



Daniel Alfradique Leite

**Definição das metas regulatórias de perdas não técnicas
para as distribuidoras de energia elétrica no Brasil:
proposição de um modelo baseado na Análise de Fronteira
Estocástica (SFA)**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Metrologia (Área de concentração: Metrologia para Qualidade e Inovação) da PUC-Rio.

Orientador: Prof. Rodrigo Flora Calili
Co-orientador: Prof. José Francisco Moreira Pessanha

Rio de Janeiro
Outubro de 2020



Daniel Alfradique Leite

**Definição das metas regulatórias de perdas não técnicas
para as distribuidoras de energia elétrica no Brasil:
proposição de um modelo baseado na Análise de Fronteira
Estocástica (SFA)**

Dissertação apresentada como requisito parcial para
obtenção do grau de Mestre pelo Programa de Pós-
Graduação em Metrologia da PUC-Rio. Aprovada pela
Comissão examinadora abaixo.

Rodrigo Flora Calili
Orientador
PUC-Rio

José Francisco Moreira Pessanha
Co-orientador
CEPEL

Guilherme de Azevedo Dantas
Essenz Soluções

Reinaldo Castro Souza
PUC-Rio

Rio de Janeiro, 02 de Outubro de 2020

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Daniel Alfradique Leite

Graduou-se em Ciências Econômicas pela Universidade Federal do Rio de Janeiro em 2000. Trabalha no setor elétrico desde de 2002, tendo passado por diversas empresas do setor. Atual trabalha na ENEL Brasil como Coordenador de Comercialização e é mestrando do Programa de Pós-Graduação em Metrologia pela PUC Rio.

Ficha Catalográfica

Leite, Daniel Alfradique

Definição das metas regulatórias de perdas não técnicas para as distribuidoras de energia elétrica no Brasil: proposição de um modelo baseado na Análise de Fronteira Estocástica (SFA) / Daniel Alfradique Leite ; orientador: Rodrigo Flora Calili ; co-orientador: José Francisco Moreira Pessanha. – 2020.

147 f. : il. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Centro Técnico Científico, Programa de Pós-Graduação em Metrologia, 2020.

Inclui bibliografia

1. Metrologia - Teses. 2. Metrologia. 3. Perdas não técnicas. 4. Distribuidoras de energia elétrica. 5. Fronteira Estocástica. 6. Dados em painel. I. Calili, Rodrigo Flora. II. Pessanha, José Francisco Moreira. III. Pontifícia Universidade Católica do Rio de Janeiro. Centro Técnico Científico. Programa de Pós-Graduação em Metrologia. IV. Título.

CDD: 389.1

A minha família

Agradecimentos

Ao meu orientador Prof. Rodrigo Flora Calili e ao meu co-orientador Prof. José Francisco Moreira Pessanha pelo estímulo e parceria para a realização deste trabalho.

A PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

A todos os professores e funcionários do Departamento de Metrologia pelos ensinamentos e pela ajuda.

Aos amigos da equipe de Regulação da Enel, pela amizade e apoio ao longo do desenvolvimento desse trabalho.

Aos professores que participaram da Comissão examinadora.

A Sabrina, Lucca e Davi pelo amor e compreensão a longo dessa jornada.

O presente trabalho foi desenvolvido dentro do Projeto de P&D PUC-Rio e Enel Brasil que busca estabelecer modelos alternativos para definição de metas de inadimplência e perdas não técnicas regulatórias. Assim agradeço a equipe de P&D da Enel e a todos os pesquisadores envolvidos no projeto pelo apoio.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

Leite, Daniel Alfradique; Calili, Rodrigo Flora (orientador); Pessanha, José Francisco Moreira (co-orientador). **Definição das metas regulatórias de perdas não técnicas para as distribuidoras de energia elétrica no Brasil: proposição de um modelo baseado na Análise de Fronteira Estocástica (SFA)**. Rio de Janeiro, 2020. 115p. Dissertação de Mestrado - Programa de Pós-graduação em Metrologia, Pontifícia Universidade Católica do Rio de Janeiro.

A redução das perdas não técnicas de energia elétrica (PNT) constitui atualmente um dos principais desafios enfrentados pelas concessionárias de distribuição de eletricidade nos países em desenvolvimento. Por depender em parte de fatores não gerenciáveis pelas distribuidoras, as PNT no Brasil são repassadas aos consumidores através de tarifas de eletricidade, seguindo os limites definidos pela Agência Nacional de Energia Elétrica (ANEEL). Apesar do notável progresso realizado pela ANEEL, à metodologia atualmente empregada pela Agência para definição desses limites apresenta uma série de limitações o que impõem ao regulador a adoção de medidas *ad hoc* que acaba resultando em metas de perdas pouco condizentes com a realidade das distribuidoras. Nas últimas três décadas diversos modelos matemáticos de Análise de Eficiência vêm sendo desenvolvidos com intuito de mensurar a eficiência relativa de grupos de agente. A adoção desses modelos, no entanto, tem se limitado a temas relacionados a custos operacionais e investimento. Essa dissertação tem por objetivo cobrir essa lacuna, propondo um modelo alternativo para a definição das metas regulatórias de PNT a partir de uma família particular de modelos de Análise de Eficiência – os modelos de Análise de Fronteira Estocástica (SFA). O modelo proposto foi aplicado a um painel de dados contendo observações anuais de 62 concessionárias de distribuição de energia elétrica brasileira no período de 2007 a 2017, resultando em um conjunto de metas de PNT mais aderentes às PNT reais das distribuidoras brasileiras.

Palavras-chaves

Metrologia, Perdas não técnicas; Distribuidoras de energia elétrica; Análise de Fronteira Estocástica; Dados em Painel

Abstract

Leite, Daniel Alfradique; Calili, Rodrigo Flora (orientador); Pessanha, José Francisco Moreira (co-orientador). **Defining Non-technical losses regulatory targets for electricity distributors in Brazil: proposing a model based on Stochastic Frontier Analysis (SFA)**. Rio de Janeiro, 2020. 115p. Dissertação de Mestrado - Programa de Pós-graduação em Metrologia, Pontifícia Universidade Católica do Rio de Janeiro.

The reduction of non-technical losses of electrical power (NTL) are currently one of the main challenges faced by electricity utility companies in developing countries. Because it depends in part on factors not manageable by the utilities, in Brazil is passed on to consumers through electricity tariffs, following the limits defined by the National Electric Energy Agency (ANEEL). Despite the notable progress made by ANEEL, the currently used methodology by this Agency to define these limits has a number of limitations, which have impose the adoption of *ad hoc* measures that end up resulting in loss targets that are inconsistent with the reality of the utilities. In the last three decades, a several mathematical models of Efficiency Analysis have been developing in order to measure the relative efficiency of groups of agents. However, the adoption of these models by Regulator Bodies around the world has been limited to subjects related to operating costs and investment. This dissertation aims to fulfill this gap, proposing an alternative model for the definition of NTL regulatory targets from a particular family of Efficiency Analysis models - the Stochastic Frontier Analysis (SFA) models. The proposed model has been applied to a data panel containing annual observations of 62 distribution concessionaires in the Brazilian electrical system from 2007 to 2017, resulting in a set of NTL targets more adherent to the real PNT of Brazilian electricity utility distributors.

Keywords

Metrology, Non-technical losses; Electricity utilities; Stochastic frontier analysis; Panel data

Sumário

| | | |
|-------|---|----|
| 1 | Introdução | 14 |
| 1.1 | Definição do problema de pesquisa | 17 |
| 1.2 | Objetivos: geral e específicos | 18 |
| 1.3 | Metodologia | 18 |
| 1.3.1 | Fase exploratória e descritiva | 20 |
| 1.3.2 | Fase de pesquisa aplicada | 22 |
| 1.3.3 | Fase conclusiva-propositiva | 22 |
| 1.3.4 | Estrutura da dissertação | 22 |
| 2 | Perdas não técnicas de energia elétrica e a regulação por incentivo: uma análise crítica do modelo brasileiro | 24 |
| 2.1 | O desafio regulatório e a teoria da competição por padrões (<i>Yardstick Competition</i>) | 24 |
| 2.2 | O problema de Perdas Não Técnicas no Brasil | 27 |
| 2.3 | O Tratamento regulatório das Perdas Não Técnicas no Brasil | 30 |
| 2.3.1 | Histórico da metodologia empregada pela ANEEL | 31 |
| 2.3.2 | A metodologia atual | 33 |
| 2.3.3 | A Consulta Pública nº 029 de 2020 e as possíveis mudanças na metodologia atual de definição das metas de perdas não técnicas | 41 |
| 2.4 | Críticas ao tratamento regulatório das Perdas Não Técnicas no Brasil | 44 |
| 2.4.1 | O problema das variáveis explicativas | 45 |
| 2.4.2 | Os problemas de especificação do modelo | 52 |
| 2.4.3 | O problema da exequibilidade das metas regulatórias de Perdas Não Técnicas definidas pela ANEEL | 57 |
| 2.5 | Considerações sobre o capítulo | 59 |
| 3 | Modelos de Análise de Eficiência e suas aplicações a definição de metas regulatórias para concessionárias de serviço público | 61 |
| 3.1 | Modelos de Análise de Eficiência | 66 |
| 3.2 | Análise de Envoltória de Dados (DEA) | 67 |
| 3.2.1 | O modelo CRS (<i>Constant Returns to Scale</i>) | 67 |
| 3.2.2 | O modelo VRS (<i>Variable Return Scale</i>) | 69 |

| | | |
|-------|--|-----|
| 3.2.3 | Vantagens e desvantagens do uso do DEA e seu uso pelos reguladores de serviço público no mundo | 71 |
| 3.3 | Mínimos Quadrados Corrigidos (COLS) | 72 |
| 3.3.1 | Vantagens e desvantagens do uso do COLS e seu uso pelos reguladores de serviço público no mundo | 73 |
| 3.4 | Modelo de Fronteira Estocástica (SFA) | 74 |
| 3.4.1 | Vantagens e desvantagens do uso do SFA e seu uso pelos reguladores de serviço público no mundo | 79 |
| 3.5 | Considerações sobre o capítulo | 80 |
| 4 | Desenvolvimento de um modelo alternativo para a definição das metas regulatórias de Perdas Não Técnicas (PNT) | 82 |
| 4.1 | Modelo de Fronteira Estocástica de custo para modelo de dados cross section | 83 |
| 4.2 | Modelo de fronteira estocástica de custo para modelo de dados em painel | 88 |
| 4.3 | Proposição de um modelo alternativo para a definição das metas regulatórias de Perdas Não Técnicas (PNT) | 89 |
| 4.3.1 | Definição das metas regulatórias de PNT | 90 |
| 4.3.2 | Definição de uma trajetória de redução das PNT | 91 |
| 4.3.3 | Definição de limites para as metas regulatórias de PNT | 92 |
| 4.4 | Considerações sobre o capítulo | 93 |
| 5 | Aplicação de um modelo alternativo para definição das metas regulatórias de Perdas Não Técnicas (PNT) às distribuidoras de energia elétricas brasileiras | 94 |
| 5.1 | Base de dados e variáveis selecionadas | 94 |
| 5.2 | Modelo e processo de estimação | 96 |
| 5.3 | Critérios para seleção dos modelos | 97 |
| 5.4 | Resultados do Modelo de Eficiência | 98 |
| 5.5 | Variáveis críticas e resultados do modelo de eficiência para as distribuidoras de grande porte | 101 |
| 5.6 | Definição das metas regulatórias a partir modelo proposto e a comparação com as metas regulatórias definidas pela ANEEL | 104 |
| 5.7 | Discussão dos resultados | 110 |
| 6 | Conclusões | 112 |

| | | |
|--|----------------------------|-----|
| 7 | Referências bibliográficas | 115 |
| | | |
| Anexo I – Variáveis exógenas utilizadas nos modelos de eficiência | | 123 |
| Anexo II – Resultados dos 29 modelos de eficiência selecionados | | 128 |
| Anexo III – Distribuição da eficiência média das distribuidoras nos 164 modelos estimados de 2007-17 | | 143 |
| Anexo IV – Ganho de eficiência das distribuidoras no combate as PNT de 2007 a 2017 | | 146 |

Lista de Figuras

| | |
|--|-----|
| Figura 1.1 - Desenho da pesquisa, seus componentes e métodos | 19 |
| Figura 1.2 - Mapa conceitual da pesquisa | 21 |
| Figura 2.1 - Perdas Não Técnicas no Brasil (% da Energia Injetada Total) | 28 |
| Figura 2.2 - Perdas Não Técnicas no Brasil (% da Energia Injetada Total) | 29 |
| Figura 2.3 - Percentual de Domicílios subnormais por distribuidora de energia elétrica em 2010 | 49 |
| Figura 2.4 - Hierarquia de seleção entre os modelos | 53 |
| Figura 2.5 - Perdas Não Técnicas no Brasil vs. Metas Regulatórias (% da Energia Injetada Total) | 57 |
| Figura 2.6 - Percentual das distribuidoras que não conseguiram atingir as metas de PNT regulatórias de 2008 a 2018 | 58 |
| Figura 2.7 - Perdas Não Técnicas no Brasil – Realizado vs. Meta regulatória em 2018 | 59 |
| Figura 3.1 - Fronteira de Produção | 61 |
| Figura 3.2 - Fronteira de Custo | 63 |
| Figura 4.1 - Fronteira de Custo | 82 |
| Figura 4.2 - Fronteira de Custos de PNT | 89 |
| Figura 5.1 - Frequência das variáveis socioeconômicas nos modelos selecionados | 98 |
| Figura 5.2 - Mediana da eficiência das distribuidoras de energia elétrica no combate as perdas não técnicas (PNT) | 99 |
| Figura.5.3 - Distribuição das eficiências médias calculadas para 2017 nos 29 modelos | 100 |
| Figura 5.4 - Eficiências médias calculadas para 2017 a modelos que possuem as variáveis descrita no eixo das abscissas | 102 |
| Figura 5.5 - Eficiências médias calculadas para LIGHT em 2017 a partir dos modelos que possuem as variáveis descrita no eixo das abscissas | 103 |
| Figura 5.6 - Evolução da Performance da LIGHT de 2007 a 2017 | 103 |
| Figura 5.7 - Comparação entre as metas regulatórias de PNT para 2018 da ANEEL e as propostas pelo modelo alternativo – dados em % do mercado BT faturado | 109 |

Lista de Quadros

| | |
|--|----|
| Quadro 2.1 - Variáveis socioeconômicas utilizadas pela ANEEL na construção do IC | 34 |
| Quadro 2.2 - Variáveis socioambientais utilizadas pela ANEEL nos modelos C, G e K | 35 |
| Quadro 2.3 - Definição do Ponto de Partida das PNT regulatórias sobre o mercado de baixa tensão (BT) | 39 |
| Quadro 2.4 - Limites anuais de redução das PNT regulatórias sobre o mercado de baixa tensão (BT) | 40 |
| Quadro 2.5 - Variáveis Socioeconômicas testadas pela ANEEL na CP 029/2020 | 41 |
| Quadro 2.7 - Critério de seleção de modelos adotado pela ANEEL na CP 029/2020 | 42 |
| Quadro 2.8 - Critério para a definição do Ponto de Partida das PNT regulatórias sobre o mercado de baixa tensão (BT) proposto na CP 029/2020 | 43 |
| Quadro.2.9 - Critério para a definição dos limites anuais de redução das PNT regulatórias sobre o mercado de baixa tensão (BT) proposto na CP 029/2020 | 44 |
| Quadro 2.9 - Variáveis Socioeconômicas utilizadas pela ANEEL no cálculo do IC | 47 |
| Quadro 5.1 - Variáveis socioeconômicas aplicadas ao caso brasileiro | 95 |
| Quadro 5.2 - Critério de seleção de modelos adotados na simulação | 97 |

Lista de Tabelas

| | |
|--|-----|
| Tabela 2.1 - Inadimplência total no setor de crédito do Sistema Financeiro Nacional (SFN) – Abril de 2019 | 52 |
| Tabela 2.2 - Resultado dos Testes de Breusch-Pagan e Hausman aplicados aos modelos C, G e K da ANEEL | 55 |
| Tabela 5.1 - Eficiência no combate a PNT das distribuidoras 10 distribuidoras com maior ganho de eficiência entre 2007 e 2017 | 101 |
| Tabela 5.2 - As 10 Distribuidora que mais ganham ou perdem eficiência no combate a PNT com a inclusão da variável % de Domicílios subnormais | 102 |
| Tabela 5.3 - Comparação entre as metas regulatórias de PNT para 2018 da ANEEL e as propostas pelo modelo alternativo – dados em % do mercado BT faturado | 106 |

1

Introdução

No fornecimento de energia elétrica, as perdas referem-se a diferença entre o volume de eletricidade injetadas nas redes de transmissão e distribuição e o volume efetivamente entregue e faturado dos consumidores finais (ANEEL, 2015b). Essas perdas podem ser divididas em dois componentes principais: as perdas técnicas e as perdas não técnicas (Smith, 2004).

As Perdas Técnicas (PT) são aquelas inerentes ao processo de transmissão e distribuição de energia e ocorrem em função da dissipação de energia através dos componentes do sistema elétrico, como linhas de transmissão e distribuição, transformadores e sistemas de medição.

As Perdas Não Técnicas (PNT), por outro lado, são aquelas causadas por ações externas ao processo de geração e distribuição de energia elétrica, e ocorrem em função principalmente do furto de eletricidade (seja através da adulteração de medidores ou através de conexões ilegais à rede de distribuição) e dos erros na medição e contabilização do consumo de energia.

Diferentemente das perdas técnicas cuja origem se deve quase que exclusivamente a fenômenos físicos, as perdas não técnicas estão mais associadas a um conjunto amplo de fenômenos econômicos e sociais presentes em cada região. Características como a força das instituições públicas locais, o nível de renda da população, a qualidade da infraestrutura disponível, o nível de corrupção, a aceitação social do furto de energia, o nível de escolaridade, o nível de desemprego, o nível de violência, entre outros fenômenos, que vem sendo apontados por diversos autores com fatores fundamentais para explicar a maior ou menor incidência das perdas não técnicas em diferentes regiões (Smith, 2004; Winther, 2012; Yurtseven, 2015; Yakubu *et al.* 2018)¹.

Atualmente as perdas não técnicas de energia constituem um dos maiores desafios enfrentado por muitas concessionárias de distribuição de energia elétrica

¹ Alguns autores, como Aneel (2020a), tem apontados para importância também variáveis associadas a gestão da própria distribuidora como a eficiência operacional das empresas.

em países em desenvolvimento, em particular em alguns países da África, América Latina e sul da Ásia (Penin, 2008; Smith, 2004; Amin *et al.* 2015; WORLD BANK, 2009; Millard & Emmerton e Emmerton, 2009; Fourie 2005; Arango *et al.*, 2017; Chandel *et al.*, 2016; Yakubuab, 2018), onde os níveis atuais dessas perdas representam um riscos não apenas para a solvência das empresas de distribuição de eletricidade, mas também para a própria segurança do fornecimento de energia (Depuru *et al.*, 2011; Smith, 2004).

A Índia talvez seja desse grupo o caso mais emblemático, com perdas não técnicas da ordem de 30%, os custos associados ao furto de energia no país correspondem a aproximadamente 1,5 % do PIB (Depuru *et al.*, 2011; Singh, 2006).

Embora a situação no Brasil seja menos dramática do que a observada em outros países em desenvolvimento - e em média até próxima dos valores observados nos países da OCDE (Yurtseven, 2015)² - as perdas não técnicas têm um impacto financeiro significativo tanto para empresas de distribuição de energia elétrica quanto para os consumidores brasileiros. Em 2018 as perdas não técnicas representaram um custo adicional para as distribuidoras e demais consumidores brasileiros de R\$ 6,4 bilhões somente com a aquisição de energia elétrica e outros R\$ 3,7 bilhões aproximadamente em custos associados à infraestrutura de distribuição³.

Por depender de fatores em geral não gerenciáveis pelas distribuidoras, as perdas não técnicas de energia elétrica no Brasil são parcialmente repassadas aos consumidores através das tarifas de energia elétricas (ANEEL, 2015b). Dessa forma, um dos desafios da Agência Nacional de Energia Elétrica (ANEEL) é determinar volume de perdas não técnicas a serem reconhecidas nas tarifas.

Se por um lado a regulação deve garantir que as distribuidoras não sejam oneradas pelos custos não gerenciáveis associados ao serviço prestado, por outro, a regulação deve também assegurar que os consumidores não sejam onerados pelos custos decorrente da ineficiência das distribuidoras.

² Segundo Yurtseven (2015) o percentual médio de PNT nos EUA e nos países da Europa Ocidental se situam entre 1-2% da energia total injetada no sistema de distribuição, enquanto a média dos países da OECD estaria em 7%.

³ Considerando o valor preço médio de compra de energia das distribuidoras em 2019 (PMIX) de R\$ 191,47 por MWh, uma perda não técnica total em 2019 de 33,3 TWh e uma TUSD Fio B média de aproximadamente R\$ 111,09 por MWh. Estimativa baseada nos dados ANEEL (2019a, 2019b).

Uma alternativa proposta na literatura para resolver esse dilema, e amplamente utilizada por agências regulatórias no mundo, se baseia na teoria do *Yardstick Competition* (ou competição por padrões) proposto originalmente por Shleifer (1985). A teoria proposta por este autor preconiza que o regulador exponha as empresas reguladas aos mesmos desafios que enfrentariam em um ambiente onde houvesse concorrência direta entre elas, promovendo a comparação, premiando as empresas mais eficientes e punindo as menos eficientes.

A ANEEL (2010, 2011, 2013, 2014a, 2014b, 2020a) tem tratado o tema das perdas não técnicas seguindo os princípios preconizados na teoria do *Yardstick Competition* de forma sistemática ao longo da última década. Para tentar distinguir as ineficiências das distribuidoras nas ações de combate a perdas não técnicas das limitações impostas pelas condições socioeconômicas de cada área de concessão, a agência desenvolveu um conjunto de modelos estatísticos baseado em dados longitudinais que permite, entre outras coisas, comparar as perdas não técnicas das diversas distribuidoras do país e estabelecer qual volume deve ser reconhecido para cada distribuidora e repassado aos consumidores nas tarifas de energia elétrica. No entanto, como apontam diversos trabalhos recentes como Calili *et al.* (2017), a metodologia empregada pela ANEEL apresenta diversas limitações que acabam por resultar em distorções relevantes no processo de reconhecimento tarifário das perdas não técnicas.

A principal crítica apontada pelos autores do referido estudo (Calili *et al.*, 2017) é que a metodologia atualmente empregada pela ANEEL é excessivamente complexa, frágil do ponto de vista do metodológico e, em certa medida incompleta, o que impõem ao órgão regulador a adoção de uma série de medidas *ad hoc* que acaba resultando em metas de perdas pouco condizentes com a realidade das distribuidoras.

Em abril de 2020 a ANEEL abriu uma consulta pública (ANEEL, 2020a) com o intuito de reformular e atualizar os modelos atualmente empregados na definição das metas regulatórias de PNT. Embora reconheça as limitações do modelo atual, a proposta da ANEEL altera apenas de forma marginal a metodologia, mantendo-se validas as críticas aprontadas por Calili *et al.* (2017).

Como apresenta Cooper (2011), nas últimas três décadas diversos modelos matemáticos de Análise de Eficiência - ou como denominado Bogetoft & Otto (2010) modelos de Análise de Fronteira - vêm sendo desenvolvidos com intuito de

mensurar a eficiência relativa de grupos de agente, sendo eles empresas, governos ou mesmo países. Esses modelos tentam lidar de forma objetiva com um dos maiores desafios impostos a qualquer análise de *benchmarking*: como comparar agentes distintos diretamente, resguardando nessa comparação as diferentes características de cada um dos agentes e as condições mercadológicas nas quais eles estão inseridos.

O relativo sucesso desses modelos tem levado, desde o final da década de 90, a sua adoção por diversos reguladores no mundo. Esse processo, no entanto, tem se mostrado relativamente restrito aos temas de custo operacional e remuneração de investimento, com poucos trabalhos explorando a aplicação desses mesmos modelos a outros temas regulatórios relevantes.

A presente pesquisa tem por objetivo cobrir essa lacuna, propondo um modelo alternativo para a definição das metas regulatórias de perdas não técnicas para as distribuidoras de energia elétrica no Brasil, avaliando a aplicabilidade, nesse processo, de uma família particular de modelos de Análise de Eficiência – os modelos de Análise de Fronteira Estocástica (SFA).

Esta dissertação insere-se na linha de pesquisa “Redes Inteligentes (*Smart Grid*)” do Programa de Pós-graduação em Metrologia (PósMQI) da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio).

1.1

Definição do problema de pesquisa

Considerando a relevância do tema de perdas não técnicas, o desafio de tratar o tema do ponto de vista regulatório e os avanços recentes nos modelos de Análise de Eficiência, definiu-se a seguinte questão principal a ser respondida ao longo da pesquisa:

“Como estabelecer metas regulatórias de redução de perdas não técnicas para as distribuidoras de energia elétrica brasileiras equilibrando o princípio da regulação por incentivo e a complexidade da área de concessão de cada distribuidora?”

1.2

Objetivos: geral e específicos

Essa pesquisa busca dar uma contribuição para o processo de definição metas regulatórias de redução de perdas não técnicas para as distribuidoras de energia elétrica brasileiras.

Desse modo, o objetivo geral da dissertação é propor um modelo alternativo ao modelo regulatório atual, avaliando a aplicabilidade de uma família particular de modelos de Análise de Eficiência – os modelos de Análise de Fronteira Estocástica (SFA).

Em termos específicos, a dissertação pretende:

- Identificar as principais limitações no modelo atual para metas regulatórias de redução de perdas não técnicas para as distribuidoras de energia elétrica brasileiras;
- Realizar uma revisão da Literatura sobre modelos de Análise de Eficiência, em particular os modelos de Análise de Fronteira Estocástica (SFA), avaliando a aplicabilidade desses modelos à definição metas regulatórias de redução de perdas não técnicas para empresas de serviços públicos;
- Desenvolver um modelo alternativo ao modelo da ANEEL para definição metas regulatórias de redução de perdas não técnicas para as distribuidoras de energia elétrica no Brasil;
- Demonstrar a aplicabilidade do modelo alternativo no contexto do setor de distribuição de energia elétrica brasileiro, explicitando seus diferenciais em relação ao modelo atual adotado pela ANEEL.

1.3

Metodologia

Segundo a taxonomia proposta por Vergara (2002), a pesquisa pode ser considerada aplicada, descritiva e metodológica (quanto aos fins).

Quanto aos meios de investigação, a metodologia compreende:

- Pesquisa bibliográfica, documental, de estudo de caso e experimental sobre os temas centrais da pesquisa, como indicado na fase exploratória e descritiva da Figura 1.1;

- Desenvolvimento de um modelo estatístico para a definição das metas regulatórias de redução de perdas não técnicas das distribuidoras de energia elétrica;
- Aplicação de modelos de Análise de Fronteira Estocástica (SFA) para permitir a comparação e a mensuração da eficiência das distribuidoras em relação ao tema;
- Estudo empírico visando demonstrar a aplicabilidade do modelo no contexto brasileiro.

A Figura 1.1 apresenta o desenho da pesquisa, destacando-se seus componentes e métodos, de acordo com três fases principais: (i) exploratória e descritiva; (ii) pesquisa aplicada; e (iii) conclusiva-propositiva.

Detalham-se, a seguir, o desenvolvimento de cada fase e os resultados esperados em cada bloco da Figura 1.1.

Destaca-se que o desenho da pesquisa foi desenvolvido em total alinhamento com os objetivos enunciados na seção 1.2 deste capítulo.

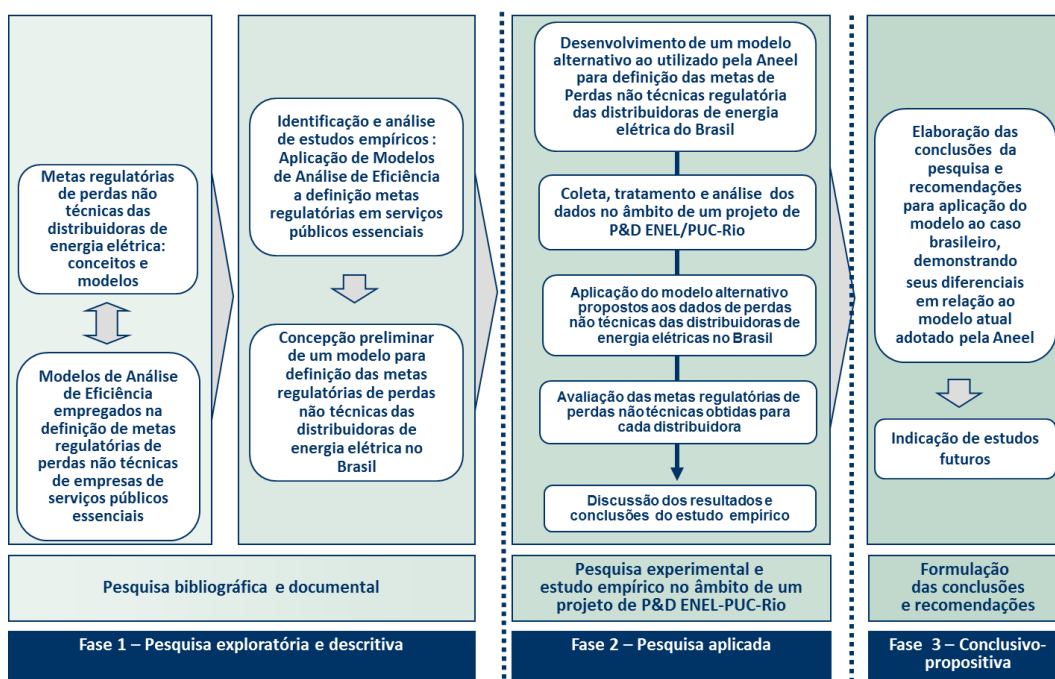


Figura 1.1 - Desenho da pesquisa, seus componentes e métodos

Fonte: Elaboração própria.

1.3.1

Fase exploratória e descritiva

A fase exploratória e descritiva foi iniciada com pesquisa bibliográfica e documental, com o objetivo de levantar trabalhos conceituais e documentos de referência para delimitação do tema central da pesquisa – definição metas regulatórias de redução de perdas não técnicas para as distribuidoras de energia elétrica brasileiras.

Em seguida, aprofundou-se a revisão bibliográfica, buscando identificar estudos empíricos sobre esse tema, focalizando-se nos métodos de Análise de Eficiência que estavam sendo adotados com esse propósito e os objetos de análise.

A partir da constatação de que os métodos de Análise de Eficiência (DEA, SFA dentre outros) eram os mais adequados ao problema, pesquisou-se se os estudos voltados à definição de metas regulatórias de perdas não técnicas estavam adotando modelos de Análise de Eficiência, particularmente Análise de Fronteira Estocástica (SFA). Identificou-se, assim, uma lacuna na literatura que foi explorada na etapa de pesquisa aplicada.

Apresenta-se na Figura 1.2 uma visão geral e esquemática dos resultados desta primeira fase.

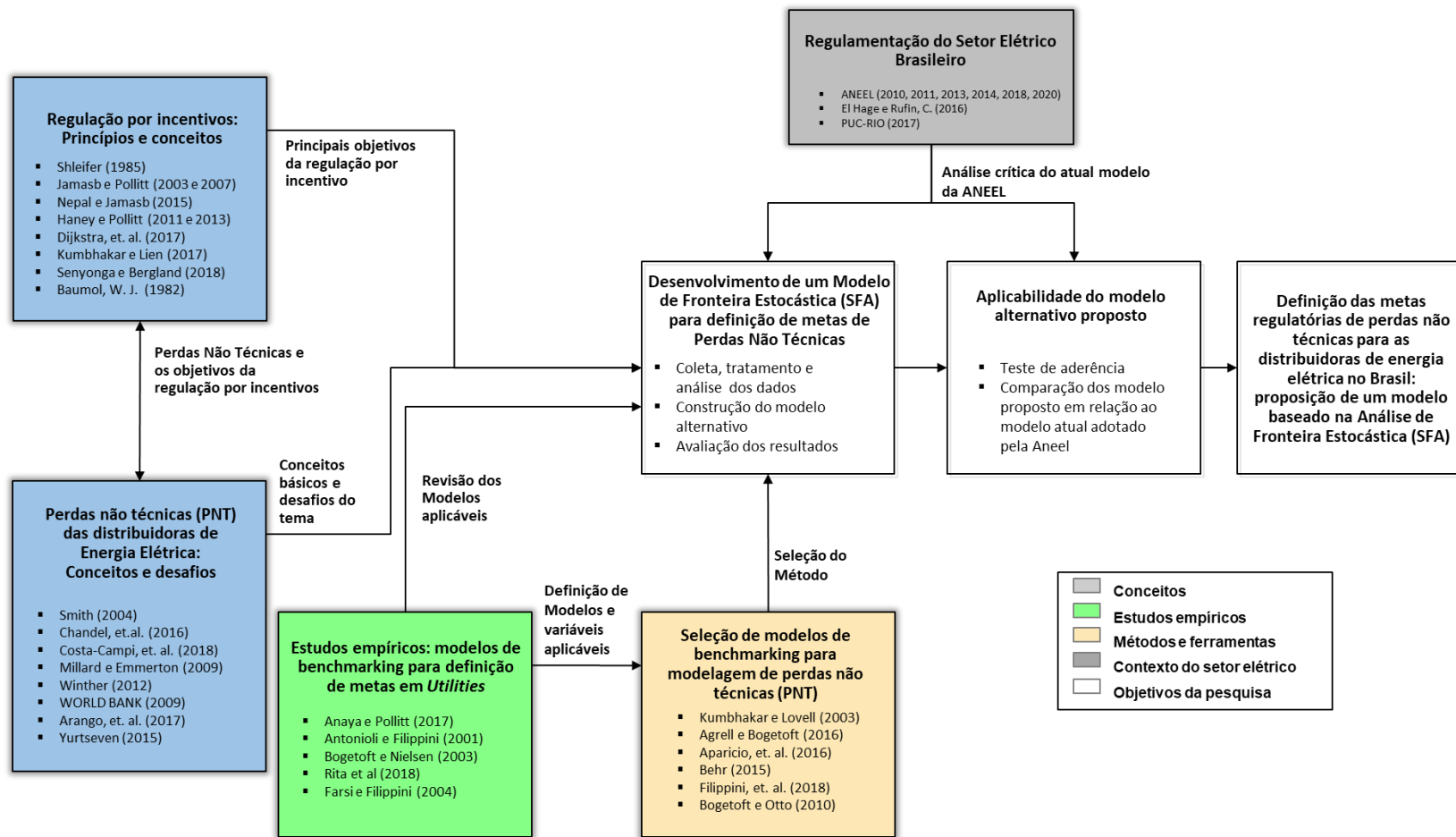


Figura 1.2 - Mapa conceitual da pesquisa

Fonte: Elaboração própria.

O referencial teórico constituiu uma orientação conceitual para a pesquisa, ao compor o vocabulário especializado e organizar o conhecimento sobre a definição de metas regulatórias de perdas não técnicas.

1.3.2

Fase de pesquisa aplicada

Desenvolveu-se a fase da pesquisa aplicada propriamente dita, a partir dos resultados da revisão bibliográfica e documental, que nortearam a definição dos objetivos da pesquisa e o desenvolvimento de um modelo alternativo ao modelo regulatório brasileiro atual, para a definição das metas regulatórias de perdas não técnicas para as distribuidoras de energia elétrica.

1.3.3

Fase conclusiva-propositiva

Na terceira fase, elaboraram-se a conclusão geral e as específicas em relação a cada um dos objetivos enunciados na seção 1.2, formulando um conjunto de recomendações aos diversos atores interessados na aplicação do modelo alternativo proposto para determinar das metas regulatórias de perdas não técnicas para as distribuidoras de energia elétrica brasileiras.

1.3.4

Estrutura da dissertação

A dissertação encontra-se estruturada em seis capítulos, incluindo esta introdução.

O capítulo 2 apresenta uma breve revisão do conceito que norteia a regulação por incentivo assim como os principais conceitos e desafios em relação ao tema de perdas não técnicas, em particular no Brasil. É apresentado também o modelo utilizado pela Agência Nacional de Energia Elétrica (ANEEL) na definição das metas regulatórias de perdas não técnicas, assim como suas principais limitações.

No capítulo 3, apresenta-se o referencial teórico que fundamentou a proposição de um modelo alternativo para a definição das metas regulatórias de perdas não técnicas para as distribuidoras de energia elétrica. Expõem-se a definição dos conceitos e critérios das metodologias de Análise de Fronteira, assim

como os principais estudos empíricos realizados sobre a aplicação desses modelos à definição de metas regulatórias em serviços públicos.

O capítulo 4 apresenta o modelo alternativo proposto para a definição das metas regulatórias de perdas não técnicas para as distribuidoras de energia elétrica, construído a partir do referencial teórico apresentado no capítulo 3.

No capítulo 5, apresentam-se os principais resultados da aplicação do modelo proposto no capítulo 4 ao caso brasileiro, destacando-se a aderência desses resultados à realidade local, assim como uma comparação com os resultados obtidos a partir da aplicação do modelo atual da ANEEL.

No capítulo 6, apresentam-se as conclusões do estudo, demonstrando a aplicabilidade do modelo proposto ao Brasil, assim como suas vantagens em relação ao modelo atual aplicado pela ANEEL. Além disso, nesse capítulo, são apresentadas recomendações para estudos futuros.

2

Perdas não técnicas de energia elétrica e a regulação por incentivo: uma análise crítica do modelo brasileiro

Nesse capítulo inicialmente apresenta-se os princípios teóricos que norteiam a regulação por incentivo e os desafios do agente regulador. Na segunda parte é feita uma contextualização do cenário de perdas não técnicas (PNT) no Brasil, sua evolução recente e os principais desafios para o seu enfrentamento. Por fim, é apresentado o modelo com o qual o ANEEL vem enfrentando o tema de PNT no Brasil e suas principais limitações.

2.1

O desafio regulatório e a teoria da competição por padrões (*Yardstick Competition*)

Em qualquer atividade econômica uma parte das perdas no processo produtivo é repassada ao preço final dos produtos ou serviço. Em um ambiente concorrencial, no entanto, a capacidade de um determinado agente de repassar os seus custos individuais com as perdas é limitada pelo nível médio de perdas praticado pelos demais agentes.

Essa limitação decorre do simples fato de que em um ambiente concorrencial os consumidores podem, a princípio, optar entre diversos fornecedores, inclusive novos entrantes (Baumol, 1982; Baumol & Sidak, 1995; ANEEL, 2010).

A “ameaça” imposta pela concorrência (*perfect contestable markets*) cria um incentivo permanente para que os agentes busquem um aumento da eficiência produtiva e, conseqüentemente, a redução do nível de perdas, assegurando a maximização da eficiência econômica (eficiência produtiva e alocativa) e conseqüentemente do bem-estar social (Baumol, 1982; Baumol & Sidak, 1995).

A Teoria dos Mercados Contestáveis desenvolvida por Baumol (1982), no entanto, pressupõem uma série de condições para que de fato esse mecanismo de mercado funcione, entre elas a inexistência de custos irrecuperáveis (*sunk cost*), a inexistência de barreiras à entrada/saída de novos agentes, e a inexistência de economias de escala e escopo (Baumol, 1982; Baumol & Sidak, 1995).

Em uma indústria com características de “rede”, como é o caso do setor elétrico, a introdução da competição para toda a extensão da cadeia produtiva não é viável (Baumol, 1982).

A presença de elevadas economias de escala (custos fixos altos e custos marginais baixos) e de custos irrecuperáveis (*sunk cost*), principalmente nas etapas de transmissão e distribuição de energia elétrica, impedem a entrada de novos concorrentes, inviabilizando o funcionamento adequando do mecanismo de “auto regulação” do mercado exercido pela concorrência.

Em cenários como esse, segundo Baumol (1982), é necessário a criação de mecanismos adicionais que estimulem as empresas a se tornarem mais eficientes, de forma a se obter um resultado próximo ao que seria obtido em um mercado competitivo (Haney e Pollitt, 2009, 2011 e 2013). Assim a regulação do setor, através da intervenção do Estado, seria então fundamental (Baumol et al, 1982; Shleifer, 1985; Baumol & Sidak, 1995, J; Haney & Pollitt, 200, 2011 e 2013).

A regulação por incentivo busca introduz a componente produtividade associado à redução nos custos como um dos principais objetivos das entidades reguladoras. Há um grande número de mecanismos de regulação que buscam atender a esse objetivo. Alguns mecanismos usualmente utilizados estão mais direcionados para a eficiência produtiva, como o *Price Cap* (onde são estabelecidos preços máximos a serem cobrados dos consumidores finais), e outros mais para a eficiência alocativa, como *Revenue Cap* (onde é definido uma receita máxima a ser auferida pelo agente detentor da concessão). Na maior parte dos países, inclusive o Brasil, no entanto, o que tem ocorrido é adoção de um modelo híbrido, por parte do regulador, conjugando características tanto do modelo de *Price Cap* e *Revenue Cap*.

No entanto, independente do mecanismo, cabe ao regulador definir as tarifas máximas que serão cobradas dos consumidores finais, definindo em última instância o nível máximo de custos que pode ser repassado aos consumidores (Haney & Pollitt, 2009).

O desafio nesse processo é como definir esses preços de forma a manter os incentivos econômicos que sejam capazes de estimular os agentes regulados a reduzirem custos, inovar e melhorar a qualidade dos serviços (Haney & Pollitt, 2009, 2011 e 2013).

O reconhecimento tarifário integral dos custos incorridos pelas empresas reguladas nas tarifas a serem cobradas dos consumidores finais reduz os estímulos

das empresas reguladas na busca da maior eficiência (Shleifer, 1985; Kumbhakar & Lien, 2017). Por outro lado, o não reconhecimento dos custos incorridos pelas empresas reguladas, particularmente aqueles não gerenciáveis por essas empresas, pode reduzir a atratividade do setor, conduzindo a um cenário de redução de investimento e consequentemente perda da qualidade do serviço prestado pela concessionárias do mesmo (Shleifer, 1985, Kumbhakar & Lien, 2017).

A Teoria de competição por padrões (*Yardstick Competition*) desenvolvida por Shleifer (1985) sugere que a melhor forma de definir de forma eficiente o repasse de custos das empresas para os consumidores seria através da comparação, seja entre os diversos agentes regulados do setor, seja na comparação com uma empresa “fictícia”.

O modelo de comparação emularia a dinâmica de mercado em regime de concorrências, evitando o problema clássico de assimetria de informações entre reguladores e agentes regulados assegurando a maximização da eficiência econômica (Baumol, 1982; Shleifer, 1985; Baumol & Sidak, 1995; Haney & Pollitt, 2009, 2011 e 2013).

Ao definir os custos a serem repassados ao consumidor com base nos custos médios obtidos através de um modelo de comparação (*Benchmarking*), o regulador criaria um mecanismo de incentivo claro a eficiência, punindo as empresas menos eficientes (com a redução dos seus ganhos) e beneficiando as empresas mais eficientes (com ampliação dos seus ganhos). Shleifer (1985) demonstra que mesmo onde haja uma relativa heterogeneidade entre as empresas reguladas, o modelo de competição por comparação ainda pode ser adotado desde que essas heterogeneidades sejam tratadas pelo regulador dentro do processo (Shleifer, 1985)

Os métodos de regulação por comparação (*Benchmarking*) vem sendo bastante empregados por reguladores de todo mundo durante os processos de definição ou revisão de tarifas dos mais diversos setores (Haney & Pollitt, 2009, 2011 e 2013).

A forma, no entanto, como os resultados desses processos vem sendo aplicados varia muito de regulador para regulador (Haney & Pollitt, 2011, Jamasb & Pollitt, 2002, 2007; Haney & Pollitt, 2009). Há também uma grande variedade de técnicas de comparação empregadas nesse processo.

Segundo Haney & Pollitt (2009 e 2011) os métodos mais que vem sendo mais empregados pelos reguladores no mundo são: o uso do *Data Envelopment Analysis*

(DEA), Mínimos Quadrados Corrigidos (COLS), Análise de Fronteira Estocástica (SFA), Análise de Dados em Painel e o método de *benchmarking* por processo/atividade (ex. modelo de Empresa de Referência utilizado no Brasil até 2010).

A escolha do método é, no entanto, bastante complexa e depende em geral do contexto do setor e do país onde ele está sendo aplicado (Agrell & Bogetoft, 2016; Haney & Pollitt, 2011). Diferentes métodos de comparação podem gerar diferentes resultados.

Uma discussão sobre a metodologia empregada pela Agência Nacional de Energia Elétrica (ANEEL) e suas implicações para o contexto das PNT no Brasil será apresentada nas próximas seções desse capítulo.

Outros métodos alternativos de comparação e suas aplicações ao contexto desse trabalho, serão discutidas no capítulo 3.

2.2

O problema de Perdas Não Técnicas no Brasil

No Brasil, as perdas não técnicas de energia elétrica (PNT) tem se mantido relativamente estáveis ao longo da última década apesar dos fortes investimentos realizados pelas distribuidoras do país para combatê-las.

Em 2018 as PNT no país totalizaram 6,5% da energia total injetada no sistema de distribuição, valor bastem similar aos 6,4% observados em 2008, uma década antes (Figura 2.1).

Embora a situação do Brasil seja bem menos dramática que a observada em outros países em desenvolvimento, as PNT têm um impacto financeiro bastante relevante tanto para as distribuidoras quanto para os consumidores brasileiros.

Isso porque além dos custos da própria energia gerada e não paga, os consumidores e/ou as distribuidoras de energia tem que arcar com todo o custo da infraestrutura de transmissão e distribuição associado a essas PNT⁴.

⁴ Considerando o valor preço médio de compra de energia das distribuidoras em 2019 (PMIX) de R\$ 191,47 por MWh, uma perda não técnica total em 2019 de 33,3 TWh e uma TUSD Fio B média de aproximadamente R\$ 111,09 por, tem-se que o custo médio das PNT somente com a compra de energia foi de R\$ 6,4 bilhões e os custos indiretos de R\$ 3,7 bilhões, perfazendo um total de R\$ 10,2 bilhões (ANEEL, 2019a; ANEEL, 2019b).

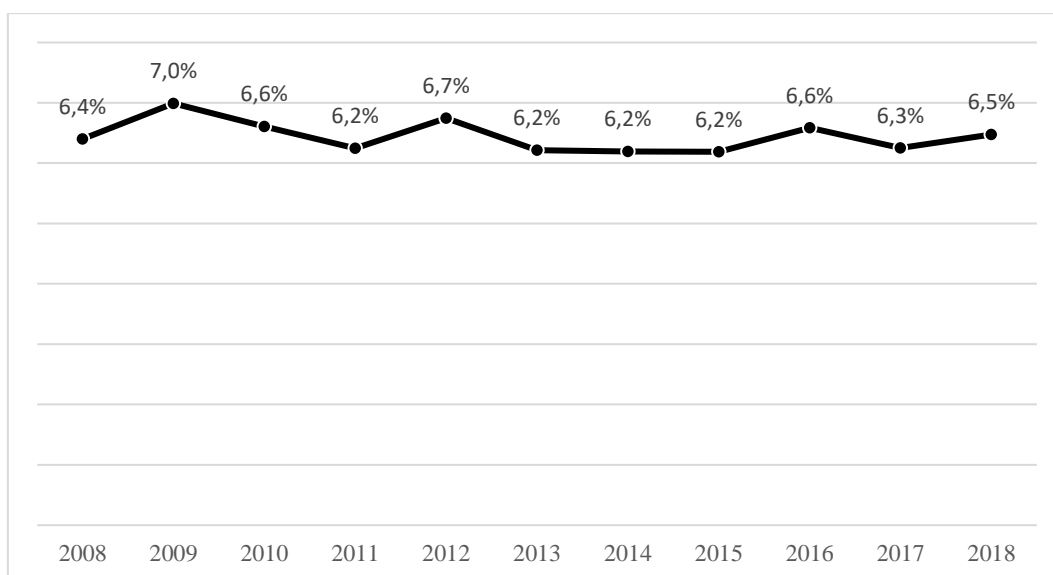


Figura 2.1 - Perdas Não Técnicas no Brasil (% da Energia Injetada Total)

Fonte: ANEEL (2019a)

Um ponto que torna o tratamento regulatório do tema bastante complexo no Brasil, é a relativa heterogeneidade das PNT observada entre as distribuidoras de energia elétricas do país. Em 2018, por exemplo, das 54 principais distribuidoras de energia elétrica, 25 registraram PNT inferiores a 3% da energia injetada total, valores bastante próximos aos valores de PNT observado em países como EUA e nos países da Europa Ocidental (Yurtseven, 2015)⁵. Por outro lado, 11 dessas distribuidoras apresentaram PNT superiores a 10% da energia injetada total (ANEEL, 2019a), chegando em alguns casos a atingir mais de 30%, valores comparáveis aos observado em países como Índia e Bangladesh (Yurtseven, 2015), países que registram os piores resultados de PNT no mundo.

⁵ Segundo Yurtseven (2015) o percentual médio de PNT nos EUA e nos países da Europa Ocidental se situam entre 1-2% da energia total injetada no sistema de distribuição.

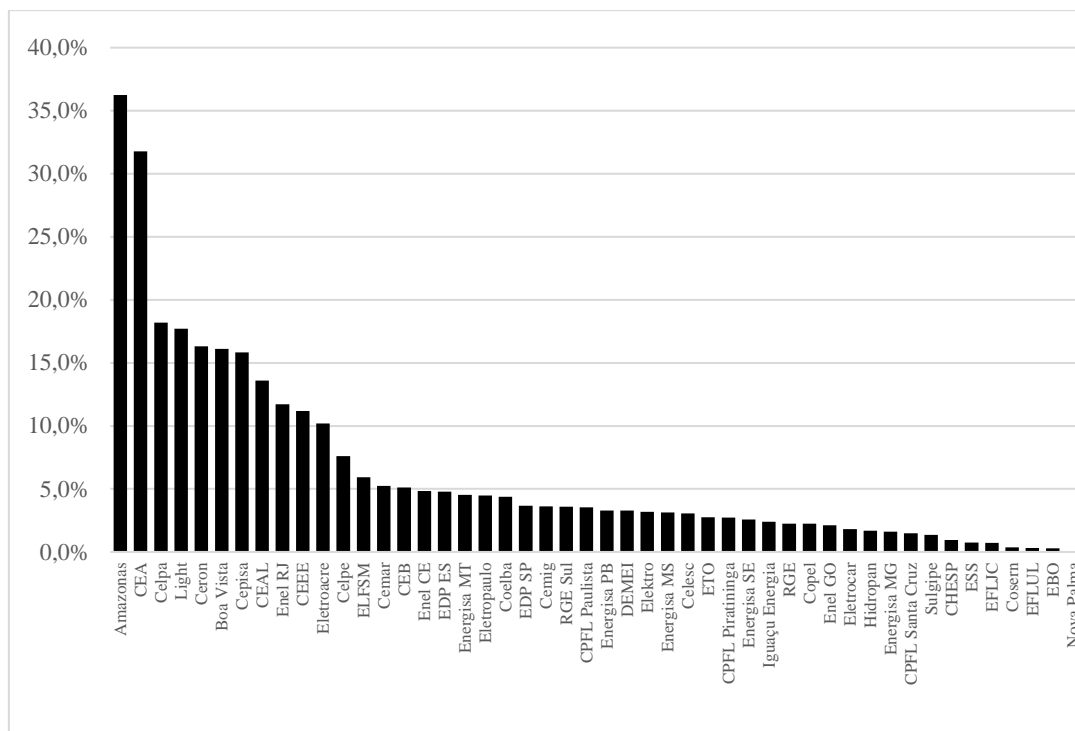


Figura 2.2 - Perdas Não Técnicas no Brasil (% da Energia Injetada Total)

Fonte: ANEEL (2019a)

Um ponto que se destaca nessa heterogeneidade do nível de PNT observados entre as distribuidoras brasileira, parece ser a baixa correlação entre as PNT observada e a renda *per capita* de cada região, o que sugere que o problema de PNT no Brasil é mais complexo que um problema meramente de nível renda dos consumidores.

Na Figura 2.2 é possível observar, por exemplo, que entre as dez distribuidoras de energia elétrica do país com maior nível de PNT, encontram-se as duas principais distribuidoras de energia elétrica do estado do RJ, o estado com o segundo maior PIB *per capita* do País. Por outro lado, distribuidoras de estados entre os menores níveis de PIB *per capita* do país, como por exemplo as distribuidoras da Paraíba e do Rio Grande do Norte (respectivamente Energisa Paraíba e Cosern), apresentam níveis de PNT bastante baixos (inferiores a 4% da energia injetada total).

De fato, essa baixa correlação entre o nível de PNT e renda *per capita*, não é uma particularidade apenas do Brasil (Smith, 2004; Millard & Emmerton & Emmerton, 2009). Países como a Indonésia, Perú e a Tailândia, em certa medida a China, apesar de apresentarem PIB *per capita* relativamente baixos frente a média

mundial, apresentam nível de PNT bastante próximos aos observados em países desenvolvidos como EUA e Canadá (Millard & Emmerton, 2009).

Esse fato tem levando alguns autores a proporem um espectro mais amplo de variáveis para explicar o fenômeno das PNT. Variáveis como a força das instituições públicas locais, o nível de corrupção, a aceitação social do furto de energia, o nível de escolaridades, o nível de desemprego, o nível de violência, entre outras variáveis vem sendo propostas como variáveis relevantes ao estudo do fenômeno (Smith, 2004; Winther, 2012; Yurtseven, 2015; Yakubu *et al.*, 2018; Chandel *et al.*, 2016; Yakubuab, 2018).

Apesar desse relativo consenso entorno do tema, a forma, de medir esses fenômenos e estimar o seu real impacto sobre as PNT parece continuar sendo um tema em aberto na literatura especializada.

No Brasil, como será visto nas próximas seções deste capítulo, a ANEEL tem tentado tratar esse tema com a adoção de um conjunto mais amplos de variáveis que buscam capturar o mais amplo conjunto possível de fatores que possam impactar a capacidade das distribuidoras de combater as PNT.

2.3

O Tratamento regulatório das Perdas Não Técnicas no Brasil

O modelo regulatório brasileiro divide as tarifas de energia elétrica a serem cobradas pelas distribuidoras em duas partes: parcela A e parcela B.

A parcela A corresponde aos custos não gerenciáveis pela distribuidora, onde em geral, a distribuidora exerce um papel de mero “arrecadador”. Os custos da parcela A englobam os custos de compra de energia elétrica dos geradores, os custos com o uso do sistema de transmissão, os encargos setoriais, as perdas de energia elétrica e, em alguns casos, a inadimplência.

A parcela B corresponde, por outro lado, aos custos gerenciáveis pelas distribuidoras, tipicamente os custos com a operação e manutenção do sistema de distribuição de energia e os investimentos na rede.

Por se tratarem de custos não gerenciáveis pelas distribuidoras, os custos relativos aos itens da parcela A são em geral repassados integralmente para os consumidores, com duas importantes exceções: as PNT e a Inadimplência (ou Receita Irrecuperáveis na denominação adotada pela ANEEL).

Nesta seção será discutido o tratamento regulatório dado as PNT.

As PNT nunca serão totalmente erradicadas (Smith, 2004). Há um limite mínimo tolerável para as PNT, a partir do qual os custos para reduzi-las superam os benefícios de sua redução. É até certo ponto consensual que as características da área na qual a distribuidora de energia elétrica opera podem influenciar dramaticamente a sua capacidade de combater às PNT. Características socioeconômicas da região, como a presença de criminalidade, acentuada desigualdade, precariedade de infraestrutura, ausência de presença do Estado e suas instituições, dentre outras (ANEEL, 2010; Leite *et al.*, 2020; Simões *et al.*, 2020; Smith, 2004; Winther, 2012; Yurtseven, 2015; Yakubu *et al.* 2018), tornam o combate PNT em algumas áreas bem mais complexo que em outras.

Isso significa que esse limite mínimo tolerável para as PNT tende a variar de distribuidora para distribuidora em função da complexidade da sua área de atuação.

Por outro lado, o esforço empreendido pela concessionária no combate às PNT, afeta o seu nível. Algumas concessionárias têm alcançado substanciais reduções de PNT mesmo em condições mais adversas por meio de inovação tecnológica, melhorias de gestão e investimentos (ANEEL, 2013).

Se fosse possível, através da observação, decompor os níveis de PNT de cada concessão em duas parcelas, uma decorrente das ações realizadas ou não pela distribuidora (parcela gerenciável) e uma segunda relacionados ao ambiente no qual a distribuidora está inserida (parcela não gerenciável), o problema do regulador se resumiria a determinar o nível ótimo de PNT da parcela gerenciável assegurando o repasse integral da parcela não gerenciável. Entretanto não é possível efetuar essa decomposição de forma direta e precisa, portanto, o regulador precisa definir um método que permita-lhe estimar essas parcelas de forma separada.

Na seção seguinte será apresentado um breve histórico da evolução do tratamento dado as PNT pela ANEEL. Será apresentado também um detalhamento da metodologia atual empregada pela ANEEL e as discussões recentes para sua atualização propostas na Consulta Pública 029/2020 promovida pela ANEEL.

2.3.1

Histórico da metodologia empregada pela ANEEL

Até 2003 as perdas elétricas consideradas no processo tarifário das distribuidoras de energia elétrica brasileiras correspondiam às perdas reais praticadas pelas distribuidoras. Apesar da parcela de perdas nem sempre ser

explicitada no processo tarifário, tais perdas estavam embutidas nos custos de compra de energia reconhecidos nas tarifas (ANEEL, 2010).

A partir do ano 2003, com o início do Primeiro Ciclo de Revisões Tarifárias Periódicas – 1ºCRTP, o reconhecimento tarifário das perdas de energia passou a ser limitado pelos patamares de perdas observados historicamente pela própria distribuidoras. Distribuidoras que vinham registrando elevações dos seus níveis de perdas passaram a enfrentar restrições no repasse desses custos à tarifa (ANEEL, 2010).

A partir de 2007, com o início do 2º ciclo de Revisões Tarifárias Periódicas – 2ºCRTP, a definição das perdas regulatórias passou a ser segmentada, com o desenvolvimento de uma metodologia específica para a definição de perdas técnicas (PT) e outra para as perdas não técnicas (PNT). Para as PNT foi empregado um modelo de *benchmarking* para definição do referencial regulatório. A meta de cada concessionária passou a ser definida a partir da observação dos níveis praticados por outras distribuidoras mais eficientes e diretamente comparáveis. A comparação foi viabilizada pela criação de um “índice de complexidade”, para cada distribuidora baseado em um conjunto de variáveis socioeconômicas relacionadas com o fenômeno PNT (ANEEL, 2013).

No Terceiro Ciclo de Revisões Tarifária Periódicas – 3ºCRTP, iniciado em 2011, foram incorporadas algumas alterações na forma de construção do “índice de complexidade”. A principal mudança foi em relação à avaliação da velocidade da trajetória de redução das PNT resultante da identificação da empresa *benchmark*. O dimensionamento dos recursos necessários para cumprimento das metas regulatórias de PNT utilizado no 2ºCRTP foi abandonado, e os limites de redução passaram a ser definidos também com base na análise de *benchmarking* (ANEEL, 2013 e 2014a).

Em 2014/15 essa metodologia foi atualizada, sofrendo pequenas alterações. Essa metodologia definida em 2014/15 vem sendo aplicada até o presente momento e será detalhada na sessão 2.3.2.

Em abril de 2020 a ANEEL (2020a) abriu uma consulta pública propondo algumas alterações pontuais na metodologia atualmente empregada, principalmente associada a inclusão/substituição de algumas variáveis explicativas e nos ajustes dos limites máximos de redução anual.

Uma breve análise das mudanças propostas será apresentada na seção 2.3.3.

2.3.2

A metodologia atual

A metodologia atual utilizada pela ANEEL para a definição das PNT regulatória consiste na realização de um processo de comparação da performance das distribuidoras de energia elétrica brasileiras ponderadas por um Índice de Complexidade Socioeconômica (IC). O IC atua como fator de “correção”, permitindo uma comparação mais equânime das distribuidoras. Essa metodologia foi aprovada no início de 2015 (ANEEL, 2015).

A construção do IC começa com a estimação através de um modelo de dados em painel com efeitos aleatórios dos coeficientes de correlação entre o percentual de PNT sobre o mercado de baixa tensão das distribuidoras e um conjunto de variáveis socioeconômicas das áreas de concessão.

No modelo estimado em 2014 foram usados os dados históricos de 2003 a 2013.

O modelo inicial proposto pela ANEEL foi estimado de acordo com a equação 2.1:

$$PNT_i = C_i + X_i\beta + IG_i \quad (2.1)$$

Onde:

PNT_i é a perda não técnica da empresa “ i ” em proporção do mercado baixa tensão;

X_i são as variáveis socioeconômicas da área de concessão;

β são os coeficientes a serem estimados;

IG_i é a parcela das perdas não técnicas da empresa “ i ” devido à ineficiência gerencial;

C_i é a variável específica da empresa “ i ” que influenciam no seu nível de perdas não técnicas não consideradas (e não observáveis) nos demais termos.

As variáveis socioeconômicas utilizadas pela ANEEL foram:

Quadro 2.1 - Variáveis socioeconômicas utilizadas pela ANEEL na construção do IC

| Código | Variáveis socioeconômicas | Fonte |
|---------------|--|--------------|
| Vio | Violência - Óbitos por Agressão a cada 100 mil habitantes | DATASUS |
| pob2 | % de pessoas com renda <i>per capita</i> inferior a ½ salário mínimo | IBGE/IPEA |
| Gini | Gini | IBGE |
| sub2 | Precariedade - % de pessoas em domicílios subnormais (Censo 2010) | IBGE |
| lixo.u | % Domicílios com Coleta de lixo – urbano | IBGE |
| Inad | Inadimplência do setor de crédito – BACEN | BACEN |
| Mbr.Mb1Mbr | Mercado baixa renda/(Mercado de baixa tensão residencial) | ANEEL |
| Mbr.Mbt | Mercado baixa renda/(Mercado de baixa tensão total) | ANEEL |

Fonte: ANEEL (2014a e 2014b)

Segundo a ANEEL, a opção pelo método de dados em painel com efeitos aleatórios se deu por três razões: (i) pela estrutura em painel em que se apresentam os dados, (ii) pela óbvia vantagem de se ganhar um maior número de observações na análise temporal, e (iii) mais importante, pelo fato de esses modelos oferecerem um relativo controle para heterogeneidade presente nos indivíduos, que no caso específico de perdas é interpretado pela ANEEL como um fator gerenciável pelas distribuidoras (ANEEL, 2010 e 2014b).

Inicialmente foram apresentados 12 modelos pela ANEEL, a partir dos quais foram selecionados 3 modelos que, segundo a agência, melhor representariam o fenômeno (ANEEL, 2014b). Os critérios considerados para a seleção desses 3 modelos, segundo a ANEEL (2014b) foram: (i) a maior significância estatística dos coeficientes estimados; (ii) a coerência do sinal dos coeficientes em relação ao esperado; (iii) a maior relevância das variáveis para explicação das PNT; e (iv) as contribuições recebidas durante a Audiência Pública nº 23/2014 (ANEEL, 2014b). Cada um dos 3 modelos foi estimado a partir de uma combinação distintas de variáveis socioeconômicas, como pode ser visto no Quadro 2.2.

Quadro 2.2 - Variáveis socioambientais utilizadas pela ANEEL nos modelos C, G e K

| Modelos | Variáveis socioeconômicas |
|---------|--------------------------------------|
| C | sub2, lixo.u, pob2, Mbr.Mb1Mbr, inad |
| G | sub2, lixo.u, gini, Mbr.Mb1Mbr, inad |
| K | sub2, lixo.u, pob2, Mbr.Mbt, vio |

Fonte: ANEEL (2014b)

As especificações dos modelos utilizados (C, G e K) para a estimação dos coeficientes das variáveis que explicam as PNT, são apresentadas nas Equações 2.2 e 2.3.

$$y_2(C) = 0,144 + 1,98sub2_{it} - 0,17lixo.u_{it} + 0,24pob2_{it} - 0,10Mbr.Mb1Mbr_{it} + 1,39inad_{it} + \varepsilon_{it} + \mu_i \quad (2.2)$$

$$y_2(G) = 0,160 + 1,96sub2_{it} - 0,26lixo.u_{it} + 0,18gini_{it} + 1,47inad_{it} + \varepsilon_{it} + \mu_i \quad (2.3)$$

$$y_2(K) = 0,175 + 2,21sub2_{it} - 0,18lixo.u_{it} + 0,20pob2_{it} - 0,19Mbr.Mbt_{it} + 0,001vio_{it} + \varepsilon_{it} + \mu_i \quad (2.4)$$

Onde:

$y_2(i)$ é o índice de perdas não técnicas estimados pelos modelos C, G ou K;

$sub2_{it}$ é o percentual de pessoas que vivem em domicílios subnormais, na região de concessão da empresa i , no ano t ;

$lixo.u_{it}$ é o percentual dos domicílios que possuem coleta de lixo, na região de concessão da empresa i , no ano t ;

$pob2_{it}$ é o percentual da população de que vive com menos de meio salário mínimo, residentes na região de concessão da empresa i , no ano t ;

$Mbr.Mb1Mbr_{it}$ é o mercado de baixa renda em relação ao mercado B1 (residencial) total, na região de concessão da empresa i , no ano t ;

$inad_{it}$ é a inadimplência do setor de crédito, dos residentes da região de concessão da empresa i , no ano t ;

$gini_{it}$ é o índice de Gini, na região de concessão da empresa i , no ano t ;

$Mbr.Mbt_{it}$ é o mercado de baixa renda em relação ao mercado BT (Baixa Tensão) total, na região de concessão da empresa i , no ano t ;

vio_{it} é o índice de mortes por agressão a cada cem mil habitantes, na região de concessão da empresa i , no ano t ;

α é o intercepto da regressão, ou seja, o valor das perdas não técnicas estimadas se todas as variáveis forem iguais a zero; equivale a 0,144 no modelo C, 0,159 no modelo G e 0,174 no modelo K;

ε_{it} é o desvio idiossincrático da regressão; e

μ_i é a heterogeneidade de cada região de concessão i , invariante no tempo.

Após a estimação desses parâmetros, os β estimados em cada um dos modelos (C, G e K) foram utilizados para construir o que a ANEEL denominou de Índice de Complexidade Socioeconômica (IC). O índice para cada uma das distribuidoras foi construído multiplicando-se os parâmetros β estimados em cada modelo pelos valores das variáveis socioeconômicas de cada distribuidora para o ano de 2013, conforme Equação 2.5:

$$IC^A = \sum_{i=1}^n X_i^A \beta_i \quad (2.5)$$

Onde:

IC^A é o Índice de Complexidade Socioeconômica da distribuidora “A”;

X_i^A é o valor da variável “ i ” para a distribuidora “A” no ano de 2013;

β_i é o valor do coeficiente de regressão estimado para a variável “ i ”.

A heterogeneidade estimada no modelo de dados em painel (μ_i) foi desconsiderada no cálculo do IC. Segundo a ANEEL a heterogeneidade estimada (μ_i) estaria associada em sua maior parte a ineficiência das distribuidoras e portanto não deveria compor o IC.

Como será discutida na seção 2.5 deste capítulo, essa é uma hipótese forte e que pode estar por trás de parte dos problemas observados na aplicação do modelo da ANEEL.

A partir dos resultados da estimação dos modelos C, G e K foi construída uma matriz probabilidade para medir a incerteza sobre os IC. Essa matriz contém a probabilidade de uma determinada distribuidora “A” ter um IC inferior a empresa “B” ($Prob(\blacksquare)$). Essa probabilidade foi calculada para cada combinação de 2 a 2 das distribuidoras contidas na amostra da ANEEL através de um teste $t - Student$ conforme a equação 2.6⁶:

$$Prob(IC_A < IC_B) = Prob\left(\tau_{critico} > \frac{\hat{Y}_A - \hat{Y}_B}{\sqrt{V(\hat{Y}_A) + V(\hat{Y}_B)}}\right) \quad (2.6)$$

Onde:

$Prob(IC_A < IC_B)$ é a probabilidade da distribuidora “A” ter um IC inferior a distribuidora “B”

\hat{Y}_A é o valor estimado de PNT para distribuidora “A”

\hat{Y}_B é o valor estimado de PNT para distribuidora “B”

$V(\hat{Y}_A)$ é variância da estimativa de \hat{Y}_A

$V(\hat{Y}_B)$ é variância da estimativa de \hat{Y}_B

$\tau_{critico}$ é o valor da estatística $t - student$

A partir dos três ICs construídos para cada uma das distribuidoras e da matriz de probabilidade construída a partir da equação 2.6, foi proposta uma metodologia para a definição dos percentuais de PNT a serem aplicados nos processos tarifários (ANEEL, 2015).

⁶ No caso, infere-se que existe um teste de hipótese do tipo: $\begin{cases} H_0: \hat{Y}_{Bench} \geq \hat{Y}_{Empr} \\ H_a: \hat{Y}_{Bench} < \hat{Y}_{Empr} \end{cases}$

Nos testes de hipóteses são gerados p-valores, que informalmente é a probabilidade de ocorrência da hipótese nula (H_0). Desta forma, tal p-valor, ou o valor da probabilidade do Benchmarking estar em uma área realmente mais complexa é dada pelo seguinte teste $t - Student$: $t_{cal} = (\hat{Y}_{Bench} - \hat{Y}_{emp}) / \left(\sqrt{V(\hat{Y}_{Bench}) + V(\hat{Y}_{emp})} \right)$. Desta forma, calcula-se, via Tabela de distribuição de t de $student$ a $Pr(t_{cri} > t_{cal})$.

Segundo essa metodologia, no processo de Revisão Tarifária Periódica a PNT real mais recente disponível da distribuidora que estiver passando por Revisão Tarifária (por exemplo a “Distribuidora A”), é comparada com as PNT observadas nas demais distribuidora no mesmo período.

Essa comparação é ponderada pela probabilidade das demais distribuidora terem um IC superior à da distribuidora em processo de Revisão. A meta regulatória de PNT da distribuidora é então definida conforme a Equação 2.6, sendo selecionada com empresa *Benchmark*, aquela distribuidora que aplicando-se a Equação 2.6 resulta no menor valor de *Meta PNT_A*.

$$Meta\ PNT_{A,j} = Prob(\blacksquare) \times PNT_{Benchmark} + (1 - Prob(\blacksquare)) \times PNT_A \quad (2.7)$$

Onde:

Meta PNT_{A,j} é a meta de PNT medida a ser estabelecida pela ANEEL para a distribuidora “A” em percentual do mercado de BT medido da distribuidora “A” pelo modelo j;

Prob(■) é a Probabilidade que a empresa *Benchmarking* de fato possua um IC maior que a empresa “A”;

PNT_{Benchmark,j} é a PNT medida da empresa em percentual do mercado de BT medido, que resulte no menor valor de *Meta PNT_{A,j}* para a distribuidora “A” em percentual do mercado de BT pelo modelo j;

PNT_A é a PNT da distribuidora “A” medida no ano imediatamente anterior à data de realização do cálculo da meta em percentual do mercado de BT medido.

São calculados três valores para *Meta PNT_A* da distribuidora A, cada um considerando os resultados de um dos três modelos das equações 2.2, 2.3 e 2.4 (modelos C, G e K). Em seguida a meta fina é estabelecida a partir da média aritmética desses três valores conforme a Equação 2.8 (ANEEL, 2015b).

$$Meta\ Final\ PNT_A = \frac{\sum_{j=1}^n Meta\ PNT_{A,j}}{3} \quad (2.8)$$

Onde:

Meta Final PNT_A = potencial de redução da empresa A [%];

$Meta\ PNT_{A,j}$ = meta de perdas da empresa “A” [%] conforme modelo j ;

Na metodologia estabelecida pela ANEEL, na definição das $Meta\ PNT_{A,j}$ as distribuidoras são segregadas conforme o seu porte em dois grupos: grupo 1 e grupo 2. O grupo 1 consiste nas distribuidoras de maior porte definidas como as distribuidoras que possuem mercado de baixa tensão maior que 1.000 GWh/ano e que atendem mais que 500 mil unidades consumidoras ou que possuem mais do que 15.000 km de rede elétrica. As demais distribuidoras são consideradas no Grupo 2.

Uma vez estabelecida a meta regulatória de $Meta\ Final\ PNT_A$, segundo a metodologia da ANEEL é definido um ponto de partida (meta a ser aplicada para o primeiro ano) e uma trajetória de redução de PNT anual de 4/5 anos para que a distribuidora possa se adequar à nova meta de PNT estabelecida. O ponto de partida é definido conforme o Quadro 2.3.

Quadro 2.3 - Definição do Ponto de Partida das PNT regulatórias sobre o mercado de baixa tensão (BT)⁷

| Grupo | Condição Analisada | Ponto de partida |
|-------|---|---|
| Todas | Meta do novo Ciclo > Meta do Ciclo anterior | Meta do novo Ciclo |
| 1 | Média 4 anos (PNT_A) > 7,5% | Máximo [7,50%; Mínimo (Meta regulatória anterior medida, Média realizada dos últimos 4 anos)] |
| | Média 4 anos (PNT_A) < 7,5% | Média dos últimos 4 anos |
| 2 | Média 4 anos (PNT_A) > 2,5% | Máximo [2,50%; Mínimo (Meta regulatória anterior medida, Média realizada dos últimos 4 anos)] |
| | Média 4 anos (PNT_A) < 2,5% | Média dos últimos 4 anos |

Fonte: ANEEL (2015b)

Por fim é definida uma trajetória anual de redução do valor definido como ponto de partida da PNT regulatórias de forma que o valor de PNT regulatório se aproxime do valor da $Meta\ Final\ PNT_A$ definido na equação 2.8 ao longo do ciclo

⁷ Embora de aplicação geral, o Procedimentos de Regulação Tarifária – PRORET – submódulo 2.6 – Perdas de Energia, que regulamenta o processo de definição das PNT regulatórias, prevê a possibilidade de adoção de outros critérios para definição do ponto de partida para as distribuidoras com maior IC dentro de cada grupo.

tarifário. Essa trajetória de redução é limitada conforme os parâmetros definidos no Quadro 2.4.

Quadro 2.4 - Limites anuais de redução das PNT regulatórias sobre o mercado de baixa tensão (BT)

| Grupo | Condição Analisada | Limite de redução |
|-------|--|--------------------------------|
| 1 | $Meta\ Final\ PNT_A > 7,50\%$ | $Meta\ Final\ PNT_A/8 - 15/16$ |
| | $Meta\ Final\ PNT_A < 7,50\%$ | Sem Trajetória |
| 2 | $Meta\ Final\ PNT_A > 11,50\%$ | $Meta\ Final\ PNT_A/8 - 15/16$ |
| | $2,5\% < Meta\ Final\ PNT_A < 11,50\%$ | 0,50% ao ano |
| | $Meta\ Final\ PNT_A < 2,50\%$ | Sem Trajetória |

Fonte: ANEEL (2015b)

Após o estabelecimento dos valores as metas de PNT regulatória anuais da distribuidora, esses valores são convertidos para percentuais de PNT sobre o mercado BT faturado, conforme a equação 2.9.

$$Meta\ Final\ PNTfat_{A,t} = Meta\ Final\ PNT_{A,t} + \% Ajuste \quad (2.9)$$

Onde:

$Meta\ Final\ PNT_{A,t}$ são as metas anuais de PNT regulatórias definidas por empresa A para o ano t, metas essas construídas a partir da diferença entre a $Meta\ Final\ PNT_A$ e a PNT_A , definidas nas equações 2.8 e 2.7 respectivamente, aplicando-se as os limites estabelecidos nos Quadros 2.3 e 2.4;

$\% Ajuste$ é o percentual resultante da mediana das diferenças entre $PNTfat_A$ e PNT_A da distribuidora observado nos últimos quatro anos civis disponíveis;

$PNTfat_A$ é a PNT faturada da distribuidora “A” em percentual do mercado BT faturado da distribuidora;

PNT_A é a PNT medido da distribuidora “A” em percentual do mercado BT medido da distribuidora.

Os valores para cada um dos 4/5 anos, após o ajuste, passam a ser considerados nos processos tarifários dos anos subsequentes. Esses valores representam, em última instância, o volume de PNT que será repassado à tarifa de energia elétrica.

Os ICs vigentes, a matriz de probabilidade calculada para todas as distribuidoras, assim como a metodologia completa de cálculo das metas de PNT são apresentadas pela ANEEL no regulamento técnico da agência denominado Procedimentos de Regulação Tarifária – PRORET – submodulo 2.6 – Perdas de Energia (ANEEL, 2015b).

2.3.3

A Consulta Pública nº 029 de 2020 e as possíveis mudanças na metodologia atual de definição das metas de perdas não técnicas

Em 10 de abril de 2020 a ANEEL pôs em consulta pública uma proposta de alteração da metodologia atual utilizada pela agência para a definição das PNT regulatória.

Com o intuito de reformular e atualizar os modelos atualmente empregados no cálculo do Índice de Complexidade Socioeconômico (IC), a ANEEL estimou 127 modelos de dados em painel com efeitos aleatórios correlacionado o percentual de PNT sobre o mercado de baixa tensão medido das distribuidoras e um conjunto de 7 variáveis socioeconômicas das áreas de concessão. As variáveis socioeconômicas testadas foram:

Quadro 2.5 - Variáveis Socioeconômicas testadas pela ANEEL na CP 029/2020

| Código | Variáveis socioeconômicas | Fonte |
|----------|---|-----------|
| Vio | Violência - Óbitos por Agressão a cada 100 mil habitantes | DATASUS |
| Pob3 | % de domicílios com renda <i>per capita</i> inferior a ½ salário mínimo | IBGE/IPEA |
| Inad | Inadimplência do setor de crédito – BACEN | BACEN |
| sub2 | Precariedade - % de pessoas em domicílios subnormais (Censo 2010) | IBGE |
| Admd | Nº de moradores por dormitório por domicílio | IBGE |
| PIB.PC | PIB <i>per capita</i> | IBGE |
| EficOper | % de eficiência operacional das Distribuidoras | ANEEL |

Fonte: ANEEL (2020a)

Os 127 modelos foram avaliados posteriormente pela ANEEL de forma sequencial por 4 critérios previamente estabelecidos e listados no Quadro 2.6, de forma que ao final a ANEEL obtivesse 3 modelos (X, Y e Z).

Quadro 2.6 – Critério de seleção de modelos adotado pela ANEEL na CP 029/2020

| Filtro | Critérios Utilizados | Nº de Modelos Resultantes |
|---------------|---|----------------------------------|
| 1º | Poder Explicativo $R^2 > 0,10$ | 90 |
| 2º | Significância Estatística $p\text{-valor} < 0,10$ | 31 |
| 3º | Sinal do Coeficiente coerente | 14 |
| 4º | Ranking (R^2) | 3 |

Fonte: ANEEL (2020a)

As especificações dos 3 modelos propostos (X, Y e Z) para a estimação dos coeficientes das variáveis que explicam as PNT, são apresentadas nas Equações 2.10, 2.11 e 2.12.

$$y_2(X) = 0,3756 + 2,926 \text{ sub}2_{it} - 1,441 \text{ Admd}_{it} - 0,00000303 \text{ PIB.pc}_{it} - 0,449 \text{ EficOper}_{it} + \varepsilon_{it} + \mu_i \quad (2.10)$$

$$y_2(Y) = 0,3214 + 2,735 \text{ sub}2_{it} - 1,502 \text{ Admd}_{it} + 0,001 \text{ vio}_{it} - 0,454 \text{ EficOper}_{it} + \varepsilon_{it} + \mu_i \quad (2.11)$$

$$y_2(Z) = 0,3482 + 2,928 \text{ sub}2_{it} - 1,449 \text{ Admd}_{it} - 0,4629 \text{ EficOper}_{it} + \varepsilon_{it} + \mu_i \quad (2.12)$$

Onde:

$y_2(i)$ é o índice de perdas não técnicas estimados pelos modelos X, Y ou Z;

$\text{sub}2_{it}$ é o percentual de pessoas que vivem em domicílios subnormais, na região de concessão da empresa i , no ano t ;

Admd_{it} é o N° de moradores por dormitório por domicílio, na região de concessão da empresa i , no ano t ;

EficOper_{it} é o percentual % de eficiência operacional da empresa i , no ano t ;

$PIB.pc$ é o PIB *per capita* da região de concessão da empresa i , no ano t ;

vio_{it} é o índice de mortes por agressão a cada cem mil habitantes, na região de concessão da empresa i , no ano t ;

α é o intercepto da regressão, ou seja, o valor das perdas não técnicas estimadas se todas as variáveis forem iguais a zero;

ε_{it} são os desvios idiossincráticos da regressão; e

μ_i são as heterogeneidades de cada região de concessão i , invariante no tempo.

Assim pela proposta da ANEEL, esses modelos substituiriam os modelos atuais C, G e K no cálculo do IC, mantendo-se, no entanto, as demais etapas da metodologia atual.

Nos novos modelos, três novas variáveis socioeconômicas (ADMD, PIB.pc e EficOper) passam a compor o IC substituindo 4 variáveis atualmente empregadas (lixo.u, pob2, Mbr.Mb1Mbr e o Mbr.Mbt).

Além da alteração dos parâmetros utilizados no cálculo do IC, ANEEL também propões uma alteração no critério de separação das empresas, passando a considerar no grupo 1 (grande porte) as distribuidoras com mercado de baixa tensão superior à 700 GWh/ano (ANEEL, 2020).

A regra para definição do ponto de partida (meta do primeiro ano) também seria alterada passando a ser definida a partir de um único critério independente do porte das distribuidoras. Esse novo critério proposto é apresentado no Quadro 2.7.

Quadro 2.7 – Critério para a definição do Ponto de Partida das PNT regulatórias sobre o mercado de baixa tensão (BT) proposto na CP 029/2020

| Grupo | Condição Analisada | Ponto de partida |
|-------|---|--|
| Todas | Meta do novo Ciclo > Meta do Ciclo anterior | Meta do novo Ciclo |
| | Média 3 anos (PNT_A) > 7,5% | Mínimo (Meta Ciclo Anterior, Média dos últimos 3 anos) |

Fonte: ANEEL (2020a)

Por fim, a proposta da ANEEL também altera os limites de redução anual das metas de PNT regulatórias o que é apresentado o Quadro 2.8.

Quadro.2.8 - Critério para a definição dos limites anuais de redução das PNT regulatórias sobre o mercado de baixa tensão (BT) proposto na CP 029/2020

| Grupo | Condição Analisada | Limite de redução |
|-------|--------------------------------------|--|
| 1 | $Meta\ Final\ PNT_A > 5,0\%$ | $\% \ 7 * Meta\ Final\ PNT_A / 50 - 7 / 10$ |
| | $Meta\ Final\ PNT_A < 5,0\%$ | <i>Sem Trajetória</i> |
| 2 | $Meta\ Final\ PNT_A > 8,6\%$ | $\% \ 7 * P\ Meta\ Final\ PNT_A / 50 - 7 / 10$ |
| | $2,5\% < Meta\ Final\ PNT_A < 8,6\%$ | 0,50% ao ano |
| | $Meta\ Final\ PNT_A < 2,50\%$ | <i>Sem Trajetória</i> |

Fonte: ANEEL (2020a)

A nova proposta ainda inclui a possibilidade de se incorporar um tratamento *ad hoc* para as Áreas de Risco das distribuidoras de grande porte que apresentem o maior nível de perdas não técnicas (terceiro quartil considerando as perdas não técnicas sobre o mercado de BT medido).

Embora a proposta da ANEEL altera alguns parâmetros da metodologia atual empregada na definição das metas de PNT regulatórias, a Agência não altera a metodologia que vêm sendo aplicada desde de 2011. Assim, as mesmas críticas em relação a metodologia atual permanecem validas para o modelo proposto. Na seção 2.4 são apresentadas as principais críticas ao tratamento regulatório das PNT no Brasil.

Apesar da metodologia proposta pela ANEEL ainda estar em consulta pública e, portanto, a sua aplicação ainda seja incerta, neste trabalho optou-se por testar e incorporar algumas das melhorias propostas na consulta pública. Assim no capítulo 5 foram testados no modelo proposto nesta dissertação as variáveis ADMD e PIB *per capita*, assim como foram utilizados os dados históricos de PNT e das variáveis socioeconômicas mais atuais disponibilizadas pela ANEEL na referida consulta pública.

2.4

Críticas ao tratamento regulatório das Perdas Não Técnicas no Brasil

Há três conjuntos de críticas à metodologia atualmente empregada pela ANEEL na definição das metas de PNT das distribuidoras. O primeiro deles se refere a seleção das variáveis explicativas utilizada no processo de estimação do Índice de Complexidade Socioeconômica (IC). O segundo se refere a problemas

técnicos no processo de definição e estimação dos modelos econométricos utilizados. O terceiro se refere a metodologia adotada pela ANEEL para a definição das metas a partir do cálculo do IC e a exequibilidade das próprias metas estabelecidas pela ANEEL a partir dessa metodologia. Nessa seção serão apresentadas as principais críticas desses três grupos.

2.4.1

O problema das variáveis explicativas

O primeiro conjunto de crítica refere-se à seleção das variáveis explicativas utilizadas pelas ANEEL na elaboração do modelo proposto. Para construção do Índice de Complexidade Socioeconômica (IC) a ANEEL utilizou 8 variáveis socioeconômicas. Embora as variáveis selecionadas busquem retratar um conjunto amplo de fenômenos que podem influenciar a operação das distribuidoras (ANEEL 2010, 2014a e 2014b), a disponibilidade dessas variáveis e a suas respectivas abrangências são limitações relevantes da metodologia utilizada pela ANEEL.

O primeiro problema, amplamente discutido durante as Audiências Públicas 040/2010 e 023/2014, se refere a abrangência geográfica das variáveis. Para a correta estimação do IC é necessário que as variáveis explicativas reflitam a realidade de cada área de concessão de cada distribuidora. Embora em muitos casos a área de concessão das distribuidoras corresponda a totalidade da área de um estado da federação, isso não é uma regra geral. Há diversas distribuidoras, particularmente no Sul e Sudeste do Brasil, cujas áreas de concessão se restringem a apenas parte dos municípios de um estado da federação. Essa é a realidade por exemplo, das distribuidoras que atuam no Estado de São Paulo (7 distribuidoras), Rio de Janeiro (3 distribuidoras), Minas Gerais (3 distribuidoras) e Rio Grande do Sul (3 distribuidoras). Nesses casos para que fosse possível construir um conjunto variáveis que refletisse a realidade de cada distribuidora seria necessário que as variáveis socioeconômicas utilizadas estivessem disponíveis por municípios.

Como pode ser observado no Quadro 2.9, ao menos duas variáveis (inad e Gini) só estão disponíveis por estados, e outras três (pob2, sub2 e lixo.u), embora disponíveis por município para o ano de 2010, tiveram suas séries temporais construídas a partir de dados agregados por estado⁸. Desta forma, estas variáveis

⁸ Nesse caso, embora as variáveis sejam construídas de forma individualizada para cada área de concessão, a evolução relativa da variável ao longo do tempo entre as distribuidoras é parcial ou

podem não refletir de maneira correta as situações relativas da complexidade socioeconômica das distribuidoras distorcendo o cálculo do IC e consequentemente o processo de comparação entre elas.

integralmente ignorada. Um dos problemas dessa abordagem é apresentado por Miranda, A. P. M. *et al.* (2018) em um estudo sobre a evolução da violência no estado do RJ.

Quadro 2.9 - Variáveis Socioeconômicas utilizadas pela ANEEL no cálculo do IC

| Código | Variáveis socioeconômicas | Fonte | Abrangência | Disponibilidade do dado |
|---------------|--|--------------|--------------------------|--|
| vio | Violência - Óbitos por Agressão a cada 100 mil habitantes | DATASUS | Municípios | Anualmente por município |
| pob2 | % de pessoas com renda <i>per capita</i> inferior a ½ salário mínimo | IBGE/IPEA | Municípios | Ano de 2010 - Dados por municípios (Censo) Demais anos – atualização dos dados do Censo (Municípios) com base nos dados da PNAD (Estados) |
| Gini | Gini | IBGE | Estados | Anual por estados |
| sub2 | Precariedade - % de pessoas em domicílios subnormais (Censo 2010) | IBGE | Municípios | Ano de 2010 - Dados por municípios (Censo) Demais anos – atualização dos dados do Censo (Municípios) com base nos dados da PNAD de crescimento da população (Estados) |
| lixo.u | % Domicílios com Coleta de lixo – urbano | IBGE | Municípios | Ano de 2010 - Dados por municípios (Censo) Demais anos – atualização dos dados do Censo (Municípios) com base nos dados da PNAD (Estados) |
| inad | Inadimplência do setor de crédito – BACEN | BACEN | Estados | Anual por estados |
| Mbr.Mb1Mbr | Mercado baixa renda/(Mercado B1 total) | SAMP | Distribuidora de Energia | Anual por distribuidora de Energia |
| Mbr.Mbt | Mercado baixa renda/(Mercado BT total) | SAMP | Distribuidora de Energia | Anual por distribuidora de Energia |

Fonte: ANEEL (2014a e 2014b)

O segundo problema se refere a abrangência conceitual das próprias variáveis, combinado ou não, aos seus respectivos processos de atualização (como no caso da variável sub2). Esse problema é particularmente relevante em três variáveis chaves: o percentual de domicílios subnormais (sub2), a variável violência (vio) e a variável inadimplência (inad).

A variável percentual de domicílios subnormais (sub2) talvez seja a variável mais crítica nesse processo. O percentual de domicílios subnormais é calculado pela divisão dos números de domicílios presentes nos aglomerados subnormais em uma determinada região (setor censitário, bairro, município, estado ou país) pelo total de domicílios nesta mesma região. Esse cálculo é realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) a partir dos dados coletados no Censo Demográfico. O último dado disponibilizado pelos IBGE foi calculado durante o Censo Demográfico de 2010.

Segundo o IBGE (2010) um aglomerado subnormal é um conjunto constituído por 51 ou mais unidades habitacionais caracterizadas por ausência de título de propriedade e pelo menos uma das características: (a) irregularidade das vias de circulação e do tamanho e forma dos lotes e/ou (b) carência de serviços públicos essenciais (como coleta de lixo, rede de esgoto, rede de água, energia elétrica e iluminação pública).

Segundo a ANEEL (2010) essa variável refletiria a precariedade dos domicílios em determinadas áreas de atuação da distribuidora. A Figura 2.3 apresenta o percentual dos domicílios subnormais para as 64 distribuidoras estimado pela ANEEL com base nos dados do IBGE para 2010.

O problema dessa variável está, como reconhece a própria ANEEL (2010), principalmente na primeira condição adotada pelo IBGE para classificar um aglomerado subnormal: a ausência de título de propriedade. Essa premissa é fortemente dependente da colaboração entre o IBGE e as Prefeituras (ANEEL, 2010), além disso ela ignora as políticas de regularização fundiária e urbanização adotada por diversas prefeituras, o que compromete a qualidade da comparação entre diferentes municípios e/ou regiões. Esse problema da qualidade da classificação dos aglomerados subnormais foi reconhecido pelo próprio Governo Federal, no estudo realizado pelo Ministério das Cidades em 2007 (Marques *et al.*, 2007) denominado “Assentamentos precários no Brasil urbano”.

Nesse sentido, Miranda *et al.* (2018), em um estudo sobre a evolução da violência no estado do RJ e seus impactos sobre as perdas não técnicas de energia elétrica na Enel Distribuição Rio argumentam que ao utilizar essa variável a ANEEL tende a subestimar as dificuldades de atuação de muitas distribuidoras.

Por fim, um outro problema da adoção da variável subnormal decorre do próprio processo como a variável foi construída para os anos nos quais não há dados do Censo Demográfico. Ao atualizar o percentual de domicílio subnormais por município exclusivamente pelo crescimento da população, a ANEEL reconheceu, implicitamente, a hipótese que a participação dos domicílios subnormais nos municípios se mantem constante ao longo do tempo (ao menos a partir de 2010). Essa hipótese parece pouco plausível se comparadas a evolução de outras variáveis como a violência (Miranda *et al.* 2018).

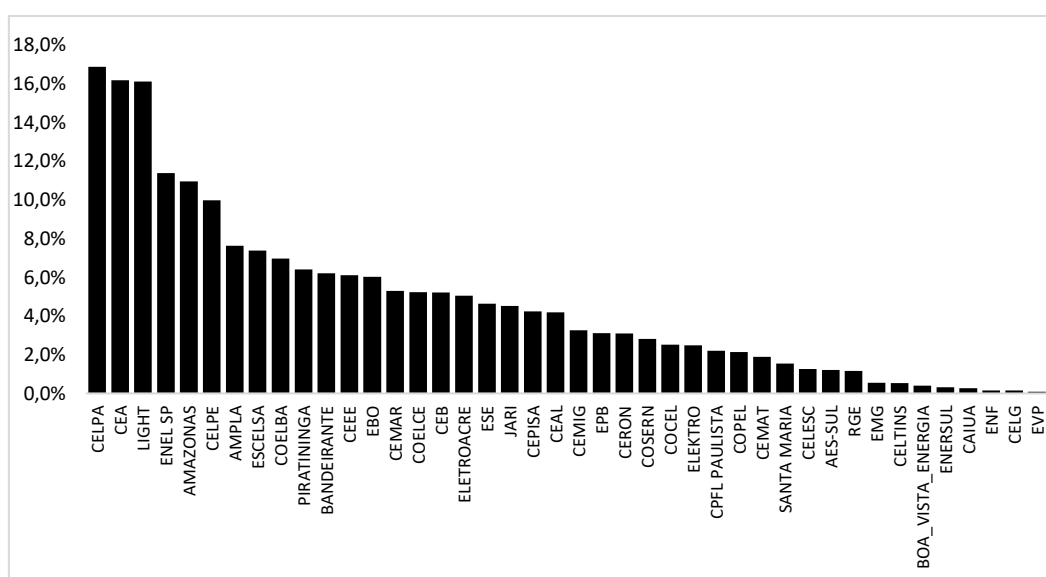


Figura 2.3 - Percentual de Domicílios subnormais por distribuidora de energia elétrica em 2010⁹

Fonte: ANEEL (2020b)

Miranda *et al.* (2018) também criticam a adoção pela ANEEL da variável óbitos por agressão (vio) como uma variável adequada para refletir a violência. As autoras argumentam que, embora a variável óbitos por agressão (vio) possa indicar um maior ou menor grau de violência em uma determinada região, ele é incapaz de refletir de forma clara os aspectos da violência mais relevantes para atuação das

⁹ As demais distribuidoras tiveram percentuais de domicílios inferiores a 0,4%.

distribuidoras que é a existência de uma espécie de poder paralelo exercido por grupos criminosos nessas regiões.

O principal problema do indicador estaria em sua própria abrangência, que incluiria uma gama ampla de homicídios. Homicídios decorrentes de crimes passionais, por exemplo, afetariam o indicador de óbitos por agressão, mas seriam irrelevantes para a atuação das distribuidoras de energia elétricas. Segundo as autoras, essa poderia ser uma das razões para baixa participação do tema da violência nos IC calculado pela ANEEL.

Segundo as autoras, a variável chave para explicar a menor ou maior capacidade de atuação das distribuidoras estaria associada mais ao controle territorial exercido por grupos armados que propriamente a tipos específicos de violência, como os homicídios. Esses grupos, mediante a ameaça e violência, limitariam a atuação das distribuidoras sobre determinadas áreas da concessão impedindo as ações de combate a perda.

As autoras sugerem, que a identificação da presença desses grupos, se daria de forma mais adequada pela adoção de um conjunto mais amplo de variáveis de violência, como por exemplo, os dados de latrocínios, lesão corporal seguida de morte, número de policiais militares e civis mortos em situação de confronto (em serviço) e morte em decorrência de atividade policial, entre outras¹⁰.

A crítica a variável inadimplência (inad) concentra-se em um aspecto um pouco diferente das críticas as variáveis de domicílios subnormais e violência. Há um relativo consenso de que as perdas de energia elétricas e a inadimplência são problemas correlacionados (Smith, 2004). A maior ou menor capacidade de pagamento dos consumidores parece ser uma variável explicativa para ambos os fenômenos.

No cálculo do Índice de Complexidade Socioeconômica (IC) a ANEEL utilizou entre as variáveis socioeconômicas o percentual de inadimplência total no setor de crédito do Sistema Financeiro Nacional (SFN) por estado calculado pelo Banco Central do Brasil como uma proxy para o problema da inadimplência. Segundo o Banco Central (2019) esse indicador inclui operações contratadas, tanto

¹⁰ Calili et al. (2018) propõem, com base em um modelo de componentes principais, a construção de um indicador de áreas de risco a partir de um conjunto de 7 variáveis de violência.

por pessoas físicas quanto jurídicas. Entretanto o problema de PNT parece estar pouco relacionado as pessoas jurídicas.

Como reconhece a própria ANEEL (2010), a maior parte das irregularidades que produzem PNT encontra-se concentradas no mercado de baixa tensão¹¹. Segundo dados da própria agência 70% do consumo desse segmento encontra-se nas classes residencial e rural, sendo, portanto, a maior parte do consumo no mercado de baixa tensão associados a pessoas físicas e não Jurídicas.

Os dados do Banco Central, no entanto, demonstram que o percentual de inadimplência no mercado de crédito no Brasil tende a ser mais elevado no segmento de pessoas físicas que no de pessoas jurídicas, sendo tanto essa relação quanto a participação das operações de crédito entre pessoas físicas e jurídicas bastante distinta entre os estados. Isso pode ser demonstrado por exemplo analisando-se os dados de inadimplência dos dados de abril de 2019 do Brasil e dos estados do RJ e SP.

Comparando-se os dados de inadimplência total do Brasil e dos estados RJ e SP (Tabela 2.1), observa-se que o estado do RJ apresenta uma inadimplência global inferior tanto a média do Brasil, quanto a do estado de SP. Entretanto, quando se analisa apenas a inadimplência das pessoas físicas, esse Quadro se inverte.

Isso se deve a dois fatores: (a) a maior participação das pessoas jurídicas no total de operações de crédito realizadas no estado do RJ vis a vis o observado em SP e na média do Brasil, e (b) a menor taxa da inadimplência das pessoas jurídicas nas operações de crédito no estado do RJ frente ao observado em SP e na média do Brasil.

¹¹ ANEEL (2010): “73. Complementarmente, para fins de comparação, pode ser verificada a evolução de perdas não técnicas em termos percentuais em relação ao mercado de baixa tensão da distribuidora, visto que é neste nível de tensão que se encontra a grande maioria das irregularidades que produzem perdas não técnicas.”

Tabela 2.1 - Inadimplência total no setor de crédito do Sistema Financeiro Nacional (SFN) – Abril de 2019

| | Pessoa Física | | Pessoa Jurídica | | Total | |
|--------|--------------------|------------|--------------------|------------|--------------------|------------|
| | % de inadimplência | % do total | % de inadimplência | % do total | % de inadimplência | % do total |
| Brasil | 3,36% | 56% | 2,56% | 44% | 3,01% | 100% |
| SP | 3,03% | 51% | 2,5% | 49% | 2,77% | 100% |
| RJ | 4,15% | 37% | 1,09% | 63% | 2,22% | 100% |

Fonte: Banco Central do Brasil (2019)

Assim, ao utilizar os dados de inadimplência global e não os dados de inadimplência apenas das pessoas físicas, a ANEEL acaba por distorcer o cálculo do IC.

2.4.2 Os problemas de especificação do modelo

A primeira crítica relativa a especificação dos modelos se refere a opção da ANEEL, pela estimação de um modelo de Dados em Painel com Efeitos Aleatórios, em detrimento de um modelo de Dados em Painel com Efeitos Fixo (Calili *et al.*, 2017). O modelo geral para dados em painel pode ser representado pela equação:

$$y_{it} = \beta_{0it} + \beta_{1it}x_{1it} + \dots + \beta_{nit}x_{kit} + \mu_i + e_{it} \quad (2.13)$$

Onde;

i denota os diferentes indivíduos;

t denota o período de tempo em análise;

β_0 é o intercepto da equação;

β_k denota os coeficientes de cada uma das k -ésima variável explicativa do modelo;

μ_i representa as heterogeneidades observadas entre diferentes indivíduos;

e_{it} representa o erro de estimação.

Há na literatura três tipos básicos de modelos que combinam dados de séries temporais e dados em corte transversal: (a) os Modelo de Dados Empilhado (*pooled*), (b) os Modelos de Dados em Painel com Efeitos Fixos, e (c) os Modelos de Dados em Painel com Efeitos Aleatórios. A opção entre esses modelos, no entanto deve se basear na observação de dois comportamentos específicos nas séries de dados utilizados: (a) a presença ou não de heterogeneidade entre as unidades de análise (ex. as distribuidoras) ao longo do tempo, e (b) a presença ou não de correlações entre essas eventuais heterogeneidades e as variáveis explicativa (regressores) do modelo (Wooldridge, 2010).

O modelo empilhado – ou *pooled* – pressupõe que a heterogeneidade (μ_i), de cada unidade i , não existe. Já os modelos de efeitos fixos e aleatórios assumem a existência da heterogeneidade. Porém, a diferença entre os dois modelos é que, enquanto o modelo de efeitos aleatórios assume que não existe correlação da heterogeneidade com nenhuma variável explicativa ($E(\mu_i x b) = 0$), o modelo de efeitos fixos assume tal correlação exista ($E(\mu_i x b) \neq 0$) (Greene, 2012).

A Figura 2.4 ilustra os modelos que melhor se adequam a cada uma das situações.

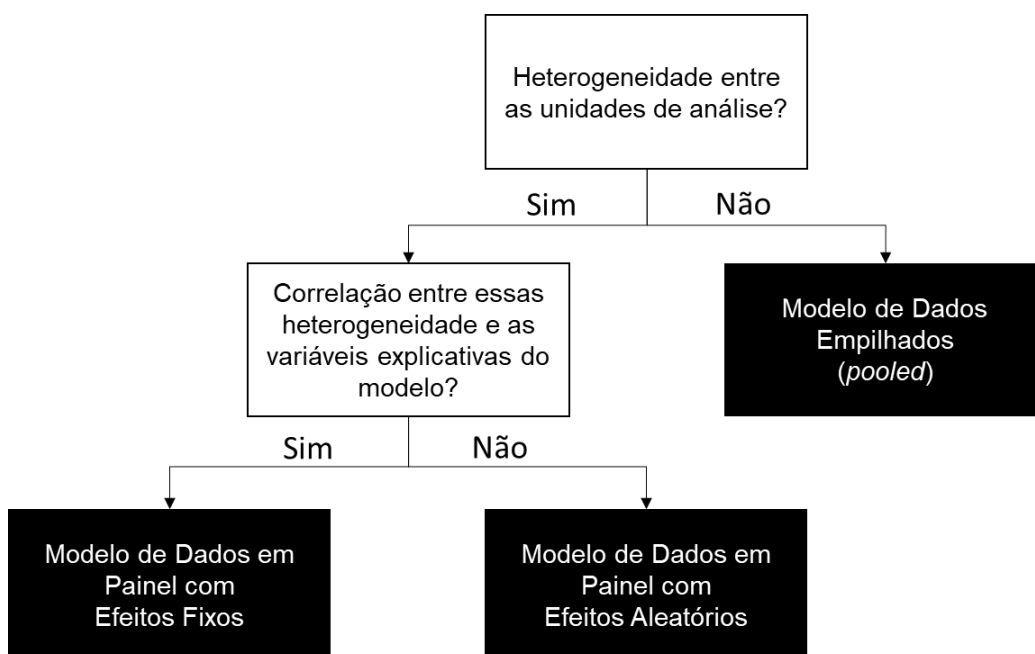


Figura 2.4 - Hierarquia de seleção entre os modelos

Fonte: Elaboração própria

Tanto a heterogeneidades entre as unidades de análise, quando a correlações entre essas eventuais heterogeneidades e as variáveis explicativa podem ser testadas a partir da aplicação de dois testes de hipóteses: o teste de Breusch-Pagan e Teste Hausman (Green, 2012).

O teste de hipótese proposto por Breusch-Pagan, baseado no multiplicador de Lagrange, busca avaliar a existência de heterogeneidades entre as unidades de análise (efeito painel). O teste define as seguintes hipóteses nula e alternativa:

$$\begin{cases} H_0: \mu_i = 0 \\ H_1: \mu_i \neq 0 \end{cases}$$

Onde:

μ_i é a heterogeneidade de cada unidade de análise i , invariante no tempo.

Já o teste de hipótese proposto por Hausman (Green, 2012) busca avaliar a existência de correlações entre essas eventuais heterogeneidades e as variáveis exógenas utilizadas no modelo. O teste defini as seguintes hipóteses nula e alternativa:

$$\begin{cases} H_0: E(\mu_i | X\beta) = 0 \\ H_1: E(\mu_i | X\beta) \neq 0 \end{cases}$$

Onde:

μ_i é as heterogeneidade de cada unidade de análise i , invariante no tempo.

$X\beta$ são as variáveis exógenas consideradas na estimação do modelo.

Conforme demonstra Calili *et al.* (2017) a opção da ANEEL de estimar um modelo de dados em painel com efeitos aleatórios parece equivocada. No referido trabalho os autores realizaram o teste de Breusch-Pagan e o teste Hausman para os modelos C, G e K atualmente propostos.

Os autores demonstram que pelos resultados do teste de Breusch-Pagan, há a rejeição da hipótese nula (H_0) para os três modelos, indicando que as empresas são diferentes e que a estimação deve ser conduzida por modelos de dados em painel e não pelo modelo clássico de Mínimos Quadrados Ordinários. Entretanto, pelos resultados do teste de Hausman, os três modelos adotados pela ANEEL (C, G e K)

deveriam ter sido estimados considerando efeitos fixos e não efeitos aleatórios como realizados pela ANEEL (Calili *et al.*, 2017).

A Tabela 2.2 apresenta os resultados dos referidos testes realizados por Calili *et al.*, 2017 para os três modelos propostos pela ANEEL.

Tabela 2.2 – Resultado dos Testes de Breusch-Pagan e Hausman aplicados aos modelos C, G e K da ANEEL

| Modelo | Teste de Breusch-Pagan χ^2 (p – valor) | Teste de Hausman χ^2 (p – valor) |
|---------------|--|--|
| C | 1861,308 (0,000) | 168,084 (0,000) |
| G | 2244,313 (0,000) | 142,673 (0,000) |
| K | 1928,675 (0,000) | 16,106 (0,000) |

Fonte: Calili *et al.* (2017)

A estimação dos modelos de dados em painel com efeitos aleatórios, quando o teste de Hausman indica que o correto seria com efeitos fixos, tende a gerar parâmetros inconsistentes e ineficientes, o que pode resultar em distorções relevantes no cálculo do Índice de Complexidade Socioeconômico (IC) da ANEEL.

A segunda crítica relativa a especificação dos modelos se refere à habilidade explicativa dos modelos econométricos estimados pela ANEEL e uma possível omissão de variáveis. Os três modelos estimados pela ANEEL (C, G e K) apresentam um R^2 ajustado bastante baixo (todos inferiores a 0,20), o que demonstra uma baixa capacidade dos modelos estimados de explicar o fenômeno das PNT. Valores de R^2 ajustado baixos podem indicar tanto baixa correlação das variáveis explicativas e o fenômeno que se pretende explicar, como também uma possível omissão de variáveis relevantes no modelo.

O problema é ainda mais relevante, considerando-se que a ANEEL, no cálculo do Índice de Complexidade Socioeconômico (IC), ignora a componente μ_i estimada (heterogeneidade ou especificidade da empresa).

A ANEEL ignora essa componente μ_i estimada com o argumento de que essa componente representaria em sua maior parte a ineficiência das distribuidoras. A componente μ_i busca a capturar as especificidades da empresa não explicados

pelas demais variáveis do modelo, incluindo eventuais ineficiências, mas também eventuais problemas decorrentes da omissão de variáveis relevantes no modelo. Como nos modelos C, G e K propostos pela ANEEL, há evidências fortes de que possa se estar incorrendo em um problema de omissão de variáveis relevantes para a explicação do problema, portanto, essa exclusão pode incutir distorções relevantes no cálculo do IC e consequentemente no processo de comparação entre as distribuidoras (*Benchmarking*)¹².

Essas críticas aos modelos adotados foram levantadas em diversas contribuições durante as Audiências Pública 040 de 2010 e 023 de 2014. Na Nota Técnica 271 de 2010 a ANEEL defendeu que a importância do R^2 seria reduzida devido a três razões básicas: (a) existência de variáveis não observadas, particularmente a ineficiência das empresas, o que naturalmente reduziria o R^2 ; (b) o fato de que elevados valores de R^2 não indicam necessariamente modelos “melhores” ou “melhores” especificados (risco de correlação espúria) e (c) as próprias limitações do R^2 como uma medida eficiente de ajuste dos modelos (ANEEL, 2010).

Uma outra crítica relevante, ainda sobre a especificação dos modelos, consiste na não utilização da variável tempo de forma explícita dentro do modelo. Calili *et al.* (2017) sugere por exemplo que, ao ignorar a variável tempo, a ANEEL não estaria considerando na definição das metas de PNT das distribuidoras a eficiência relativa das distribuidoras no combate as PNT ao longo do tempo. Os autores apresentam como exemplo uma comparação entre a evolução das PNT da Enel Rio e da Light onde, embora ambas as empresas permaneçam registrando elevados níveis de PNT ao longo de todo o período, o percentual de PNT da Enel Distribuição Rio vem decaindo ao longo dos anos, enquanto que o da Light, tem se mantido praticamente inalterado.

Ainda no que se refere à especificação dos modelos, uma crítica de considerável relevância, se refere ao não tratamento prévio dos *outliers*. Em qualquer modelo estatístico a não retirada de *outliers*, pode interferir no resultado, provocando vieses.

¹² Esse risco foi posteriormente reconhecido pela própria ANEEL na Nota Técnica nº 179/2018-SRM/ANEEL (ANEEL, 2018a) e parcialmente tratado pela Agência no decorrer do processo 48500.005123/2016-06 que tratou da Definição dos níveis regulatórios de perdas não técnicas de energia elétrica no sistema de distribuição da Enel-Rio no período de 2019 a 2022.

A última crítica relevante ainda sobre a especificação dos modelos refere-se as inúmeras etapas *ad hoc* da metodologia utilizada pela ANEEL na definição das PNT. Como pontua Calili *et al.* (2018), as metas de perdas não técnicas não são obtidas diretamente de um modelo estatístico-matemático, mas sim a partir de uma série de passos e etapas metodológicas para as quais há pouco ou nenhum embasamento teórico e rigor científico.

2.4.3

O problema da exequibilidade das metas regulatórias de Perdas Não Técnicas definidas pela ANEEL

O terceiro grupo de críticas a modelo atualmente empregado pela ANEEL na definição das metas regulatórias de PNT das distribuidoras, refere-se a efetividade e a eficácia das metas estabelecidas pela ANEEL. Como visto nesse capítulo, o uso de modelos de *Benchmarking* vem sendo utilizado pela ANEEL desde 2007. Entretanto, quando comparadas as metas percentuais de PNT regulatórias e os percentuais efetivamente realizados ao longo da última década, observa-se que as metas definidas pela ANEEL não foram capazes de induzir um ciclo de reduções de PNT gerais no país (Figura 2.5).

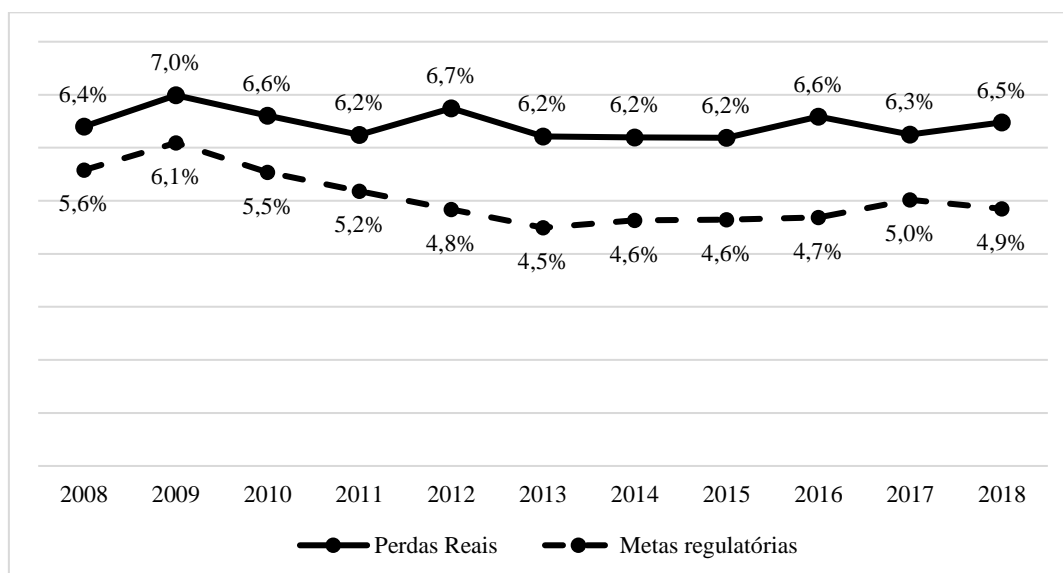


Figura 2.5 - Perdas Não Técnicas no Brasil vs. Metas Regulatórias (% da Energia Injetada Total)

Fonte: ANEEL (2019b)

Esse tema é ainda mais grave quando se analisa as distribuidoras individualmente. De 2008 a 2018, segundo dados da ANEL (2019), em média 68% das 54 principais distribuidoras do país foram incapazes de atingir as metas de perdas regulatórias definidas pela ANEEL (Figura 2.6).

A situação é ainda mais dramática se considerarmos apenas os resultados dos últimos 3 anos, quando 77% das distribuidoras foram incapazes de atingir as suas metas regulatórias de PNT.

A Figura 2.7 apresenta uma comparação entre a meta regulatória de PNT em percentual da energia injetada total e os valores realizados para as distribuidoras brasileiras.

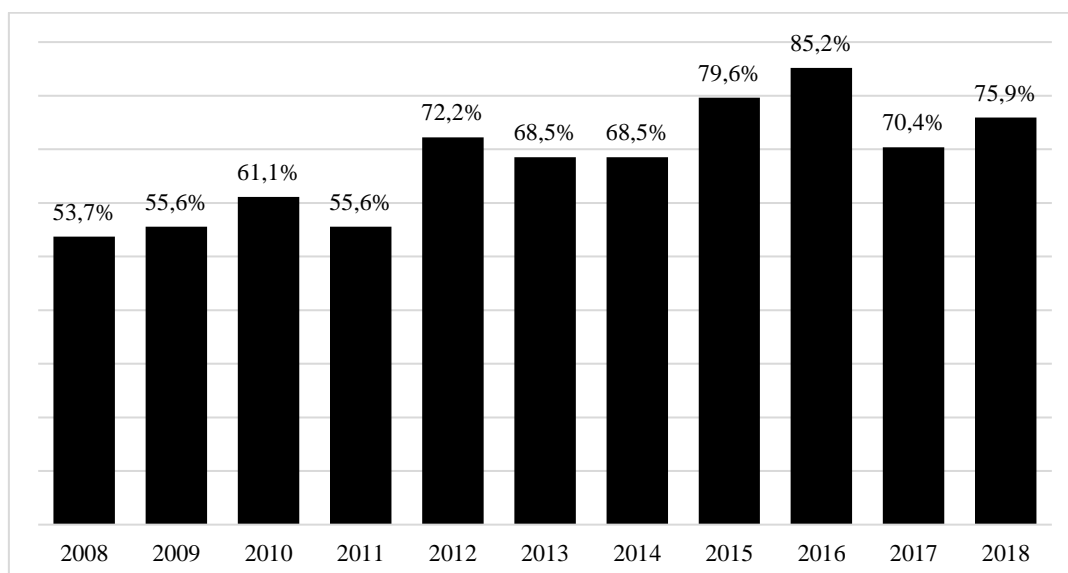


Figura 2.6 - Percentual das distribuidoras que não conseguiram atingir as metas de PNT regulatórias de 2008 a 2018

Fonte: ANEEL (2019b)

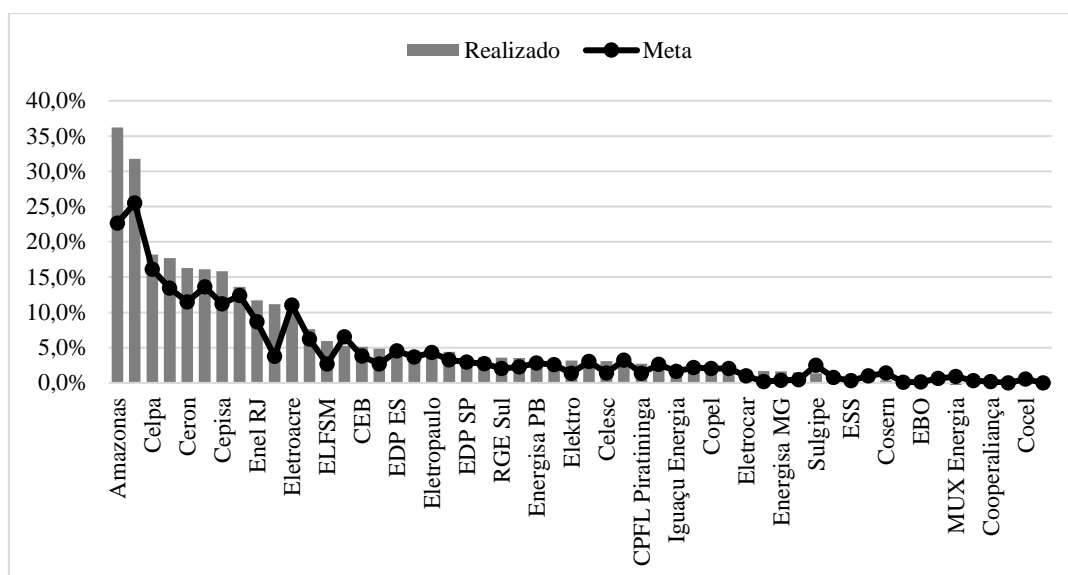


Figura 2.7 - Perdas Não Técnicas no Brasil – Realizado vs. Meta regulatória em 2018
(% da Energia Injetada Total)

Fonte: ANEEL (2019b)

2.5 Considerações sobre o capítulo

Como demonstrado nesse capítulo, a excessiva complexidade da metodologia atualmente empregada pela ANEEL na definição das metas de PNT regulatória, assim como a adoção de um conjunto grande critérios *ad hoc*, põem em risco um dos principais parâmetros que norteiam a regulação por incentivo: o estabelecimento de incentivos econômicos corretos para busca contínua pela melhoria da eficiência.

Os problemas apontados, em particular os problemas relativos à especificação do modelo, apresentam indícios claros de que há falhas no processo de comparação das PNT entre as distribuidoras no país que podem estar conduzindo tanto a um não reconhecimento tarifário das PNT de algumas distribuidoras quanto o sobre reconhecimento da PNT em outras.

A incapacidade dos modelos adotados pela ANEEL de induzir uma redução contínua das PNT técnicas na última década, além do grande número de distribuidoras que não tem conseguido adequar as suas PNT reais aos parâmetros regulatórios, reforça a necessidade de buscar metodologias alternativas para lidar com o tema.

Uma alternativa capaz de resolver parte dos problemas aqui apresentados no modelo da ANEEL, parece ser possível adotando-se os Modelos de Análise de Eficiência.

Nos capítulos 3 será feita uma breve revisão dos modelos de Análise de Eficiência e no capítulo 4 a proposição de uma metodologia alternativa para a definição das metas de PNT que busca superar alguns dos problemas do modelo regulatório atual e que foram apontados nessa seção.

3

Modelos de Análise de Eficiência e suas aplicações a definição de metas regulatórias para concessionárias de serviço público

A produção de um bem ou serviço é um processo no qual um conjunto inicial de insumos (X) é transformado em um conjunto de produtos (Y). O Conjunto de todos os processos de transformação tecnologicamente possíveis de X em Y é denominado Conjunto de Possibilidades de Produção (CPP) que pode ser definido como:

$$T(X, Y) = \{(X, Y) | \text{é viável produzir } Y \text{ a partir de } X\} \quad (3.1)$$

Um processo de produção é tecnicamente eficiente se utiliza a menor quantidade de insumos (x), que os demais processos, para produzir um mesmo nível de produto (y). O subconjunto da CPP, que representa a menor quantidade de insumos (x) que se pode consumir para gerar determinado volume de produto (y) é denominado de Fronteira de Possibilidades de Produção (FPP). A CPP e a FPP são ilustradas na Figura 3.1.

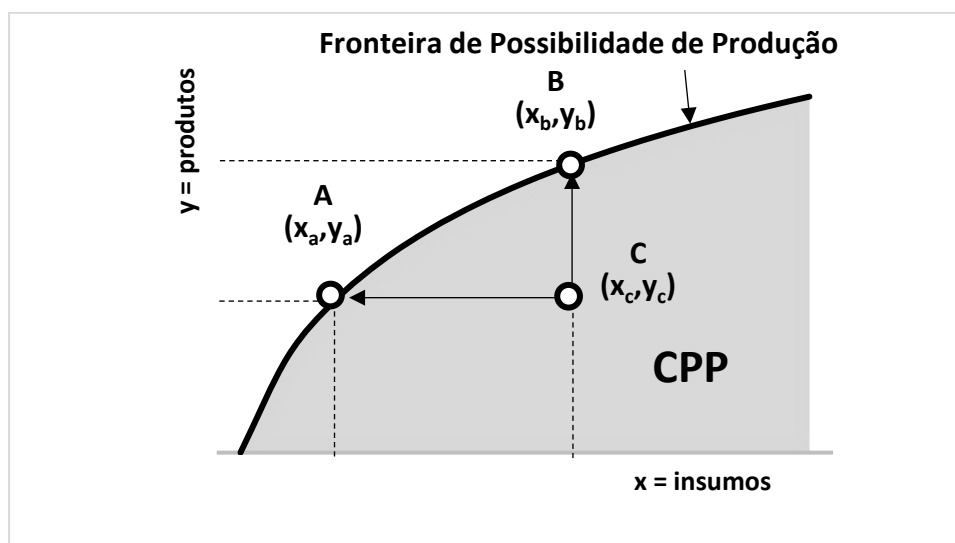


Figura 3.1 – Fronteira de Produção
Fonte: Elaboração própria

Na Figura 3.1 o ponto $(x,y) \in \text{CPP}$ indica uma maneira tecnologicamente viável de transformar uma quantidade x de insumo em uma quantidade y de produto, entretanto apenas aquelas que encontram-se localizados na Fronteira de Possibilidades de Produção (FPP) podem ser classificados como tecnicamente eficiente.

Por exemplo, suponha-se que na Figura 3.1 os pontos A, B e C representem diferentes produtores, ou DMU's (*Decision Makings Units*) como comumente são denominados na literatura sobre Modelos de Análise de Eficiência (Bogetoft & Otto, 2011; Behr, 2015). Nesse caso, é possível afirmar que as DMUs A e B são tecnicamente eficientes, dados que para produzir os volumes produzido por A e B (y_a e y_b), não existe nenhuma maneira tecnologicamente viável de se consumir um volume menor de insumo que aqueles consumidos pelas duas DMUs (x_a e x_b). Entretanto, o mesmo não pode ser dito sobre C, já que A consegue obter o mesmo nível de produção de C ($y_a = y_c$) com uma quantidade de insumo inferior a de C ($x_a < x_c$)¹³.

Assim, a FPP fornece implicitamente um primeiro padrão contra o qual a performance de uma determinada DMU pode ser avaliada. No exemplo da Figura 3.1 a eficiência de C poderia ser definida a partir de uma comparação direta com A como exposto na equação 3.2.

$$E_C = \frac{\frac{y_C}{x_C}}{\frac{y_A}{x_A}} = \frac{x_A y_C}{x_C y_A}$$

Sendo $y_A = y_C$

(3.2)

$$E_C = \frac{x_A}{x_C}$$

¹³ Analogamente poderia se dizer que C não pode ser considerado eficiente porque consumindo o mesmo volume de insumos ($x_b = x_c$), B consegue produzir um volume de produto superior a C ($y_b > y_c$).

Um processo de produção ser tecnicamente eficiente não garante, entretanto, que ele seja necessariamente eficiente do ponto de vista econômico. Isso porque o conjunto utilizado de insumos pode ser mais caro que um conjunto de insumos alternativo resultando em *trade off* entre custo de produção vs. volume produzido inferior ao obtido por um processo de produção alternativo. Para ser economicamente eficiente, portanto, um processo de produção deve ser além de tecnicamente eficiente, ser também eficiente do ponto de vista alocativo (Kumbhakar e Lovell, 2000; Behr, 2015).

Nesse sentido uma DMU racional buscaria produzir um conjunto de produtos $y \in R_+^n$ não a partir do menor uso dos insumos $x \in R_+^n$, mas sim da combinação de preços $w \in R_+^n$ e $x \in R_+^n$ que resultasse no menor custo de produção possível (Kumbhakar & Lovell, 2000; Zanini, 2004; Behr, 2015).

A Fronteira de Custo (FC) é uma representação alternativa da FPP e indica o mínimo custo necessário para produzir um conjunto de produtos (Y), dados o nível de produção, os preços dos insumos (W) e a tecnologia existente (T). A Fronteira de Custos é ilustrada na Figura 3.2.

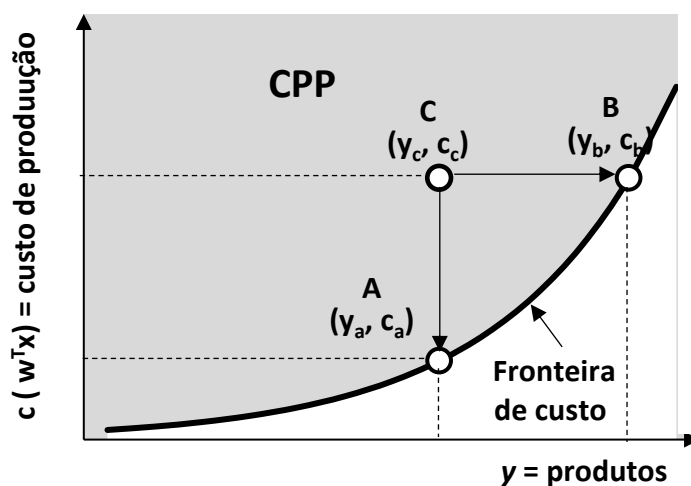


Figura 3.2 – Fronteira de Custo
Fonte: Elaboração própria

De forma análoga ao que foi dito em relação a Figura 3.1, na Figura 3.2 qualquer ponto $(y, c) \in \text{CPP}$ indica uma maneira tecnologicamente viável de produzir y unidades do produto a um custo c , entretanto apenas aquelas que encontram-se localizados na Fronteira de Custo podem ser classificados como

economicamente eficiente. Na Figura 3.2 é possível afirmar que as DMUs A e B são economicamente eficientes, já que não existe nenhuma maneira tecnologicamente viável de se produzir a um custo menor, os volumes produzidos por A e B (y_a e y_b). Entretanto, o mesmo não pode ser dito sobre C, já que A consegue obter o mesmo nível de produção de C ($y_a = y_c$) com um custo de produção bem inferior ao de C ($c_a < c_c$) e B consegue obter um volume de produção maior com um custo produção igual a C.

Assim o critério de eficiência definido na equação 3.2 pode ser reescrito como sendo:

$$EE_C = \frac{\frac{y_C}{c_C}}{\frac{y_A}{c_A}} = \frac{c_A y_C}{c_C y_A} \quad (3.3)$$

Sendo $y_A = y_C$

$$EE_C = \frac{c_A}{c_C}$$

A Fronteira de Produção e a Fronteira de Custo funcionam assim como uma referência contra o qual a performance de uma determinada DMU pode ser avaliada. Isto sugere que o grau de eficiência relativa de uma DMU pode ser avaliado por meio da distância entre ela e a Função Fronteira, conforme a métrica radial proposta por Debreu (Kumbhakar & Lovell, 2000; Zanini, 2004; Farrell, 1957). 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.

A eficiência técnica de uma determinada DMU “k” que produz um vetor de produtos y , a partir de um vetor de insumos x pode ser definida como:

$$TE_k = \frac{y_{ef}}{y_k} \quad \text{dado } x_{ef} = x_k \quad (3.4)$$

Onde:

TE_k : Eficiência Técnica da DMU “k”

y_{ef} : nível máximo de produto que se pode obter a partir de um determinado volume de insumos x_k

x_k : Volume de insumos utilizados pela DMU “ k ” par produzir um determinado volume de produto y_k

De forma análoga, admitindo que os preços dos insumos (vetor W) sejam conhecidos, o custo de produção a eficiência econômica do produtor é:

$$EE_k = \frac{w_{ef} \times x_{ef}}{w_k \times x_k} \quad \text{dado } y_{ef} = y_k \quad (3.5)$$

Onde:

EE_k : Eficiência Econômica da DMU “ k ”

y_{ef} : nível máximo de produto que se pode obter a partir de um determinado volume de insumos x_k

x_k : Volume de insumos utilizados pela DMU “ k ” para produzir um determinado volume de produto y_k

x_{ef} : Volume de insumos utilizados pela DMU na fronteira para produzir um determinado volume de produto y_{ef}

w_{ef} : Preço dos insumos utilizados pela DMU na fronteira para produzir um determinado volume de produto y_{ef}

w_k : Preço dos insumos utilizados pela DMU “ k ” para produzir um determinado volume de produto y_{ef}

Na prática, no entanto, nem a Fronteiras de Produção (FPP), nem a Fronteira de Custo de uma determinada indústria são conhecidas *a priori*. Assim, é necessário estimar a Fronteiras de Produção (FPP) ou Fronteira de Custo antes de se calcular a eficiência de uma determinada DMU ou de um determinado conjunto de DMUs. Na seção a seguir, serão apresentados os principais métodos atualmente propostos na literatura para estimação da FPP e da Fronteira de Custo.

3.1

Modelos de Análise de Eficiência

Os modelos de Análise de Eficiência buscam estimar a eficiência de uma determinada DMU, ou de um determinado conjunto de DMUs, aplicando métricas análogas as da equação 3.5 usando como referência Fronteiras de Produção ou de Custo estimadas.

Em geral, todos os modelos comportam tanto uma abordagem orientada para a estimação da fronteira de produção (modelo orientado ao produto), quanto para a estimação da fronteira de custo (modelo orientado aos insumos). A principal divergência entre os diferentes modelos está na técnica empregada para estimar a fronteira.

Segundo Agrell e Bogetoft (2016) e Sickles *et al.* (2018), os três principais métodos de Análise de Eficiência proposto na literatura são Análise de Envoltória de Dados (DEA), Mínimos Quadrados Corrigidos (COLS) e Análise de Fronteira Estocástica (SFA).

O DEA consiste em um método não paramétrico formulado como um problema de programação linear (PPL) onde busca-se estimar a função de produção ou de custos de uma determinada indústria a partir de dados sobre insumos e produtos utilizados/produzidos por um conjunto de DMUs dessa indústria (Souza, *et al.*, 2010; Bogetoft & Otto, 2011; Agrell & Bogetoft, 2016).

Já COLS e SFA são ambos métodos paramétricos derivados modelos tradicionais de regressão linear, que se diferenciam na forma pela qual estimam a fronteira de produção ou custo (Kumbhakar e Lovell, 2000; Zanini, 2004; Sickles *et al.*, 2018).

No método dos mínimos quadrados corrigidos, a equação da fronteira é um modelo de regressão linear estimado por mínimo quadrados, onde o intercepto da equação é “ajustado” de forma que os desvios de todas as DMUs frente a fronteira estimada sejam sempre positivos. Já no modelo de Fronteira Estocástica, a fronteira é um modelo de regressão com erro assimétrico, cujos coeficientes são estimados por máxima verossimilhança (Kumbhakar e Lovell, 2000; Zanini, 2004; Sickles *et al.*, 2018).

3.2

Análise de Envoltória de Dados (DEA)

O *Data Envelopment Analysis* ou Análise de Envoltória de Dados (DEA) é um método não-paramétrico utilizado para medir a eficiência relativa de um conjunto de unidades tomadoras de decisão (DMUs) similares. Desenvolvido originalmente por Charnes *et al.* (1978), como uma extensão do trabalho de Farrell (1957), o DEA utiliza programação matemática para estimar a Fronteira de Produção ou de Custo de uma determinada indústria a partir de um conjunto predefinido de insumos e produtos.

No DEA, cada unidade é comparada com uma unidade virtual obtida por meio de uma combinação linear de todas as unidades eficiente. O modelo de otimização busca os pesos para os insumos e para os produtos que maximizem o escore de eficiência da unidade sob análise. Diferentemente dos métodos paramétricos, a identificação da Fronteira de Produção se dá através da resolução de um modelo de programação matemática.

Uma hipótese fundamental do método é a convexidade da Função de Produção, o que implica assumir o ‘livre descarte’ (*free disposal*) dos insumos.

Os modelos DEA podem ser classificados em dois grupos em função da forma como o problema da eficiência é formulado:

- **Modelos orientados para a maximização de produtos:** a eficiência da DMU é calculada através da máxima expansão dos produtos, dada uma quantidade de insumos utilizados.
- **Modelos orientados para a minimização de insumo:** a eficiência da DMU é calculada através da minimização do uso de insumos, dada uma quantidade de produtos.

A seguir são apresentadas as duas formulações clássicas dos modelos DEA.

3.2.1

O modelo CRS (*Constant Returns to Scale*)

O CRS (*Constant Returns to Scale*), proposto originalmente por Charnes *et al.* (1978), trabalha com a hipótese de retornos constantes de escala. Nesse modelo os autores definem a eficiência como sendo a razão entre a soma ponderada dos produtos e a soma ponderada dos insumos selecionados.

No modelo proposto, um conjunto individual de pesos (ou ponderadores) é estimado para cada DMU de forma a maximização a sua eficiência individual. A única condição imposta pelo modelo é que todas as DMU's tenham uma eficiência inferior ou igual a 1.

A formulação matemática dessas condições é apresentada no modelo básico CRS orientado ao insumo a seguir:

$$Max Ef_0 = \frac{\sum_{j=1}^s u_j y_{j0}}{\sum_{i=1}^r v_i x_{i0}} \quad (3.6)$$

Sujeito a:

$$\frac{\sum_{j=1}^s u_j y_{jk}}{\sum_{i=1}^r v_i x_{ik}} \leq 1, k = 1, \dots, n$$

$$u_j \text{ e } v_i \geq 0 \forall j, i$$

Onde:

Ef_0 : eficiência da DMU₀;

u_j e v_i : pesos de produtos e insumos respectivamente a serem estimados;

y_{jk} e x_{ik} : insumos “i” e produtos j da DMU_k;

y_{j0} e x_{i0} : insumos “i” e produtos j da DMU₀.

Este modelo (3.7) pode ser transformado em um problema de programação linear (PPL), fazendo o denominador da função objetivo igual a um.

A partir do modelo primal é possível desenvolver o modelo dual, ou Modelo Envelope que, pelo teorema da dualidade forte apresentará o mesmo valor ótimo para a função objetivo (Lachtermacher, 2016).

Assim o modelo Dual do modelo da equação 3.6 pode ser reescrito como sendo:

$$Min \theta \quad (3.7)$$

Sujeito a:

$$\theta x_{i0} - \sum_{k=1}^n x_{ik} \lambda_k \geq 0, \quad i = 1, \dots, r$$

$$-y_{j0} + \sum_{k=1}^n y_{jk} \lambda_k \geq 0, \quad j = 1, \dots, s$$

$$\lambda_k \geq 0 \forall k$$

Onde:

θ : eficiência;

u_j e v_i : pesos de produtos e insumos respectivamente;

y_{jk} e x_{ik} : insumos “i” e produtos j da DMU_k;

y_{j0} e x_{i0} : insumos “i” e produtos j da DMU₀;

λ_k : k-ésima coordenada da DMU₀ em uma base formada pelas DMU’s de referência.

O modelo dual busca valores para λ_k que minimizem θ , sendo λ_k a contribuição da DMU_k na formação do alvo da DMU₀. As DMUs com λ_k não nulo são os *benchmarks*.

3.2.2

O modelo VRS (*Variable Return Scale*)

O modelo VRS (*Variable Return Scale*), proposto por Banker *et al.* (1984), adicionaram uma restrição de combinação convexa [$\lambda_1 + \lambda_2 + \dots + \lambda_N = 1$ no PPL] ao CRS, criando um modelo que contempla a hipótese de rendimentos variáveis de escala. Assim, o modelo pressupõe que as DMU’s avaliadas apresentem retornos variáveis de escala (retornos crescentes ou decrescentes). Em outras palavras, o VRS considera que um acréscimo no input poderá promover um acréscimo no output, não necessariamente proporcional.

O modelo VRS surgiu da divisão da eficiência do modelo CCR em duas componentes: eficiência técnica e a eficiência de escala. A medida de eficiência técnica, resultante do modelo VRS, identifica a correta utilização dos recursos dada a escala de operação da DMU sob análise. A eficiência de escala, por outro lado, identifica a distância da DMU em análise em relação a uma DMU fictícia, que opera com a escala mais eficiente.

A formulação matemática dessas condições é apresentada no modelo básico VRS orientado ao insumo a seguir:

$$Max h_0 = \sum_{j=1}^s u_j y_{j0} - u_* \quad (3.8)$$

Sujeito a:

$$\begin{aligned} \sum_{i=1}^r v_i x_{ik} &= 1 \\ \sum_{j=1}^s u_j y_{jk} - \sum_{i=1}^r v_i x_{ik} - u_i &\leq 0, k = 1, 2, \dots, n \\ u_j \text{ e } v_i &\geq 0 \forall j, i \\ u_* &\in R \end{aligned}$$

Onde:

h_0 : eficiência;

u_j e v_i : pesos de produtos e insumos respectivamente;

y_{jk} e x_{ik} : insumos “i” e produtos j da DMU_k;

y_{j0} e x_{i0} : insumos “i” e produtos j da DMU₀;

Assim como no modelo CRS, o modelo Dual do modelo da equação 3.8 pode ser rescrito como sendo:

$$Min \theta \quad (3.9)$$

Sujeito a:

$$\begin{aligned} \theta x_{i0} - \sum_{k=1}^n x_{ik} \lambda_k &\geq 0, \quad i = 1, \dots, r \\ -y_{j0} + \sum_{k=1}^n y_{jk} \lambda_k &\geq 0, \quad j = 1, \dots, s \\ \sum_{k=1}^n \lambda_k &= 1 \\ \lambda_k &\geq 0 \forall k \end{aligned}$$

Onde:

θ : eficiência;

u_j e v_i : pesos de produtos e insumos respectivamente;

y_{jk} e x_{ik} : insumos “i” e produtos j da DMU_k;

y_{j0} e x_{i0} : insumos “i” e produtos j da DMU₀;

λ_k : k-ésima coordenada da DMU₀ em uma base formada pelas DMU’s de referência.

3.2.3

Vantagens e desvantagens do uso do DEA e seu uso pelos reguladores de serviço público no mundo

A principal vantagem do DEA é a sua simplicidade de implementação, quando comparado as metodologias paramétricas de análise de fronteira, particularmente o SFA. Por se tratar de uma técnica não-paramétrica, no DEA não é necessário fazer nenhuma suposição da distribuição das variáveis. Além disso, pode-se gerar um indicador único de eficiência considerando diversos insumos e produtos, sem que haja a necessidade de predefinir uma forma funcional para Função de Produção ou Custo. Ademais, como a FPP ou a Fronteira de Custos no DEA é formada a partir das DMUs observadas, os *benchmankings* são mais claramente identificados o que torna o processo de aprendizado mais direto. (Bogetoft & Otto, 2011; Agrell & Bogetoft, 2016; Sickles *et al.*, 2018)

As principais desvantagens do método estão associadas a fato de por ser um método não paramétrico não é possível testar diretamente a qualidade do modelo e o impacto relativo da introdução ou exclusão de variáveis e/ou DMUs. Outro ponto destacado por Bogetoft & Otto (2011) é a relativa sensibilidade dos modelos DEA a escolha dos insumos e produtos considerados no processo de modelagem. No modelo DEA todo o eventual erro de especificação, como a omissão de variáveis, é considerado como ineficiência.

Como afirma Agrell & Bogetoft (2016), o DEA é atualmente o modelo de Análise de Eficiência mais empregada por reguladores de serviços públicos em todo o mundo, particularmente pelos reguladores de serviços de energia elétrica e gás (Harney e Pollitt, 2009, 2011; Bogetoft & Otto ,2011; Agrell & Bogetoft, 2016; Sickles *et al.* 2018).

Um estudo realizado por Harney & Pollitt em 2009 e atualizado por Agrell & Bogetoft (2016), revela que de uma amostra de 24 reguladores de diferentes países da Europa, 83% utilizavam modelo de *benchmarking* baseados na metodologia DEA nos seus processos tarifário para o setor de transmissão de energia elétrica, e 40% também utilizavam essa metodologia nos processos tarifários associados ao setor de distribuição de energia elétrica e gás.

No Brasil o DEA vem sendo empregado pela Agência Nacional de Energia Elétrica (ANEEL) desde de 2010 como principal instrumento para a definição do custo operacional regulatório das distribuidoras de energia elétrica (ANEEL, 2018b).

Apesar da larga adoção do DEA por parte dos reguladores, o seu emprego, tem se limitado, em geral, temas como a avaliação dos custos operacionais e dos investimentos desses agentes (Harney e Pollitt, 2009, 2011; Bogetoft & Otto, 2011; Agrell & Bogetoft, 2016; Sickles *et al.* 2018).

3.3 Mínimos Quadrados Corrigidos (COLS)

O método dos mínimos quadrados corrigidos (COLS – *Corrected OLS*) é um método paramétrico de estimação da eficiência. A equação da Fronteira de Produção ou de Custo é obtida a partir de um modelo de regressão linear estimado por mínimo quadrados ordinários, onde o intercepto é corrigido para garantir que os resíduos sejam todos não positivos (Olson *et al.*, 1980).

Como propõem Sickles *et al.* (2018) o método pode ser subdividido em três etapas. O primeiro passo consiste em estimar a função de produção (assumindo por uma forma funcional do tipo Cobb-Douglas) a partir dos dados de insumo e produto disponíveis para as k DMUs:

$$\ln(y_i) = \beta_0 + \beta_1 \ln(x_{i1}) + \dots + \beta_N \ln(x_{iN}) + \varepsilon_i, \quad i = 1, \dots, k \quad (3.10)$$

Onde:

y_i : produto da DMU _{i}

x_{iN} : representa os N insumo utilizados pela DMU_i
 β_N : Os parâmetros da função de produção a serem estimados;
 ε_i : o erro de estimação da DMU_i.

O segundo passo consiste em ajustar o intercepto, conforme a equação 3.11, transformando a equação de regressão estimada na função de produção em uma Fronteira de Produção:

$$\hat{\beta}_0^{COLS} = \hat{\beta}_0 + \max_i\{\hat{\varepsilon}_i\}, \quad i = 1, \dots, k \quad (3.11)$$

Onde:

$\hat{\beta}_0^{COLS}$: intercepto corrigido;
 $\hat{\beta}_0$: representa o intercepto estimado no primeiro passo;
 $\max_i\{\hat{\varepsilon}_i\}$: maior erro estimado para uma das k DMUs.

O terceiro passo consiste em calcular a eficiência técnica de cada uma das DMUs como sendo:

$$TE_i = \varepsilon_i^{COLS} = \max_i\{\hat{\varepsilon}_i\} - \hat{\varepsilon}_i, \quad i = 1, \dots, k \quad (3.12)$$

Onde:

TE_i : Eficiência técnica da DMU_i;
 $\max_i\{\hat{\varepsilon}_i\}$: maior erro estimado para uma das k DMUs;
 $\hat{\varepsilon}_i$: erro estimado para uma da DMU_i.

3.3.1

Vantagens e desvantagens do uso do COLS e seu uso pelos reguladores de serviço público no mundo

A principal vantagem do COLS é a simplicidade de implementação. Na comparação com os modelos DEA, o COLS tem a vantagem de possibilitar a

mensuração da significância estatística de cada um dos insumos e produtos considerados no processo de modelagem, o que garante uma maior robustez ao processo de seleção dessas variáveis. Por outro, é necessário assumir algumas hipóteses sobre a distribuição das variáveis utilizadas e sobre a forma funcional da função de produção, o que não são necessárias nos modelos DEA.

Ademais como apenas o intercepto é corrigido, a Fronteira de Produção estimada pelo COLS é paralela à linha de tendência central estimada pelo OLS, uma hipótese pouco plausível em para muitas indústrias e muito restrita (Kumbhakar & Lovell, 2003; Zanini, 2004).

Como afirma Agrell & Bogetoft (2016), o COLS vem sendo cada vez menos empregado por reguladores de serviços públicos em todo o mundo (Harney e Pollitt, 2009; Bogetoft & Otto, 2011; Agrell & Bogetoft, 2016; Sickles *et al.*, 2018).

Um estudo realizado por Harney & Pollitt em 2009 e atualizado por Agrell & Bogetoft (2016), revela que de uma amostra de 24 reguladores de diferentes países da Europa, apenas 12% utilizavam modelo de *benchmarking* baseados na metodologia COLS nos seus processos tarifário associados ao setor de distribuição de energia elétrica e gás e nenhum utilizavam essa metodologia nos processos tarifários associados ao setor de transmissão de energia elétrica.

No Brasil o COLS não é atualmente empregado pela Agência Nacional de Energia Elétrica (ANEEL) em nenhum de seus processos tarifários (ANEEL, 2014; Harney e Pollitt, 2009).

3.4 Modelo de Fronteira Estocástica (SFA)

Nos modelos de fronteira determinística, como nos modelos DEA, qualquer desvio em relação à fronteira de produção é atribuído à ineficiência técnica da DMU sob análise. Tais modelos ignoram o fato de que a produção pode ser afetada por choques aleatórios fora do controle das DMU, como por exemplo, greves, eventos climáticos entre outros.

Os modelos de Fronteira Estocásticas (*Stochastic Frontier Analysis*) são modelos paramétricos introduzidos na literatura simultaneamente, em 1977, por Aigner *et al.* (ALS) e Meeusen & van den Broeck (MB) (Kumbhakar & Lovell, 2000). Esses modelos tentam superar essa limitação incorporando no processo de estimação um termo de erro composto, cuja soma é o resultado de duas

componentes: um erro aleatório puro e um termo de ineficiência estocástico (Behr, 2015).

A principal virtude desses modelos está no fato deles reconhecerem que os desvios em relação à Fronteira de Produção podem ser causados por choques aleatórios fora do controle das DMUs (Zanini, 2004; Sickles *et al.*, 2018).

Seguindo a notação de Sickles *et al.* (2018), os modelos de Fronteira Estocásticas podem descritos com na equação 3.13 e 3.14.

$$y_i = f(x_i, \beta) e^{\varepsilon_i} \quad (3.13)$$

$$\varepsilon_i = v_i + u_i \quad (3.14)$$

Onde:

y_i : é a produção da DMU_i;

$f(x_i, \beta) e^{\varepsilon_i}$: é a função de produção estocástica;

ε_i : é uma variável aleatória irrestrita em sinal;

v_i : é uma variável aleatória irrestrita;

e^{-u_i} : é ineficiência da DMU_i

A fronteira estocástica é composta de duas partes : a fronteira determinística $f(x_i, \beta)$, comum a todos os produtores, e um termo e^{ε_i} , cuja finalidade é capturar o efeito de choques aleatórios que afeta especificamente a i-ésima DMU. Considerando que ε_i pode ser decomposto em $\varepsilon_i = v_i + u_i$, a ineficiência pode ser descrita como:

$$e^{-u_i} = \frac{y_i}{f(x_i, \beta) e^{v_i}} \quad (3.15)$$

Onde:

e^{-u_i} : é ineficiência da DMU_i;

$f(x_i, \beta) e^{v_i}$: é a função de produção estocástica sem a ineficiência da DMU_i;

v_i : é uma variável aleatória irrestrita.

Admitindo que $f(x_i, \beta)$ seja linear nos logaritmos, as equação 3.13 e 3.14 pode ser rescrita como:

$$\ln(y_i) = \ln(f(x_i, \beta)) + v_i - u_i \quad (3.16)$$

Admitindo-se:

$v_i \sim i.i.d N(0, \sigma_v^2)$ e mede o erro da estimação
 $u_i \sim i.i.d N^+(0, \sigma_u^2)$ (*Half-Normal*¹⁴) e representa a ineficiência da DMU_i
 u_i e v_i são independentes entre si
 u_i e v_i não são correlacionadas com as variáveis explicativas do modelo

Assim, a função densidade do erro composto é dada por:

$$f(\varepsilon) = \int_0^\infty f(u, \varepsilon) du = \frac{2}{\sqrt{2\pi}\sigma} \left[1 - \Phi\left(\frac{-\varepsilon\lambda}{\sigma}\right) \right] \cdot e^{-\frac{\varepsilon^2}{2\sigma^2}} = \frac{2}{\sigma} \phi\left(\frac{\varepsilon}{\sigma}\right) \Phi\left(\frac{\varepsilon\lambda}{\sigma}\right) \quad (3.17)$$

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 $N(0,1)$.

O logaritmo da função de verossimilhança dado uma amostra com n DMU é escrito como:

¹⁴ Seja Y definido como o valor absoluto da variável aleatória $X \sim (N, \sigma^2)$. Se Y segue uma distribuição *Half-Normal*, então:

$$\text{Função de densidade de } Y: \quad f_Y(y) = \begin{cases} \frac{2}{\sqrt{2\pi}\sigma} \exp\left(-\frac{y^2}{2\sigma^2}\right) & \text{se } y \geq 0 \\ 0 & \text{se } y < 0 \end{cases}$$

$$\text{Esperança de } Y: \quad E(Y) = \sqrt{\frac{2}{\pi}} \sigma$$

$$\text{Variança de } Y: \quad Var(Y) = \frac{\pi - 2}{\pi} \sigma^2$$

$$\ln L = \text{constante} - n \ln \sigma + \sum_{i=1}^n \ln \Phi \left(\frac{\lambda \varepsilon_i}{\sigma} \right) - \frac{1}{2\sigma^2} \sum_{i=1}^n \varepsilon_i^2 \quad (3.18)$$

As estimativas de máxima verossimilhança correspondem aos valores de σ_u , σ_v e β que maximizam a equação 3.20. Estas estimativas são assintoticamente consistentes (Kumbhakar e Lovell, 2000).

Vale ressaltar que o logaritmo da verossimilhança foi parametrizado em termos de $\sigma^2 = \sigma_u^2 + \sigma_v^2$ e $\lambda = \sigma_u/\sigma_v$. Tal parametrização permite uma melhor interpretação do modelo, por exemplo, a estatística $\gamma = \sigma_u^2/(\sigma_u^2 + \sigma_v^2) = \lambda^2/(1 + \lambda^2) \in [0,1]$ e permite avaliar qual das duas componentes do erro composto é predominante.

Quando $\gamma=0$ a ineficiência é inexistente, pois σ_v^2 representa a maior parte do erro. Desta forma, os desvios entre a fronteira e as DMUs são representados, em sua maior parte, por ruídos aleatórios. Contrariamente, quando $\gamma=1$, o erro é dominado por σ_u^2 , logo, os desvios em relação à fronteira demonstram ineficiência. Assim, por meio do teste de razão de máxima verossimilhança pode-se testar a hipótese $H_0 \gamma=0$ contra a hipótese alternativa $H_1 \gamma \neq 0$ para avaliar se a ineficiência está presente no conjunto de dados analisados. Como γ pertence ao intervalo $[0,1]$, a distribuição da estatística teste pode ser aproximada por uma χ_1^2 .

Conforme indicado na equação 3.20, para estimar a eficiência de cada DMU é preciso ter uma estimativa de u_i , a componente do erro que captura o efeito da ineficiência. Esta estimativa pode ser obtida a partir dos resíduos $\hat{\varepsilon}_i$, pois $\varepsilon_i = v_i + u_i$. Como u_i e v_i são independentes entre si, a distribuição conjunta dessas variáveis é o produto das respectivas densidades marginais:

$$f(u, v) = \frac{2}{2\pi\sigma_u\sigma_v} e^{\left(\frac{-u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2}\right)} \quad (3.19)$$

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

$$f(u, \varepsilon) = \frac{2}{2\pi\sigma_u\sigma_v} e^{\left(\frac{-u^2}{2\sigma_u^2} - \frac{(\varepsilon-u)^2}{2\sigma_v^2}\right)} \quad (3.20)$$

Para obter a distribuição marginal de ε_i , basta integrar a função de densidade conjunta (Equação 3.22) em função de u_i :

$$f(\varepsilon) = \int_0^{\infty} f(u, \varepsilon) du = \frac{2}{\sqrt{2\pi} \sigma} \left[1 - \Phi \left(-\frac{\varepsilon \lambda}{\sigma} \right) \right] e^{\frac{-\varepsilon^2}{2\sigma^2}} = \frac{2}{\sigma} \phi \left(\frac{\varepsilon}{\sigma} \right) \Phi \left(\frac{\varepsilon \lambda}{\sigma} \right) \quad (3.21)$$

em que ϕ e Φ são a densidade acumulada e a distribuição acumulada de uma normal padrão respectivamente.

Por meio da densidade conjunta em (equação 3.20) e a densidade de ε_i em (equação 3.21) consegue-se definir a densidade de probabilidade condicional de u_i dado ε_i :

$$f(u|\varepsilon) = \frac{f(u, \varepsilon)}{f(\varepsilon)} = \frac{1}{\sqrt{2\pi} \sigma_*} e^{\left[\frac{-(u-\mu_*)^2}{2\sigma_*^2} \right]} \left/ \left[1 - \Phi \left(\frac{-\mu_*}{\sigma_*} \right) \right] \right. \Rightarrow u_i | \varepsilon_i \sim N^+(\mu_*, \sigma_*^2) \quad (3.22)$$

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

Uma estimativa pontual para u_i pode ser a média ou a moda da distribuição condicionada em (equação 3.22). Em qualquer um dos casos, a estimativa \hat{u}_i define a estimativa da eficiência $\hat{\theta}_i = \exp(-\hat{u}_i)$. Por exemplo, a medida de eficiência proposta por Battese & Coelli (Kumbhakar & Lovell, 2000), obtida a partir da fronteira estocástica, é dada por:

$$\theta_i = E(e^{-u_i} | \varepsilon_i) = \left[\frac{1 - \Phi \left(\sigma_* - \frac{\mu_{*i}}{\sigma_*} \right)}{1 - \Phi \left(-\frac{\mu_{*i}}{\sigma_*} \right)} \right] e^{-\mu_{*i} + \frac{1}{2} \sigma_*^2} \quad (3.23)$$

Os resultados apresentados baseiam-se na premissa de que a componente aleatória μ_{*i} tenha distribuição *Half-Normal*, mas outras premissas para a densidade de probabilidade de μ_{*i} podem ser admitidas. Sickles *et al.* (2018) cita algumas outras distribuições consideradas para o termo da ineficiência (μ_{*i}) presente na literatura de SFA como a exponencial, normal truncado, gama e duplamente normal truncado (Stevenson, 1980; Greene 1980; Qian & Sickles, 2008; Almanidis & Sickles, 2012; Almanidis *et al.*, 2014).

Outro ponto relevante se refere a forma funcional da função de produção. Os resultados apresentados baseiam-se na premissa de uma função de produção do tipo Cobb-Douglas.

Os modelos SFA exigem a adoção *a priori* de uma premissa sobre a forma funcional da Fronteira de Produção. As formas funcionais mais usualmente adotadas são as de Cobb-Douglas (CD) e a Translog (TL) (Coelli *et al.*, 2005; Behr, 2015). A função Translog apresenta como vantagem o fato de ser mais flexível, porém a um custo alto em termos de parcimônia. Na função Translog a um conjunto maior de parâmetros a serem estimados que aumenta de forma exponencial em função do número de insumos considerados. Isso além de acarretar problemas como multicolinearidade, conduz a necessidade de uso de amostras bastante grandes no processo de estimação.

A função Cobb-Douglas, por outro lado, é menos flexível que a Translog, porém é mais parcimoniosa. Ao exigir a estimação de um número menor de parâmetros, a função Cobb-Douglas se adapta melhor a amostras menores (Coelli *et al.*, 2005).

3.4.1

Vantagens e desvantagens do uso do SFA e seu uso pelos reguladores de serviço público no mundo

A principal vantagem do SFA em relação ao DEA está na possibilitar de testar estatisticamente tanto a especificação do modelo quanto suas conclusões sobre a eficiência das DMUs. A significância dos insumos e produtos considerados no processo de modelagem, assim como seus impactos, pode ser avaliados de forma mais criteriosa e direta.

Além disso o modelo SFA, por decompor o resíduos em duas componentes ($\varepsilon_i = v_i + u_i$), consegue lidar melhor com eventuais choques aleatórios sendo menos sensível a o problema de omissões de variáveis, uma vez que apenas parte do resíduo é considerada como ineficiência (Kumbhakar & Lovell, 2000; Zanini, 2004; Bogetoft & Otto, 2011; Agrell & Bogetoft, 2016; Sickles *et al.*, 2018).

Por outro, no SFA, assim como no COLS é necessário fazer suposição da distribuição das variáveis, em particular da componente u_i , e sobre a forma função de produção o que torna naturalmente o modelo mais complexo de implementar.

(Kumbhakar & Lovell, 2000; Zanini, 2004; Bogetoft & Otto, 2011; Agrell & Bogetoft, 2016; Sickles *et al.*, 2018).

Como afirma Agrell & Bogetoft (2016), embora os resultados obtidos em um modelo bem especificado através do SFA e do DEA tendam a convergir, o SFA ainda é pouco empregado por reguladores de serviços públicos em todo o mundo (Harney e Pollitt, 2009, 2011; Bogetoft & Otto, 2011; Agrell & Bogetoft, 2016; Sickles *et al.*, 2018).

Um estudo realizado por Harney & Pollitt em 2009 e atualizado por Agrell & Bogetoft (2016), revela que de uma amostra de 24 reguladores de diferentes países da Europa, apenas 12% utilizavam modelo de *benchmarking* baseados na metodologia SFA nos seus processos tarifário associados ao setor de distribuição de energia elétrica e gás e nenhum utilizavam essa metodologia nos processos tarifários associados ao setor de transmissão de energia elétrica. Harney & Pollitt em um estudo anterior (2009) também apontaram o uso do SFA pelos reguladores de dois países fora da Europa, México e Nova Zelândia.

No Brasil o SFA não é empregado pela Agência Nacional de Energia Elétrica (ANEEL) em nenhum de seus processos tarifários (ANEEL, 2014; Harney e Pollitt, 2009) embora alguns trabalhos, como Calili *et al.* (2017), tenham sugerido a sua aplicação ao caso brasileiro.

3.5

Considerações sobre o capítulo

Como discutido nesse capítulo os modelos de Análise de Eficiência demonstram ser uma ferramenta mais adequada para comparação e mensurar a eficiência relativas dos agentes ou DMUs. Por essa razão, esses modelos vêm sendo largamente utilizados por reguladores de todo o mundo na definição dos níveis ótimos de custo operacionais e investimento (Agrell & Bogetoft, 2016).

Dos três principais modelos de Análise de Fronteira mais frequentes na literatura, o DEA é o mais empregado pelo Reguladores no mundo. Isso se deve principalmente à sua simplicidade de implementação. Por se tratar de uma técnica não-paramétrica, no DEA não é necessário fazer nenhuma suposição da distribuição das variáveis. Além disso, pode-se gerar um indicador único de eficiência considerando diversos insumos e produtos, sem que haja a necessidade de predefinir uma forma funcional da Função de Produção ou Custo.

Apesar disso, o DEA traz uma desvantagem bastante relevante para a sua aplicação ao problema de PNT. Por ser um método não paramétrico não é possível testar diretamente nem a qualidade do modelo, nem o impacto relativo da introdução ou exclusão de variáveis e/ou DMUs nos resultados dos modelos.

Além disso, no modelo DEA todo o eventual erro de especificação, como a omissão de variáveis, é interpretado como ineficiência. Como foi exposto no capítulo 2, há no tema de PNT um elevado risco de omissão de variáveis, seja porque as próprias relações de causalidade de PNT essas variáveis ainda não estão bem estabelecidas, seja pelas limitações das próprias variáveis disponíveis.

Como foi demonstrado nesse capítulo a principal vantagem do SFA em relação ao DEA está na possibilitar de testar estatisticamente tanto a especificação do modelo quanto suas conclusões sobre a eficiência das DMUs. A significância dos insumos e produtos considerados no processo de modelagem, assim como seus impactos, pode ser assim avaliados de forma mais criteriosa e direta.

Por fim, o modelo SFA, por decompor os resíduos em duas componentes ($\varepsilon_i = v_i + u_i$), consegue lidar melhor com eventuais choques aleatórios e com o problema de omissões de variáveis. Nesse sentido, para o problema de PNT, o SFA parece se apresentar como o método mais adequado.

No capítulo 4 será apresentado um modelo para definição de metas regulatórias de PNT a partir dos modelos de Fronteira Estocástica (SFA) e no capítulo 5 será apresentada uma demonstração da aplicabilidade desse modelo ao caso brasileiro.

4

Desenvolvimento de um modelo alternativo para a definição das metas regulatórias de Perdas Não Técnicas (PNT)

Com exposto no capítulo 3, a fronteira de custo indica o custo mínimo necessário para produzir uma quantidade de produto, dados os preços e a tecnologia dos insumos. Assim, os produtores ineficientes estão localizados acima da fronteira, enquanto os produtores eficientes estão na Fronteira de Custo, como mostra a Figura 4.1:

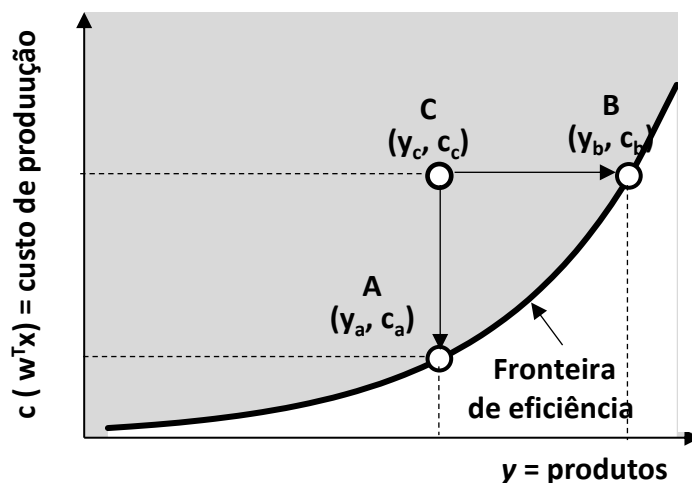


Figura 4.1 – Fronteira de Custo
Fonte: Elaboração própria

A Fronteira de Custo é assim uma referência com a qual o desempenho dos diferentes agentes de um determinado setor industrial (DMUs) podem ser comparados. Nesse trabalho, as DMUs são as concessionárias de distribuição de energia elétrica e o custo a ser minimizado é o percentual de PNT.

O problema das PNT de fato não se enquadra como problema clássico da teoria da produção. Não há insumos que podem ser livremente selecionados pelas distribuidoras. Os insumos nesse problema são características em certa medida estanques representadas pelas características socioeconômicas das áreas onde essas

distribuidoras de energia atuam e que condicionam a sua capacidade combater as PNTs.

Apesar disso, o problema continua sendo um problema de minimização de custos e é possível adaptar o conceito clássico dos modelos de Análise de Eficiência baseados nas fronteiras de custo para o problema de PNT. Assim como nos problemas clássicos de minimização de custo, é possível comparar diferentes distribuidoras a partir da estimação de uma fronteira de eficiência que indica o menor nível PNT que uma determinada distribuidora consegue atingir dada um determinado contexto socioeconômico.

Nesse capítulo é apresentado um modelo alternativo para definição das metas regulatórias de Perdas Não Técnicas (PNT) a partir de um modelo de Fronteira Estocástica (SFA) com dados em painel.

4.1

Modelo de Fronteira Estocástica de custo para modelo de dados cross section¹⁵

Para uma DMU produzir uma quantidade y de produto a partir de uma quantidade x de insumos com preços unitários w , a eficiência θ é a razão do custo potencial mínimo definido pela fronteira de eficiência $c(y, w)$ e o custo de produção $E = wx \geq c(y, w)$:

$$\theta = c(y, w)/E \quad (4.1)$$

Reorganizando os termos da equação (4.1), obtém-se que $E = c(y, w) \theta^{-1}$. Aplicando-se a transformação logarítmica a esse termo obtém-se:

$$\log(E) = \log(c(y, w)) - \log(\theta) \quad (4.2)$$

Supondo que a função de produção seja linear nos logaritmos das variáveis (por exemplo, assumindo uma especificação Cobb-Douglas para a função de

¹⁵ Esse capítulo se baseou no estudo realizado por Leite et al. (2020).

produção) e tornando $\varepsilon = -\text{Log}(\theta)$, que é um termo aleatório, tem-se a seguinte equação de regressão linear para cada DMU i em um conjunto com n DMU:

$$\log(E_i) = \beta_0 + \beta_1 \cdot \log(y_i) + \beta_2 \cdot \log(w_i) + \varepsilon_i \quad \forall I = 1, n \quad (4.3)$$

Na Equação (4.3), o termo aleatório ε_i expressa o desvio entre o custo verificado ($\log(E_i)$) e o custo mínimo definido pela fronteira de eficiência ($\beta_0 + \beta_1 \cdot \log(y_i) + \beta_2 \cdot \log(w_i)$). Vale ressaltar que, diferentemente do modelo de regressão linear convencional, o termo aleatório na Equação (4.3) tem uma média diferente de zero ($E(\varepsilon_i) > 0$) e não é normalmente distribuído (Kumbhakar *et al.*, 2015).

Como apresentado no capítulo 3, um avanço dos modelos de Fronteira Estocástica (SFA) reside no reconhecimento de que os desvios da fronteira podem se originar tanto da ineficiência das DMUs quanto de choques aleatórios não gerenciáveis pelas DMUs. Para acomodar as duas fontes de desvio, o SFA decompõe o termo aleatório em dois componentes ($\varepsilon_i = v_i + u_i$). Assim, a equação 4.3 pode ser rescrita como:

$$\log(E_i) = \beta_0 + \beta_1 \cdot \log(y_i) + \beta_2 \cdot \log(w_i) + v_i + u_i \quad \forall I = 1, n \quad (4.4)$$

Na Equação (4.4), v_i é um componente aleatório, com distribuição Normal e média zero que representa os choques aleatórios não gerenciáveis pela i -ésima DMU. Já u_i , representa um componente aleatório não-negativo que captura o grau de ineficiência da i -ésima DMU. A soma dos componentes aleatórios define o erro composto $\varepsilon_i = v_i + u_i$ que é positivamente assimétrico.

Assim, a medida de eficiência da i -ésima DMU pode ser definida como sendo:

$$\theta_i = \exp(-u_i) \quad (4.5)$$

A equação (4.4) pode ser estimada por máxima verossimilhança. Como as variáveis aleatórias u e v são não observáveis é necessário a adoção de hipóteses

sobre as suas respectivas distribuições de probabilidade. Geralmente, as hipóteses mais comumente adotadas são uma distribuição normal para v e uma distribuição *Half-Normal* para u , embora como observado no capítulo 3, outras distribuições sejam admitidas.

Essa especificação é conhecida na literatura como o modelo SFA normal / *Half-Normal* (Von Hirschhausen *et al.*, 2006) e pode ser descrito a partir das seguintes condições:

1. $v_i \sim i. i. d N(0, \sigma_v^2)$
2. $u_i \sim i. i. d N^+(0, \sigma_u^2)$ (*Half-Normal*)
3. u_i e v_i são independentes
4. u_i e v_i não estão correlacionados com as variáveis explicativas

Conforme indicado, como admite-se u_i e v_i são independentes, a distribuição conjunta dessas variáveis pode ser obtida pelo produto de suas respectivas funções densidades marginais:

$$f(u, v) = \frac{2}{2\pi\sigma_u\sigma_v} e^{\left(\frac{-u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2}\right)} \quad (4.6)$$

Dado que $\varepsilon_i = v_i + u_i$, a distribuição conjunta de u_i e ε_i pode ser descrita como sendo:

$$f(u, \varepsilon) = \frac{2}{2\pi\sigma_u\sigma_v} e^{\left(\frac{-u^2}{2\sigma_u^2} - \frac{(\varepsilon-u)^2}{2\sigma_v^2}\right)} \quad (4.7)$$

A partir da integração da função de densidade descrita pela equação 4.7 é possível obter-se a distribuição marginal de ε_i :

$$f(\varepsilon) = \int_0^\infty f(u, \varepsilon) du = \frac{2}{\sqrt{2\pi}\sigma} \left[1 - \Phi\left(\frac{-\varepsilon\lambda}{\sigma}\right) \right] \cdot e^{\frac{\varepsilon^2}{2\sigma^2}} = \frac{2}{\sigma} \phi\left(\frac{\varepsilon}{\sigma}\right) \Phi\left(\frac{\varepsilon\lambda}{\sigma}\right) \quad (4.8)$$

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 cumulativa de um $N(0,1)$.

Com base na equação (4.8), é possível calcular o logaritmo da função de probabilidade para uma amostra com n DMUs:

$$\log(L) = \text{constante} - n\log(\sigma) + \sum_{i=1}^n \log\left(\Phi\left(\frac{\lambda\varepsilon_i}{\sigma}\right)\right) - \frac{1}{2\sigma^2} \sum_{i=1}^n \varepsilon_i^2 \quad (4.9)$$

As estimativas de máxima verossimilhança nesse caso correspondem aos valores de σ_u , σ_v e β que maximizam a equação (4.9). Como demonstrado por Coelli *et al.* (2005) e Kumbhakar *et al.*, (2015) essas estimativas são assintoticamente consistentes.

Deve-se notar que o logaritmo da probabilidade foi parametrizado em termos de $\sigma^2 = \sigma_u^2 + \sigma_v^2$ e $\lambda = \sigma_u / \sigma_v$. Essa parametrização permite uma melhor interpretação do modelo. Por exemplo, a estatística $\gamma = \sigma_u^2 / (\sigma_u^2 + \sigma_v^2) = \lambda^2 / (1 + \lambda^2) \in [0,1]$ permite avaliar qual dos dois componentes do erro composto é predominante.

No modelo, $\gamma = 0$, a ineficiência é inexistente, pois σ_v^2 representa a maior parte do erro. Em outras palavras, os desvios entre a fronteira e as DMUs são ruídos aleatórios. Por outro lado, quando $\gamma = 1$, o erro é dominado por σ_u^2 , portanto os desvios da fronteira e as DMUs são devidos oriundo exclusivamente da ineficiência dessas DMUs.

Assim, através do teste da razão de máxima verossimilhança, a hipótese $H_0 \gamma = 0$ pode ser testada contra a hipótese alternativa $H_1 \gamma \neq 0$ para avaliar se a ineficiência está presente no conjunto de dados analisados. No caso do modelo de SFA *Half-Normal*, o mesmo adotado neste trabalho, a distribuição da estatística do teste pode ser aproximada por χ_1^2 , ou seja, uma distribuição qui-quadrado com 1 grau de liberdade (Coelli *et al.*, 2005).

Conforme indicado na equação (4.5), para estimar a eficiência de cada DMU, é necessário ter uma estimativa de u_i , o componente de erro que captura o efeito da ineficiência. Esta estimativa pode ser obtida a partir dos resíduos $\varepsilon_i = v_i + u_i$ da função de densidade conjunta dos parâmetros de u_i e ε_i descrita na equação (4.7) e

da função de densidade de ε_i descrita na equação (4.8). Assim pode-se definir a função densidade de probabilidade condicional da u_i dada ε_i como sendo:

$$f(u|\varepsilon) = \frac{f(u,\varepsilon)}{f(\varepsilon)} = \frac{1}{\sqrt{2\pi}\sigma_*} e^{\left[\frac{-(u-\mu_*)^2}{2\sigma_*^2}\right]} \Big/ \left[1 - \Phi\left(\frac{-\mu_*}{\sigma_*}\right)\right] \Rightarrow u_i|\varepsilon_i \sim N^+(\mu_*, \sigma_*^2) \quad (4.10)$$

Onde

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

Uma estimativa pontual \hat{u}_i para u_i pode ser a média ou a moda da função distribuição condicional descrita na equação (4.10). Em qualquer um dos casos, a estimativa de eficiência é igual a $\exp(-\hat{u}_i)$. Por exemplo, o valor esperado da probabilidade de densidade condicional $f(u_i|\varepsilon_i)$ (Coelli *et al.*, 2005; Kumbhakar *et al.*, 2015; Jondrow *et al.*, 1982) é uma estimativa pontual de u_i :

$$\begin{aligned} \hat{u}_i &= E(u_i|\varepsilon_i) = \mu_{*i} + \sigma_* \left[\frac{\phi(-\mu_{*i}/\sigma_*)}{1 - \Phi(-\mu_{*i}/\sigma_*)} \right] \\ &= \sigma_* \left[\frac{\phi(\lambda\varepsilon_i/\sigma)}{1 - \Phi(-\lambda\varepsilon_i/\sigma)} + \frac{\lambda\varepsilon_i}{\sigma} \right] \end{aligned} \quad (4.11)$$

Uma alternativa a 4.11 é a medida de eficiência proposta por Battese & Coelli (Coelli *et al.*, 2005) que é descrita como:

$$\theta_i = E(e^{-u_i}|\varepsilon_i) = \left[\frac{1 - \Phi\left(\sigma_* - \frac{\mu_{*i}}{\sigma_*}\right)}{1 - \Phi\left(-\frac{\mu_{*i}}{\sigma_*}\right)} \right] e^{-\mu_{*i} + \frac{1}{2}\sigma_*^2} \quad (4.12)$$

4.2

Modelo de fronteira estocástica de custo para modelo de dados em painel

Como pontuado por Simões *et al.* (2020), um avanço importante introduzido pela ANEEL no processo de definição das metas regulatórias de PNT foi a especificação de um modelo econométrico de dados em painel para a definição do Índice de Complexidade Socioeconômica (IC).

Os modelos de dados em painel trabalham com um conjunto mais amplo de observações o que, a princípio, permite a obtenção de estimadores mais eficientes para os parâmetros do modelo (Coelli *et al.*, 2005; Kumbhakar, 1990). No caso particular das PNT, a adoção de dados em painel é fundamental para incorporar na análise a evolução temporal da PNT e da evolução individual de cada uma das distribuidoras (DMUs) em relação a PNT.

A especificação de um modelo Análise de Fronteira Estocástica (SFA) com dados em painel é semelhante à especificação para o modelo de dados em corte transversal descrito na equação 4.4, adicionando-se apenas a indexação das variáveis no tempo de acordo com a equação (4.13):

$$\log(y_{it}) = \beta_0 + \beta_1 \log(x_{1t}) + \dots + \beta_k \log(x_{kt}) + v_{it} + u_{it}, \quad (4.13)$$

$$\forall i = 1, n, \forall t = 1, T$$

Sendo u_{it} e v_{it} mutuamente independentes e ambos independentes das variáveis explicativas x_{it} , os parâmetros do modelo de dados do painel podem ser estimados da mesma maneira que os parâmetros para o modelo de dados da seção transversal.

Entretanto, diferentemente dos modelos com dados em seção transversal, no modelo Análise de Fronteira Estocástica (SFA) com dados em painel é fundamental a incorporação de alguma estrutura para capturar a evolução temporal do termo u_{it} na Equação (4.13), por exemplo, assumindo que $u_{it} = f(t) \cdot u_i$, onde $f(t)$ é uma função que determina como a eficiência evolui ao longo do tempo (Coelli *et al.*, 2005; Kumbhakar, 1990; Battese & Coelli, 1992).

Battese & Coelli (1992) propuseram $f(t) = \exp[\eta(t - T)]$, onde η é um parâmetro a ser estimado. Nos modelos de dados em painel, o teste da razão de verossimilhança e o teste z podem ser aplicados para avaliar as hipóteses de efeitos

de eficiência invariantes no tempo $H_0: \eta = 0$, ou seja, o efeito de ineficiência é constante ao longo do tempo (Coelli *et al.*, 2005; Behr, 2015). Esta especificação para $f(t)$ implica que a ordem de classificação dos escores de eficiência para as empresas permaneça inalterada ao longo do tempo (Coelli *et al.*, 2005).

Além disso, na estimativa do modelo de dados em painel, os termos u_{it} podem ser tratados como efeitos fixos ou aleatórios, sendo a última opção recomendada por Kumbhakar (1990) e Battese & Coelli (1992).

4.3

Proposição de um modelo alternativo para a definição das metas regulatórias de Perdas Não Técnicas (PNT)

Se a PNT são custo a serem minimizados, o grau de eficiência relativa de uma determinada distribuidoras de energia elétrica pode ser medido pela distância entre a respectiva distribuidora e Fronteira de Custo. Entretanto, diferentemente da Fronteira de Custos tradicional, a Função de Custo aqui seria dada por uma relação entre o nível de PNT e a Complexidade Socioeconômica. Assim, como apresentado na Figura 4.2, a fronteira representaria o menor nível de PNT possível a ser alcançada, dada uma determinada complexidade socioeconômica representada por uma ou mais variáveis.

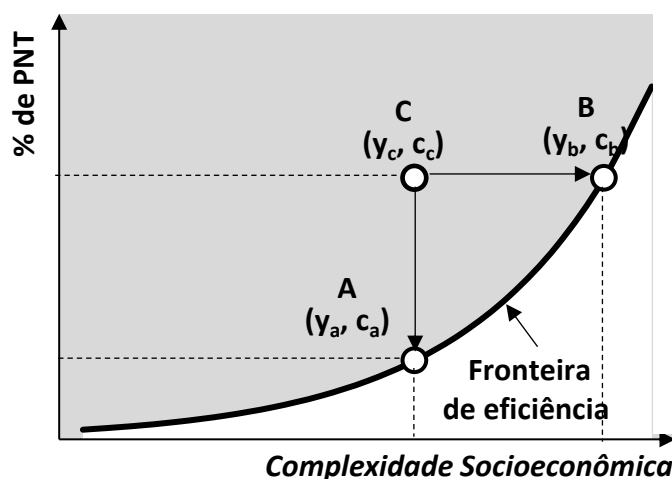


Figura 4.2 – Fronteira de Custos de PNT
Fonte: Elaboração própria

4.3.1

Definição das metas regulatórias de PNT

Para a definição das metas regulatórias de PNT, o primeiro passo seria medir a eficiência relativa das distribuidoras no combate a PNT. Uma forma de mensurar essas eficiências seria a especificação de um modelo Análise de Fronteira Estocástica (SFA) com dados em painel para eficiência. Assumindo a forma funcional de Cobb-Douglas para a função de custo, o modelo poderia ser descrito como sendo:

$$\log(PPNT_{it}) = \beta_0 + \beta_1 \log(x_{1it}) + \dots + \beta_k \log(x_{kit}) + v_{it} + u_{it}, \quad (4.14)$$

$$\forall i = 1, n, \forall t = 1, T$$

Onde:

$PPNT_{it}$: é o percentual de PNT sobre o mercado de BT da distribuidora “i” no período “t”;

x_{kit} : a variável socioeconômica “k” da distribuidora “i” no período “t”;

v_{it} : é uma variável aleatória e onde $v_i \sim i.i.d N(0, \sigma_v^2)$;

u_{it} : é o termo que representa a ineficiência e onde $u_i \sim i.i.d N^+(0, \sigma_u^2)$ (*Half-Normal*).

Admitindo-se que:

$$v_i \sim i.i.d N(0, \sigma_v^2)$$

$$u_i \sim i.i.d N^+(0, \sigma_u^2) \text{ (Half-Normal)}$$

u_i e v_i são independentes

u_i e v_i não estão correlacionados com as variáveis explicativas

Como o percentual de perdas não técnicas (PPNT) é geralmente um número no intervalo [0,1] a distribuição de probabilidade do logaritmo dessa variável pode não ser compatível com a suposição de assimetria positiva para o erro composto assumido nos modelos de Fronteira Estocástica de Custos. Para superar esse problema, o $\log(PPNT)$ deve ser substituído por $\log(1 + PPNT)$.

Uma vez estimado o modelo descrito na equação 4.14, a eficiência da i -ésima distribuidora no período “t” pode ser definida obtida por:

$$\theta_{it} = \exp(-u_{it}) \quad (4.15)$$

Assim a meta regulatória de PPNT seria definido como sendo:

$$PPNTmeta_{i0} = PPNT_{i0} \times \theta_{i0} \quad (4.16)$$

Onde:

$PPNTmeta_i$: meta regulatória de PNT da distribuidora “i” no período “0” em percentual do mercado BT;

θ_{i0} : eficiência da distribuidora “i” no período “0”;

$PPNT_{i0}$: PNT realizado pela distribuidora “i” no período “0”; em percentual do mercado de BT.

Ao invés de escolher um modelo específico e tomar decisões com base apenas no modelo escolhido, pode-se considerar o *ensemble* de um conjunto modelos selecionados, uma abordagem capaz de aproveitar a complementariedade entre os modelos (Page, 2008). Essa abordagem é especialmente valida no caso do tema de PNT da complexidade do fenômeno.

4.3.2

Definição de uma trajetória de redução das PNT

Uma vez estabelecida a $PPNTmeta_i$ é preciso estabelecer um período de tempo de transição para que a distribuidora consiga ajustar o seu patamar atual de PNT a meta regulatória definida na equação 4.15. Esse período de transição poderia ser definido, pelo agente regulador, por exemplo, como base na duração do ciclo de revisão tarifária de quatro ou cinco anos ($\Delta = 4$ ou 5).

Assim a equação 4.16 poderia ser rescrita como sendo:

$$PPNTMeta_t = PPNT_0(\theta^{t/\Delta}) \quad (4.17)$$

Onde:

$PPNTMeta_t$: é a meta de PNT da distribuidora “i” para o ano “t” em percentual do mercado de BT;

$PPNT_{i0}$: PNT realizado pela distribuidora “i” no período “0” em percentual do mercado de BT;

θ_{i0} : eficiência da distribuidora “i” no período “0”

Δi : prazo (em anos) para que a distribuidora “i” alcance a meta regulatória de PNT estabelecida no período “0”

4.3.3

Definição de limites para as metas regulatórias de PNT

Um ponto relevante da adoção de um critério de eficiência como base para definição de metas regulatórias é a definição de qual o nível de exigência deve ser adotar pelo Regulador nesse processo. Há um natural *trade off* entre a maior ou menor flexibilidade do regulador nesse processo e a manutenção dos incentivos econômicos a melhoria continua por parte dos agentes. No modelo proposto nas seções anteriores apenas as distribuidoras com $\theta_0 = 1$ teriam a totalidade das suas PNT reconhecidas ao final do período de transição. Distribuidoras com níveis de PNT muito baixos poderiam ter parte da suas PNT não reconhecidas ou terem que buscar reduções inexecutáveis na pratica.

Há uma série de fatores que corroboram uma análise mais pormenorizada desse processo. O primeiro deles é que o que os modelos de Analise de Eficiência fazem é estimam da Fronteira de Custos, e essa estimativa traz em si algum grau de incerteza. Não é possível, por exemplo, afirmar com absoluta certeza que uma distribuidora com $\theta_0 = 0,95$ não tenha de verdade um parâmetro $\theta_0 = 1$. A segunda incerteza é em relação as variáveis socioeconômicas utilizadas no processo de estimação. Como explorado no capítulo 2, essas variáveis têm diversas limitações que podem gerar distorções no processo de análise da eficiência das distribuidoras.

Há duas formas de adequar proposta apresentada nas seções anteriores para a definição da meta regulatória de PNT e a trajetória de redução. O primeiro deles é estabelecendo valores mínimos de PNT a partir dos quais as metas de PNT regulatórias não seriam mais reduzidas. O desafio nesse caso é como definir esses parâmetros.

A segunda alternativa seria considerar ajustar a equação 4.17 de forma:

$$PPNTMeta_t = PPNT_0 \left[\left(\frac{\theta_{i0}}{\theta_{ref}} \right)^{t/\Delta} \right] \quad (4.18)$$

Onde:

$PPNTMeta_t$: meta de PNT da Distribuidora “ i ” para o ano “ t ”

$PPNT_{i0}$: percentual PNT realizado sobre o mercado de BT da distribuidora “ i ” no período “0”

θ_{i0} : eficiência da distribuidora “ i ” no período “0”

θ_{ref} : nível de eficiência de referência estabelecido pelo regulador

Δt : prazo (em anos) para que a distribuidora “ i ” alcance a meta regulatória de PNT estabelecida no período “0”

Nessa segunda alternativa, se o Regulador adota-se por exemplo $\theta_{ref} = 0,9$ todas as distribuidoras com $\theta = 0,9$ teriam a totalidade da suas PNT reconhecidas enquanto nas distribuidoras com $\theta \geq 0,9$ as PNT reconhecidas seriam superiores aos níveis corrente, i.e. as distribuidoras mais eficientes seriam beneficiadas.

4.4

Considerações sobre o capítulo

Nesse capítulo foi proposto um modelo para definição de metas regulatórias de PNT partindo-se de um modelo SFA com dados em painel. Foi proposto também um conjunto de métricas auxiliares para a sua aplicação nos processos tarifários, como uma metodologia para definição de uma trajetória anual de redução de perdas e a possibilidade da adoção de limites adicionais para as metas regulatórias de PNT.

O modelo proposto tem como vantagens não só tornar o processo de definição das metas regulatórias mais simples e direto, evitando uma série de etapa *ad hoc* adotadas pelas ANEEL no modelo atual, como também torna o processo mais robusto em relação a seleção ao processo de seleção/omissão de variáveis.

O capítulo 5 apresenta uma aplicação desse modelo às distribuidoras brasileiras, comparando os seus resultados ao modelo atualmente aplicado pela ANEEL como forma de demonstrar não só a viabilidade da aplicação do modelo proposto, como também demonstrar as suas vantagens em relação ao modelo atualmente adotado pela ANEEL.

5

Aplicação de um modelo alternativo para definição das metas regulatórias de Perdas Não Técnicas (PNT) às distribuidoras de energia elétricas brasileiras

O modelo proposto no capítulo 4 foi aplicado as distribuidoras brasileiras como forma de demonstrar não só a viabilidade de sua aplicação como também as suas vantagens sobre o modelo atualmente aplicado pela ANEEL.

5.1

Base de dados e variáveis selecionadas

O modelo proposto no capítulo 4 foi estimado para o caso brasileiro a partir da base de dados de PNT disponibilizada pela ANEEL para a Consulta Pública nº 029 de 2020 (ANEEL, 2020b). Foram utilizados nesse processo de estimação os dados de perdas não técnicas (PNT) divididos pelo mercado de baixa tensão de cada uma das distribuidoras para o período de 2007 a 2017.

Foram selecionadas 10 variáveis socioeconômicas das 22 disponibilizada pela ANEEL na referida consulta pública (ANEEL, 2020b) (Quadro 5.1). A seleção das variáveis tomou como base o estudo de Takamine (2020) e a seleção de variáveis realizada pela ANEEL na Nota técnica 046/2020.

Takamine (2020) realizou um estudo exaustivo a partir do ajuste de modelos 1.097.789 de modelos de regressão linear para dados em painel, resultante do total de combinações das 22 variáveis disponibilizadas pela ANEEL (2020b) e avaliou painéis capacidade de cada modelo em estimar as PNT das distribuidoras.

Tomando por base as 11 variáveis previamente selecionadas por Takamine (2020), bem como as 7 selecionadas pela ANEEL e descartando as variáveis com lacuna de dados, ao final forma selecionadas 10 variáveis socioeconômicas para o processo de estimação dos modelos de Fronteira Estocástica nessa dissertação, conforme indicado no Quadro 5.1.

Quadro 5.1 –Variáveis socioeconômicas aplicadas ao caso brasileiro

| Código | Variáveis socioeconômicas | Fonte |
|------------|--|-----------|
| Vio | Violência - Óbitos por Agressão a cada 100 mil habitantes | DATASUS |
| Pob3 | % de pessoas com renda <i>per capita</i> inferior a ½ salário mínimo | IBGE/IPEA |
| Gini | Gini | IBGE |
| Sub2 | Precariedade - % de pessoas em domicílios subnormais (Censo 2010) | IBGE |
| lixo.u | % Domicílios com Coleta de lixo – urbano | IBGE |
| Inad | Inadimplência do setor de crédito – BACEN | BACEN |
| Inad.pf | Inadimplência do setor de crédito das pessoas físicas – BACEN | BACEN |
| Mbr.Mb1Mbr | Mercado baixa renda/(Mercado B1 total) | ANEEL |
| PIB.PC | PIB <i>per capita</i> da população | IBGE |
| Admd | Nº de moradores por dormitório por domicílio | IBGE |

Fonte: ANEEL (2014, 2020b)

Foram considerados no processo de estimação os dados relativos a 62 distribuidora de energia elétrica. As distribuidoras que passaram por processo de consolidação da área de concessões nos últimos anos (distribuidoras da Energisa no estado de SP e Paraná consolidadas na Energisa Sul Sudeste, a distribuidoras pequenas da CPFL no interior de SP consolidadas na CPF Santa Cruz e a RGE SUL e RGE no RS) foram tratadas no processo de estimação como distribuidoras independentes em função da maior disponibilidade de dados. A distribuidora CER foi desconsiderada do processo de estimação em função do grande número de dados faltantes.

As variáveis Vio, ADMD e Lixo.u não apresentavam dados disponíveis para 2016 (ADMD e Lixo.u) e 2017 (todas). Para essas variáveis, os valores para os anos de 2016 e 2017 foram estimados através de um modelo de regressão simples considerando a série de dados disponível dos últimos 6 anos.

5.2

Modelo e processo de estimação

O modelo estimado proposto, representado pela equação 5.1, foi estimado assumindo-se uma forma funcional para Função de Custo do tipo Cobb-Douglas e as componentes v_{it} (erro aleatório) e u_i (ineficiência) apresentam repectivamente distribuição Normal e *Half-Normal*.

$$\log(PPNT_{it}) = \beta_0 + \beta_1 \log(x_{1t}) + \dots + \beta_k \log(x_{kt}) + v_{it} + u_{it}, \quad (5.1)$$

$$\forall i = 1, n, \forall t = 1, T$$

Onde:

$PPNT_{it}$: é o percentual de PNT sobre o mercado de BT da distribuidora “i” no período “t”

x_{kt} : a variável socioeconômica “k” da distribuidora “i” no período “t”

v_{it} : é uma variável aleatória, onde $v_i \sim i. i. d N(0, \sigma_v^2)$

u_{it} : é o termo que representa a ineficiência, onde $u_i \sim i. i. d N^+(0, \sigma_u^2)$ (*Half-Normal*)

Admitindo-se que:

u_i e v_i são independentes

u_i e v_i não estão correlacionados com as variáveis explicativas

Foram estimados 164 modelos combinando as variáveis apresentadas no Quadro 5.1. Como as variáveis *Inad* e *Inad.pf* conceitualmente representam um mesmo fenômeno da inadimplência, optou-se por não estimar modelos que contivesse simultaneamente as duas variáveis. Em todos os modelos considerou-se a presença de uma componente que captura a evolução temporal da eficiência (variável tempo) (Coelli *et al*, 2005). Um resumo dos modelos estimados é apresentado no Anexo I desta dissertação.

Os modelos foram estimados utilizando-se o pacote “Frontier” desenvolvido por Coelli & Henningsen (2020) versão 1.1-18 para o *software* R (R Core Team, 2020).

5.3 Critérios para seleção dos modelos

Uma vez estimados os 164 modelos, eles foram avaliados de forma sequencial por 3 critérios previamente estabelecidos e listados no Quadro 5.2.

Quadro 5.2 – Critério de seleção de modelos adotados na simulação

| Filtro | Critérios Utilizados | Número de Modelos aprovados |
|---------------|--|------------------------------------|
| 1º | Gama estatisticamente significativo ao nível de 10% | 155 |
| 2º | Coefficientes da regressão estatisticamente significativos ao nível de 10% | 29 |
| 3º | Estimativa dos coeficientes de regressão com os sinais esperados | 29 |

Fonte: Elaboração própria

Inicialmente, entre os modelos analisados foram retidos apenas os modelos com todos os coeficientes de regressão estatisticamente significativo ao nível de 10% (Bressan *et. al.*, 2011).

Dos 164 modelos analisados, 155 apresentaram gama estatisticamente significativo ao nível de 10%. Desse subconjunto, 11 não apresentaram nenhum dos coeficientes da regressão estatisticamente significativos ao nível de 10% e outros 115 apresentaram ao menos um coeficiente da regressão estatisticamente não significativos no mesmo critério (significância de 10%). Os 29 modelos restantes apresentaram todos os coeficientes de regressão com os sinais esperados (um detalhamento dos modelos selecionados é apresentado no Anexo II dessa dissertação).

Das dez variáveis inicialmente testadas, as variáveis mais frequentes nos 29 modelos selecionados foram as variáveis relativas a percentual de domicílios subnormais (Sub2), o PIB *per capita* (PIB.PC), o coeficiente de Gini (Gini) e o percentual de domicílios beneficiários da tarifa social de energia (Mbr.Mb1Mbr). Apenas as variáveis violência (Vio) e % domicílios com Coleta de lixo – urbano (lixo.u) não foram observadas nos 29 modelos selecionados. A Figura 5.1 apresenta a frequência de participação das variáveis selecionadas inicialmente nos 29 modelos selecionados.

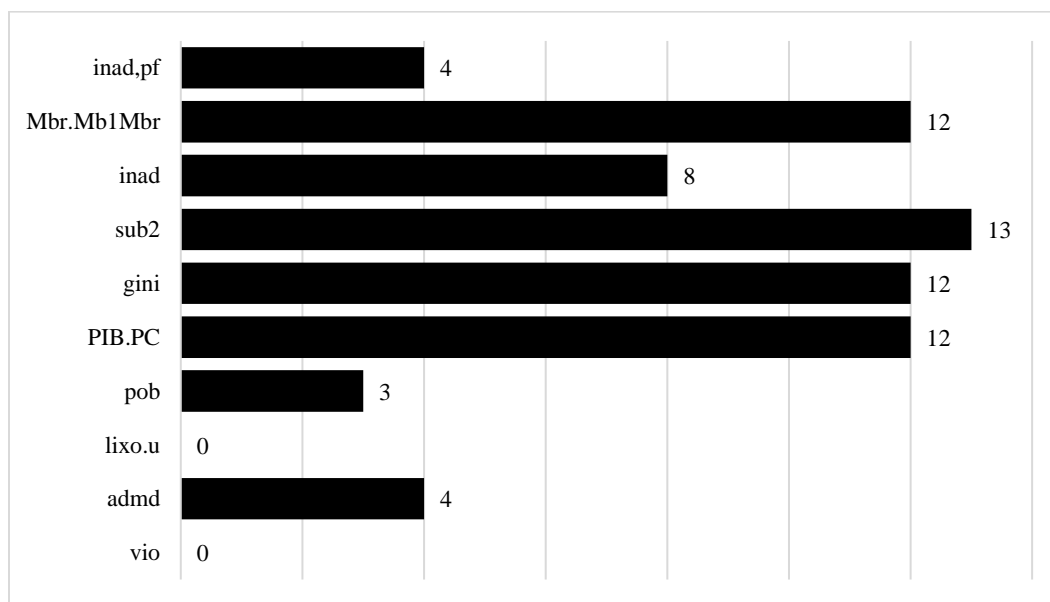


Figura 5.1 – Frequência das variáveis socioeconômicas nos modelos selecionados

Fonte: Elaboração própria

Ao invés de escolher um modelo específico e tomar decisões com base apenas no modelo escolhido, pode-se considerar o *ensemble* de 29 modelos selecionados, uma abordagem capaz de aproveitar a complementariedade entre os modelos (Page, 2008).

Assim nesta dissertação optou-se por trabalhar com a mediana dos resultados dos 29 modelos selecionados.

5.4

Resultados do Modelo de Eficiência

A Figura 5.2 apresenta a mediana das eficiências das 62 distribuidoras em cada ano, obtidas a partir dos resultados dos 29 modelos selecionados na fase anterior para cada ano.

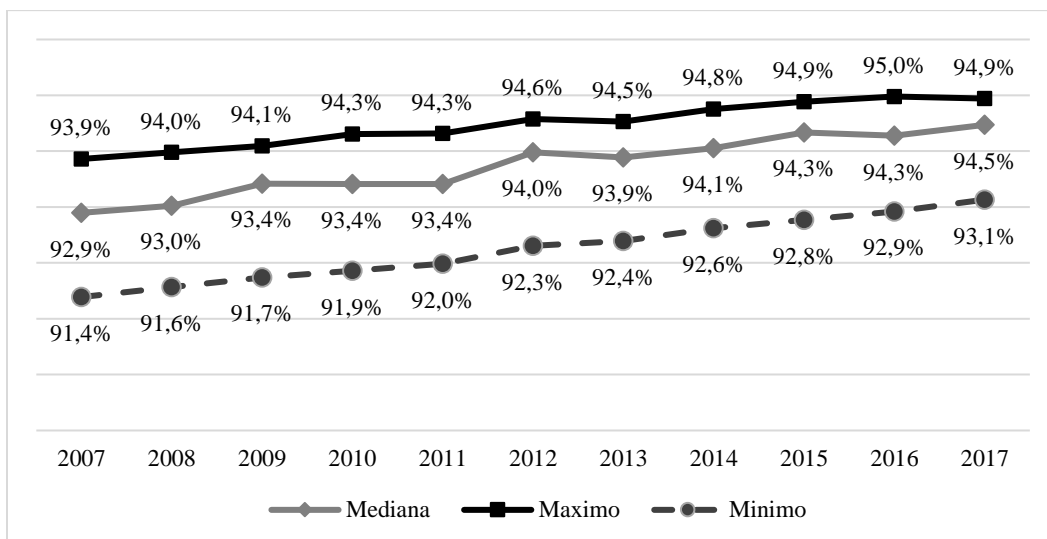


Figura 5.2 - Mediana da eficiência das distribuidoras de energia elétrica no combate às perdas não técnicas (PNT)

Fonte: Elaboração própria

Um primeiro ponto em relação a esse resultado é que eles apontam para um aparente ganho de eficiência, por parte das distribuidoras, ao longo dos últimos 10 anos no combate a PNT, o que se contrapõem a ideia de estabilidade do nível de PNT no Brasil comentado no capítulo 2 (Figura 2.1). Isso pode indicar que a manutenção dos níveis de PNT no país pode estar mais associado a uma deterioração do ambiente aonde essas distribuidoras estão inseridas, do que a própria performance das distribuidoras. Este seria um assunto bastante relevante a ser discutido pelo regulador no processo de definição das metas regulatórias de PNT.

O segundo ponto em relação a esses resultados, se refere a dispersão desses ao longo dos anos. Embora a Figura 5.1 mostre, por exemplo, uma amplitude de dois pontos percentuais entre a eficiência máxima das distribuidoras em 2017 (94,9%) e a mínima (93,1%), entre os 29 modelos selecionados, a distribuição desses resultados parece apontar para uma incerteza muito menor. Dos 29 modelos, por exemplo, apenas 8 estimam uma eficiência média do setor inferior a 94,2% para o ano de 2017. Ou seja, 73% dos modelos estimaram uma eficiência média para o setor em 2017 entre 94,2% e 94,9%.

Portanto, há uma robustez dos resultados globais frente a combinação das variáveis socioeconômicas selecionadas. Essa é sem dúvida vantagem do modelo proposto já que está menos exposto ao processo de seleção de variáveis. A Figura

5.3 apresenta a distribuição desses resultados para ano de 2017. Resultado similar é observado para todos os anos (os histogramas dos resultados dos demais anos são apresentados no Anexo III dessa dissertação).

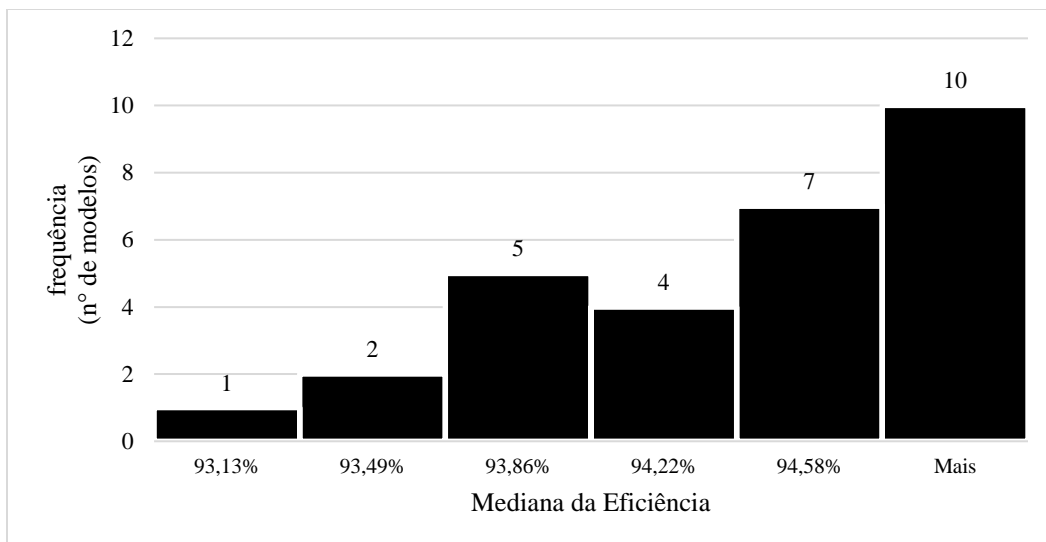


Figura.5.3 - Distribuição das eficiências médias calculadas para 2017 nos 29 modelos
Fonte: Elaboração própria

O terceiro ponto em relação a esses resultados, refere-se as performances individuais das distribuidoras. Das 10 distribuidoras com maior ganho de eficiência em termos de PNT de 2007 a 2017, 9 apresentam PNT bastante elevadas. Embora a performance média desses agentes continue sendo abaixo da mediana das demais distribuidoras, os modelos mostram que elas têm avançado no combate a PNT.

Na Tabela 5.1 é apresentada a mediana das eficiências estimadas para as 10 distribuidoras com maior ganho de eficiência entre 2007 e 2017. Os resultados das eficiências estimadas para todas as distribuidoras são apresentados no Anexo IV dessa dissertação.

Tabela 5.1 - Eficiência no combate a PNT das distribuidoras 10 distribuidoras com maior ganho de eficiência entre 2007 e 2017

| Distribuidora | 2007 | 2017 | Variação da eficiência |
|----------------------|-------------|-------------|-------------------------------|
| Amazonas | 43% | 50% | 7% |
| CEA | 51% | 57% | 7% |
| CEAL | 65% | 70% | 5% |
| Light | 66% | 71% | 5% |
| Cepisa | 70% | 75% | 5% |
| Celipa | 67% | 72% | 5% |
| Energisa RO | 70% | 75% | 5% |
| Enel RJ | 77% | 80% | 4% |
| CEEE | 77% | 80% | 4% |
| Eletroacre | 81% | 84% | 3% |

Fonte: Elaboração própria

5.5

Variáveis críticas e resultados do modelo de eficiência para as distribuidoras de grande porte

Embora os resultados globais tenham mostrado uma certa convergência entorno da mediana dos resultados dos 29 modelos, apresentando uma baixa sensibilidade em relação a seleção das variáveis socioeconômicas selecionadas, o mesmo não pode ser dito quando se isolam apenas as distribuidoras de grande porte.

Nesse caso, observa-se claramente uma mudança na performance média das distribuidoras quando se inclui ou não duas variáveis subnormal e coeficiente de Gini, mostrando que essas duas variáveis tem um impacto relevante na performance das distribuidoras de grande porte em relação a PNT.

A Figura 5.4 apresenta a mediana das eficiências das distribuidoras apenas dos modelos que contem a variável descrita no eixo das abscissas. Observa-se a partir da Figura que a mediana das eficiências das distribuidoras se altera pouco quando são considerados os 29 modelos selecionados (eficiência = 94,4%) ou apenas aqueles que contem a variável subnormal (eficiência = 94,9%).

Entretanto, quando se isolam apenas as distribuidoras de grande porte, esse cenário se altera. A mediana da eficiência das distribuidoras de grande porte que é 91,6% para 2017, quando se consideram os 29 modelos, passa para 92,8% quando são considerados apenas os modelos que possuem a variável subnormal e para 92,7% quando se consideram apenas os modelos que possuem a variável Gini.

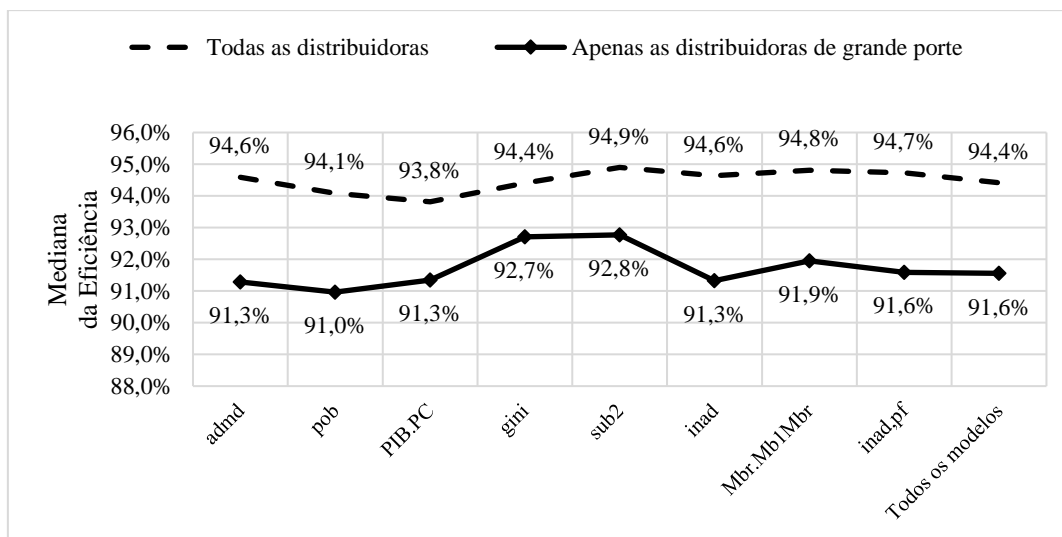


Figura 5.4 - Eficiências médias calculadas para 2017 a modelos que possuem as variáveis descritas no eixo das abscissas
Fonte: Elaboração própria

Isso é ainda mais relevante quando se observam os impactos sobre as distribuidoras individualmente. A Tabela 5.2 apresenta respectivamente as 10 distribuidoras que mais ganham e a 10 que mais perdem com a inclusão da variável “subnormal”.

Tabela 5.2 - As 10 Distribuidoras que mais ganham ou perdem eficiência no combate a PNT com a inclusão da variável % de Domicílios subnormais

| Distribuidora | Ganho de eficiência | Distribuidora | Perda de eficiência |
|------------------|---------------------|-----------------|---------------------|
| Light | 3,2% | Sulgipe | -1,2% |
| CPFL Piratininga | 3,0% | CHESP | -1,2% |
| Enel SP | 3,0% | EFLJC | -1,1% |
| CEB | 2,9% | MOCOCA | -1,0% |
| Enel RJ | 2,9% | CSPE | -1,0% |
| EDP ES | 2,7% | Cooperaliança | -1,0% |
| Celpe | 2,7% | CPEE | -1,0% |
| CEEE | 2,7% | CPFL Santa Cruz | -1,0% |
| EDP SP | 2,5% | Iguaçu Energia | -0,9% |
| Celipa | 2,4% | NACIONAL | -0,8% |

Fonte: Elaboração própria

O impacto relativo na performance das distribuidoras pode ser melhor exemplificado a partir da análise da distribuidora Light. Embora o ganho de eficiência ao longo dos últimos 10 anos se altere pouco, o patamar da performance

da empresa se altera de forma relevante quando se inclui a variável domicílios subnormais. As Figuras 5.5 e 5.6 ilustram esse impacto.

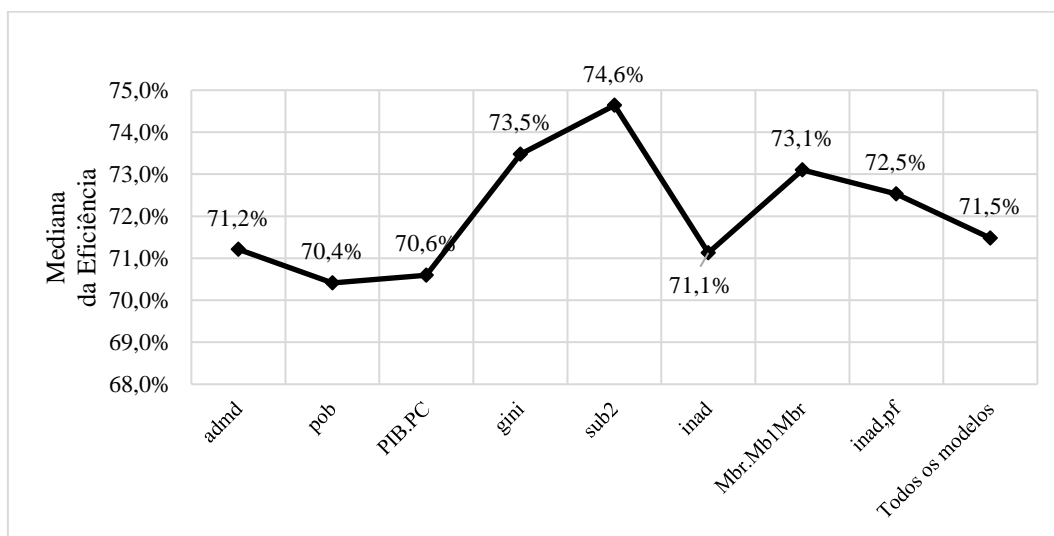


Figura 5.5 - Eficiências médias calculadas para LIGHT em 2017 a partir dos modelos que possuem as variáveis descritas no eixo das abscissas

Fonte: Elaboração própria

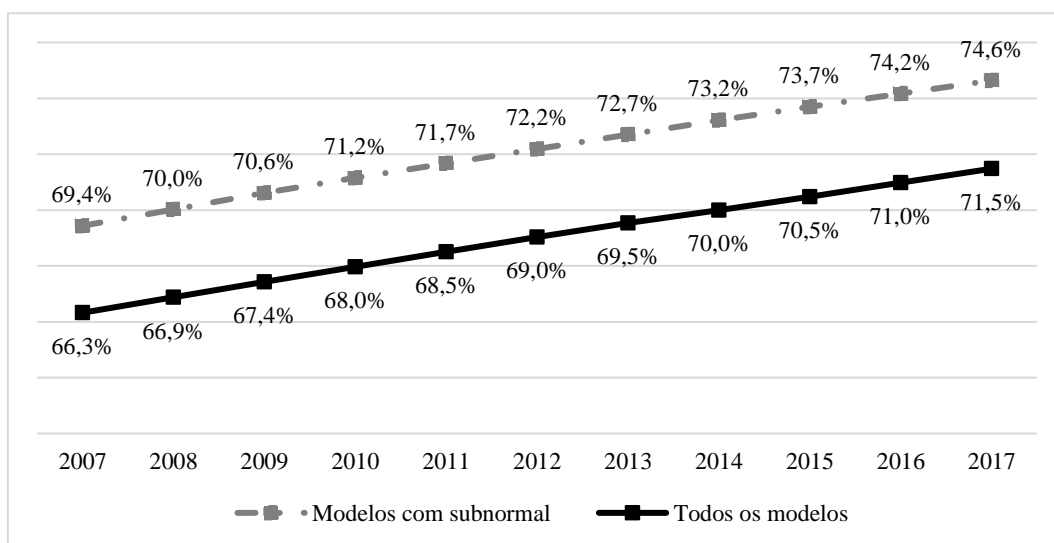


Figura 5.6 - Evolução da Performance da LIGHT de 2007 a 2017

Fonte: Elaboração própria

Um ponto a se destacar nessa análise sobre o impacto da variável dos domicílios subnormais é que o modelo parece indicar que a omissão da variável pode levar o modelo a subestimar a performance das distribuidoras de grande porte e superestimar a performance das distribuidoras de pequeno porte.

Assim, apesar das restrições e críticas apontadas no capítulo 2, a variável percentual de domicílios subnormais parece capturar, ainda que de forma precária, um fenômeno socioeconômico relevante para o combate as PNT.

5.6

Definição das metas regulatórias a partir modelo proposto e a comparação com as metas regulatórias definidas pela ANEEL

Para exemplificar a aplicação dos modelos e permitir uma comparação com as metas estabelecidas pela ANEEL, foram estimadas aqui as metas regulatórias que teriam sido definidas a partir do modelo proposto nesse trabalho para o ano de 2018 e as metas definidas pela ANEEL. As metas definidas pela ANEEL e as PNT realizadas para 2018 em termos de percentual do mercado de baixa tensão faturado das distribuidoras foi obtido através da base de dados de perdas disponibiliza pela ANEEL (2019b).

Para determinar as metas referentes ao ano 2018, a partir do modelo proposto, tomou-se a mediana das eficiências de cada uma das distribuidoras, obtida a partir dos 13 modelos de fronteira estocástica que incluíam a variável percentual de domicílios subnormais e estimados com dados até o ano de 2017. Foi considerado também um ciclo padrão de 5 anos para todas as distribuidoras ($\Delta_i = 5$). Assim o ano de 2018 corresponderia ao ano 1 da trajetória de PNT regulatória ($t = 1$).

As metas foram inicialmente calculadas conforme a equação 5.2 (que corresponde a equação 4.16 descrita no capítulo 4).

$$PMeta_t = \%PNT_0 \left[\left(\frac{\theta_{i0}}{\theta_{ref}} \right)^{t/\Delta} \right] \quad (5.2)$$

Onde:

$PMeta_t$: meta de PNT da distribuidora “i” para o ano 2018

$\%PNT_{i0}$: % de PNT faturadas pela distribuidora “i” no ano de 2017

θ_{i0} : eficiência da distribuidora “i” para o ano de 2017

θ_{ref} : nível de eficiência que foi considerado igual a 1

Δ_i : prazo (em anos) para que a distribuidora “i” alcance a meta regulatória de PNT estabelecida no ano de 2017. Nesse caso considerou-se $\Delta_i = 5$ para todas as distribuidoras.

Há algumas simplificações nessa análise que foram tomadas de forma a tornar mais simples e direta a comparação.

A primeira delas é que foi estabelecida uma meta para o ano civil. Na prática as metas de PNT regulatórias são definidas em termos de anos tarifários o que varia de distribuidora para distribuidora. A adoção do ano tarifário, no entanto, tornaria a comparação extremamente complexa. Para se evitar uma eventual distorção na análise, as metas regulatórias propostas aqui foram comparadas com as metas regulatória ajustada para o ano civil que foram divulgadas pela ANEEL (2019b).

A segunda simplificação é que as metas regulatórias propostas nesse trabalho foram definidas como se todas as distribuidoras estivessem passando por revisão tarifária em 2017. As distribuidoras de fato passam por processo de Revisão tarifária em anos distintos e as metas de 2018 definidas pela ANEEL foram definidas em anos distintos. A opção por se adotar essa simplificação foi tornar a estimativa das metas pelo modelo aqui proposto mais simples e o mais equânime possível entre as distribuidoras.

Essa é a razão para uma terceira simplificação adotada aqui. A adoção de um ciclo tarifário padrão de 5 anos para todas as distribuidoras, de forma a tornar o tratamento entre as distribuidoras mais isonômico. A adoção de um ciclo tarifário padrão de 5 anos para todas as distribuidoras. A opção pelos 5 anos foi equalizar o período de adaptação das distribuidoras a nova meta, tomando-se como base o ciclo mais longo como padrão.

Por último, como as PNT faturadas podem resultar em PNT inferior a zero. Adotou-se como premissa, que as distribuidoras com PNT inferior a zero teriam metas regulatórias de PNT no novo modelo iguais a zero.

A Base de dados de PNT faturadas para 2018 disponibilizada pela ANEEL trata as distribuidoras que foram consolidadas na Energisa Sul Sudeste (5 distribuidoras) e na CPFL Santa Cruz (5 distribuidoras). Ademais, não foi possível estimar a eficiência da Boa Vista para 2017 por falta de dados. Assim, adotou-se nesse caso a eficiência estimada para 2016. Assim ao final a comparação conta com 53 distribuidoras.

A Figura 5.7 compara as metas de PNT das distribuidoras obtidas a partir do modelo proposto nesse trabalho com as definidas pela ANEEL para o ano de 2018.

Das 53 distribuidoras para as quais foi possível realizar a estimativa da meta de PNT pelo novo modelo, 60,4% apresentariam um incremento no reconhecimento

de PNT regulatório com adoção do modelo proposto. Com consequência disso, o percentual de distribuidoras que teriam metas regulatórias de PNT iguais ou superiores as PNTs realizadas 2018, subiria de 24,5% para 41,5%. Mesmo entre aquelas que ainda não teriam as suas PNT plenamente reconhecidas (56,4% restantes), haveria uma melhora no reconhecimento regulatório em 61,3% dos casos. A Tabela 5.3 apresenta os resultados para as todas as distribuidoras.

Tabela 5.3 - Comparação entre as metas regulatórias de PNT para 2018 da ANEEL e as propostas pelo modelo alternativo – dados em % do mercado BT faturado

| Distribuidora | PNT realizado | Meta ANEEL | Meta do modelo Alternativo Proposto | Houve incremento no reconhecimento de PNT ou Redução? |
|----------------------|----------------------|-------------------|--|--|
| Amazonas | 124,3% | 77,7% | 106,4% | incremento |
| CEA | 76,5% | 61,4% | 62,9% | incremento |
| Light | 47,5% | 36,1% | 35,5% | redução |
| Celpe | 38,4% | 34,0% | 33,5% | redução |
| Energisa RO | 32,1% | 22,6% | 32,0% | incremento |
| Cepisa | 30,5% | 21,6% | 28,4% | incremento |
| Boa Vista | 28,7% | 24,2% | 27,4% | incremento |
| CEAL | 29,8% | 27,2% | 26,0% | redução |
| Enel RJ | 25,2% | 18,7% | 23,4% | incremento |
| CEEE | 22,9% | 7,7% | 22,3% | incremento |
| Eletroacre | 16,9% | 18,4% | 20,2% | incremento |
| Celpe | 16,7% | 13,6% | 15,0% | incremento |
| EDP ES | 12,1% | 11,5% | 14,4% | incremento |
| EDP SP | 10,9% | 8,7% | 12,1% | incremento |
| Cemig | 9,8% | 7,4% | 11,6% | incremento |
| CEB | 9,5% | 7,0% | 9,2% | incremento |
| Energisa MT | 8,8% | 7,1% | 8,9% | incremento |
| RGE Sul | 8,6% | 4,9% | 8,8% | incremento |
| Enel CE | 9,3% | 5,1% | 8,7% | incremento |
| Enel SP | 8,8% | 8,4% | 8,2% | redução |
| CPFL Paulista | 8,6% | 5,5% | 8,1% | incremento |
| Cemar | 8,1% | 10,1% | 7,9% | redução |
| Celesc | 8,1% | 3,7% | 7,7% | incremento |
| Iguaçu Energia | 5,3% | 3,6% | 7,1% | incremento |
| CPFL Piratininga | 7,9% | 3,9% | 6,8% | incremento |
| Coelba | 8,9% | 6,6% | 6,8% | incremento |
| Energisa MS | 5,8% | 5,7% | 6,7% | incremento |
| Elektro | 7,7% | 3,3% | 5,7% | incremento |
| RGE | 5,7% | 5,5% | 5,6% | incremento |

| Distribuidora | PNT realizado | Meta ANEEL | Meta do modelo Alternativo Proposto | Houve incremento no reconhecimento de PNT ou Redução? |
|----------------------|----------------------|-------------------|--|--|
| CHESP | 1,4% | 1,1% | 5,6% | incremento |
| Energisa PB | 6,0% | 5,1% | 5,3% | incremento |
| Eletrocar | 3,1% | 1,7% | 5,1% | incremento |
| Energisa SE | 5,5% | 5,6% | 4,8% | redução |
| Sulgipe | 2,7% | 5,0% | 4,7% | redução |
| Enel GO | 4,2% | 4,0% | 4,3% | incremento |
| Copel | 5,2% | 4,7% | 4,0% | redução |
| CPFL Santa Cruz | 3,3% | 1,0% | 3,9% | incremento |
| Energisa TO | 4,5% | 5,3% | 3,5% | redução |
| Hidropan | 2,9% | 0,3% | 2,4% | incremento |
| DEMEI | 4,6% | 3,7% | 2,3% | redução |
| Energisa MG | 2,7% | 0,6% | 2,1% | incremento |
| ELFSM | 8,5% | 3,8% | 2,0% | redução |
| EFLUL | 1,4% | 0,4% | 1,5% | incremento |
| EFLJC | 1,1% | 1,5% | 1,4% | redução |
| Cosern | 0,7% | 2,6% | 0,8% | redução |
| Energisa SS | 1,4% | 0,6% | 0,2% | redução |
| Energisa BO | 0,5% | 0,2% | 0,0% | redução |
| MUX Energia | -0,9% | 1,9% | 0,0% | redução |
| DME-PC | -1,5% | 0,9% | 0,0% | redução |
| Energisa NF | -2,3% | 0,0% | 0,0% | - |
| Cooperaliança | -2,8% | 0,3% | 0,0% | redução |
| Cocel | -7,8% | 1,3% | 0,0% | redução |
| Forcel | -9,8% | 0,0% | 0,0% | - |

Fonte: Elaboração própria e ANEEL (2019b)

Das 10 distribuidoras que teriam o maior incremento com adoção do modelo proposto, três passariam a ter metas regulatórias iguais ou superiores a suas PNTs realizadas, zerando a posição das PNT sem reconhecimento tarifário e outras seis teriam uma redução média da parcela de PNT sem reconhecimento tarifário em médio de 81%.

Por outro lado, das 10 distribuidoras que teriam a maior redução das metas de PNT regulatórias, apenas duas passariam a ter metas regulatórias inferiores a PNT realizada, e a redução média de reconhecimento seria de 1,5 pontos percentuais. Isto revela que o impacto sobre as distribuidoras que teriam a suas metas de PNT elevadas seriam bem superiores ao daquelas que teriam suas metas regulatórias reduzidas. Ademais, dentre as distribuidoras que teriam suas metas de PNT

regulatórias reduzidas, 6 possuem níveis de PNT inferiores a 5% do mercado de baixa tensão.

Um exemplo do impacto do novo modelo é o caso da Enel Rio. A distribuidora teria um incremento do reconhecimento de PNT regulatórias de 4,9 pontos percentuais se o modelo proposto fosse adotado. Considerando um preço médio de compra de energia da distribuidora em 2018 de R\$ 198,62 e um mercado de baixa tensão fatura de 6,7 GWh (ANEEL, 2019), esse incremento representaria um adicional de receita para a distribuidora de R\$ 65 milhões.

Se for considerado que ao menos em parte o não reconhecimento de PNT regulatório da companhia pode estar associado aos problemas modelos atualmente utilizados pela ANEEL (como os apontados no capítulo 2), a distribuidora teria até 20% das suas PNTs indevidamente não reconhecidas.

O caso da Enel Rio é emblemático porque a própria ANEEL reconheceu em parte o problema na definição das metas de PNT daquela distribuidora e promoveu um incremento de 4 pontos percentuais nas metas de PNT regulatória da distribuidora para 2018 em relação ao percentual regulatório original (de 14,7% prevista na REH 1.703/2014 para 18,54%). Para fazer isso, a ANEEL teve que tratar a Enel Rio como um caso excepcional adotando uma metodologia, em certa medida *ad hoc* para definir as PNT regulatórias da distribuidora (ANEEL, 2019b).

Outros dois casos interessantes são os das distribuidoras de energia elétrica recentemente privatizadas, particularmente a Amazonas, Ceron (atualmente Energisa RO) e a CEPISA. Para viabilizar o processo de privatização dessas distribuidoras uma série de flexibilizações nas metas regulatórias dessas companhias tiveram que ser adotadas pela ANEEL, incluindo as metas de PNT. Entretanto, se considerados os modelos propostos nesse trabalho, em 2018 as 3 distribuidoras teriam incrementos significativos nas suas metas regulatórias de PNT o que dispensaria, ao menos em parte, a adoção de um tratamento diferenciado por parte da Agência Reguladora para essas distribuidoras. A Amazonas teria um incremento de 28,7 pontos percentuais no reconhecimento de perdas, enquanto a CEPISA e a CERON, respectivamente de 6,9 e 9,5 pontos percentuais, respectivamente.

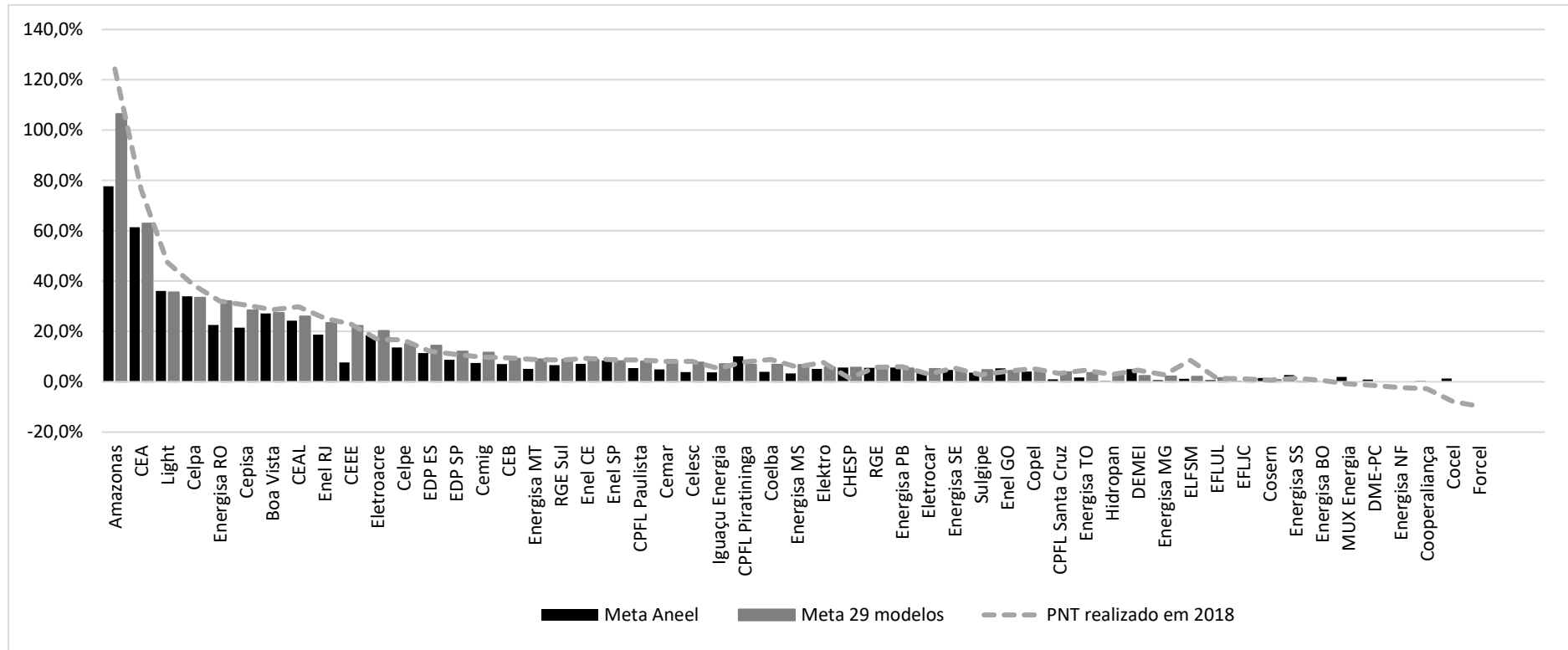


Figura 5.7 - Comparação entre as metas regulatórias de PNT para 2018 da ANEEL e as propostas pelo modelo alternativo – dados em % do mercado BT faturado

Fonte: ANEEL(2019b)

5.7

Discussão dos resultados

Os resultados obtidos a partir da aplicação do modelo proposto para o caso brasileiro fomentam algumas discussões importantes, e mostram um pouco do potencial do modelo.

A primeira delas é que ao segregar, de forma clara, a eficiência das distribuidoras no combate a PNT dos demais fatores, o modelo proposto lança luz a sobre os avanços no combate a PNT no Brasil nos últimos 10 anos.

De forma global, os resultados apontam para um aparente ganho de eficiência por parte das distribuidoras na última década, sugerindo que a manutenção dos níveis de PNT no país pode estar mais associado a uma deterioração do ambiente socioeconômico, do que uma inação por parte das distribuidoras.

Por outro lado, os resultados também sugerem que as distribuidoras com maior nível de PNT foram as que mais investiram esforços no combate as PNT e consequentemente as que mais avançaram em termos de eficiência. Das 10 distribuidoras com maior ganho de eficiência em termos de PNT de 2007 a 2017, 9 apresentam PNT bastante elevadas. Embora a performance média desses agentes continue sendo abaixo da mediana das demais distribuidoras, esse resulta reforça a ideia que o setor vem avançando no tema.

Uma segunda discussão importante, refere-se à pertinência das variáveis socioeconômicas para o tema da PNT. Ao simplificar e tornar o processo de definição das metas regulatórias de PNT mais direto, o modelo proposto torna essa discussão mais objetiva. Os resultados sugerem, por exemplo, que embora a seleção das variáveis tenha pouco impacto sobre a eficiência global do setor no combate a PNT, ela afeta as eficiências relativas das empresas.

Os resultados, sugerem por exemplo, que a introdução das variáveis percentual de domicílios subnormais e coeficiente de Gini, aumenta a eficiência relativa das distribuidoras de grande porte. Isso sugerem que esses fenômenos talvez estejam mais concentrados em algumas regiões específicas. A variável subnormal, por exemplo, busca retrata a ocupação irregular do espaço urbano, um fenômeno típico das grandes cidades e que tem forte implicações na atuação das distribuidoras no combate ao PNT. É natural, por exemplo, que a introdução dessa

variável aumente a eficiência das distribuidoras que atuam nessas regiões em detrimento das distribuidoras que atuam em regiões menores.

Os resultados apresentados também fomentam uma discussão importante sobre a exequibilidade das metas de PNT regulatórias atualmente estabelecidas pela ANEEL.

Os resultados do modelo proposto reforçam um ponto que foi evidenciado no capítulo 2, que há um aparente subreconhecimento tarifário do fenômeno do PNT, particularmente relevante para algumas distribuidoras. Alguns desses casos, inclusive já reconhecidos, mesmo que de forma parcial, pela própria ANEEL, como no caso das Enel Rio e das distribuidoras recentemente privatizadas.

O modelo proposto sugere que das 53 distribuidoras, para as quais foi possível realizar a estimativa da meta de PNT, 60,4% tiveram um subreconhecimento de PNT em 2018.

Há duas consequências deletérias desse fenômeno. Ao não reconhecer parte dos custos das distribuidoras com PNT, o regulador limita a capacidade de investimento da concessionária, inclusive no próprio processo de combate a PNT. Por outro lado, o próprio regulador passa a ter que adotar uma série de procedimento *ad hoc* para tentar lidar com essa situação, tornando o processo mais complexo e menos previsível.

Nesse sentido o modelo proposto demonstra que é possível corrigir essas distorções, sem que se tenha que lançar mão de procedimento *ad hoc*. Isso foi evidenciado para o caso da Enel Rio e das distribuidoras recentemente privatizadas.

6 Conclusões

O objetivo geral da dissertação foi propor um modelo alternativo ao modelo regulatório atual de definição das metas regulatórias de PNT, avaliando a aplicabilidade de uma família particular de modelos de Análise de Eficiência – os modelos de Análise de Fronteira Estocástica (SFA).

Foi realizada uma breve explanação sobre o problema das PNT e dos principais conceitos que norteiam a Regulação do setor de distribuição de energia elétrica. Foram também identificadas as principais limitações no modelo atual adotado pela ANEEL para definição das metas regulatórias de PNT. Como demonstrado, a excessiva complexidade da metodologia atualmente empregada pela ANEEL na definição das metas de PNT regulatória, assim como a adoção de um conjunto grande critérios *ad hoc*, põem em risco um dos principais parâmetros que norteiam a regulação por incentivo: o estabelecimento de incentivos econômicos corretos para busca contínua pela melhoria da eficiência.

Os problemas apontados, apresentam indícios claros que há falhas no processo de comparação das PNT entre as distribuidoras o que pode estar conduzindo tanto a um não reconhecimento da PNT de algumas distribuidoras quanto ao reconhecimento indevido da PNT em outras.

Com o objeto de formular um modelo alternativo, foi apresentado e analisado uma família de modelos matemáticos denominados modelos de Análise de Eficiência, que demonstram ser uma ferramenta mais adequada para comparação e mensurar a eficiência relativas dos agentes. Por essa razão, esses modelos vêm sendo largamente utilizados por reguladores de todo o mundo na definição dos níveis ótimos de custo operacionais e investimento.

Dos três principais modelos de Análise de Fronteira de Eficiência mais frequentes na Literatura, o DEA é o mais empregado pelo Reguladores no mundo. Isso se deve principalmente à sua simplicidade de implementação.

Apesar disso, o DEA traz uma desvantagem bastante relevante para a sua aplicação ao tema de PNT. Por ser um método não paramétrico não é possível testar

diretamente nem a qualidade do modelo, nem o impacto relativo da introdução ou exclusão de variáveis e/ou distribuidoras nos resultados dos modelos. Além disso no modelo DEA todo o eventual erro de especificação, como a omissão de variáveis, é interpretado como ineficiência. Como foi exposto nesse trabalho, há no tema de PNT um elevado risco de omissão de variáveis, seja porque as próprias relações de causalidade de PNT ainda não estão bem estabelecidas, seja pelas limitações das próprias variáveis disponíveis.

Por outro lado, foi demonstrado que a principal vantagem do SFA em relação ao DEA está exatamente na possibilitar de testar estatisticamente tanto a especificação do modelo quanto suas conclusões sobre a eficiência das distribuidoras. Além disso o modelo SFA, por decompor o resíduos em duas componentes ($\varepsilon_i = v_i + u_i$), consegue lidar melhor com eventuais choques aleatórios e ou omissões de variáveis, uma vez que apenas parte do resíduo é considerada como ineficiência. Nesse sentido, para o problema de PNT, a utilização do SFA parece apresentar como o método mais adequado.

Foi então formulado um modelo para definição metas regulatórias de PNT a partir dos modelos de Fronteira Estocástica (SFA). Embora o problema das PNT de fato não se enquadre como problema clássico da teoria da produção, o problema em tela continua sendo um problema de minimização de custos e é possível adaptar o conceito clássico dos modelos de análise de eficiência baseados nas fronteiras de custo para o problema de PNT.

Foi proposto então, um modelo para definição de metas regulatórias de PNT partindo-se de um modelo SFA com dados em painel. Além disso, um conjunto de métricas auxiliares para a sua aplicação nos processos tarifários foi proposto, como uma metodologia para definição de uma trajetória anual de redução de perdas e a possibilidade da adoção de limites adicionais para as metas regulatórias de PNT. O modelo proposto tem como vantagens, não só tornar o processo de definição das metas regulatórias mais simples e direto, evitando uma série de etapa *ad hoc* adotadas pela ANEEL com o uso do seu modelo atual, como também tornaria o processo mais robusto em relação a seleção/omissão de variáveis. Um ponto bastante relevante é que o modelo proposto permitiria avaliar de uma forma mais criteriosa e objetiva o comportamento das distribuidoras no combate as PNT, separando a eficiência operacional das distribuidoras dos demais efeitos exógenos.

Foi feita uma aplicação do modelo proposto ao contexto do setor de distribuição de energia elétrica brasileiro para definir as metas regulatórias das distribuidoras no ano de 2018. As metas obtidas foram comparadas com as metas de 2018 definidas pela ANEEL. Foi possível observar que o modelo proposto não só é aplicável ao caso brasileiro, como sugere que em função das limitações do modelo do modelo atual da ANEEL possa estar havendo um não reconhecimento de uma parte relevante das PNT das distribuidoras.

A proposta apresentada nesse trabalho, no entanto, não é exaustiva. Uma comparação mais ampla, incluindo um conjunto maior de anos, com os modelos da ANEEL poderia confirmar ou não alguns pontos levantados e poderia ser objeto de um trabalho futuro sobre o tema. Uma série de críticas levantadas em relação ao modelo da ANEEL, especialmente em relação as variáveis socioeconômicas, continua em aberto e também poderia ser objeto de futuros trabalho.

Adoção de critérios alternativos no processo de seleção dos modelos SFA, como a introdução de critérios de informação Akaike (AIC), daria maior robustez também poderia ser objeto de um trabalho futuro sobre o tema.

Ademais a possibilidade de aplicação de outros modelos de Análise Fronteira poderia contribuir ainda mais para o desenvolvimento do tema. O modelo *Stochastic Non-Smooth Envolvment of Data* (StoNED) proposto por Kuosmanen e Kortelainen (2010) pode vir a ser um concorrente natural ao SFA num futuro próximo (Kuosmanen, 2012; Kuosmanen *et. al*, 2013). O modelo proposto busca superar algumas das limitações do DEA apresentados nesse trabalho e uma eventual comparação entre modelo proposto nesta dissertação e um modelo desenvolvido a partir da aplicação do método StoNED pode vir a ser um avanço importante nesse tema.

AGRELL, P.; BOGETOFT, P. Regulatory benchmarking: Models, analyses and applications. **Data Envelopment Analysis Journal**, v.3, n. 1, p. 49-91. 2016.

ALMANIDIS, P.; QIAN, J.; SICKLES, R. C. Stochastic frontier with bounded ineciency. In: Sickles, R. C.; Horrace, W. C. (Ed.). **Festschrift in Honor of Peter Schmidt: Econometric Methods and Applications**. 1. ed. New York: Springer, 2014. p. 47-82.

ALMANIDIS, P.; SICKLES, R. C. The skewness problem in stochastic frontier models: Fact or fiction? In: Van Keilegom, I.; Wilson, P. W. (Ed.) **Exploring Research Frontiers in Contemporary Statistics and Econometrics: A Festschrift in Honor of Leopold Simar**. 1. ed. New York: Springer, 2012. p.201-227.

AMIN, S., SCHWARTZ, G. A.; CARDENAS, A. A.; SASTRY, S. S. Game-theoretic models of electricity theft detection in smart utility networks: Providing new capabilities with advanced metering infrastructure. **IEEE Control Systems**, v.35, p.66–81. 2015.

ANAYA, K. L.; POLLITT, M. G. Using stochastic frontier analysis to mesure the impact of weather on the efficiency of electricity distribution businesses in the developing economics. **European Journal of Operational Research**, v.263, n. 3, p. 1078-1094. 2017.

ANEEL – AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA . **Nota Técnica 271/2010-SRE/ANEEL – Metodologia de tratamento regulatório para perdas não técnicas de energia elétrica**. 2010. Disponível em: <http://www.ANEEL.gov.br/audiencias-publicas> >. Acesso em: 07 nov. 2018.

_____. **Nota Técnica 0312011-SRE/ANEEL – Metodologia de tratamento regulatório para perdas não técnicas de energia elétrica**. 2011. Disponível em: <http://www.ANEEL.gov.br/audiencias-publicas> >. Acesso em: 07 nov.2018.

_____. **Nota Técnica 453/2013-SGT/SRM/ANEEL – Discussão conceitual sobre a metodologia de definição do nível regulatório de perdas na distribuição, a qualidade dos serviços e os incentivos regulatórios**. 2013. Disponível em: <http://www.ANEEL.gov.br/consultas-publicas>>. Acesso em: 07 nov.2018.

_____. **Nota Técnica 188/2014-SRE/ANEEL – Atualização do Índice de Complexidade Socioeconômica adotado para definição do nível regulatório de Perdas Não Técnicas na Distribuição**. 2014. Disponível em: <http://www.ANEEL.gov.br/audiencias-publicas> >. Acesso em: 07 nov.2018.

_____. **Nota Técnica 406/2014-SRE/ANEEL – Metodologia de tratamento regulatório para perdas não técnicas de energia elétrica**. 2014b. Disponível em: <http://www.ANEEL.gov.br/audiencias-publicas> >. Acesso em: 07 nov.2018.

_____. **Nota Técnica 106/2015-SGT/SRM/ANEEL – Metodologia de tratamento regulatório para perdas não técnicas de energia elétrica.** 2015a. Disponível em: <http://www.ANEEL.gov.br/audiencias-publicas> >. Acesso em: 07 nov.2018.

_____. **Procedimentos de Regulação Tarifária – PRORET - Submodulo 2.6 – Perdas de Energia.** 2015b. Disponível em: < http://www2.ANEEL.gov.br/cedoc/aren2015660_Proret_Submod_2_6_V3.pdf >. Acesso em: 07 nov.2018.

_____. **Nota Técnica no 179/2018-SRM/ANEEL – Definição dos níveis regulatórios de perdas não técnicas de energia elétrica no sistema de distribuição da Enel-Rio no período de 2019 a 2022.** 2018a. Disponível em: < https://www.aneel.gov.br/audiencias-publicas-antigas?p_p_id=participacaopublica_WAR_participacaopublicaportlet&p_p_lifecycle=2&p_p_state=normal&p_p_mode=view&p_p_cacheability=cacheLevelPage&p_p_col_id=column-2&p_p_col_pos=1&p_p_col_count=2&participacaopublica_WAR_participacaopublicaportlet_idDocumento=31670&participacaopublica_WAR_participacaopublicaportlet_tipoFaseReuniao=fase&participacaopublica_WAR_participacaopublicaportlet_jspPage=%2Fhtml%2Fpp%2Fvisualizar.jsp >. Acesso em: 17 jan.2019.

_____. **Nota Técnica no 38/2018-SRM/ANEEL – Resultado da Audiência Pública nº 52/2017, realizada para colher subsídios à atualização dos parâmetros relacionados à definição dos Custos Operacionais Regulatórios (Submódulos 2.2 e 2.2A dos Procedimentos de Regulação Tarifária - PRORET).** 2018b. Disponível em: < https://www.aneel.gov.br/audiencias-publicas-antigas?p_p_id=participacaopublica_WAR_participacaopublicaportlet&p_p_lifecycle=2&p_p_state=normal&p_p_mode=view&p_p_cacheability=cacheLevelPage&p_p_col_id=column-2&p_p_col_pos=1&p_p_col_count=2&participacaopublica_WAR_participacaopublicaportlet_idDocumento=30138&participacaopublica_WAR_participacaopublicaportlet_tipoFaseReuniao=fase&participacaopublica_WAR_participacaopublicaportlet_jspPage=%2Fhtml%2Fpp%2Fvisualizar.jsp >. Acesso em: 17 jan. 2019.

_____. **Relatório de indicadores de sustentabilidade econômico-financeira das distribuidoras – Base Set/2019.** 2019a. Disponível em: < <https://www.ANEEL.gov.br/documents/656815/14887148/Base+em+Excel+2019+3T.zip/1e460055-5a34-e558-0275-57fc68c89fb5> >. Acesso em: 15 jan. 2020.

_____. **Base de dados – Perdas.** 2019b. Disponível em: < <https://www.ANEEL.gov.br/documents/15188137/0/Base+Perdas+Internet+-05-2019.xlsx/2e0051a3-93c9-a5b9-7f7b-2b0977adb00c> >. Acesso em: 03 jun. 2019.

_____. **Nota Técnica nº 35/2019-SGT/ANEEL.** 2019c. Disponível em: <https://www2.ANEEL.gov.br/aplicacoes/tarifa/arquivo/NT35-2019.pdf>>. Acesso em: 15 jan. 2020.

_____. **Nota Técnica Nº 46/2020 – SRM/SGT/ANEEL – Proposta de Consulta Pública para revisão da metodologia e atualização dos parâmetros dos Submódulos 2.2/2.2 A (Receitas Irrecuperáveis) e 2.6 (Perdas de Energia) dos Procedimentos de Revisão Tarifária - Proret.** 2020a. Disponível em: < https://www.ANEEL.gov.br/consultas-publicas?p_p_id=participacaopublica_WAR_participacaopublicaportlet&p_p_lifecycle=2&p_p_state=normal&p_p_mode=view&p_p_cacheability=cacheLevelPage&p_p_col_id=column-

[2&p_p_col_pos=1&p_p_col_count=2& participacaopublica_WAR_participacaopublicaportlet ideDocumento=39824& participacaopublica_WAR_participacaopublicaportlet tipoFaseReuniao=fase& participacaopublica_WAR_participacaopublicaportlet jspPage=%2Fhtml%2Fpp%2Fvisualizar.jsp](https://www.ANEEL.gov.br/consultas-publicas?p_auth=nJuT0z50&p_p_id=participacaopublica_WAR_participacaopublicaportlet ideDocumento=39824& participacaopublica_WAR_participacaopublicaportlet tipoFaseReuniao=fase& participacaopublica_WAR_participacaopublicaportlet jspPage=%2Fhtml%2Fpp%2Fvisualizar.jsp) >. Acesso em: 20 abr. 2020.

_____. **Base de Perdas disponibilizada no ambito da Consulta Pública 029/2020 da ANEEL.** 2020b. Disponível em: https://www.ANEEL.gov.br/consultas-publicas?p_auth=nJuT0z50&p_p_id=participacaopublica_WAR_participacaopublicaportlet&p_p_lifecycle=1&p_p_state=normal&p_p_mode=view&p_p_col_id=column-

[2&p_p_col_pos=1&p_p_col_count=2& participacaopublica_WAR_participacaopublicaportlet ideParticipacaoPublica=3431& participacaopublica_WAR_participacaopublicaportlet javax.portlet.action=visualizarParticipacaoPublica](https://www.ANEEL.gov.br/consultas-publicas?p_auth=nJuT0z50&p_p_id=participacaopublica_WAR_participacaopublicaportlet ideParticipacaoPublica=3431& participacaopublica_WAR_participacaopublicaportlet javax.portlet.action=visualizarParticipacaoPublica) >. Acesso em: 20 abr. 2020.

ANTONIOLI, B.; FILIPPINI, M. The use of a variable cost function in the regulation of the Italian water industry. **Journal of Productivity Analysis**, v. 10, n. 3, pp. 181-187. 2001.

APARICIO, J.; LOVELL, C.A.K.; PASTOR, J.T. **Advances in Efficiency and Productivity**. 1 ed. New York: Springer. 2016. 421 p.

ARANGO, L. G.; DECCACHE, E.; BONATTO, B. D. ; ARANGO, H.; PAMPLONA, E. O. Study of Electricity Theft Impact on the Economy of a Regulated Electricity Company. **Journal Of Control Automation And Electrical Systems**, v.28, n.4, p. 567-575. 2017.

BANCO CENTRAL DO BRASIL. SGS - Sistema Gerenciador de Séries Temporais - v2.1. 2018. Disponível em: https://www.ANEEL.gov.br/consultas-publicas?p_auth=nJuT0z50&p_p_id=participacaopublica_WAR_participacaopublicaportlet&p_p_lifecycle=1&p_p_state=normal&p_p_mode=view&p_p_col_id=column-

[2&p_p_col_pos=1&p_p_col_count=2& participacaopublica_WAR_participacaopublicaportlet ideParticipacaoPublica=3431& participacaopublica_WAR_participacaopublicaportlet javax.portlet.action=visualizarParticipacaoPublica](https://www.ANEEL.gov.br/consultas-publicas?p_auth=nJuT0z50&p_p_id=participacaopublica_WAR_participacaopublicaportlet ideParticipacaoPublica=3431& participacaopublica_WAR_participacaopublicaportlet javax.portlet.action=visualizarParticipacaoPublica) >. Acesso em: 20 jul. 2019.

BANKER R.D.; CHARNES, A.; COOPER, W.W. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. **Management Science** v. 30, n. 9, p. 1078-1092. 1984.

BATTESE, G.E.; COELLI, T.J. Frontier production function, technical efficiency and panel data: With application to paddy farmers in India. **J. Productivity Analysis**, v.3, p.153-169. 1992.

BAUMOL, W. Contestable Markets: An Uprising in the Theory of Industry Structure. **The American Economic Review**, v. 72, p. 1-15. 1982.

BAUMOL, W.; SIDAK, J. **Transmission Pricing and Stranded Costs in the Electric Power Industry**. 1 ed. The AEI Press. 1995. 200 p.

BEHR, A. **Production and Efficiency Analysis with R**. 1 ed. New York: Springer, 2015. 230 p.

BOGETOFT, P.; NIELSEN, K. DEA based yardstick competition in natural resource management. In: Helle, F.; Strange, N.; Wichmann, L. (Ed.) **Recent Accomplishments in Applied Forest Economics Research**, Kluwer Academic Publisher, 2003.

BOGETOFT, P.; OTTO, L. **Benchmarking with DEA, SFA, and R**. 1 ed. New York: Springer, 2011. 352 p.

BRESSAN, V.; BRAGA, M. J.; BRESSAN, A. A.; RESESENDO FILHO, M. A. Avaliação de insolvência em cooperativas de crédito: uma aplicação do sistema Pearls. **Revista de Administração Mackenzie**, v. 1.12, n.2, p.113-144. 2011.

CALILI, R. F.; SOUZA, R. C.; SIMÕES, P.F.M. **Audiência 095/2016 -Termo aditivo ao Contrato de Concessão de Distribuição de Energia Elétrica da Enel Distribuição Rio Energia e Serviços S.A.:** contribuições e críticas à metodologia de perdas não técnicas. 2017. Disponível em: <
https://www.ANEEL.gov.br/audiencias-publicas-antigas?p_p_id=participacaopublica_WAR_participacaopublicaportlet&p_p_lifecycle=2&p_p_state=normal&p_p_mode=view&p_p_cacheability=cacheLevelPage&p_p_col_id=column-2&p_p_col_pos=1&p_p_col_count=2&participacaopublica_WAR_participacaopublicaportlet_idDocumento=16530&participacaopublica_WAR_participacaopublicaportlet_tipoFaseReuniao=fase&participacaopublica_WAR_participacaopublicaportlet_jspPage=%2Fhtml%2Fpp%2Fvisualizar.jsp>. Acesso em: 08 nov. 2018.

CHANDEL, P.; THAKUR, T.; SAWALE, B. A.; SHARMA, R. Power Theft: Major Cause of Non-Technical Losses in Indian Distribution Sector. In: 2016 IEEE 7th Power India International Conference (PIICON). **Proceedings...** Bikaner: IEEE, 2016. P. 1-6.

CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision making units. **European Journal of Operational Research**, v.2, n.3, p. 429-444. 1978.

COELLI, T. J.; RAO, D. S. P.; O'DONNELL, C.J.; BATTESE, G.E. **A Introduction to Efficiency and Productivity Analysis**. 2 ed. New York: Springer, 2005. 349 p.

COELLI, T.J.; HENNINGSEN, A. **Stochastic Frontier Analysis – R Package ‘frontier’**. Viena: R Foundation for Statistical Computing, 2020. Disponível em: <<https://www.r-project.org>>.

COOPER, W.; LOVELL, C. A. K. History lessons. **Journal of Productivity Analysis**, v.36, p. 193–200. 2011.

COSTA-CAMPI, M. T.; DAVÍ-ARDERÍUS, D.; TRUJILL-BAUTE, E. The economic impact of electricity losses. **Energy Economics**, v.75, p. 309-322. 2018.

DEPURU, S.; WANG, L.; DEVABHAKTUNI, V. Electricity theft: overview, issues, prevention and smart meter based approach to control theft. **Energy Policy**, v.39, p.1007–15. 2011.

DE SOUZA, M. V. P.; DIALLO, M.; SOUZA, R. C.; BAIDYA, T. K. N. The Cost Efficiency of the Brazilian Electricity Distribution Utilities: A Comparison of

Bayesian SFA and DEA Models. **Mathematical Problems in Engineering**, v.2010, p. 1-20. 2010.

DIJKSTRA, P. T.; HAAN, M. A.; MULDER, M. Design of yardstick competition and consumer prices: Experimental evidence. **Energy Economics**, v.66, p. 261-271. 2017.

EL HAGE, F. S.; RUFIN, C. Context analysis for a new regulatory model for electric utilities in Brazil. **Energy Policy**, v.97, p. 145-154. 2016.

FARRELL, M. J. The Measurement of Productive Efficiency. **Journal of the Royal Statistical Society**, v.120, p. 253-290. 1957.

FARSI, M.; FILIPPINI, M. Regulation and measuring cost-efficiency with panel data models: Application to electricity distribution utilities. **Review of Industrial Organization**, v.25, p.1 – 19. 2004.

FOURIE, J. W. **A strategy for the management of energy losses in a local electricity distribution network**. Petróvia, Africa do Sul, 2004. 126p. Dissertação de Mestrado – Faculdade de Engenharia, University of Petroria, 2005.

GREENE, W. **Econometric Analysis**. 7 ed. New jersey: Pearson-prentice Hall, 2012. 1232 p.

HANEY, A. B.; POLLITT, M. Efficiency analysis of energy networks: An international survey of regulators. **Energy Policy**, v.37, n.12, p. 5814–5830. 2009.

_____. Exploring the determinants of “best practice” *benchmarking* in electricity network regulation. **Energy Policy**, v.39, .n.12, p. 7739-7746. 2011.

_____. International benchmarking of electricity transmission by regulators: A contrast between theory and practice? **Energy Policy**, v.37, p. 267–281. 2013.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTASTICA. **Censo Demográfico de 2010 – Aglomerados Subnormais**. 2010. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/92/cd_2010_aglomerados_subnormais.pdf>

JAMASB, T.; POLLITT, M. International *Benchmarking* and Yardstick Regulation: An Application to European Electricity Utilities. **DAE Working Paper**, n. 0115, Department of Applied Economics, University of Cambridge. 2002.

_____. International benchmarking and regulation: an application to European electricity distribution utilities. **Energy Policy**, v.31, n.12, p. 1609-1622. 2003.

_____. Incentive regulation of electricity distribution networks: Lessons of experience from Britain. **Energy Policy**, v.35, n.12, p. 6163-6187. 2007.

JAMASB, T.; POLLITT, M.; TRIEBES, T. Productivity and efficiency of US gas transmission companies: A European regulatory perspective. **Energy Policy**, v.36, n.9, p. 3398-3412. 2008.

JONDROW, J.; LOVELL, C. A. K.; MATEROV, L.S.; SCHMIDT, P. On the estimation of technical inefficiency in the stochastic frontier production function model. **Journal of Economics**, v.19, p. 233–238. 1982.

KUMBHAKAR, S. C. Production frontiers, panel data and time-varying technical inefficiency, **Journal of Econometrics**, v.46, p. 201-211. 1990.

KUMBHAKAR, S. C.; LOVELL, C. A. K. **Stochastic Frontier Analysis**. 1 ed. Cambridge: Cambridge University Press, 2000. 346 p.

KUMBHAKAR, S. C.; WANG, H. J.; HORNCastle, A. P. **A Practitioner's Guide to Stochastic Frontier Analysis Using Stata**. 1 ed. Cambridge: Cambridge University Press: Cambridge, 2015. 374 p.

KUMBHAKAR, S.C.; LIEN, G. Yardstick Regulation of Electricity Distribution-Disentangling Short-run and Long-run Inefficiencies. **Energy Journal**, v.38, n.5, p. 17-38. 2017.

KUOSMANEN, T; KORTELAINE, M. Stochastic non-smooth envelopment of data: Semi-parametric frontier estimation subject to shape constraints. **Journal of Productivity Analysis**, v.38, p. 1-18. 2010.

KUOSMANEN, T. Stochastic semi-nonparametric frontier estimation of electricity distribution networks: Application of the StoNED method in the Finnish regulatory model. **Economic Energy**, v.34, p. 2189-2199. 2012.

KUOSMANEN, T.; SAASTAMOINEN, A.; SIPIILAINEN, T. What is the best practice for *benchmark* regulation of electricity distribution? Comparison of DEA, SFA and StoNED methods. **Energy Policy**, v.61, p. 740-750, 2013.

LACHTERMACHER, G. **Pesquisa Operacional na Tomada de Decisões: Modelagem em Excel**. 5 ed. Rio de Janeiro: LTC, 2016. 204 p.

LEITE, D.; PESSANHA, J.; SIMÕES, P.; CALILI, R.; SOUZA, R. A Stochastic Frontier Model for Definition of Non-Technical Loss Target. **Energies**, v.13, n. 12 p.3227. 2020. Disponível em: < <https://www.mdpi.com/1996-1073/13/12/3227/htm>> Acesso em: 20 jul. 2020.

LOWRY, M. N.; GETACHEW, L. Statistical benchmarking in utility regulation: Role, standards and methods. **Energy Economics**, v.37, n. 4, p. 1323-1330. 2009.

MARQUES, E.; GOMES, S.; GONÇALVES, R.; MOYA, D. T. E.; CAZZOLATO, D.; FERREIRA, M. P. **Assentamentos precários no Brasil urbano**. 1 ed. Brasília: Ministério das Cidades, 2007.

MILLARD, R.; EMMERTON, M. Non-Technical Losses – How do other countries tackle the problem? In: 2009 AMEU Convention. **Proceedings...** Porto Elizabeth, Africa do Sul: AMEU, 2009.

MIRANDA, A. P. M.; MUNIZ, J. O.; LIMA, R. K.; PAES, V. G. F.; MAYOR, A. S.; MOTA, F. R.; PLATERO, K. A. S.; MARTINS, L. P. B. **Desenvolvimento de Metodologia “Mapas de Percepção de Riscos” para a análise da variável socioeconômica ‘violência’ no diagnóstico de perdas não técnicas de energia nos municípios de São Gonçalo e Duque de Caxias (RJ)”** – Relatório Final PD-0383-0059/2016. Niterói: Universidade Federal Fluminense, 2016.

NEPAL, R.; MENEZES, F.; JAMASB, T. Network regulation and regulatory institutional reform: Revisiting the case of Australia. **Energy Policy**, v.73, p. 259-268. 2014.

OLSON, J. A.; SCHMIDT, P.; WALDMAN, D. M. A Monte Carlo study of estimators of the stochastic frontier production function. **Journal of Econometrics**, v.13, n.1, p. 67-82. 1980.

PAGE, S. E. **The Model Thinker: What You Need to Know to Make Data Work for You**. New York: Basic Books, 2018. 448p.

PENIN, C. **Combate, prevenção e otimização das perdas comerciais de energia elétrica**. São Paulo, 2008. 227p. Tese de Doutorado – Escola Politecnica, Universidade São Paulo, 2008.

QIAN, J., SICKLES, R. C. **Stochastic frontiers with bounded inefficiency**. Working Papers. Shanghai: Shanghai Jiao Tong University, 2008.

R CORE TEAM. **R: A language and environment for statistical computing**. Viena : R Foundation for Statistical Computing, 2020. Disponível em: <<https://www.r-project.org>>. Acesso em: 20 jul. 2020.

RITA, R.; MARQUES, V.; COSTA, A. L.; CHAVES, I. M.; GOMES, J.; PAULINO, P. Efficiency performance and cost structure of Portuguese energy “utilities” – Non-parametric and parametric analysis. **Energy**, v.155, p. 35-45. 2018.

SENYONGA, L.; BERGLAND, O. Impact of High-Powered Incentive Regulations on Efficiency and Productivity Growth of Norwegian Electricity Utilities. **Energy Journal**, v.39, n.5, p. 231-255. 2018.

SHLEIFER, A. A Theory of yardstick competition. **Randon Jornal of Economics**, v.16, n.3, p. 319-327. 1985.

SIMÕES, P. F. M.; SOUZA, R. C.; CALILI, R. F.; PESSANHA, J. F. M. Analysis and Short-term Predictions of non-Technical Loss of Electric Power Based on Mixed Effects Models. **Socio-Economic Planning Sciences**. v.71, 2020.

SICKLES, R. C.; SONG, W.; ZELENYUK, V. **Econometric Analysis of Productivity: Theory and Implementation in R**. Working Paper - Centre for Efficiency and Productivity Analysis. Brisbane: University of Queensland, 2018.

SINGH A. Power sector reform in India: current issues and prospects. **Energy Policy**, v. 34, p.2480-90. 2006.

SMITH, T.B. Electricity theft: a comparative analysis. **Energy Policy**, v.32, n.18, p. 2067-2076, 2004.

SOUZA, M. V. P.; SOUZA, R. C.; PESSANHA, J. F. M. The cost efficiency of the Brazilian electricity distribution utilities: A comparative study with DEA and SFA models. **Gestão Produção**, v.17, p.653–667. 2010.

STEVENSON, R. E. Likelihood functions for generalized stochastic frontier estimation. **Journal of Econometrics**, v.13, n.1, p. 57-66. 1980.

TAKAMINE, E. **Seleção de modelos automáticos para dados em painel: aplicação as perdas não técnicas das distribuidoras de energia elétricas brasileiras**. Rio de Janeiro, 2020. Dissertação de Mestrado - Departamento de Metrologia, Pontificia Universidade Católica do Rio de Janeiro (PUC-Rio), 2020 - *Mimeo*.

VERGARA, S. C. **Metodologia do trabalho científico**. São Paulo: Cortez, 2002.

VON HIRSCHHAUSEN, C. R.; CULLMANN, A.; KAPPELER, A. Efficiency analysis of German electricity distribution utilities–non–parametric and parametric tests. **Applied Economics**, v.38, n. 21, p. 2553–2566. 2006.

WINTHER, T. Electricity theft as a relational issue: A comparative look at Zanzibar, Tanzania, and the Sunderban Islands, India. **Energy for Sustainable Development**, v.16, p. 111-119. 2012.

WOOLDRIDGE, J.M. **Econometric Analysis of Cross Section and Panel Data**. 2 ed. Boston: MIT Press, 2010. 1064 p.

WORLD BANK. **Reducing Technical and Non-Technical Losses in the Power Sector**. Background Paper for the World Bank Group Energy Sector Strategy. Washington: World Bank. 2009. Disponível em: < <https://openknowledge.worldbank.org/bitstream/handle/10986/20786/926390WP0Box3800in0the0power0sector.pdf?sequence=1&isAllowed=y> >. Acesso em: 09 nov.18.

YAKUBUAB, O.; BABU, N. C.; ADJEIB, O. Electricity theft: Analysis of the underlying contributory factors in Ghana. **Energy Policy**, v.123, n.9, p. 611-618. 2018.

YURTSEVEN, C. The causes of electricity theft: An econometric analysis of the case of Turkey. **Utilities Policy**, v.37, p. 70-78. 2015.

ZANINI, A. **Regulação Econômica no Setor Elétrico Brasileiro: Uma Metodologia Para Definição de Fronteiras de Eficiência e Cálculo do Fator X para Empresas Distribuidoras de Energia Elétrica**. Rio de Janeiro, 2004. 147p. Tese de Doutorado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). 2004.

Anexo I – Variáveis exógenas utilizadas nos modelos de eficiência

| | vio | admd | lixo.u | pob | PIB.PC | gini | sub2 | inad | Mbr.Mb1Mbr | inad,pf | Tempo | Constante |
|----|-----|------|--------|-----|--------|------|------|------|------------|---------|-------|-----------|
| 1 | X | | | | | | | | | | X | X |
| 2 | X | X | | | | | | | | | X | X |
| 3 | X | | X | | | | | | | | X | X |
| 4 | X | | | X | | | | | | | X | X |
| 5 | X | | | | X | | | | | | X | X |
| 6 | X | | | | | X | | | | | X | X |
| 7 | X | | | | | | X | | | | X | X |
| 8 | X | | | | | | | X | | | X | X |
| 9 | X | | | | | | | | X | | X | X |
| 10 | X | | | | | | | | | X | X | X |
| 11 | X | X | X | | | | | | | | X | X |
| 12 | X | X | | X | | | | | | | X | X |
| 13 | X | X | | | X | | | | | | X | X |
| 14 | X | X | | | | X | | | | | X | X |
| 15 | X | X | | | | | X | | | | X | X |
| 16 | X | X | | | | | | X | | | X | X |
| 17 | X | X | | | | | | | X | | X | X |
| 18 | X | X | | | | | | | | X | X | X |
| 19 | X | X | X | X | | | | | | | X | X |
| 20 | X | X | X | | X | | | | | | X | X |
| 21 | X | X | X | | | X | | | | | X | X |
| 22 | X | X | X | | | | X | | | | X | X |
| 23 | X | X | X | | | | | X | | | X | X |
| 24 | X | X | X | | | | | | X | | X | X |
| 25 | X | X | X | | | | | | | X | X | X |
| 26 | X | X | X | X | X | | | | | | X | X |
| 27 | X | X | X | X | | X | | | | | X | X |
| 28 | X | X | X | X | | | X | | | | X | X |
| 29 | X | X | X | X | | | | X | | | X | X |
| 30 | X | X | X | X | | | | | X | | X | X |
| 31 | X | X | X | X | | | | | | X | X | X |
| 32 | X | X | X | X | X | X | | | | | X | X |
| 33 | X | X | X | X | X | | X | | | | X | X |
| 34 | X | X | X | X | X | | | X | | | X | X |
| 35 | X | X | X | X | X | | | | X | | X | X |
| 36 | X | X | X | X | X | | | | | X | X | X |
| 37 | X | X | X | X | X | X | X | | | | X | X |
| 38 | X | X | X | X | X | X | | X | | | X | X |
| 39 | X | X | X | X | X | X | | | X | | X | X |
| 40 | X | X | X | X | X | X | | | | X | X | X |

| | vio | admd | lixo.u | pob | PIB.PC | gini | sub2 | inad | Mbr.Mb1Mbr | inad,pf | Tempo | Constante |
|-----|-----|------|--------|-----|--------|------|------|------|------------|---------|-------|-----------|
| 81 | | | X | X | | | | | | | X | X |
| 82 | | | X | | X | | | | | | X | X |
| 83 | | | X | | | X | | | | | X | X |
| 84 | | | X | | | | X | | | | X | X |
| 85 | | | X | | | | | X | | | X | X |
| 86 | | | X | | | | | | X | | X | X |
| 87 | | | X | | | | | | | X | X | X |
| 88 | | | X | X | X | | | | | | X | X |
| 89 | | | X | X | | X | | | | | X | X |
| 90 | | | X | X | | | X | | | | X | X |
| 91 | | | X | X | | | | X | | | X | X |
| 92 | | | X | X | | | | | X | | X | X |
| 93 | | | X | X | | | | | | X | X | X |
| 94 | | | X | X | X | X | | | | | X | X |
| 95 | | | X | X | X | | X | | | | X | X |
| 96 | | | X | X | X | | | X | | | X | X |
| 97 | | | X | X | X | | | | X | | X | X |
| 98 | | | X | X | X | | | | | X | X | X |
| 99 | | | X | X | X | X | X | | | | X | X |
| 100 | | | X | X | X | X | | X | | | X | X |
| 101 | | | X | X | X | X | | | X | | X | X |
| 102 | | | X | X | X | X | | | | X | X | X |
| 103 | | | X | X | X | X | X | X | | | X | X |
| 104 | | | X | X | X | X | X | | X | | X | X |
| 105 | | | X | X | X | X | X | | | X | X | X |
| 106 | | | X | X | X | X | X | X | X | | X | X |
| 107 | | | X | X | X | X | X | | X | X | X | X |
| 108 | | | | X | | | | | | | X | X |
| 109 | | | | X | X | | | | | | X | X |
| 110 | | | | X | | X | | | | | X | X |
| 111 | | | | X | | | X | | | | X | X |
| 112 | | | | X | | | | X | | | X | X |
| 113 | | | | X | | | | | X | | X | X |
| 114 | | | | X | | | | | | X | X | X |
| 115 | | | | X | X | X | | | | | X | X |
| 116 | | | | X | X | | X | | | | X | X |
| 117 | | | | X | X | | | X | | | X | X |
| 118 | | | | X | X | | | | X | | X | X |
| 119 | | | | X | X | | | | | X | X | X |
| 120 | | | | X | X | X | X | | | | X | X |

| | vio | admd | lixo.u | pob | PIB.PC | gini | sub2 | inad | Mbr.Mb1Mbr | inad,pf | Tempo | Constante |
|-----|-----|------|--------|-----|--------|------|------|------|------------|---------|-------|-----------|
| 121 | | | | X | X | X | | X | | | X | X |
| 122 | | | | X | X | X | | | X | | X | X |
| 123 | | | | X | X | X | | | | X | X | X |
| 124 | | | | X | X | X | X | X | | | X | X |
| 125 | | | | X | X | X | X | | X | | X | X |
| 126 | | | | X | X | X | X | | | X | X | X |
| 127 | | | | X | X | X | X | X | X | | X | X |
| 128 | | | | X | X | X | X | | X | X | X | X |
| 129 | | | | | X | | | | | | X | X |
| 130 | | | | | X | X | | | | | X | X |
| 131 | | | | | X | | X | | | | X | X |
| 132 | | | | | X | | | X | | | X | X |
| 133 | | | | | X | | | | X | | X | X |
| 134 | | | | | X | | | | | X | X | X |
| 135 | | | | | X | X | X | | | | X | X |
| 136 | | | | | X | X | | X | | | X | X |
| 137 | | | | | X | X | | | X | | X | X |
| .38 | | | | | X | X | | | | X | X | X |
| .39 | | | | | X | X | X | X | | | X | X |
| .40 | | | | | X | X | X | | X | | X | X |
| .41 | | | | | X | X | X | | | X | X | X |
| .42 | | | | | X | X | X | X | X | | X | X |
| .43 | | | | | X | X | X | | X | X | X | X |
| .44 | | | | | | X | | | | | X | X |
| .45 | | | | | | X | X | | | | X | X |
| .46 | | | | | | X | | X | | | X | X |
| .47 | | | | | | X | | | X | | X | X |
| .48 | | | | | | X | | | | X | X | X |
| .49 | | | | | | X | X | X | | | X | X |
| 150 | | | | | | X | X | | X | | X | X |
| 151 | | | | | | X | X | | | X | X | X |
| 152 | | | | | | X | X | X | X | | X | X |
| 153 | | | | | | X | X | | X | X | X | X |
| 154 | | | | | | | X | | | | X | X |
| 155 | | | | | | | X | X | | | X | X |
| 156 | | | | | | | X | | X | | X | X |
| 157 | | | | | | | X | | | X | X | X |
| 158 | | | | | | | X | X | X | | X | X |
| 159 | | | | | | | X | | X | X | X | X |
| 160 | | | | | | | | X | | | X | X |

| | vio | admd | lixo.u | pob | PIB.PC | gini | sub2 | inad | Mbr.Mb1Mbr | inad,pf | Tempo | Constante |
|-----|-----|------|--------|-----|--------|------|------|------|------------|---------|-------|-----------|
| 161 | | | | | | | | X | X | | X | X |
| 162 | | | | | | | | | X | | X | X |
| 163 | | | | | | | | | X | X | X | X |
| 164 | | | | | | | | | | X | X | X |

Fonte: Elaboração Própria

Anexo II – Resultados dos 29 modelos de eficiência selecionados

```
> summary (modelo45)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 14 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error   z value   Pr(>|z|)
(Intercept) 0.0510051 0.0183351   2.7818 0.005405 **
log(admd)    0.0105695 0.0040449   2.6131 0.008973 **
sigmaSq      0.0298753 0.0053609   5.5728 2.507e-08 ***
gamma        0.9622574 0.0071843 133.9393 < 2.2e-16 ***
time         0.0180764 0.0024529   7.3695 1.713e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1133.758

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014
0.8837244 0.8851143 0.8896472 0.8906690 0.8905818 0.8963690 0.8966774 0.8994244
      2015      2016      2017
0.9012293 0.9022501 0.9011935

> summary (modelo48)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 18 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error   z value   Pr(>|z|)
(Intercept) 0.4038395 0.1040524   3.8811 0.0001040 ***
log(admd)    0.0069840 0.0041497   1.6830 0.0923707 .
log(pib)     -0.0398971 0.0117914  -3.3836 0.0007155 ***
sigmaSq      0.0302708 0.0054509   5.5534 2.802e-08 ***
gamma        0.9633400 0.0069622 138.3671 < 2.2e-16 ***
time         0.0175393 0.0023800   7.3693 1.715e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1138.163

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8816992 0.8832348 0.8867254 0.8876905 0.8878951 0.8935647 0.8937356 0.8962807 0.8981171
      2016      2017
0.8990621 0.8989379
```



```

> summary (modelo50)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 16 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept) 0.1232850  0.0208878   5.9022 3.586e-09 ***
log(admd)    0.0107363  0.0039916   2.6897 0.0071507 **
log(sub)     0.0118481  0.0034200   3.4644 0.0005315 ***
sigmaSq      0.0239408  0.0043734   5.4742 4.396e-08 ***
gamma        0.9526999  0.0088342 107.8424 < 2.2e-16 ***
time         0.0198620  0.0028067   7.0768 1.476e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1139.194

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8977301 0.8990420 0.9026051 0.9040710 0.9042801 0.9088273 0.9094805 0.9117928 0.9135575
      2016      2017
0.9144321 0.9141227

> summary (modelo51)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 18 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept) 0.1644349  0.0359920   4.5687 4.909e-06 ***
log(admd)    0.0107965  0.0040458   2.6686 0.0076177 **
log(inad)    0.0319580  0.0086716   3.6854 0.0002284 ***
sigmaSq      0.0284490  0.0050268   5.6595 1.518e-08 ***
gamma        0.9610498  0.0072635 132.3119 < 2.2e-16 ***
time         0.0186457  0.0024422   7.6348 2.262e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1140.381

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8862206 0.8875356 0.8920849 0.8932586 0.8932189 0.8988368 0.8992231 0.9018675 0.9037534
      2016      2017
0.9047074 0.9037052

```

```

> summary(modelo112)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 14 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  0.1395146   0.0318105   4.3858 1.156e-05 ***
log(pob)      0.0188188   0.0075696   2.4861 0.012915 *
log(inad)     0.0285092   0.0087062   3.2746 0.001058 **
sigmaSq       0.0289432   0.0052807   5.4809 4.231e-08 ***
gamma         0.9616716   0.0073046  131.6536 < 2.2e-16 ***
time          0.0189708   0.0024578   7.7188 1.175e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1139.855

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8843698 0.8856722 0.8898184 0.8911098 0.8911604 0.8967966 0.8971354 0.8997773 0.9018444
      2016      2017
0.9028219 0.9025258

> summary(modelo113)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 19 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  0.0323480   0.0161375   2.0045 0.045014 *
log(pob)      0.0314002   0.0080193   3.9156 9.018e-05 ***
log(mblmbr)  -0.0122654   0.0041171  -2.9791 0.002891 **
sigmaSq       0.0286258   0.0047360   6.0443 1.500e-09 ***
gamma         0.9611524   0.0069953  137.4006 < 2.2e-16 ***
time          0.0205632   0.0025451   8.0796 6.496e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1138.987

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8823969 0.8839073 0.8885563 0.8899951 0.8901415 0.8959772 0.8964745 0.8994054 0.9015383
      2016      2017
0.9027309 0.9022600

```

```

> summary(modelol18)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 22 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  0.3825693  0.1289282   2.9673 0.0030042 **
log(pob)     0.0181039  0.0095078   1.9041 0.0568949 .
log(pib)     -0.0407844  0.0150902  -2.7027 0.0068777 **
log(mblmbr)  -0.0131200  0.0039850  -3.2924 0.0009935 ***
sigmaSq      0.0294257  0.0052289   5.6275 1.828e-08 ***
gamma        0.9627818  0.0069859  137.8186 < 2.2e-16 ***
time         0.0208158  0.0024840   8.3800 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1142.399

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8782046 0.8800763 0.8842082 0.8855045 0.8859483 0.8920221 0.8925149 0.8955815 0.8976462
      2016      2017

```

```

> summary(modelol129)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 18 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  0.4290410  0.1172977   3.6577 0.0002545 ***
log(pib)     -0.0460032  0.0128525  -3.5793 0.0003445 ***
sigmaSq      0.0315891  0.0056789   5.5626 2.658e-08 ***
gamma        0.9648858  0.0065820  146.5954 < 2.2e-16 ***
time         0.0189029  0.0021057   8.9770 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1136.88

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8760690 0.8779302 0.8815982 0.8825261 0.8828865 0.8890644 0.8893520 0.8923013 0.8941927
      2016      2017
0.8954428 0.8956796

```

```

> summary(modelo130)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 19 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  0.3899108  0.1218054   3.2011 0.0013691 **
log(pib)     -0.0314954  0.0138964  -2.2664 0.0234239 *
log(gini)     0.1283984  0.0338790   3.7899 0.0001507 ***
sigmaSq      0.0312428  0.0056534   5.5264 3.268e-08 ***
gamma        0.9653453  0.0065457 147.4772 < 2.2e-16 ***
time         0.0174564  0.0021659   8.0595 7.658e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1144.484

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8806916 0.8821016 0.8857993 0.8867500 0.8868527 0.8925204 0.8926605 0.8953536 0.8971833
      2016      2017
0.8981572 0.8978593

> summary(modelo131)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 22 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  0.4504444  0.1125742   4.0013 6.299e-05 ***
log(pib)     -0.0396284  0.0127226  -3.1148 0.001841 **
log(sub)     0.0135699  0.0046268   2.9329 0.003358 **
sigmaSq      0.0249939  0.0046388   5.3880 7.125e-08 ***
gamma        0.9551346  0.0086553 110.3529 < 2.2e-16 ***
time         0.0209772  0.0027609   7.5979 3.010e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1140.907

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8901816 0.8920051 0.8947407 0.8961675 0.8969264 0.9017332 0.9024429 0.9049664 0.9068258
      2016      2017
0.9079969 0.9089846

```

```
> summary(modelol133)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 19 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
              Estimate Std. Error   z value   Pr(>|z|)
(Intercept)  0.5009432  0.1197621    4.1828 2.879e-05 ***
log(pib)     -0.0567415  0.0132994   -4.2665 1.986e-05 ***
log(mblmbr)  -0.0109527  0.0040366   -2.7133 0.006661 **
sigmaSq      0.0311053  0.0056120    5.5427 2.979e-08 ***
gamma        0.9648100  0.0066081  146.0037 < 2.2e-16 ***
time         0.0215322  0.0023422    9.1931 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1140.646

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8719586 0.8742658 0.8786388 0.8797139 0.8803365 0.8869791 0.8875760 0.8910486 0.8930103
      2016      2017
0.8946578 0.8948354
1
```

```
> summary(modelol135)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 18 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
              Estimate Std. Error   z value   Pr(>|z|)
(Intercept)  0.3971798  0.1100099    3.6104 0.0003057 ***
log(pib)     -0.0240924  0.0128604   -1.8734 0.0610166 .
log(gini)    0.1306123  0.0324878    4.0203 5.811e-05 ***
log(sub)     0.0124553  0.0039525    3.1513 0.0016256 **
sigmaSq      0.0251885  0.0047075    5.3508 8.758e-08 ***
gamma        0.9566787  0.0081877  116.8437 < 2.2e-16 ***
time         0.0192774  0.0025268    7.6292 2.362e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1148.762

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8941117 0.8954629 0.8983372 0.8997453 0.9001838 0.9045912 0.9051134 0.9074261 0.9092171
      2016      2017
0.9101108 0.9104593
```

```

> summary (modelol136)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 20 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  0.4447167  0.1127163   3.9455 7.965e-05 ***
log(pib)     -0.0281134  0.0126820  -2.2168 0.026637 *
log(gini)     0.1214107  0.0281073   4.3195 1.564e-05 ***
log(inad)     0.0259129  0.0088401   2.9313 0.003376 **
sigmaSq       0.0301495  0.0054310   5.5514 2.834e-08 ***
gamma         0.9645085  0.0067236  143.4508 < 2.2e-16 ***
time         0.0181769  0.0022128   8.2144 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1148.967

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8823138 0.8836991 0.8875399 0.8886153 0.8887573 0.8943669 0.8946006 0.8972831 0.8991834
      2016      2017
0.9001558 0.8998460

> summary (modelol137)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 19 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  0.4685502  0.1292779   3.6244 0.0002897 ***
log(pib)     -0.0422245  0.0146266  -2.8868 0.0038914 **
log(gini)     0.1501965  0.0337155   4.4548 8.396e-06 ***
log(mblmbr)  -0.0142522  0.0041161  -3.4625 0.0005351 ***
sigmaSq       0.0308155  0.0056831   5.4223 5.882e-08 ***
gamma         0.9656443  0.0065951  146.4183 < 2.2e-16 ***
time         0.0205498  0.0023463   8.7584 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1150.843

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8752040 0.8771043 0.8817378 0.8828757 0.8832590 0.8894528 0.8899647 0.8933034 0.8952053
      2016      2017
0.8966492 0.8961670

```

```

> summary (modelol139)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 16 iterations:
cannot find a parameter vector that results in a log-likelihood value
larger than the log-likelihood value obtained in the previous step

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  0.4369346  0.1042204   4.1924  2.76e-05 ***
log(pib)     -0.0207266  0.0121995  -1.6990  0.0893240 .
log(gini)     0.1217460  0.0318848   3.8183  0.0001344 ***
log(sub)      0.0112015  0.0037931   2.9531  0.0031462 **
log(inad)     0.0242158  0.0085261   2.8402  0.0045084 **
sigmaSq       0.0249029  0.0045236   5.5051  3.69e-08 ***
gamma         0.9567090  0.0083732  114.2578 < 2.2e-16 ***
time          0.0199014  0.0025062   7.9407  2.01e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1152.657

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8938055 0.8951503 0.8982954 0.8997839 0.9002198 0.9047229 0.9053130 0.9076779 0.9095433
      2016      2017
0.9104600 0.9107289

> summary (modelol140)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 17 iterations:
cannot find a parameter vector that results in a log-likelihood value
larger than the log-likelihood value obtained in the previous step

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  0.5304857  0.1119181   4.7399  2.138e-06 ***
log(pib)     -0.0395697  0.0127502  -3.1035  0.0019127 **
log(gini)     0.1553761  0.0344287   4.5130  6.392e-06 ***
log(sub)      0.0143560  0.0037786   3.7993  0.0001451 ***
log(mblmbr)  -0.0158449  0.0041304  -3.8361  0.0001250 ***
sigmaSq       0.0235771  0.0043777   5.3858  7.214e-08 ***
gamma         0.9547482  0.0086780  110.0189 < 2.2e-16 ***
time          0.0235093  0.0026715   8.8001 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1156.428

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8913114 0.8932066 0.8969207 0.8986070 0.8994743 0.9041567 0.9051536 0.9080617 0.9099306
      2016      2017
0.9113552 0.9116578
|

```



```

> summary (modelol42)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 18 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  0.5406520  0.1121244   4.8219 1.422e-06 ***
log(pib)     -0.0348655  0.0129950  -2.6830 0.0072966 **
log(gini)     0.1467316  0.0331235   4.4298 9.431e-06 ***
log(sub)      0.0136100  0.0038979   3.4916 0.0004801 ***
log(inad)     0.0170891  0.0085490   1.9990 0.0456116 *
log(mblmbr)  -0.0140014  0.0041245  -3.3947 0.0006871 ***
sigmaSq       0.0234053  0.0041958   5.5782 2.430e-08 ***
gamma         0.9546546  0.0086598  110.2402 < 2.2e-16 ***
time          0.0235566  0.0026874   8.7656 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1158.294

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8918121 0.8936380 0.8974351 0.8991617 0.8999854 0.9046680 0.9056729 0.9085404 0.9104520
      2016      2017
0.9118311 0.9120915

> summary (modelol45)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 14 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  0.1881965  0.0252672   7.4483 9.459e-14 ***
log(gini)     0.1471407  0.0290682   5.0619 4.151e-07 ***
log(sub)      0.0126046  0.0038253   3.2950 0.0009841 ***
sigmaSq       0.0251445  0.0046613   5.3943 6.879e-08 ***
gamma         0.9563241  0.0083513  114.5114 < 2.2e-16 ***
time          0.0199249  0.0024778   8.0413 8.889e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1146.773

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8950096 0.8962883 0.8998344 0.9012722 0.9015242 0.9060649 0.9066911 0.9092096 0.9109773
      2016      2017
0.9119689 0.9117893

```



```

> summary(modelol46)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 14 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error   z value   Pr(>|z|)
(Intercept) 0.2044432 0.0365493  5.5936 2.224e-08 ***
log(gini)    0.1405324 0.0314329  4.4709 7.790e-06 ***
log(inad)    0.0274128 0.0086213  3.1796 0.001475 **
sigmaSq      0.0301759 0.0053950  5.5933 2.228e-08 ***
gamma        0.9642628 0.0066863 144.2152 < 2.2e-16 ***
time         0.0187715 0.0021869  8.5837 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1146.651

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8827280 0.8840713 0.8887129 0.8897845 0.8897552 0.8955407 0.8958825 0.8987758 0.9006556
      2016      2017
0.9017390 0.9008100
|

> summary(modelol47)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 18 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error   z value   Pr(>|z|)
(Intercept) 0.0996070 0.0248246  4.0124 6.010e-05 ***
log(gini)    0.1740531 0.0320999  5.4222 5.886e-08 ***
log(mblmbr) -0.0116721 0.0039414 -2.9614 0.003062 **
sigmaSq      0.0319115 0.0060040  5.3151 1.066e-07 ***
gamma        0.9663925 0.0064828 149.0704 < 2.2e-16 ***
time         0.0205023 0.0022774  9.0025 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1146.072

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8738658 0.8756309 0.8812380 0.8823566 0.8824372 0.8888144 0.8894325 0.8929799 0.8948528
      2016      2017
0.8963748 0.8950717

```

```

> summary (modelol150)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 15 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error   z value   Pr(>|z|)
(Intercept)  0.1678416  0.0263196   6.3771 1.805e-10 ***
log(gini)    0.1679996  0.0309518   5.4278 5.706e-08 ***
log(sub)     0.0120241  0.0036862   3.2619 0.001107 **
log(mblmbr) -0.0116667  0.0040373  -2.8898 0.003855 **
sigmaSq      0.0256004  0.0047042   5.4420 5.269e-08 ***
gamma        0.9578899  0.0080889  118.4205 < 2.2e-16 ***
time         0.0227482  0.0026020   8.7427 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1151.184

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8875054 0.8892134 0.8939260 0.8955425 0.8959896 0.9010829 0.9020915 0.9052686 0.9071222
      2016      2017
0.9086019 0.9080499

> summary (modelol152)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 18 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error   z value   Pr(>|z|)
(Intercept)  0.2453364  0.0389031   6.3063 2.857e-10 ***
log(gini)    0.1564702  0.0302910   5.1656 2.397e-07 ***
log(sub)     0.0123161  0.0036432   3.3805 0.0007234 ***
log(inad)    0.0222514  0.0087148   2.5533 0.0106713 *
log(mblmbr) -0.0095104  0.0039196  -2.4264 0.0152512 *
sigmaSq      0.0245428  0.0031952   7.6813 1.575e-14 ***
gamma        0.9563395  0.0064340  148.6393 < 2.2e-16 ***
time         0.0229352  0.0025969   8.8319 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1154.332

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8904409 0.8920523 0.8965934 0.8982904 0.8987575 0.9036498 0.9046686 0.9076804 0.9095782
      2016      2017
0.9109552 0.9105427

```

```

> summary (modelol54)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 13 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept) 0.0895341  0.0207021   4.3249 1.526e-05 ***
log(sub)     0.0144739  0.0045152   3.2056 0.001348 **
sigmaSq      0.0247707  0.0045265   5.4724 4.441e-08 ***
gamma        0.9539090  0.0086322  110.5061 < 2.2e-16 ***
time         0.0226808  0.0024726   9.1728 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1135.749

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8902773 0.8920983 0.8959971 0.8975326 0.8980578 0.9031769 0.9041542 0.9071097 0.9089644
      2016      2017
0.9103902 0.9106424

> summary (modelol56)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 13 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept) 0.0697326  0.0229065   3.0442 0.002333 **
log(sub)     0.0147528  0.0046368   3.1817 0.001464 **
log(mblmbr) -0.0071829  0.0040863  -1.7578 0.078781 .
sigmaSq      0.0250841  0.0045673   5.4921 3.971e-08 ***
gamma        0.9548436  0.0086553  110.3195 < 2.2e-16 ***
time         0.0247236  0.0025528   9.6847 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1137.33

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8842589 0.8864336 0.8910834 0.8928005 0.8935288 0.8990133 0.9003099 0.9037231 0.9056736
      2016      2017
0.9074675 0.9075965

```

```

> summary (modelol57)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 18 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept) 0.1572640  0.0331977   4.7372 2.167e-06 ***
log(sub)     0.0142707  0.0036540   3.9055 9.402e-05 ***
log(inadpf)  0.0206093  0.0086271   2.3889  0.0169 *
sigmaSq      0.0244993  0.0043519   5.6296 1.806e-08 ***
gamma        0.9537286  0.0082701  115.3223 < 2.2e-16 ***
time         0.0201742  0.0027585   7.3136 2.601e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1138.418

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8945790 0.8961101 0.8998323 0.9009507 0.9012331 0.9062645 0.9069431 0.9094657 0.9111769
      2016      2017
0.9122250 0.9120790

> summary (modelol59)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 19 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept) 0.1397878  0.0345381   4.0474 5.180e-05 ***
log(sub)     0.0134573  0.0036024   3.7357 0.0001872 ***
log(inadpf)  0.0234521  0.0089290   2.6265 0.0086269 **
log(mblmbr) -0.0083093  0.0039143  -2.1228 0.0337704 *
sigmaSq      0.0247915  0.0045110   5.4958 3.890e-08 ***
gamma        0.9547434  0.0086863  109.9140 < 2.2e-16 ***
time         0.0222181  0.0028258   7.8627 3.759e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1140.659

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8896139 0.8914669 0.8960825 0.8972582 0.8976405 0.9031864 0.9041114 0.9071078 0.9088684
      2016      2017
0.9102466 0.9097890

```

```

> summary(modelol60)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 16 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept) 0.1153043  0.0316141   3.6472 0.0002651 ***
log(inad)    0.0317121  0.0087978   3.6045 0.0003127 ***
sigmaSq      0.0304256  0.0051338   5.9266 3.094e-09 ***
gamma        0.9634122  0.0067078  143.6250 < 2.2e-16 ***
time         0.0210140  0.0021919   9.5869 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1136.947

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8772441 0.8790832 0.8842083 0.8853158 0.8854777 0.8919714 0.8925858 0.8959409 0.8979077
      2016      2017
0.8993895 0.8987761

> summary(modelol63)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 16 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept) 0.0468362  0.0297466   1.5745  0.11537
log(mblmbr) -0.0082587  0.0039256  -2.1038  0.03540 *
log(inadpf)  0.0201874  0.0088712   2.2756  0.02287 *
sigmaSq      0.0321334  0.0058936   5.4523 4.973e-08 ***
gamma        0.9652959  0.0067455  143.1020 < 2.2e-16 ***
time         0.0201020  0.0024760   8.1187 4.711e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1134.436

panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8727047 0.8746582 0.8803340 0.8810641 0.8811034 0.8880416 0.8885944 0.8921131 0.8939316
      2016      2017
0.8954720 0.8942357

```

```

> summary (modelol64)
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged
Iterative ML estimation terminated after 16 iterations:
log likelihood values and parameters of two successive iterations
are within the tolerance limit

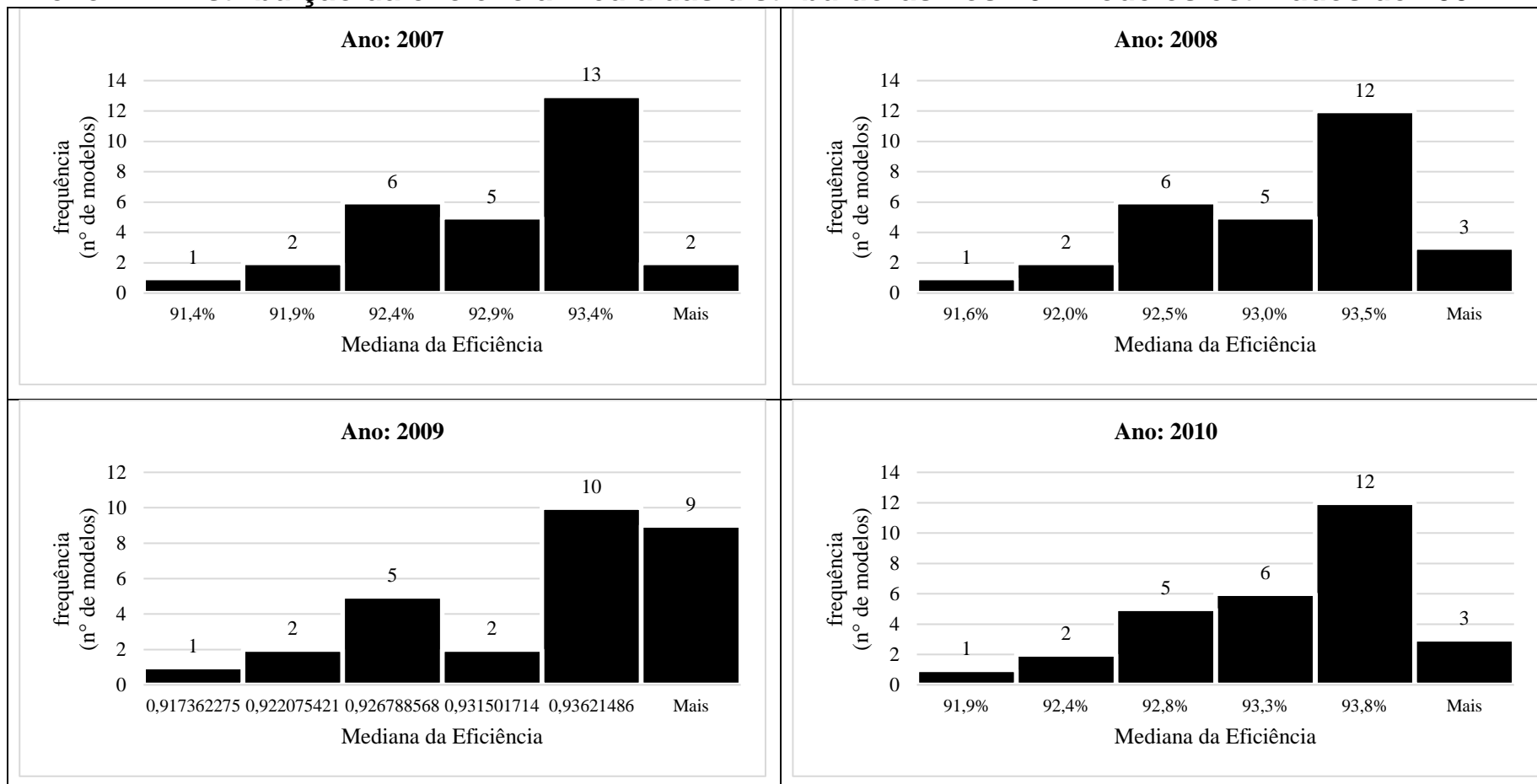
final maximum likelihood estimates
      Estimate Std. Error  z value  Pr(>|z|)
(Intercept) 0.0574955  0.0289076   1.9889   0.04671 *
log(inadpf) 0.0163994  0.0087402   1.8763   0.06061 .
sigmaSq      0.0318009  0.0057679   5.5134 3.520e-08 ***
gamma        0.9645792  0.0067237 143.4592 < 2.2e-16 ***
time         0.0185109  0.0024294   7.6195 2.547e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 1132.256

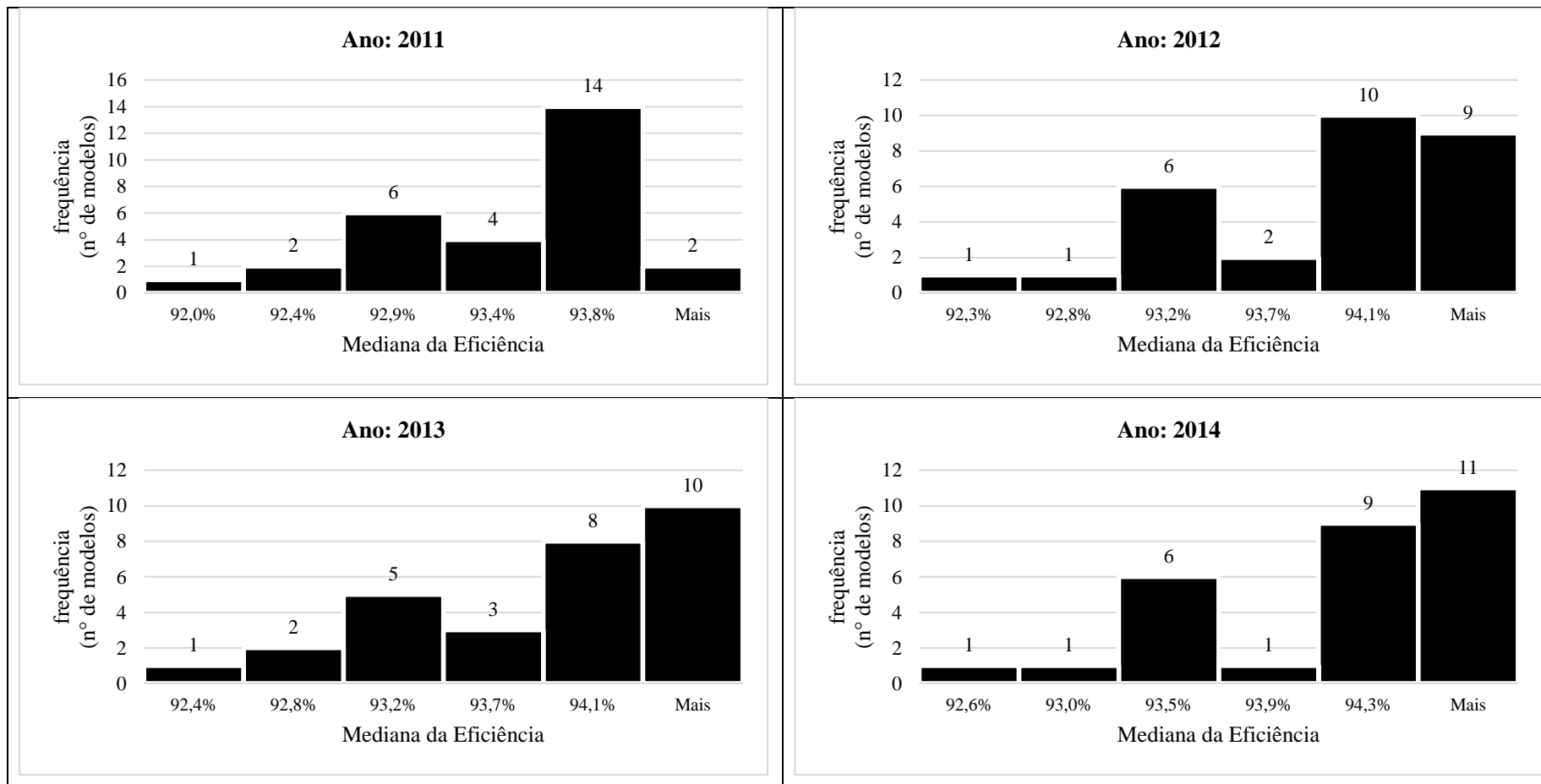
panel data
number of cross-sections = 62
number of time periods = 11
total number of observations = 644
thus there are 38 observations not in the panel

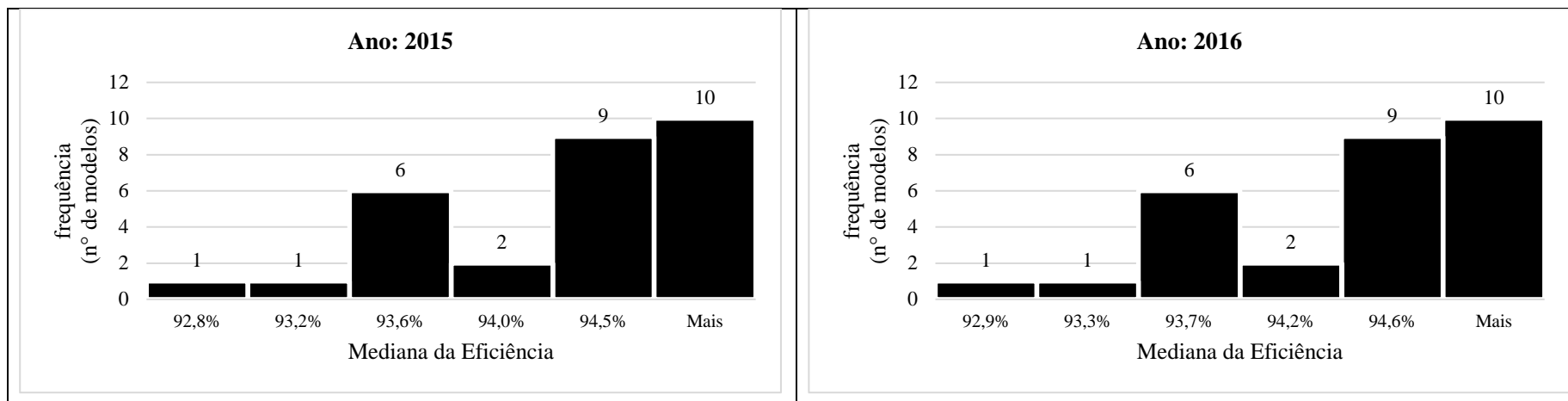
mean efficiency of each year
      2007      2008      2009      2010      2011      2012      2013      2014      2015
0.8775351 0.8791982 0.8841271 0.8847911 0.8847394 0.8912500 0.8915634 0.8946709 0.8964514
      2016      2017
0.8977224 0.8967266

```

Anexo III – Distribuição da eficiência média das distribuidoras nos 164 modelos estimados de 2007-17







Fonte: Elaboração própria

Anexo IV – Ganho de eficiência das distribuidoras no combate as PNT de 2007 a 2017

| Distribuidora | 2007 | 2017 | Variação da eficiência |
|----------------------|-------------|-------------|-------------------------------|
| Amazonas | 43% | 50% | 7% |
| CEA | 51% | 57% | 7% |
| CEAL | 65% | 70% | 5% |
| Light | 66% | 71% | 5% |
| Cepisa | 70% | 75% | 5% |
| Celipa | 67% | 72% | 5% |
| Energisa RO | 70% | 75% | 5% |
| Enel RJ | 77% | 80% | 4% |
| CEEE | 77% | 80% | 4% |
| Eletroacre | 81% | 84% | 3% |
| EDP ES | 82% | 85% | 3% |
| Cemar | 84% | 87% | 3% |
| Energisa MS | 85% | 88% | 3% |
| Celpe | 83% | 86% | 2% |
| EDP SP | 84% | 87% | 2% |
| Iguaçu Energia | 88% | 90% | 2% |
| Sulgipe | 89% | 91% | 2% |
| Energisa MT | 88% | 90% | 2% |
| Energisa PB | 89% | 91% | 2% |
| Cemig | 89% | 91% | 2% |
| Enel SP | 90% | 92% | 2% |
| Energisa SE | 91% | 93% | 2% |
| Enel GO | 92% | 93% | 1% |
| RGE Sul | 92% | 93% | 1% |
| Coelba | 92% | 93% | 1% |
| DEMEI | 93% | 94% | 1% |
| Cooperaliança | 92% | 93% | 1% |
| Eletrocar | 93% | 94% | 1% |
| CEB | 93% | 94% | 1% |
| Enel CE | 94% | 95% | 1% |
| CSPE | 94% | 95% | 1% |
| RGE | 94% | 95% | 1% |
| CPEE | 94% | 95% | 1% |
| CPFL Paulista | 94% | 95% | 1% |
| Celesc | 94% | 95% | 1% |
| ELFSM | 95% | 96% | 1% |
| CPFL Piratininga | 95% | 96% | 1% |
| Energisa TO | 96% | 96% | 1% |
| UHENPAL | 95% | 96% | 1% |
| Energisa BO | 95% | 96% | 1% |
| DME-PC | 96% | 96% | 1% |

| Distribuidora | 2007 | 2017 | Variação da eficiência |
|----------------------|-------------|-------------|-------------------------------|
| Cosern | 96% | 97% | 1% |
| Copel | 96% | 96% | 1% |
| Elektro | 96% | 97% | 1% |
| CHESP | 96% | 97% | 1% |
| Cocel | 97% | 97% | 1% |
| Forcel | 97% | 97% | 1% |
| BRAGANTINA | 97% | 97% | 1% |
| Energisa SS | 97% | 98% | 1% |
| CPFL Santa Cruz | 97% | 98% | 1% |
| Hidropan | 97% | 97% | 1% |
| EVP | 98% | 98% | 0% |
| NACIONAL | 98% | 98% | 0% |
| Energisa MG | 98% | 99% | 0% |

Fonte: Elaboração Própria