



SPOLM 2007

ISSN 2175-6295

Rio de Janeiro- Brasil, 08 e 09 novembro de 2007.

AVALIAÇÃO DOS CUSTOS OPERACIONAIS EFICIENTES DAS DISTRIBUIDORAS DE ENERGIA ELÉTRICA: UM EXERCÍCIO COM MODELOS DE FRONTEIRA ESTOCÁSTICA BAYESIANA

Marcus Vinicius Pereira de Souza

PUC-RJ – Pontifícia Universidade Católica do Rio de Janeiro
Departamento de Engenharia Elétrica
Rua Marquês de São Vicente 225 – Gávea 22451-041 – Rio de Janeiro – RJ
mvinic@ele.puc-rio.br

Reinaldo Castro Souza

PUC-RJ – Pontifícia Universidade Católica do Rio de Janeiro
Departamento de Engenharia Elétrica
Rua Marquês de São Vicente 225 – Gávea 22451-041 – Rio de Janeiro – RJ
reinaldo@ele.puc-rio.br

José Francisco Moreira Pessanha

UERJ e CEPEL – Centro de Pesquisas de Energia Elétrica
Av. Um s/nº – Cidade Universitária 21941-590 – Rio de Janeiro – RJ
francisc@cepel.br

RESUMO

A avaliação dos custos operacionais eficientes das empresas de distribuição é um dos principais requisitos do mecanismo de regulação tarifária pelo preço teto, o modelo adotado desde 1995 no Setor Elétrico Brasileiro (SEB) em substituição ao controle pelo custo do serviço. Neste artigo apresenta-se um modelo de fronteira estocástica de custos, concebido sob a ótica Bayesiana, e os resultados obtidos pela aplicação do modelo em uma amostra formada pelas 45 maiores distribuidoras de energia elétrica que operam no SEB.

Palavras-chave: Fronteira estocástica bayesiana, Amostrador de Gibbs, regulação econômica.

ABSTRACT

The evaluation of the efficient operational costs is one of the main requirements to the price cap scheme adopted in Brazil since 1995 to regulate the electricity tariff. This work presents a Bayesian stochastic frontier model and the results obtained for a sample with the 45 more important Brazilian electricity distribution utilities.

Keywords: Bayesian stochastic frontiers, Gibbs Sampler, economic regulation

1. Introdução

No Setor Elétrico Brasileiro (SEB) os valores das tarifas das empresas distribuidoras são revisados periodicamente a cada quatro ou cinco anos. No ano de realização da revisão periódica as tarifas são reposicionadas em um nível compatível com a cobertura dos custos operacionais eficientes e com a remuneração adequada dos investimentos realizados pelas distribuidoras, garantindo assim a manutenção do equilíbrio econômico-financeiro (EEF) da concessão.

No período entre duas revisões tarifárias periódicas as tarifas são reajustadas anualmente¹ pelo Índice de Reajuste Tarifário (IRT) definido como:

$$IRT = \frac{VPA_t + VPB_0 \cdot (IGPM - X)}{RA_0} \quad (1)$$

onde VPA_t é a parcela referente aos custos não-gerenciáveis pela distribuidora (compras de energia e encargos setoriais) na data do reajuste, RA_0 é a receita anual calculada considerando-se as tarifas vigentes (sem ICMS) na data de referência anterior e o mercado de referência, e VPB_0 é parcela relativa aos custos gerenciáveis pela distribuidora (despesas com mão-de-obra, serviços de terceiros, material, depreciação e remuneração adequada dos ativos investidos e do capital de giro) na data de referência anterior ($VPB_0 = RA_0 - VPA_0$).

Conforme indicado em (1), os custos não gerenciáveis (VPA) são repassados integralmente às tarifas finais, enquanto a parcela relativa aos custos gerenciáveis (VPB) é atualizada nominalmente pelo IGPM, descontado do Fator X. Este fator incide apenas sobre os custos gerenciáveis e constitui o mecanismo que permite repassar aos consumidores, por meio da redução das tarifas, os ganhos de produtividade das distribuidoras de energia elétrica.

A Resolução ANEEL N^o 55/2004 define o Fator X como sendo a combinação de três componentes (X_e , X_a e X_c), conforme a expressão abaixo:

$$X = (X_e + X_c) \cdot (IGPM - X_a) + X_a \quad (2)$$

A componente X_a reflete o efeito da aplicação do Índice de Preço ao Consumidor Amplo (IPCA) para a componente mão-de-obra da Parcela B da concessionária. A componente X_c relaciona-se com a qualidade percebida pelos consumidores atendidos pela concessionária. Por fim, a componente X_e reflete os ganhos esperados de produtividade da concessionária em virtude do crescimento do mercado atendido pela concessionária.

Esta última componente é a mais importante, sendo obtida com base no método de fluxo de caixa descontado do tipo *forward looking*. A componente X_e é definida de maneira a igualar o valor presente líquido dos fluxos de caixa da concessionária no período tarifário, acrescido do valor residual, com o valor dos ativos da concessionária no início do período tarifário, conforme a expressão a seguir:

$$A_0 = \sum_{t=1}^N \left[\frac{(RO_t \cdot (1 - X_e)^{t-1} - T_t - OM_t - d_t) \cdot (1 - g) + d_t - I_t}{(1 + r_{WACC})^t} \right] + \frac{A_N}{(1 + r_{WACC})^N} \quad (3)$$

onde N é o período em anos entre duas revisões tarifárias consecutivas, A_0 é o valor dos ativos da concessionária na data da revisão tarifária, A_N é o valor dos ativos da concessionária no final do período tarifário, g são as alíquotas do Imposto de Renda e da Contribuição Social sobre o Lucro Líquido (CSLL), r_{WACC} é o custo médio ponderado do capital (*Weighted Average Cost of Capital - WACC*), RO_t é a receita operacional da concessionária, T_t são os

¹ O reajuste não é realizado no ano da revisão tarifária.

² Obtido adicionando-se o valor inicial dos ativos aos investimentos líquidos de depreciação do período tarifário

³ Calculado conforme os procedimentos descritos na Nota Técnica N^o 122/2005/SRE/ANEEL

valores dos tributos (PIS/PASEP, COFINS e P&D), OM_t são os custos de operação e manutenção da concessionária, I_t é o montante de investimentos realizados e d_t é a depreciação, todos referentes ao ano t .

Na Resolução ANEEL N^o 55/2004 encontram-se detalhados os critérios utilizados na projeção das variáveis que compõem o fluxo de caixa das concessionárias de distribuição no período entre duas revisões tarifárias. Por exemplo, a projeção da receita operacional é determinada pelo produto entre o mercado previsto e a tarifa média do reposicionamento tarifário. Já a projeção do fluxo de investimentos inclui as obras de expansão do sistema de distribuição e subtransmissão para atender o crescimento do mercado e os investimentos requeridos para a renovação das instalações de distribuição que chegaram ao fim de sua vida útil. Por sua vez os custos operacionais⁴ são projetados com base nos custos da “Empresa de Referência”, referenciados à data do reposicionamento tarifário (ANEEL, 2004).

A empresa de referência é uma empresa virtual onde se simula a prestação do serviço de distribuição de energia elétrica nas mesmas condições nas quais opera uma empresa real. Para um dado nível de qualidade exigido pelo regulador, a empresa de referência é orientada para a utilização eficiente dos recursos, estabelecendo assim uma estimativa dos custos operacionais eficientes, ou seja, um padrão de desempenho (*benchmark*) a ser perseguido pela empresa real. Assim, esta estratégia de regulação emula uma competição entre uma distribuidora real e a respectiva empresa de referência.

A metodologia de empresa de referência adotada no primeiro ciclo de revisões tarifárias periódicas segue uma abordagem *bottom-up* e envolve desde a identificação de todos os processos inerentes à atividade de distribuição de energia elétrica nas áreas comercial e técnica até o estabelecimento de custos eficientes para cada um dos processos e atividades (ANEEL, 2006b). Portanto, trata-se de uma metodologia extremamente complexa e que abre a possibilidade de “imiscuir o regulador em uma espécie de micro-gestão da empresa” (ANEEL, 2006b), justamente o contrário do que recomenda a boa prática da ação reguladora.

Para evitar a complexidade da metodologia da empresa de referência e tornar mais objetivo o processo de estabelecimento do nível dos custos operacionais eficientes, a alternativa é a utilização de técnicas de *benchmarking*, entre as quais destacam-se os modelos para identificação ou estimação de fronteiras de eficiência.

Os principais métodos para identificação de fronteiras de eficiência são a Análise Envoltória de Dados ou DEA⁵, uma técnica baseada em programação linear (COOPER *et al.*, 2000), e os modelos econométricos, entre os quais destacam-se os modelos de fronteira estocástica ou SFA⁶ (KUMBHAKAR & LOVELL, 2000).

Este trabalho faz uma breve introdução ao modelo de fronteira estocástica de custo, nas versões clássica e Bayesiana, e apresenta os resultados obtidos pelas duas abordagens para uma *cross-section* formada por 45 distribuidoras de energia elétrica que operam no SEB com dados referentes ao ano 2001.

As medidas de eficiência obtidas podem auxiliar na definição dos níveis dos custos operacionais eficientes. A escolha do modelo de fronteira estocástica deve-se ao fato de que nos modelos fronteira determinística (DEA) qualquer desvio em relação à fronteira de produção é atribuído à ineficiência das empresas. Tais modelos ignoram o fato de que a produção pode ser afetada por choques aleatórios não gerenciáveis, por exemplo, greves e condições ambientais sob a produção (KUMBHAKAR & LOVELL, 2000). Diferente destes últimos, os modelos de fronteira estocástica reconhecem que os desvios em relação à fronteira podem ter origem na ineficiência ou são causados por choques aleatórios não gerenciáveis.

A seguir, na seção 2, tem-se uma breve introdução à abordagem clássica do modelo de fronteira estocástica. A versão Bayesiana da fronteira estocástica é introduzida na seção 3, juntamente com uma breve descrição do amostrador de Gibbs, o algoritmo utilizado na estimação da fronteira. A especificação do modelo econométrico adotado e os principais

⁴ Custos de operação, manutenção, administração e gestão comercial.

⁵ Data Envelopment Analysis

⁶ Stochastic Frontier Analysis

resultados obtidos são apresentados na seção 4. Finalmente, na seção 5 são resumidas as principais conclusões do trabalho.

2. Modelos de Fronteira Estocástica (abordagem clássica)

2.1 Fronteira de Produção

Considere que cada concessionária de distribuição em uma amostra tipo *cross-section* com I DMUs (*decision making units*) 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. A medida de eficiência técnica para a i -ésima DMU ($0 \leq TE_i \leq 1$) é definida pela razão entre a sua produção (y_i) e a máxima produção possível, especificada pela fronteira estocástica de eficiência ($f(x_i, \beta)e^{v_i}$):

$$TE_i = \frac{y_i}{f(x_i, \beta)e^{v_i}} \quad (4)$$

onde β é o vetor de parâmetros da função de produção a ser estimada e v_i é uma variável aleatória irrestrita em sinal.

Arranjando os termos da igualdade (4) tem-se a equação do modelo de fronteira estocástica:

$$y_i = TE_i \cdot f(x_i, \beta)e^{v_i} \quad (5)$$

A fronteira estocástica é composta de duas partes: a fronteira determinística $f(x_i, \beta)$, comum a todos os produtores, e o termo e^{v_i} , cuja finalidade é capturar o efeito de choques aleatórios que afetam especificamente a i -ésima DMU. Seja $u_i \geq 0$ uma variável aleatória. Fazendo $TE_i = e^{-u_i}$ e substituindo em (5) tem-se que:

$$y_i = f(x_i, \beta)e^{v_i - u_i} \quad (6)$$

Admitindo que $f(x_i, \beta)$ seja linear nos logaritmos, tem-se o seguinte modelo, após a aplicação de uma transformação logarítmica em (6):

$$\ln y_i = \beta_0 + \beta_1 \ln x_i + \dots + \beta_n \ln x_n + v_i - u_i \quad (7)$$

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

Na equação (7), o desvio entre o nível de produção observado 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, e v_i , um erro simétrico que captura qualquer choque aleatório não gerenciável. A hipótese de simetria da distribuição de v_i é suportada pelo fato de que condições ambientais desfavoráveis e favoráveis são igualmente prováveis. Fazendo $\varepsilon_i = v_i - u_i$ tem-se a seguinte equação de regressão:

$$\ln y_i = \beta_0 + \beta_1 \ln x_i + \dots + \beta_n \ln x_n + \varepsilon_i \quad (8)$$

Admitindo que v_i seja *i.i.d.* com distribuição simétrica, usualmente uma distribuição normal, e que seja independente de u_i , o qual sempre assume um valor não negativo, a distribuição do erro composto ε_i é assimétrica. Portanto, métodos de estimação que admitem erros simetricamente distribuídos e com média zero não são apropriados para estimar a função fronteira estocástica em (8).

Na formulação básica dos modelos de fronteira estocástica são admitidas as seguintes hipóteses acerca dos termos aleatórios u_i e v_i :

$$(H1) \quad v_i \sim iid N(0, \sigma_v^2) \quad \Rightarrow f(v) = \left(1/\sqrt{2\pi\sigma_v}\right) e^{-v^2/2\sigma_v^2}$$

$$(H2) \quad u_i \sim iid N^+(0, \sigma_u^2) \text{ (Half-Normal não negativa)} \quad \Rightarrow f(u) = \left(2/\sqrt{2\pi\sigma_u}\right) e^{-u^2/2\sigma_u^2}$$

(H3) u_i é independente de v_i

(H4) v_i e u_i são independentes de x_i .

As hipóteses acima permitem a utilização do método de máxima verossimilhança na estimação dos parâmetros da equação (8). 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):

$$f(u, v) = \frac{2}{2\pi\sigma_u\sigma_v} e^{\left(-\frac{u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2}\right)} \quad (9)$$

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

$$f(u, \varepsilon) = \frac{2}{2\pi\sigma_u\sigma_v} e^{\left(-\frac{u^2}{2\sigma_u^2} - \frac{(\varepsilon+u)^2}{2\sigma_v^2}\right)} \quad (10)$$

Integrando a função densidade (10) em relação a u_i obtém-se a distribuição marginal de ε_i , dada por:

$$f(\varepsilon) = \int_0^{\infty} f(u, \varepsilon) du = \frac{2}{\sqrt{2\pi}\sigma} \left[1 - \Phi\left(\frac{\varepsilon\lambda}{\sigma}\right)\right] \cdot e^{-\frac{\varepsilon^2}{2\sigma^2}} = \frac{2}{\sigma} \phi\left(\frac{\varepsilon}{\sigma}\right) \Phi\left(-\frac{\varepsilon\lambda}{\sigma}\right) \quad (11)$$

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

Para diferentes valores de σ_u^2 e σ_v^2 , a distribuição de ε_i é negativamente assimétrica com média e variância dadas por:

$$E(\varepsilon) = -E(u) = -\sigma_u \sqrt{\frac{2}{\pi}} \quad (12)$$

$$V(\varepsilon) = \frac{\pi-2}{\pi} \sigma_u^2 + \sigma_v^2 \quad (13)$$

Lembrando que $\varepsilon_i = \ln y_i - (\beta_0 + \beta_1 \ln x_{1i} + \dots + \beta_n \ln x_{ni})$ e usando a função densidade em (11), o logaritmo da função de verossimilhança para uma amostra com I produtores é escrito como:

$$\ln L = \text{constante} - I \ln \sigma + \sum_i \ln \Phi\left(\frac{\varepsilon_i \lambda}{\sigma}\right) - \frac{1}{2\sigma^2} \sum_i \varepsilon_i^2 \quad (14)$$

As estimativas de máxima verossimilhança, são os valores de σ_u , σ_v e β_j , para $j=0, n$, que maximizam a equação (14). Estas estimativas são assintoticamente consistentes.

A parametrização do logaritmo da verossimilhança em termos de $\sigma^2 = \sigma_u^2 + \sigma_v^2$ e $\gamma = \sigma_u^2 / (\sigma_u^2 + \sigma_v^2) \in [0,1]$ ajuda na maximização da função de verossimilhança e facilita a interpretação do modelo. Quando $\gamma \neq 0$, 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*) para testar a hipótese nula $H_0: \gamma \neq 0$. Porém, como γ pode cair nos limites do intervalo $[0,1]$, a distribuição da razão de verossimilhança é aproximadamente uma χ_1^2 .

Para estimar a eficiência técnica de cada produtor é necessário 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 $\hat{\varepsilon}_i$. Usando as funções densidade em (10) e (11) pode-se estimar u_i como sendo a média ou a moda da seguinte distribuição condicionada:

$$f(u|\varepsilon) = \frac{f(u, \varepsilon)}{f(\varepsilon)} = \frac{1}{\sqrt{2\pi}\sigma_*} e^{\left[\frac{-(\mu - \mu_*)^2}{2\sigma_*^2} \right] / \left[1 - \Phi\left(\frac{-\mu_*}{\sigma_*} \right) \right]} \Rightarrow u_i | \varepsilon_i \sim N^+(\mu_*, \sigma_*^2) \quad (15)$$

onde $\mu_* = -\varepsilon\sigma_u^2/\sigma^2$ e $\sigma_*^2 = \sigma_u^2\sigma_v^2/\sigma^2$

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

$$\hat{u}_i = E(u_i | \varepsilon_i) = \mu_* + \sigma_* \left[\frac{\phi(-\mu_*/\sigma_*)}{1 - \Phi(-\mu_*/\sigma_*)} \right] = \sigma_* \left[\frac{\phi(\varepsilon_i \lambda / \sigma)}{1 - \Phi(\varepsilon_i \lambda / \sigma)} - \frac{\varepsilon_i \lambda}{\sigma} \right] \quad (16)$$

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

$$\begin{aligned} \hat{u}_i &= M(u_i | \varepsilon_i) = -\varepsilon_i \cdot \sigma_u^2 / \sigma^2, \text{ se } \varepsilon_i \leq 0 \\ \hat{u}_i &= 0, \text{ caso contrário} \end{aligned} \quad (17)$$

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

$$TE_i = e^{-\hat{u}_i} \quad (18)$$

2.2 Fronteira de Custo

Da mesma forma que na fronteira de produção, v_i é irrestrita em sinal e captura choques aleatórios não gerenciáveis e u_i , é 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 medida de eficiência econômica da i -ésima DMU é a razão entre o custo da i -ésima DMU (E_i) e o mínimo custo possível especificado pela função fronteira estocástica $C(y_i, w_i, \beta)$:

$$CE_i = \frac{E_i}{C(y_i, w_i, \beta)e^{v_i}} = e^{u_i} \quad (19)$$

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 pela i -ésima DMU, $C(y_i, w_i, \beta)e^{v_i}$ é o custo estabelecido pela fronteira estocástica.

Arranjando os termos da igualdade em (19) encontra-se a seguinte equação:

$$E_i = C(y_i, w_i, \beta)e^{v_i + u_i}, \quad i=1, I \quad (20)$$

Neste trabalho são apresentados modelos de fronteira com equação única. Tais modelos baseiam-se no custo total, nas quantidades dos produtos e nos preços dos insumos e não permitem decompor a medida de eficiência econômica nas componentes alocativa e técnica. A decomposição da medida de eficiência só é possível por meio de modelos de equações simultâneas. 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 ao insumo (*input*).

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

$$\ln E_i = \beta_0 + \beta_y \ln y_i + \beta_1 \ln w_{1i} \dots + \beta_n \ln w_{ni} + v_i + u_i \quad (21)$$

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

Comparando as equações (7) e (21) pode-se perceber que na fronteira de custo o termo relacionado à ineficiência, u_i , é adicionado à fronteira. A razão para tal consiste no fato de que a função custo representa o custo mínimo enquanto a função produção representa o produto máximo. Apesar destas diferenças, a estimação dos parâmetros da equação (21) e dos índices de eficiência das DMUs são realizadas da mesma forma que na fronteira de produção.

3. Fronteira estocástica (abordagem Bayesiana)

Um modelo Bayesiano pode ser definido pela combinação de uma função de verossimilhança com a distribuição de probabilidades, assumida a priori, para os parâmetros do modelo e que representa o conhecimento prévio do especialista com relação a estes parâmetros (MIGON & GAMERMAN, 1999). O produto dessas duas distribuições fornece a distribuição conjunta das observações e dos parâmetros. Matematicamente, tem-se que:

$$p(\mu|y) = \frac{p(\mu)L(\mu|y)}{p(y)} \propto p(\mu)L(\mu|y) \quad (22)$$

Onde μ é o vetor de parâmetros, y denota o vetor ou a matriz de dados, $p(\mu)$ a densidade de probabilidade assumida a priori, $L(\mu|y)$ a função de verossimilhança; e $p(\mu|y)$ a densidade a posteriori. A densidade a posteriori resume todas as informações disponíveis sobre o vetor de parâmetros (μ) condicional aos dados observados. Inferências sobre μ podem ser obtidas através da densidade a posteriori; por exemplo, a média posteriori de μ , $E(\mu|y) = \int \mu p(\mu|y) d\mu$, é frequentemente usada como uma estimativa para μ . A seguir, a Figura 1 ilustra esses conceitos:

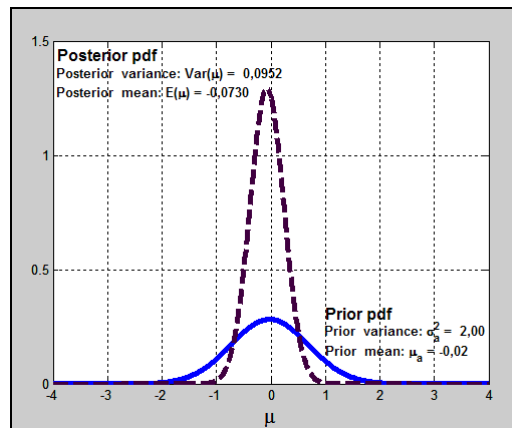


Figura 1 Densidades a priori e a posteriori para μ .

A abordagem Bayesiana é adaptativa e permite revisar a distribuição a priori dos parâmetros com novas informações obtendo-se ao final uma distribuição a posteriori. A Figura 1 ilustra claramente que a aquisição da amostra e a introdução da informação contida nesta amostra modifica a distribuição a priori com uma considerável redução com relação à incerteza sobre o parâmetro μ .

A integral $E(\mu|y) = \int \mu p(\mu|y) d\mu$ pode ser bastante complexa e não ter uma solução analítica. Neste caso a solução deve ser obtida por métodos de integração numérica, entre os quais destaca-se o MCMC (*Markov chain Monte Carlo*), em especial o Amostrador de Gibbs.

Basicamente, esta técnica permite gerar variáveis aleatórias de uma distribuição (marginal) indiretamente, sem ter de resolvê-la.

CASELLA & ROBERT (1992) descrevem que embora a maioria das aplicações do Amostrador de Gibbs tem sido em modelos Bayesianos, o mesmo pode ser perfeitamente utilizado na análise clássica. Excelentes textos sobre os métodos MCMC são, por exemplo: KOOP (1994), TIERNEY (1994), GAMERMAN (1997) e ROBERT & CASELLA (1999).

3.1 Fronteira Estocástica de Custo

Em sua versão básica, o modelo econométrico com erro composto para estimação de uma fronteira estocástica de custo pode ser escrito como:

$$y_i = h(x_i, w_i, \beta) + v_i + u_i, \quad i = 1, \dots, N \quad (23)$$

onde y_i são os custos ou as despesas totais do processo produtivo analisado, neste artigo y_i é o custo operacional da i -ésima distribuidora; x_i é vetor com as quantidades dos produtos gerados pelo processo produtivo; w_i é o vetor com os preços dos insumos; β é um vetor de k parâmetros desconhecidos que definem a função $h(x_i; w_i; \beta)$, a parte determinística da fronteira, cujo desvio em relação ao nível de produção observado é dado pela combinação de duas componentes aleatórias: u_i , um termo aleatório que assume apenas valores não negativos e que captura o efeito da ineficiência técnica ($efici\ência_i = \exp(-u_i)$), e v_i , um termo aleatório que captura choques aleatórios não gerenciáveis pelas DMUs.

A hipótese de simetria para a distribuição de v_i é natural e significa que condições ambientais desfavoráveis e favoráveis são igualmente prováveis. Por sua vez, a distribuição de probabilidade para a componente u_i ($u_i \geq 0$) não é evidente a priori e por isso pode ser especificada segundo diferentes formas, por exemplo, MEEUSEN *et al.* (1977) utilizam a distribuição exponencial, AIGNER *et al.* (1977) recomendam a distribuição *Half-Normal*, STEVENSON (1980) propõe a Normal-Truncada e, finalmente, GREENE (1990) sugere a distribuição Gama. Mais recentemente MEDRANO & MIGON, 2004 utilizaram a distribuição log Normal.

A incerteza na distribuição do termo aleatório u_i e na função $h(x_i; w_i; \beta)$ sugere a utilização de técnicas de inferência Bayesiana, conforme apresentado nos trabalhos pioneiros de VAN DEN BROECK *et al.* (1994) e KOOP *et al.*, (1995).

A formulação de um modelo Bayesiano inicia-se com a especificação das distribuições de probabilidade assumidas a priori para os parâmetros do modelo, neste caso os coeficientes β da função fronteira e os parâmetros das distribuições de probabilidade associadas com os termos aleatórios u e v . A seguir, na Tabela 1, são apresentadas duas especificações para a fronteira de custos sob a ótica Bayesiana, diferenciadas somente pela distribuição de probabilidade assumida para o termo aleatório u_i .

Tabela 1 Especificações dos modelos *Half-Normal* e exponencial

$u_i \sim \text{Half-Normal}$	$u_i \sim \text{Exponencial}$
$p(u_i \lambda) \stackrel{iid}{\sim} N^+(0, \lambda^{-1})$	$p(u_i \lambda) \stackrel{iid}{\sim} \exp(\lambda)$
$\lambda \sim \text{Gama}\left(1, \frac{1}{37,5}\right)$	$\lambda \sim \text{Gama}(1, -\log(r^*))$
	$r^* = \text{mediana da distribuição a priori da eficiência}$

$\beta_k \stackrel{iid}{\sim} N(0, \sigma_\beta^2)$	$\beta_k \stackrel{iid}{\sim} N(0, \sigma_\beta^2)$
$\sigma^{-2} \sim Gama(a, b)$	$\sigma^{-2} \sim Gama(a, b)$

3.2 Amostrador de Gibbs.

O Amostrador de Gibbs foi originalmente concebido dentro do contexto de reconstrução de imagens e está contido em uma grande classe de esquemas de simulação estocástica que utilizam cadeias de Markov (GAMERMAN, 1997). Embora ele seja um caso especial do algoritmo Metropolis-Hasting, ele apresenta duas particularidades, a saber:

- Todos os pontos gerados são aceitos;
- Existe a necessidade de conhecermos a distribuição condicional completa.

A distribuição condicional completa é a distribuição da i -ésima componente do vetor de parâmetros θ , condicionada em todas as outras componentes.

Segundo GAMERMAN (1997), o amostrador de Gibbs é, essencialmente, um esquema iterativo de amostragem de uma cadeia de Markov cujo núcleo de transição é formado pelas distribuições condicionais completas.

Para descrever este algoritmo, suponha que a distribuição de interesse seja $p(\theta)$ onde $\theta = (\theta_1, \dots, \theta_d)$. Cada uma das componentes θ_i pode ser um escalar, um vetor ou uma matriz. Ressalta-se que a distribuição p não precisa, necessariamente, ser uma distribuição a posteriori. A execução do algoritmo é efetuada conforme os passos a seguir:

- inicialize o contador de iterações da cadeia $j=1$ e arbitre valores iniciais $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})$;
- obtenha um novo valor de $\theta^{(j)} = (\theta_1^{(j)}, \dots, \theta_d^{(j)})$ a partir de $\theta^{(j-1)}$ através de sucessivas gerações de valores

$$\theta_1^{(j)} \sim p(\theta_1 | \theta_2^{(j-1)}, \dots, \theta_d^{(j-1)})$$

$$\theta_2^{(j)} \sim p(\theta_2 | \theta_1^{(j)}, \theta_3^{(j-1)}, \dots, \theta_d^{(j-1)})$$
 M

$$\theta_d^{(j)} \sim p(\theta_d | \theta_1^{(j)}, \dots, \theta_{d-1}^{(j)})$$
- mude o contador de j para $j+1$ e retorne a (ii) até a convergência.

Assim, cada iteração se completa após d movimentos ao longo dos eixos coordenados das componentes de θ . Após a convergência, os valores resultantes formam uma amostra de $p(\theta)$. Vale notar que, mesmo em problema de grandes dimensões todas as simulações podem ser univariadas, o que em geral é uma vantagem computacional (EHLERS, 2003).

No contexto dos modelos de fronteira estocástica deve-se fazer inferências não apenas sobre os parâmetros β , mas também sobre as eficiências de cada firma.

Por exemplo, considerando $u_i \stackrel{iid}{\sim} Gama(1, \lambda)$, i.e., $u_i \stackrel{iid}{\sim} \exp(\lambda)$, a distribuição conjunta de y e u , dado x e o vetor de parâmetros $\theta_* = (\beta, \sigma^2, \lambda)$ é dado por:

$$p(y_i, u_i | x_i, \theta_*) = N(y_i | h(x_i; \beta) + u_i, \sigma^2) \cdot Gama(u_i | 1, \lambda) \quad (24)$$

Dado que $p(u_i|y_i, x_i, \theta_*) = \frac{p(y_i, u_i|x_i, \theta_*)}{\int p(y_i, u_i|x_i, \theta_*) dz}$, tem-se que:

$$p(u_i|y_i, x_i, \theta_*) = \left[\Phi \left(\frac{h(x; \beta) - y_i - \lambda \sigma^2}{\sigma} \right) \right]^{-1} N(u_i | h(x; \beta) - y_i - \lambda \sigma^2 | \sigma^2) \quad (25)$$

De acordo com BKOS (1994), sendo $\Phi(\cdot)$ a distribuição normal padrão, temos que u é distribuído como uma distribuição Normal-Truncada, a mesma obtida por JONDROW et al (1982). Como a distribuição condicional completa para u é conhecida, pode-se utilizar o Amostrador de Gibbs para gerar observações da densidade posteriori conjunta. Estas observações podem ser utilizadas na inferência sobre as quantidades de interesse.

4. Resultados

Os índices de eficiência estimados pelo modelo de fronteira estocástica devem expressar o potencial de redução dos custos operacionais de cada distribuidora avaliada. Assim, a variável dependente do modelo econométrico é o custo operacional total (OPEX), enquanto as variáveis explicativas são os *cost-drivers* dos custos operacionais, representados por três agregados físicos:

- x_1 = quantidade de energia distribuída (MWh)
- x_2 = total de unidades consumidoras
- x_3 = extensão da rede de distribuição (km).

Uma empresa de distribuição tem por finalidade interligar os consumidores aos barramentos da rede de transmissão com tensões superiores ou iguais a 230 kV, portanto é uma empresa “fio”. Assim, o principal produto ofertado por uma distribuidora de energia elétrica é a capacidade de transmissão da sua rede, que tem como *proxy* o total de energia distribuída. O total de unidades consumidoras atendidas por uma distribuidora pode ser considerada como uma *proxy* da quantidade de serviços prestados pela distribuidora, tais como a conexão dos consumidores a rede, bem como a medição do consumo de cada unidade consumidora. Por fim, a extensão da rede de distribuição é uma medida da dispersão dos consumidores na área de concessão da distribuidora.

$$\text{LnOPEX}_i = \beta_0 + \beta_1 \text{Ln}x_{1i} + \beta_2 \text{Ln}x_{2i} + \beta_3 \text{Ln}x_{3i} + u_i + v_i \quad (26)$$

Considerando uma especificação tipo Cobb-Douglas para a fronteira de custos eficientes, conforme ilustrado pela equação (26), e admitindo que $u \sim \text{Half-Normal}$ e $v \sim \text{Normal}$, estimou-se a fronteira de custos eficientes sob a ótica clássica. Em seguida, o mesmo modelo foi estimado segundo a ótica Bayesiana por meio do amostrador de Gibbs. Por fim, ainda sob a ótica Bayesiana estimou-se a fronteira de custos eficientes admitindo que u tenha uma distribuição exponencial.

Para a primeira situação (modelo clássico), a estimação foi realizada pelo programa FRONTIER 4.1 disponível em www.uq.edu.au/economics/cepa/software.htm. Os modelos Bayesianos foram implementados usando o software livre WinBUGS (*Bayesian inference Using Gibbs Sampling for Windows*) que pode ser obtido em www.mrc-bsu.cam.ac.uk/bugs/Welcome.htm. Neste caso, foram realizadas 120.000 iterações, sendo que a análise de convergência dos parâmetros foi realizada por meio dos gráficos de autocorrelação serial disponibilizadas no WinBugs.

Na Tabela 2 são apresentadas as estimativas da fronteira de custos eficientes obtidas pelos três modelos considerados neste artigo. Nos modelos que assumem distribuição Half-

Normal para o termo aleatório u , as abordagens Bayesiana e Clássica convergem para as mesmas estimativas pontuais. Ressalta-se que as estimativas obtidas pela abordagem clássica pertencem aos intervalos de confiança encontrados pela abordagem Bayesiana. A substituição da distribuição Half-Normal pela distribuição exponencial não modificou de forma significativa as estimativas dos coeficientes da equação da fronteira de custos eficientes.

Tabela 2 Estimativas dos coeficientes de regressão e parâmetros do modelo de fronteira estocástica

Parâmetros	Exponencial (Bayesiano)	Half-Normal (Bayesiano)	Half-Normal (Clássico)
β_0	-1,21 (-2,03 ; -0,40)	-1,32 (-2,12 ; -0,58)	-1,26
β_1	0,37 (0,13 ; 0,61)	0,32 (0,09 ; 0,56)	0,28
β_2	0,52 (0,23 ; 0,82)	0,59 (0,28 ; 0,88)	0,63
β_3	-0,03 (-0,16 ; 0,09)	-0,05 (-0,18 ; 0,08)	-0,06
λ	7,46 (2,29 ; 25,53)	0,26 (0,10 ; 0,49)	
σ^2	0,27 (0,13 ; 0,40)	0,18 (0,07 ; 0,32)	0,27

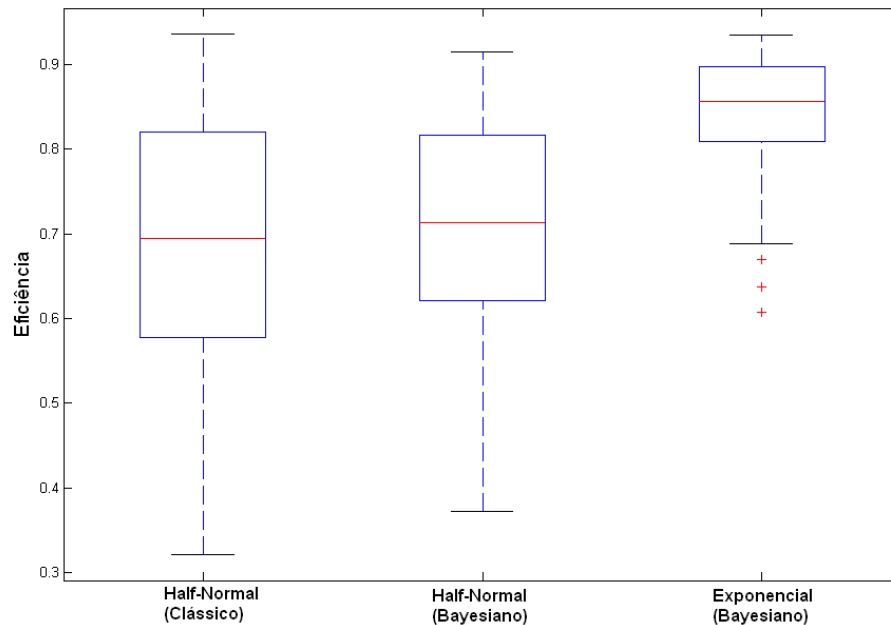


Figura 2 – Boxplots dos índices de eficiência

Embora as estimativas dos parâmetros da fronteira de eficiência (β) não sejam sensíveis à escolha da distribuição de probabilidade para o termo aleatório u , o mesmo não acontece com as estimativas dos índices de eficiência.

Conforme apresentado na Tabela 3 e, principalmente, nos *boxplots* da Figura 2, as estimativas dos índices de eficiência calculadas pelos modelos com distribuição Half-Normal são nitidamente distintas das eficiências estimadas pelo modelo com distribuição exponencial. Finalizando, o primeiro modelo (Half-Normal) apresentou um melhor desempenho segundo o critério “Deviance” (MEDRANO & MIGON, 2004).

Tabela 3 – Índice de Eficiência

DMU	Half-Normal (Clássico)	Half-Normal (Bayesiano)	Exponencial (Bayesiano)
BANDEIRANTES	0,6256	0,6743	0,8435
PIRATININGA	0,9227	0,8987	0,9277

CPFL	0,8193	0,8157	0,8997
AES-SUL	0,9358	0,9154	0,9350
CELESC	0,4806	0,5396	0,7626
ESELSA	0,7730	0,7831	0,8895
RGE	0,9252	0,9038	0,9301
CEMIG	0,5244	0,5865	0,7938
COPEL	0,8066	0,8022	0,8932
ELEKTRO	0,8230	0,8202	0,9016
ELETROPAULO	0,6426	0,6736	0,8352
LIGHT	0,6851	0,7096	0,8545
CEEE	0,3217	0,3724	0,6080
CERJ	0,7606	0,7633	0,8763
MANAUS	0,5950	0,6333	0,8138
CELG	0,4136	0,4685	0,7018
CELPE	0,8801	0,8584	0,9109
COELBA	0,6974	0,7139	0,8545
COELCE	0,7428	0,7515	0,8725
CELPA	0,4167	0,4616	0,6883
CERON	0,5210	0,5646	0,7723
CEMAT	0,3989	0,4577	0,6961
ENERSUL	0,6020	0,6524	0,8308
COSERN	0,8778	0,8582	0,9120
ENERGIPE	0,8527	0,8375	0,9045
CELTINS	0,4537	0,4965	0,7181
CEPISA	0,6674	0,6857	0,8373
CEAL	0,6919	0,7086	0,8517
CEMAR	0,6399	0,6630	0,8263
SAELPA	0,8412	0,8221	0,8963
CEB	0,3863	0,4358	0,6701
BOA VISTA	0,3622	0,4075	0,6374
CENF	0,8080	0,7946	0,8852
CAUIÁ	0,6407	0,6701	0,8351
ELETROACRE	0,7975	0,7872	0,8833
URUSSANGA	0,4874	0,5462	0,7609
XANXERÊ	0,6098	0,6475	0,8235
CPEE	0,8497	0,8348	0,9051
ELETROCAR	0,8136	0,8060	0,8936
PANAMBI	0,8155	0,7988	0,8879
SANTA MARIA	0,7519	0,7611	0,8766
CHESP	0,8873	0,8601	0,9107
POÇOS DE CALDAS	0,6786	0,7093	0,8567
CAT-LEO	0,6951	0,7158	0,8579
SULGIPE	0,9095	0,8792	0,9162

5. Conclusões

Nos próximos ciclos de revisão tarifária a ANEEL sinaliza com a possibilidade de utilizar modelos DEA e SFA na determinação dos custos operacionais eficientes, um elemento fundamental no cálculo do Fator X das distribuidoras. As duas abordagens partem de pressupostos diferentes: o método DEA é determinístico e os desvios em relação à fronteira de eficiência são tomados como sendo unicamente devido às ineficiências das distribuidoras, por sua vez, o método SFA tem uma natureza estocástica e estima uma medida de eficiência livre dos impactos dos fatores aleatórios não controláveis pelas DMUs.

Neste trabalho pôde ser observado que a análise conjunta de modelos de fronteira estocástica sob o prisma clássico e Bayesiano é de fundamental importância. A vantagem da abordagem Bayesiana reside na possibilidade de desenvolver um modelo incorporando novas informações sobre os parâmetros (prioris informativas) e até mesmo particularidades das empresas. Tal procedimento não é possível na análise clássica.

Por fim, lembra-se que na passagem dos índices de eficiência para os custos operacionais eficientes a ANEEL pode considerar um critério semelhante ao adotado para as

transmissoras, no qual são reconhecidos pelo menos 80% dos custos operacionais. Isto implica na seguinte normalização (ANEEL, 2007) dos índices de eficiência (θ_i) obtidos pelos modelos DEA e SFA, antes de serem aplicados no cálculo dos custos operacionais eficientes $OPEX_i(1 - \theta_i^*)$:

$$\theta_i^* = 0,20 \cdot [\theta_i - \min(\theta_i)] / [\max(\theta_i) - \min(\theta_i)] + 0,80 \quad (27)$$

Referências bibliográficas

ANEEL Resolução Normativa nº 55/2004, 5 de abril de 2004

ANEEL Nota Técnica nº 166/2006 – SRE/ANEEL , 19 de maio de 2006a

ANEEL Nota Técnica nº 262/2006 – SER/SFF/SRD/SFE/SRC/ANEEL , 19 de outubro de 2006b

CASELLA, G. ; GEORGE, E. I. Explaining the Gibbs Sampler. The American Statistician, v.46, n.3, 1992.

COOPER, W. W.; SEIFORD, L. M.; TONE, K. Data envelopment analysis: a comprehensive text with models, applications, references and dea-solver software. Boston, MA: Kluwer Academic, 2000.

EHLERS, R. S. Métodos computacionalmente intensivos em estatística, versão n. 2, 2004. Disponível em <<http://www.est.ufpr.br/~ehlers/notas/mci.pdf>>. Acessado em Janeiro de 2005.
KOOP, G. Recent progress in applied Bayesian econometrics. Journal of Economic Surveys, v.8, n.1, 1994.

KOOP, G. ; STEEL, M. F. J ; OSIEWALSKI, J. Posterior analysis of stochastic frontier models using Gibbs sampler. Computational Statistics, 10, 1995

KOOP, G. ; OSIEWALSKI, J. ; STEEL, M. F. J. Bayesian efficiency análise through individual effects: Hospital cost frontiers. Journal of Econometrics, 76, 1997

KUMBHAKAR, S.C., LOVELL, C.A.K. Stochastic Frontier Analysis, Cambridge University Press, 2000.

GAMERMAN, D., Markov chain Monte Carlo - Stochastic simulation for Bayesian inference. Chapman and Hall, 1997.

MEDRANO, L. A. T. ; MIGON, H. S. Critérios baseado na “Deviance” para a comparação de modelos Bayesiano de fronteira de produção estocástica. Publicação interna, 176. Universidade Federal do Rio de Janeiro.

JONDROW, J. et al. On the estimation of technical inefficiency in the stochastic frontier production function model. Journal of Econometrics, 19, 1982.

MIGON, H. S.; GAMERMAN, D. Statistical Inference – An Integrated Approach. Arnold, 1999.

ROBERT C. P.; CASELLA, G. Monte Carlo Statistical Methods. New York: Springer-Verlag, 1999.

TIERNEY, L. Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics*, 22, 1994.

VAN DEN BROECK, J.; KOOP, G.; OSIEWALSKI, J.; STEEL, M. F. J. Stochastic frontier models – A Bayesian perspective. *Journal of Econometrics*, 61, 1994.