também são importantes no procedimento de estimação da eficiência técnica dos produtores, baseado na distribuição condicional de u_i dado ε_i .

Na formulação básica dos modelos de fronteira estocástica, o termo aleatório v_i segue uma distribuição Normal, enquanto, u_i tem distribuição *Half-Normal*, isto é:

$$\begin{aligned} \text{(H1)} \quad v_i &\sim iid N(0, \sigma_v^2) &\Rightarrow f(v) &= \frac{1}{\sqrt{2\pi}\sigma_v} e^{-\frac{v^2}{2\sigma_v^2}} \\ \text{(H2)} \quad u_i &\sim iid N^+(0, \sigma_u^2) \text{ (Half-Normal não negativa)} &\Rightarrow f(u) &= \frac{2}{\sqrt{2\pi}\sigma_u} e^{-\frac{u^2}{2\sigma_u^2}} \\ \text{(H3)} \quad u_i &\text{ é independente de } v_i \end{aligned}$$

(H4) u_i e v_i são independentes de x_i

A partir da hipótese H3 tem-se que a distribuição conjunta de u_i e v_i é o produto das densidades marginais em H1 e H2 :

$$f(u, v) = \frac{2}{2\pi\sigma_u\sigma_v} e^{\left(-\frac{u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2}\right)} \tag{4.19}$$

Lembrando que $\varepsilon_i = v_i - u_i$, obtém-se facilmente a distribuição conjunta de u_i e ε_i :

$$f(u, \varepsilon) = \frac{2}{2\pi\sigma_u\sigma_v} e^{\left(-\frac{u^2}{2\sigma_u^2} - \frac{(\varepsilon+u)^2}{2\sigma_v^2}\right)} \tag{4.20}$$

Integrando a função densidade 4.20 em relação a u , obtém-se a distribuição marginal de ε_i , dada por :

$$f(\varepsilon) = \int_0^{\infty} f(u, \varepsilon) du = \frac{2}{\sqrt{2\pi}\sigma} \left[1 - \Phi\left(\frac{\varepsilon\lambda}{\sigma}\right) \right] \cdot e^{-\frac{\varepsilon^2}{2\sigma^2}} = \frac{2}{\sigma} \phi\left(\frac{\varepsilon}{\sigma}\right) \Phi\left(-\frac{\varepsilon\lambda}{\sigma}\right) \quad (4.21)$$

onde $\lambda = \sigma_u / \sigma_v$, $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$ e ϕ e Φ são respectivamente a densidade e a distribuição acumulada de uma $N(0,1)$.

Conforme mostra a figura 5, a distribuição de ε_i é negativamente assimétrica, sendo que a média e a variância são dadas por :

$$E(\varepsilon) = -E(u) = -\sigma_u \sqrt{\frac{2}{\pi}} \quad (4.22)$$

$$V(\varepsilon) = \frac{\pi - 2}{\pi} \sigma_u^2 + \sigma_v^2 \quad (4.23)$$

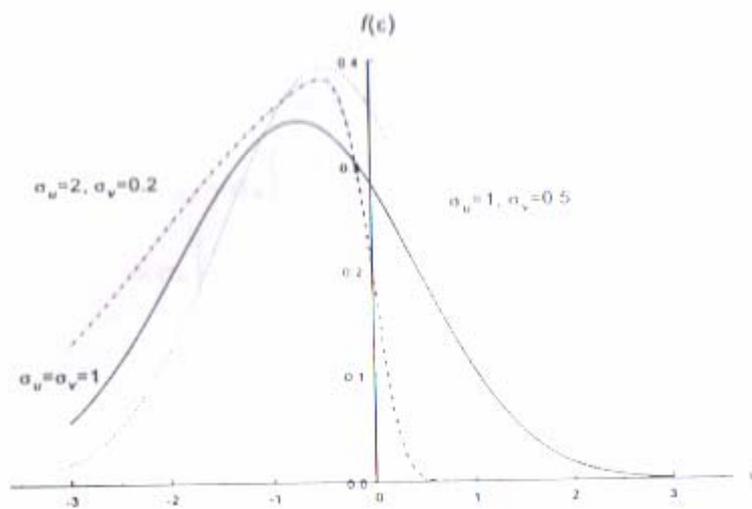


Figura 5. Função densidade do erro composto no modelo Normal/Half-Normal

Lembrando que $\varepsilon_i = Lny_i - \left(\beta_0 + \sum_{j=1}^n \beta_j Lnx_{ji} \right)$ e usando a função densidade

4.21, o logaritmo da função de verossimilhança dado uma amostra com I produtores é escrito como:

$$\ln L = \text{constante} - I \ln \sigma + \sum_i \ln \Phi\left(\frac{\varepsilon_i \lambda}{\sigma}\right) - \frac{1}{2\sigma^2} \sum_i \varepsilon_i^2 \quad (4.24)$$

As estimativas de máxima verossimilhança, são os valores de σ_u , σ_v e β_j para $j=0, n$ que maximizam a equação 4.24. Estas estimativas são assintoticamente consistentes.

Em 1977, Batters e Cora (Kumbhakar & Lovell, 2000) parametrizaram o logaritmo da verossimilhança em termos de $\sigma^2 = \sigma_u^2 + \sigma_v^2$ e $\gamma = \sigma_u^2 / (\sigma_u^2 + \sigma_v^2) \in [0,1]$. Esta nova parametrização apresenta vantagens na maximização da função de verossimilhança. Além disso, ela facilita a interpretação do modelo, pois quando $\gamma=0$, a ineficiência é inexistente, pois σ_v^2 domina a variância do erro composto e os desvios em relação à fronteira devem-se unicamente a ruídos aleatórios, por outro lado, quando $\gamma=1$, σ_u^2 domina a variância do erro composto e o desvio em relação à fronteira deve-se inteiramente a ineficiência. Portanto, pode-se utilizar o teste da razão da verossimilhança (LR *test*) (Greene, 2000) para testar a hipótese nula $H_0: \gamma=0$. Porém, como γ pode cair nos limites do intervalo $[0,1]$, a distribuição da razão de verossimilhança é uma mistura de qui-quadrados, i.e., aproximadamente uma χ_1^2 .

Para estimar a eficiência técnica de cada produtor é preciso ter uma estimativa de u_i , a componente do erro que captura o efeito da ineficiência técnica. Esta estimativa pode ser obtida a partir dos resíduos $\hat{\varepsilon}_i$, pois $\varepsilon_i = v_i - u_i$. Usando as funções densidade em 4.20 e 4.21, Jondrow *et al* (*apud* Kumbhakar & Lovell, 2000) estimaram u_i como sendo a média ou a moda da seguinte distribuição condicionada :

$$f(u|\varepsilon) = \frac{f(u, \varepsilon)}{f(\varepsilon)} = \frac{1}{\sqrt{2\pi}\sigma_*} e^{\left[\frac{-(u-\mu_*)^2}{2\sigma_*^2} \right] / \left[1 - \Phi\left(\frac{\mu_*}{\sigma_*}\right) \right]} \Rightarrow u_i|\varepsilon_i \sim N^+(\mu_*, \sigma_*^2) \quad (4.25)$$

onde $\mu_* = -\frac{\varepsilon\sigma_u^2}{\sigma^2}$ e $\sigma_*^2 = \frac{\sigma_u^2\sigma_v^2}{\sigma^2}$

Então, a estimativa pontual de u_i é definida pela média de $f(u|\varepsilon)$:

$$\hat{u}_i = E(u_i|\varepsilon_i) = \mu_* + \sigma_* \left[\frac{\phi(-\mu_*/\sigma_*)}{1 - \Phi(-\mu_*/\sigma_*)} \right] = \sigma_* \left[\frac{\phi(\varepsilon_i\lambda/\sigma)}{1 - \Phi(\varepsilon_i\lambda/\sigma)} - \frac{\varepsilon_i\lambda}{\sigma} \right] \quad (4.26)$$

ou pela moda de $f(u|\varepsilon)$

$$\hat{u}_i = M(u_i|\varepsilon_i) = -\varepsilon_i \frac{\sigma_u^2}{\sigma^2}, \text{ se } \varepsilon_i \leq 0 \quad (4.27)$$

$$\hat{u}_i = 0, \text{ caso contrário}$$

Lembrando que $u_i = -LnTE_i$, tem-se finalmente a estimativa da eficiência técnica do i -ésimo produtor :

$$TE_i = e^{-\hat{u}_i} \quad (4.28)$$

Em 1988, Battese e Coelli (Kumbhakar & Lovell, 2000) propuseram a seguinte alternativa para estimar a eficiência técnica de cada produtor:

$$TE_i = E(e^{-u_i} | \varepsilon_i) = \left[\frac{1 - \Phi(\sigma_* - \mu_{*i} / \sigma_*)}{1 - \Phi(-\mu_{*i} / \sigma_*)} \right] e^{-\mu_{*i} + \frac{1}{2}\sigma_*^2} \quad (4.29)$$

As estimativas da eficiência técnica em 4.28 e 4.29 são diferentes, pois $E(e^{u_i} | \varepsilon_i) \neq e^{-E(u_i | \varepsilon_i)}$, e em ambos os casos inconsistentes.

Explorando o fato de que a distribuição $u_i | \varepsilon_i \sim N^+(\mu_*, \sigma_*^2)$, Horrace e Schmidt (Kumbhakar & Lovell, 2000) derivaram os limites da distribuição de $u_i | \varepsilon_i$ a serem utilizados na construção de um intervalo (L_i, U_i) com $(1-\alpha)100\%$ de confiança para $e^{-u_i | \varepsilon_i}$:

$$L_i = e^{-\mu_{*i} - z_L \sigma_*} \quad U_i = e^{-\mu_{*i} - z_U \sigma_*} \quad (4.30)$$

onde

$$P(Z > z_L) = \frac{\alpha}{2} \left[1 - \Phi\left(-\frac{\mu_{*i}}{\sigma_*}\right) \right] \quad (4.31)$$

$$P(Z > z_U) = \left(1 - \frac{\alpha}{2}\right) \left[1 - \Phi\left(-\frac{\mu_{*i}}{\sigma_*}\right) \right] \quad (4.32)$$

Como $Z \sim N(0,1)$ tem-se que:

$$z_L = \Phi^{-1} \left\{ \left(1 - \frac{\alpha}{2}\right) \left[1 - \Phi\left(-\frac{\mu_{*i}}{\sigma_*}\right) \right] \right\} \quad (4.33)$$

$$z_U = \Phi^{-1} \left\{ \left[1 - \left(1 - \frac{\alpha}{2}\right) \right] \left[1 - \Phi\left(-\frac{\mu_{*i}}{\sigma_*}\right) \right] \right\} \quad (4.34)$$

A escolha da distribuição *Half-Normal* tem sido alvo de críticas, por esta razão outras distribuições tem sido sugeridas para o erro u_i , por exemplo, Meeusen & van den Broeck e também Aigner, Lovell & Schmidt (Kumbhakar & Lovell, 2000) apresentaram a verossimilhança e alguns resultados associados com a distribuição exponencial para u_i . Posteriormente, em 1980, Stevenson (Kumbhakar & Lovell, 2000) sugeriu uma distribuição Normal truncada no lugar da *Half-Normal*, e mais recentemente, em 1990, Greene sugeriu a distribuição Gama (Kumbhakar & Lovell, 2000). Apesar de tantas formulações alternativas, o modelo Normal/*Half-Normal* tem dominado as aplicações apresentadas na literatura técnica. Na verdade a escolha da distribuição de u_i tem impacto bastante

reduzido sobre as medidas de eficiência dos produtores, portanto, é natural adotar modelos mais simples como o Normal/*Half-Normal* (Kumbhakar & Lovell, 2000).

No caso do modelo da Normal truncada, proposto por Stevenson, tem-se as seguintes hipóteses para as componentes do erro composto:

$$\begin{aligned}
 \text{(H1)} \quad v_i &\sim iid N(0, \sigma_v^2) && \Rightarrow && f(v) = \frac{1}{\sqrt{2\pi}\sigma_v} e^{-\frac{v^2}{2\sigma_v^2}} \\
 \text{(H2)} \quad u_i &\sim iid N^+(\mu, \sigma_u^2) \text{ (Normal truncada)} && \Rightarrow && f(u) = \frac{1}{\sqrt{2\pi}\sigma_u\Phi(-\mu/\sigma_u)} e^{-\frac{(u-\mu)^2}{2\sigma_u^2}} \\
 \text{(H3)} \quad u_i &\text{ é independente de } v_i \\
 \text{(H4)} \quad u_i \text{ e } v_i &\text{ são independentes de } x_i
 \end{aligned}$$

A Normal truncada tem dois parâmetros (μ e σ_u^2) e é uma generalização da distribuição *Half-Normal*. A justificativa em defesa do uso da Normal truncada é que ela não impõe a restrição de moda nula, como faz a *Half-Normal*. A imposição desta restrição subestima os indicadores de eficiência.

Na função densidade $f(u)$, μ é a moda da distribuição Normal, truncada em um valor negativo. Observe que se $\mu=0$, $f(u)$ torna-se a função densidade de uma *Half-Normal*. A partir da densidade conjunta de u_i e v_i e lembrando que $\varepsilon_i = v_i - u_i$, é fácil obter a densidade conjunta de u_i e ε_i , dada por :

$$f(u, \varepsilon) = \frac{2}{2\pi\sigma_u\sigma_v\Phi(-\mu/\sigma_u)} e^{\left(-\frac{(u-\mu)^2}{2\sigma_u^2} - \frac{(\varepsilon+u)^2}{2\sigma_v^2}\right)} \quad (4.35)$$

Integrando a densidade conjunta 4.35 em relação a u , obtém-se a função densidade marginal de ε_i , dada por :

$$f(\varepsilon) = \frac{1}{\sigma} \phi\left(\frac{\varepsilon + \mu}{\sigma}\right) \Phi\left(\frac{\mu}{\sigma\lambda} - \frac{\varepsilon\lambda}{\sigma}\right) \left[\Phi\left(-\frac{\mu}{\sigma_u}\right)\right]^{-1} \quad (4.36)$$

onde, da mesma forma que no caso *Half-Normal*, $\lambda = \sigma_u/\sigma_v$ e $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$.

A partir de 4.36 tem-se a seguinte expressão para o logaritmo da função de verossimilhança, dada uma amostra com I produtores :

$$\ln L = \text{constante} - I \ln \sigma - I \ln \Phi\left(-\frac{\mu}{\sigma_u}\right) + \sum_i \ln \Phi\left(\frac{\mu}{\sigma\lambda} - \frac{\varepsilon_i\lambda}{\sigma}\right) - \frac{1}{2} \sum_i \left(\frac{\varepsilon_i + \mu}{\sigma}\right)^2 \quad (4.37)$$

onde $\sigma_u = \lambda\sigma / \sqrt{1 + \lambda^2}$.

As estimativas de máxima verossimilhança, são os valores de σ_u , μ e β_j para $j=0, n$ que maximizam a equação 4.37. De forma semelhante ao caso com distribuição *Half-Normal*, a eficiência técnica do i -ésimo produtor é estimada por $TE_i = e^{-\hat{u}_i}$, onde é dado pela média ou a moda da distribuição condicional de $u_i | \varepsilon_i \sim N^+(\hat{\mu}_i, \sigma_*^2)$:

$$\hat{u}_i = E(u_i | \varepsilon_i) = \sigma_* \left[\frac{\hat{\mu}_i}{\sigma_*} - \frac{\phi(-\hat{\mu}_i/\sigma_*)}{1 - \Phi(-\hat{\mu}_i/\sigma_*)} \right] \quad (4.38)$$

ou

$$\begin{aligned} \hat{u}_i &= M(u_i | \varepsilon_i) = \hat{\mu}_i, \text{ se } \hat{\mu}_i \geq 0 \\ \hat{u}_i &= 0, \text{ caso contrário} \end{aligned} \quad (4.39)$$

onde $\hat{\mu}_i = \frac{\mu\sigma_v^2 - \varepsilon_i\sigma_u^2}{\sigma^2}$ e $\sigma_*^2 = \frac{\sigma_u^2\sigma_v^2}{\sigma^2}$.

4.3. Estimando a Fronteira de Custos

A simples troca do sinal do erro u_i na equação 4.12 converte um modelo de fronteira estocástica de produção em modelo de fronteira estocástica de custo (Kumbhakar & Lovell, 2000), dado por :

$$E_i = C(y_i, w_i, \beta) e^{v_i + u_i}, \quad i=1, I \quad (4.40)$$

onde, $w_i = (w_{1i}, \dots, w_{ni})$ é o vetor de preços dos *inputs*, β é o vetor de parâmetros da tecnologia de produção, y_i é o vetor de *outputs* produzidos, x_i é o vetor de *inputs*, $E_i = w_i^T x_i$ é o custo total incorrido pelo i -ésimo produtor, $C(y_i, w_i, \beta) e^{v_i}$ é a fronteira estocástica de custo. Da mesma forma que na fronteira de produção, v_i é irrestrita em sinal e captura choques aleatórios fora do controle do produtor e u_i , uma variável não negativa que reflete o grau de ineficiência do produtor. Entretanto, diferentemente do modelo de fronteira estocástica de produção, na fronteira de custo o erro u_i captura o efeito da ineficiência econômica, i.e., reflete as parcelas alocativa e técnica da ineficiência.

A fronteira estocástica de custo também tem duas partes: uma parte determinística, comum a todos os produtores, e uma parte específica de cada

produtor, definida por e^{v_i} e que captura efeitos de choques aleatórios em cada produtor.

A medida de eficiência econômica do i -ésimo produtor é a razão entre o mínimo custo possível $C(y_i, w_i, \beta)$, especificado pela função fronteira estocástica de custo, e E_i , o custo do i -ésimo produtor:

$$CE_i = \frac{C(y_i, w_i, \beta)e^{v_i}}{E_i} \quad (4.41)$$

A partir das equações 4.40 e 4.41 é fácil perceber que a eficiência econômica também é dada por:

$$CE_i = e^{-u_i} \quad (4.42)$$

Neste trabalho são apresentados modelos de fronteira com equação única, baseados no custo total, na quantidade de *outputs* e nos preços dos *inputs*. Tais modelos não permitem decompor a medida de eficiência econômica nas componentes alocativa e técnica (Kumbhakar & Lovell, 2000).

Se a eficiência alocativa é assumida, o erro u_i é relacionado com a ineficiência técnica e os modelos apresentados estimam medidas de eficiência técnica, segundo uma orientação *input*. Porém se esta suposição não é feita, a interpretação de u_i não é clara, pois ele representa uma mistura de ineficiência técnica e alocativa. Por esta razão a medida de eficiência obtida a partir de uma fronteira de custo é denominada eficiência econômica.

A decomposição da medida de eficiência só é possível por meio de modelos de equações simultâneas que são mais sofisticados e necessitam de dados acerca do custo total, quantidades de *outputs*, preços e quantidades dos *inputs* de cada produtor. Tais modelos estão fora do escopo deste trabalho e mais detalhes podem ser encontrados nas referências (Kumbhakar & Lovell, 2000).

Assumindo que a parcela determinística da fronteira estocástica de custo, $C(y_i, w_i, \beta)$, tenha apenas um *output*, pode-se adotar uma forma funcional Cobb-Douglas que após a transformação logarítmica resulta na seguinte equação:

$$\ln E_i = \beta_0 + \beta_y \ln y_i + \beta_1 \ln w_{1i} + \dots + \beta_n \ln w_{ni} + v_i + u_i \quad (4.43)$$

onde $u_i = -\ln CE_i \geq 0$.

Antes de estimar o modelo 4.43 deve-se impor a homogeneidade nos coeficientes dos logaritmos dos preços:

$$\beta_1 + \dots + \beta_n = 1 \Rightarrow \beta_k = 1 - \sum_{\substack{j=1 \\ j \neq k}}^n \beta_j \quad (4.44)$$

O modelo resultante é :

$$\text{Ln} \frac{E_i}{w_{ki}} = \beta_0 + \beta_y \text{Ln} y_i + \sum_{\substack{j=1 \\ j \neq k}}^n \beta_j \frac{w_{ji}}{w_{ki}} + v_i + u_i \quad (4.45)$$

Ao contrário do modelo de fronteira estocástica de produção, no modelo de fronteira estocástica de custo, o erro composto ($\varepsilon_i = v_i + u_i$) é positivamente assimétrico. Entretanto, apesar de algumas mudanças de sinal, a estimação dos modelos de fronteira estocástica de custo é feita da mesma forma que a estimação da fronteira estocástica de produção e a técnica proposta por Jondrow *et al* pode ser usada para estimar a ineficiência econômica.

A seguir é apresentada a estimação por máxima verossimilhança do modelo 4.45 assumindo que $v_i \sim iid N(0, \sigma_v^2)$, $u_i \sim iid N^+(0, \sigma_u^2)$ e que u_i e v_i sejam independentes entre si e não correlacionados com as variáveis explicativas do modelo. De forma análoga à equação 4.21, a função densidade do erro composto, na fronteira estocástica de custo, é dada por:

$$f(\varepsilon) = \int_0^{\infty} f(u, \varepsilon) du = \frac{2}{\sqrt{2\pi}\sigma} \left[1 - \Phi\left(\frac{-\varepsilon\lambda}{\sigma}\right) \right] \cdot e^{-\frac{\varepsilon^2}{2\sigma^2}} = \frac{2}{\sigma} \phi\left(\frac{\varepsilon}{\sigma}\right) \Phi\left(\frac{\varepsilon\lambda}{\sigma}\right) \quad (4.46)$$

onde $\lambda = \sigma_u / \sigma_v$, $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$ e ϕ e Φ são respectivamente a densidade e a distribuição acumulada de uma $N(0,1)$.

A distribuição de ε_i é positivamente assimétrica com média e variância dadas por:

$$E(\varepsilon) = E(u) = \sigma_u \sqrt{\frac{2}{\pi}} \quad (4.47)$$

$$V(\varepsilon) = \frac{\pi - 2}{\pi} \sigma_u^2 + \sigma_v^2 \quad (4.48)$$

Lembrando que $\varepsilon_i = Ln \frac{E_i}{w_{ki}} - \left(\beta_0 + \beta_y Lny_i + \sum_{\substack{j=1 \\ j \neq k}}^n \beta_j Ln \frac{w_{ji}}{w_{ki}} \right)$ e usando a

função densidade 4.46, o logaritmo da função de verossimilhança dado uma amostra com I produtores é escrito como:

$$\ln L = \text{constante} - I \ln \sigma + \sum_i \ln \Phi \left(\frac{\varepsilon_i \lambda}{\sigma} \right) - \frac{1}{2\sigma^2} \sum_i \varepsilon_i^2 \quad (4.49)$$

Novamente, as estimativas de máxima verossimilhança, são os valores de σ_u , σ_v e β_j para $j=0,n$ que maximizam a equação 4.49. Estas estimativas são assintoticamente consistentes.

Para estimar a eficiência técnica de cada produtor é preciso ter uma estimativa de u_i , a componente do erro que captura o efeito da ineficiência técnica. Esta estimativa pode ser obtida a partir dos resíduos $\hat{\varepsilon}_i$, pois $\varepsilon_i = v_i - u_i$. Usando as funções densidade em 4.20 e 4.21, Jondrow *et al* (Tanure, 2000) estimaram u_i como sendo a média ou a moda da seguinte distribuição condicionada:

$$f(u|\varepsilon) = \frac{f(u, \varepsilon)}{f(\varepsilon)} = \frac{1}{\sqrt{2\pi}\sigma_*} e^{\left[\frac{-(u-\mu_*)^2}{2\sigma_*^2} \right] / \left[1 - \Phi \left(\frac{-\mu_*}{\sigma_*} \right) \right]} \Rightarrow u_i | \varepsilon_i \sim N^+(\mu_*, \sigma_*^2) \quad (4.50)$$

onde $\mu_* = \frac{\varepsilon \sigma_u^2}{\sigma^2}$ e $\sigma_*^2 = \frac{\sigma_u^2 \sigma_v^2}{\sigma^2}$

Então, a estimativa pontual de u_i pode ser a média de $f(u|\varepsilon)$:

$$\hat{u}_i = E(u_i | \varepsilon_i) = \mu_* + \sigma_* \left[\frac{\phi(-\mu_*/\sigma_*)}{1 - \Phi(-\mu_*/\sigma_*)} \right] = \sigma_* \left[\frac{\phi(\varepsilon_i \lambda / \sigma)}{1 - \Phi(\varepsilon_i \lambda / \sigma)} + \frac{\varepsilon_i \lambda}{\sigma} \right] \quad (4.51)$$

ou a moda de $f(u|\varepsilon)$

$$\hat{u}_i = M(u_i | \varepsilon_i) = \varepsilon_i \frac{\sigma_u^2}{\sigma^2}, \text{ se } \varepsilon_i \geq 0 \quad (4.52)$$

$$\hat{u}_i = 0, \text{ caso contrário}$$

Finalmente a estimativa da eficiência econômica do i -ésimo produtor é dada por:

$$CE_i = e^{-\hat{u}_i} \quad (4.53)$$

Da mesma forma que na fronteira estocástica de produção, a medida de eficiência econômica proposta por Battese e Coelli (Kumbhakar & Lovell, 2000), obtida a partir da fronteira estocástica de custo é dada por:

$$CE_i = E(e^{-u_i} | \varepsilon_i) = \left[\frac{1 - \Phi(\sigma_* - \mu_{*i} / \sigma_*)}{1 - \Phi(-\mu_{*i} / \sigma_*)} \right] e^{-\mu_{*i} + \frac{1}{2}\sigma_*^2} \quad (4.54)$$

A grande virtude da função Cobb-Douglas é a sua simplicidade. Entretanto, esta forma funcional apresenta alguns inconvenientes, como por exemplo, o fato de não acomodar múltiplos produtos sem violar as propriedades de convexidade do espaço dos produtos e não permitir a modelagem de estruturas de produção mais complexas, resultando em estimativas tendenciosas da ineficiência de custo.

Uma alternativa a função Cobb-Douglas é a função Translog, cuja principal virtude está na possibilidade de acomodar múltiplos *outputs* sem necessariamente violar as propriedades de convexidade. A seguir tem-se a equação do modelo de fronteira estocástica de custo na forma funcional Translog:

$$\begin{aligned} \ln E_i = & \beta_0 + \sum_{j=1}^m \alpha_j \ln y_{ji} + \sum_{j=1}^n \beta_j \ln w_{ji} & (4.55) \\ & + \frac{1}{2} \sum_{j=1}^m \sum_{h=1}^m \alpha_{jh} \ln y_{ji} \ln y_{hi} + \frac{1}{2} \sum_{j=1}^n \sum_{h=1}^n \beta_{jh} \ln w_{ji} \ln w_{hi} \\ & + \sum_{j=1}^m \sum_{h=1}^n \gamma_{jh} \ln w_{hi} \ln y_{ji} + v_i + u_i \end{aligned}$$

Pelo teorema de Young devem ser adicionadas as restrições $\alpha_{jh} = \alpha_{hj}$ e $\beta_{jh} = \beta_{hj}$ sobre os coeficientes de regressão do modelo 4.55 e também devem ser impostas as restrições de homogeneidade nos preços dos inputs:

$$\sum_{j=1}^n \beta_j = 1 \quad (4.56)$$

$$\sum_{j=1}^n \beta_{jh} = 0 \quad (4.57)$$

$$\sum_{j=1}^n \gamma_{jh} = 0 \quad (4.58)$$

Apesar de permitir maior flexibilidade, o modelo Translog pode apresentar problemas de multicolinearidade (Greene, 2000) se muitas variáveis explicativas são incluídas no modelo.

4.4. Diferenças entre os Modelos de Fronteira de Produção e Fronteira de Custo

Como observado, a estimação da fronteira estocástica de produção e de custo apresenta pontos semelhantes, ambas são estimadas por meio da

maximização da verossimilhança e nos dois casos a eficiência pode ser estimada por $e^{-E(u_i|\varepsilon_i)}$ ou $E(e^{u_i}|\varepsilon_i)$. Entretanto, devem ser destacadas as seguintes diferenças entre elas:

- As medidas de eficiência derivadas da fronteira de produção medem o grau de eficiência técnica de cada produtor, por outro lado, a medida de eficiência obtida a partir de uma fronteira de custo é uma medida de eficiência econômica que pode ser decomposta nas componentes alocativa e técnica.
- Na estimação da fronteira de produção adota-se uma orientação *output*, enquanto na fronteira de custo adota-se uma orientação *input*.
- A estimação da fronteira de produção necessita da informação sobre as quantidades dos *inputs* usados e dos *outputs* produzidos. Por outro lado, na fronteira de custo deve-se dispor de informações acerca do custo de produção, dos preços dos *inputs*, das quantidades dos *outputs* produzidos e das quantidades dos *inputs* usados.
- Nos dois casos o erro composto (ε_i) não é simétrico e não tem média nula. No entanto, na fronteira de produção o erro é negativamente assimétrico e no caso da fronteira de custo o erro é positivamente assimétrico.