UNIVERSIDADE ESTADUAL PAULISTA - UNESP

Faculdade de Engenharia de Ilha Solteira - FEIS- campus de Ilha Solteira

ADEMIR JOSÉ FAZZIO JUNIOR

Os impactos dos lockdowns sobre os furtos de energia elétrica no Brasil :

Uma perspectiva com análise empírica

Ilha Solteira

2026

Ademir José Fazzio Junior

Os impactos dos lockdowns sobre os furtos de energia elétrica no Brasil :
Uma perspectiva com análise empírica

Dissertação, apresentada à Universidade Estadual Paulista (UNESP), Faculdade de Engenharia de Ilha Solteira - FEIS, Ilha Solteira, para obtenção do título de Mestre em Engenharia Elétrica.

Área de Concentração: Automação

Orientador: Profª. Dra. Mara Lucia Martins Lopes

Coorientador: Profª. Dra. Thays Aparecida de Abreu Santos

Ilha Solteira

2026

FICHA CATALOGRÁFICA
Desenvolvido pelo Serviço Técnico de Biblioteca e Documentação

Fazzio Junior, Ademir José.
F287i Os impactos dos lockdowns sobre os furtos de energia elétrica no Brasil: uma perspectiva com análise empírica / Ademir José Fazzio Junior. -- Ilha Solteira: [s.n.], 2026
53 f. : il.

Dissertação (Mestrado em Engenharia Elétrica) - Universidade Estadual Paulista (UNESP), Faculdade de Engenharia, Ilha Solteira. Área de conhecimento: Automação, 2026

Orientador: Mara Lúcia Martins Lopes
Coorientador: Thays Aparecida de Abreu Santos
Inclui bibliografia

1. Covid-19. 2. Diferenças-em-diferenças. 3. Econometria. 4. Justiça tarifária. 5. Perdas não-técnicas. 6. Setor elétrico.

Sandra Montibeller - CRB-8/060/2025

IMPACTO POTENCIAL DESTA PESQUISA

O potencial impacto desta pesquisa reside na sua capacidade de subsidiar formulações de políticas públicas de justiça tarifária e incentivos federais focados na redução de tarifas para populações vulneráveis, atacando a origem socioeconômica dos furtos de energia de forma estrutural. Além disso, ao mensurar o avanço das Perdas Não-Técnicas (PNTs) diante de restrições de mobilidade, o estudo instrumentaliza distribuidoras e órgãos reguladores com evidências para construir planos de contingência em cenários de crises futuras, blindando o equilíbrio financeiro do setor e garantindo o fornecimento elétrico estável para infraestruturas críticas, como hospitais. Por fim, o trabalho gera um impacto metodológico ao fornecer um direcionamento claro de como explorar e manipular a rica base de dados do SAMP da ANEEL, abrindo caminho para que novas pesquisas ampliem o monitoramento energético e a governança regulatória no Brasil.

POTENTIAL IMPACT OF THIS RESEARCH

The potential impact of this research lies in its capacity to inform public policies on tariff justice and federal incentives aimed at reducing electricity rates for vulnerable populations, structurally addressing the socioeconomic roots of energy theft. Furthermore, by quantifying the surge in Non-Technical Losses (NTLs) under mobility restrictions, the study equips utilities and regulatory bodies with empirical evidence to construct contingency plans for future crises, protecting the sector's financial balance and ensuring a stable power supply for critical infrastructure, such as hospitals. Finally, this work delivers a significant methodological impact by providing a clear framework for exploring and manipulating ANEEL's rich SAMP database, paving the way for future research to enhance energy monitoring and regulatory governance in Brazil.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Ilha Solteira

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: Os impactos dos lockdowns sobre os furtos de energia elétrica no Brasil: uma perspectiva com análise empírica

AUTOR: ADEMIR JOSÉ FAZZIO JUNIOR
ORIENTADORA: MARA LUCIA MARTINS LOPES
COORIENTADORA: THAYS APARECIDA DE ABREU SANTOS

Aprovado como parte das exigências para obtenção do Título de Mestre em Engenharia Elétrica, área: Automação pela Comissão Examinadora:

gov.br Documento assinado digitalmente
THAYS APARECIDA DE ABREU SANTOS
Data: 11/08/2026 08:05:15-0300
Verifique em https://validar.iti.gov.br

Profa. Dra. THAYS APARECIDA DE ABREU SANTOS (Participaçao Presencial)
Departamento de Matemática / UNESP / Câmpus de Ilha Solteira - FEIS

gov.br Documento assinado digitalmente
ANDREIA DA SILVA SANTOS DE FARIA
Data: 11/08/2026 10:33:12-0300
Verifique em https://validar.iti.gov.br

Dra. ANDRÉIA DA SILVA SANTOS DE FARIA (Participaçao Presencial)
Departamento de Engenharia / UNESP / Câmpus de Rosana - FEC

gov.br Documento assinado digitalmente
CRISTIANO TOREZZAN
Data: 31/08/2026 15:56:21-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. CRISTIANO TOREZZAN (Participaçao Virtual)
Faculdade de Ciências Aplicadas / UNICAMP / Universidade Estadual de Campinas

Ilha Solteira, 03 de agosto de 2026.

AGRADECIMENTOS

Aos meus pais, por me criarem com valores, por me darem suporte na vida e nos estudos, e por sempre me motivarem a ser melhor. À minha família que sempre esteve ao meu lado e acreditou em mim. Às minhas orientadoras, que tornaram minha jornada acadêmica um caminho cheio de aprendizado e conhecimento. E, por fim, aos meus próprios esforços, que me trouxeram até aqui.

Não faltaram motivos para desistir, mas sempre houve um motivo para continuar, e este era eu mesmo.

RESUMO

As Perdas Não-Técnicas (PNTs) de energia elétrica representam um desafio complexo para concessionárias e órgãos reguladores, sendo historicamente impulsionadas por vulnerabilidades socioeconômicas e falta de justiça tarifária. Com a degradação econômica ocasionada pela pandemia de COVID-19, surgiu o risco iminente de agravamento desse fenômeno, o que poderia comprometer infraestruturas críticas em momentos de alta dependência do sistema elétrico. Esta dissertação teve como objetivo central investigar se as políticas de restrição de mobilidade (lockdowns) da COVID-19 provocaram um aumento causal nos níveis de PNTs no Brasil. Para isso, estruturou-se um painel de dados anuais e trimestrais a partir do Sistema de Acompanhamento de Informações de Mercado para Regulação Econômica (SAMP) da ANEEL. A análise econométrica empregou modelos de regressão com efeitos fixos associados a técnicas de pareamento multivariado, permitindo a estimação do impacto das restrições sanitárias por meio de um modelo de Diferenças-em-Diferenças (DiD 2x2) e um Estudo de Evento. Os resultados revelam que a defasagem do salário mínimo frente às tarifas aumenta as perdas em cerca de 13%, consolidando a justiça tarifária como determinante central. Adicionalmente, o modelo causal comprovou que os estados sob lockdown sofreram um aumento estatisticamente significativo de 2,87 pontos percentuais nas PNTs em relação à energia injetada. O estudo fornece subsídios empíricos diretos para a formulação de políticas públicas de proteção tarifária para populações vulneráveis. Ademais, demonstra a urgência de planos regulatórios de contingência que garantam o equilíbrio financeiro das distribuidoras e a segurança energética da sociedade em futuras crises.

Palavras-Chave: COVID-19; Diferenças-em-Diferenças; Econometria; Justiça Tarifária; Perdas Não-Técnicas; Setor Elétrico.

ABSTRACT

Non-Technical Losses (NTLs) in electricity represent a complex challenge for utilities and regulatory bodies, historically driven by socioeconomic vulnerabilities and a lack of tariff justice. With the economic degradation caused by the COVID-19 pandemic, there was an imminent risk of this phenomenon worsening, which could compromise critical infrastructures during periods of high reliance on the electrical system. This thesis aimed to investigate whether COVID-19 mobility restriction policies (lockdowns) caused a causal increase in NTL levels in Brazil. To achieve this, an annual and quarterly panel dataset was built using data from ANEEL's Market Information Monitoring System for Economic Regulation (SAMP). The econometric analysis employed fixed-effects regression models combined with multivariate matching techniques, allowing the estimation of the impact of sanitary restrictions through a Difference-in-Differences (2x2 DiD) model and an Event Study. The results reveal that the lag of the minimum wage against distribution tariffs increases losses by approximately 13%, consolidating tariff justice as a core determinant. Additionally, the causal model proved that states under lockdown experienced a statistically significant increase of 2.87 percentage points in NTLs relative to injected energy. The study provides direct empirical evidence for the formulation of public policies regarding tariff protection for vulnerable populations. Furthermore, it demonstrates the urgent need for regulatory contingency plans that ensure the financial balance of distribution companies and the energy security of society in future crises.

Keywords: COVID-19; Difference-in-Differences; Econometrics; Non-Technical Losses; Power Sector; Tariff Justice.

LISTA DE ILUSTRAÇÕES

Figura 1 O impacto econômico da Covid-19 de uma perspectiva Histórica 16
Figura 2 Esquema do Sistema de Energia Elétrica 18
Figura 3 Exemplo de Cálculo das Perdas de Energia 19
Figura 4 Série Temporal de 2008 até 2023 da Perdas do Sistema sobre Energia Injetada 19
Figura 5 Perdas de Energia por Região Geográfica 20
Figura 6 Valores em milhões de reais glosados das perdas não técnicas por distribuidora (2023) 20
Figura 7 Fluxograma de criação do Dataset 25
Figura 8 Matriz de Correlação 28
Figura 9 Estudo de Evento - lockdowns 44
Figura 10 Estudo de Evento - separação por intensidade IMLDS 45

LISTA DE TABELAS

Tabela 1 – Descrição dos Campos do Conjunto de Dados 23
Tabela 2 – Descrição dos Campos Finais Transformados 26
Tabela 3 – Resumo estatístico da porcentagem de PNTs por ano de referência 37
Tabela 4 – Resultados da Estimação por Efeitos Fixos (PanelOLS) 40

LISTA DE ABREVIATURAS E SIGLAS

AIE	Agência Internacional de Energia
ANEEL	Agência Nacional de Energia Elétrica
CCEE	Câmara de Comercialização de Energia Elétrica
IBGE	Instituto Brasileiro de Geografia e Estatística
IPEA	Instituto de Pesquisa Econômica Aplicada
OMS	Organização Mundial da Saúde
PNT	Perdas Não-Técnicas
DiD	Difference-in-Differences
ATT	Average Treatment Effect of the Treated

SUMÁRIO

1 INTRODUÇÃO 12

2 DESENVOLVIMENTO 15
2.1 A PANDEMIA DO COVID-19 E SEUS IMPACTOS 15
2.2 SISTEMA ELÉTRICO BRASILEIRO E AS PERDAS NÃO TÉCNICAS . 17

3 MÉTODOS E DADOS 22
3.1 DADOS 22
3.2 REGRESSÃO COM EFEITOS FIXOS 24
3.3 DIFERENÇAS-EM-DIFERENÇAS 29
3.3.1 Formalização do Diferenças-em-Diferenças e ATT 30
3.3.2 Matching 33

4 RESULTADOS E DISCUSSÕES 35
4.1 ESPECIFICAÇÃO DA REGRESSÃO DE EFEITOS FIXOS E RESULTADOS 39
4.2 DIFERENÇAS EM DIFERENÇAS PAREADO E RESULTADOS 41
4.3 DISCUSSÕES 46

5 CONCLUSÃO 49

REFERÊNCIAS 50

1 INTRODUÇÃO

Em um sistema de Energia, com processos de geração, transmissão e distribuição, existe também as perdas de energia. Essas perdas do sistemas podem ser inerente aos processos de geração, transporte e transformação da energia, o qual são conhecidas como Perdas Técnicas; ou essas perdas podem ser por agentes externos a rede, como furtos, fraudes, erros de medição e problemas de faturamento, e normalmente são incluidas sobre o mesmo nome de Perdas Não-Técnicas.

Do ponto de vista do sistema, os furtos de energia são classificados como Perdas Não-Técnicas (PNT) porque são uma energia produzida, consumida, mas não prevista como consumo e não tem origem por meios técnicos (i.e., efeito joule, corrente de fuga, transformação de energia) (Agência Nacional de Energia Elétrica, 2024; Pooya et al., 2025).

Quando fala-se sobre furtos de energia, muitos associam rapidamente ao que é popularmente conhecimento como "gato". Não existe uma origem certa para esse termo, na Engenharia Civil o termo "gato"pode ser usado para se referir a um grampo metálico usado para prender ou travar estruturas, o que se assemelha visualmente as ligações elétricas clandestinas. Pode-se advogar que venha da palavra "gatuno", o que também tem sentido.

Mesmo que a origem correta do termo não seja definida, as ligações clandestinas de energia são um problema sem precedentes dentro dos sistemas de energia elétrica (Agência Nacional de Energia Elétrica, 2024; Pooya et al., 2025).

Segundo a ANEEL (2024) os furtos de energia acontecem majoritariamente na distribuição de baixa tensão. Isso indica que os maiores causadores desse problema dentro do sistema são residências, comércios e industrias de baixa tensão (Agência Nacional de Energia Elétrica, 2024; Savian et al., 2022a).

Partindo do fato de que as PNTs são majoritariamente um problema da população consumidora final, pode-se inferir que as PNTs não são um problema exclusivamente técnico, mas estão associadas ao ambiente socioeconômico em que o indivíduo causador reside. Ou seja, não há um motivador técnico, pois o causador é o próprio consumidor e o mesmo não se submeteria a causar um delito caso não houvesse um motivador externo. O furto de energia é tipificado pelo Código Penal Brasileiro (Art. 155), onde a energia elétrica é tratada legalmente como coisa móvel e dependendo da forma como o desvio é feito, pode ser enquadrado como furto ou até mesmo estelionato, com penas que podem chegar a anos de reclusão

A literatura sobre PNTs apresenta uma ampla gama de estudos voltados à previsão e à identificação dessas perdas. De forma complementar, também se observa o interesse em compreender os fatores associados à sua ocorrência, incluindo indicadores socioeconômicos e variáveis exógenas que possam contribuir para a explicação desse fenômeno. Nesse contexto, pesquisadores e órgãos reguladores têm buscado compreender melhor os fatores relacionados às PNTs e os diferentes cenários em que medidas de mitigação podem ser adotadas, considerando os impactos dessas

perdas sobre o sistema elétrico, como degradação da infraestrutura, aumento dos custos de manutenção e distribuição de energia, entre outros (Agência Nacional de Energia Elétrica, 2024; Savian et al., 2022a; Viegas et al., 2017; Garcia et al., 2023; Silva et al., 2024; Jené-Vinuesa; Aragüés-Peñalba; Sumper, 2024).

Garcia et al. (2023) investigaram a relação entre as perdas não técnicas e variáveis socioeconômicas, empregando análise estatística avançada e a Informação Mútua para a seleção de características em relações não lineares. O estudo foi conduzido com base em unidades demográficas de uma cidade do sul brasileiro, demonstrando que o modelo possui alta replicabilidade em outras regiões. Os achados evidenciaram uma correlação significativa entre as perdas não técnicas e indicadores como renda e nível de escolaridade da população (Garcia et al., 2023).

No âmbito da mitigação de perdas não técnicas, Da Silva et al. (2024) demonstraram a eficácia de uma abordagem baseada em análise espacial e georreferenciamento. Ao aplicar seu método em um estudo de caso brasileiro, que utiliza um índice de contribuição de PNT para setores socioeconômicos, os pesquisadores conseguiram não apenas mapear as perdas, mas também refinar a seleção de variáveis preditoras. Os resultados obtidos indicaram uma melhoria substancial nos índices de Informação Mútua e Correlação de Spearman, validando o potencial da análise espacial como ferramenta robusta para o planejamento de vistorias e a compreensão das causas exógenas do setor (Silva et al., 2024).

Yépez-García et al. (2024) analisaram as perdas de energia na América Latina e no Caribe, mostrando que tanto as perdas técnicas quanto as não técnicas estão fortemente associadas a fatores estruturais e socioeconômicos. O estudo identifica que urbanização, densidade populacional, renda, desigualdade e informalidade econômica influenciam o nível de perdas, especialmente em regiões com baixa capacidade institucional e alto índice de vulnerabilidade social. Os autores destacam ainda que condições socioeconômicas adversas tendem a elevar o risco de furtos e irregularidades, reforçando a necessidade de políticas que combinem melhorias operacionais das distribuidoras com intervenções sociais voltadas à redução da pobreza e da desigualdade (Yépez-García; Mori, 2024).

Babar et al. (2022), em estudo aplicado na Turquia, evidenciam que baixa escolaridade, percepção de tarifas injustas e desconfiança na concessionária aumentam a aceitação social do furto de energia, reforçando o caráter multidimensional das PNTs (Babar; Jamil; Haq, 2022).

Savian et al. (2022) apresentam um panorama abrangente das Perdas Não Técnicas (PNTs) no Brasil, destacando seus principais desafios e as dificuldades enfrentadas pelas distribuidoras na identificação e mitigação das irregularidades. Os autores mostram que as PNTs estão profundamente relacionadas a fatores sociais e econômicos, como informalidade, criminalidade, desigualdade e limitações estruturais das concessionárias, além de ressaltar que regiões com maior vulnerabilidade social tendem a apresentar níveis mais elevados de perdas. O estudo enfatiza que estratégias eficazes de redução das PNTs exigem a combinação de ferramentas tecnológicas, aprimoramento regulatório e ações sociais, reforçando o caráter multidimensional do problema no contexto brasileiro (Savian et al., 2022b).

Briseño e Rojas (2020) realizaram um estudo com dados em painel para identificar os principais fatores associados com as PNTs usando dados de diversos países em desenvolvimento. Junto de uma revisão da literatura, os autores identificaram que educação, população urbana, crime e desemprego estão associados as PNTs, através do modelo de dados em painel com efeitos randômicos. Dos fatores identificados, educação e população urbana estão negativamente relacionados com as PNTs, enquanto crime e desemprego estão positivamente relacionados (Briseño; Rojas, 2020). O trabalho de Briseño e Rojas (2020) confirma a relação das PNTs com características socioeconômicas e sua importância no desenho de soluções para o problema (Briseño; Rojas, 2020).

Visto isso, o presente trabalho tem como objetivo entender se a situação socioeconômica e os *lockdowns* da pandemia do Covid-19 foram responsáveis por pioras no sistema elétrico de energia. Essa hipótese parte do pressuposto que as PNTs são um problema inerentemente socioeconômico, não técnico e, como consequência, choques como a pandemia podem ter motivado a população consumidora a garantir recursos básicos como energia.

Além disso, durante o ano de 2020, diversas medidas foram adotadas em âmbito federal com o objetivo de reduzir os impactos socioeconômicos da pandemia, entre elas a proibição temporária do corte de energia elétrica por inadimplência, a concessão de benefícios e descontos, além de outras ações de caráter social não diretamente relacionadas ao setor elétrico.No contexto deste trabalho, os efeitos dessas medidas não podem ser completamente dissociados daqueles decorrentes dos *lockdowns*, uma vez que muitas dessas ações foram implementadas de forma concomitante e em escala nacional. Dessa forma, os resultados obtidos devem ser interpretados considerando a possível influência conjunta dessas intervenções sobre as variáveis analisadas.

A partir das ações federais, outra hipótese surge. Essas medidas que buscaram aliviar o impacto socioeconômico, e por consequência o avanço das PNTs, foram o suficiente?

Esse trabalho busca realizar uma análise empírica para trazer evidências a essas hipóteses, avaliando o valor por estado das PNTs totais normalizada pela energia cativa injetada total (soma de todas as distribuidoras), reforçando a literatura nacional sobre o tema. A Seção 2 busca introduzir, de forma mais aprofundada, o tema das PNTs, junto da pandemia do Covid-19 e como relacionam-se. Na Seção 3 elabora-se os dados utilizados para realizar esse estudo, junto da formalização matemática e teórica dos métodos usados para estimar os esfeitos. A Seção 4 apresenta o procedimento feito na manipulação dos dados, com uma formalização algébrica das transformações, junto da aplicação dos métodos e seus resultados, finalizando com as discussões. Por fim, o trabalho encerra-se na Seção 5, concluindo o que foi apresentado.

A realização deste estudo é relevante não apenas por buscar mensurar os efeitos socioeconômicos da pandemia sobre as PNTs, mas também por incorporar métodos de inferência causal à análise das perdas de energia elétrica. Até onde se estende o conhecimento dos autores, são escassos os trabalhos que utilizam essa abordagem nesse contexto, de modo que os resultados aqui apresentados podem contribuir para ampliar essa linha de investigação e servir de referência para estudos futuros que aprofundem ou estendam as análises desenvolvidas.

2 DESENVOLVIMENTO

Em 2019 o mundo ainda iria conhecer o que veio a ser uma das maiores pandemias da historia da humanidade. No final de 2019 os primeiros casos começaram a surgir. Em março de 2020, a Organização Mundial da Saúde (OMS) declarou a Covid-19 como pandemia, após a rápida disseminação do vírus. Até maio de 2023, quando a emergência sanitária internacional foi encerrada, o mundo havia registrado mais de 765 milhões de casos e quase 7 milhões de mortes, além de fortes impactos sobre atividade econômica, educação, mobilidade e saúde mental (Pan-American Health Organization (PAHO), 2023). Esses efeitos agravaram desigualdades pré-existentes, ampliaram a pobreza e elevaram a vulnerabilidade social em diversos países, incluindo o Brasil.

Os efeitos da pandemia da Covid-19 foram diversos e amplos, impactando até o setor elétrico. As medidas de distanciamento social e os *lockdowns* estaduais reduziram a atividade econômica, fazendo com que emissões de gases do efeito diminuíssem significativamente. Na Europa, a emissão de NO2, normalmente associada a emissão por queima de combustíveis fosseis e veículos a diesel, reduziu drasticamente seus valores como resultado da mobilidade limitada e da redução do consumo de energia elétrica comercialmente, ao mesmo tempo em que aumentaram o consumo residencial de energia em até 30% (Mofijur et al., 2021). Além disso, houve deterioração indicadores sociais, como um aumento da violência domestica em alguns países e a piora nos quadros de saúde mental da população como um todo, devido ao confinamento social (Mofijur et al., 2021; Krarti; Aldubyan, 2021; Soava et al., 2021; Rouleau; Gosselin, 2021; Sánchez-Úbeda et al., 2022; Bahmanyar; Estebsari; Ernst, 2020). A piora das condições de renda, a elevação do desemprego e a maior precarização podem ter influenciado comportamentos associados à informalidade e ao uso irregular da energia elétrica, sobretudo entre consumidores de baixa tensão (Mofijur et al., 2021).

Visto isso, uma questão surgi, se houve deterioração social em vários aspectos, desde violência, educação, desemprego e renda, isso poderia fazer com que comportamentos associados à informalidade e ao uso irregular da energia elétrica, sobretudo entre consumidores de baixa tensão, aumentasse. Formulando em pergunta, "a pandemia da Covid-19, especialmente o distanciamento social e lockdowns, provocaram aumento das PNTs?".

2.1 A PANDEMIA DO COVID-19 E SEUS IMPACTOS

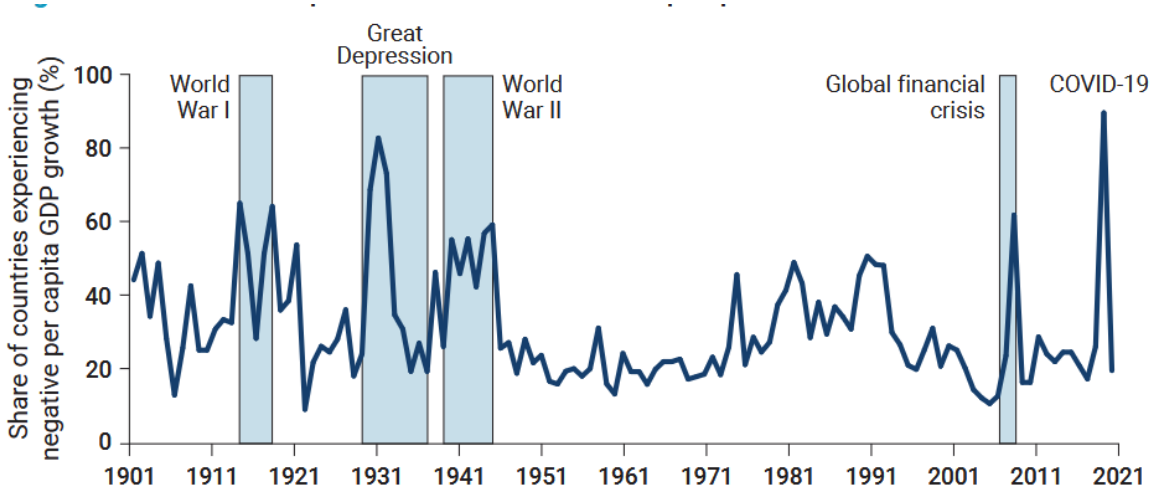
No final do ano de 2019, em dezembro, a OMS foi alertada por surtos de casos de pneumonia na cidade de Wuhan, na província de Hubei, na China. Se tratava de uma nova cepa de coronavírus que ainda não havia sido identificada em humanos. Em janeiro de 2020, as autoridades chinesas confirmaram a nova cepa. Ao todo, sete coronavírus humanos (HCoVs) já foram identificados: HCoV-229E, HCoV-OC43, HCoV-NL63, HCoV-HKU1, SARS-COV (que causa síndrome

respiratória aguda grave), MERS-COV (que causa síndrome respiratória do Oriente Médio) e o mais recente SARS-CoV-2 (que no início ficou conhecido como “novo coronavírus”). A doença causada pelo SARS-CoV-2 foi nomeada COVID-19 (Pan-American Health Organization (PAHO), 2023).

Após 3 anos da declaração inicial, em maio de 2023 a OMS declarou o fim da emergência global da Covid-19. Durante esse período a pandemia não só foi uma emergência de saúde global, mas impactou outras camadas da sociedade, como economia, ambiente, educação e saúde mental (Pan-American Health Organization (PAHO), 2023).

No World Development Report de 2022 compartilhado pelo Banco Mundial (World Bank, 2022), indicou-se que as restrições de mobilidade, bloqueios e outras medidas de saúde pública necessárias para combater a pandemia rapidamente produziram a maior crise econômica global em mais de um século. A crise econômica ainda se agravou quando a pandemia afetou o comportamento dos consumidores, levando a quedas. A atividade econômica contraiu em 2020 em cerca de 90% dos países, excedendo o número de países que viram tais declínios durante duas guerras mundiais, a Grande Depressão da década de 1930, as crises de dívida da economia emergente da década de 1980, e a crise financeira mundial de 2007-09. Em 2020, o primeiro ano da pandemia de COVID-19, a economia global encolheu em aproximadamente 3%, e a pobreza global aumentou pela primeira vez em um geração (World Bank, 2022). A Figura 1 apresenta o impacto econômico de uma perspectiva histórica, no qual mede-se o percentual de países com crescimento do PIB per capito negativo.

Figura 1 – O impacto econômico da Covid-19 de uma perspectiva Histórica



Fonte: World Bank (2022)

Além do impacto apresentado pelo Banco Mundial, Mofijur et al. (2021) mostraram que, devido a pandemia, houve um aumento de desperdício de forma generalizada guiado pelos novos hábitos de consumo e da atividade da área da saúde. Além disso, as emissão de NO2, normalmente associada a emissão por queima de combustíveis fosseis e veículos a diesel, reduziu drasticamente seus valores como resultado da mobilidade limitada e da redução do consumo

de energia elétrica comercialmente, chegando a valores de até 30% de redução (Mofijur et al., 2021).

A implementação de ações para conter a pandemia levaram a uma desaceleração das atividades econômicas, o que diminuiu o consumo de transporte público e de energia, implicando em um aumento da presença de energia renovável e diminuição de geração de energia por métodos térmicos e nucleares (Mofijur et al., 2021). Além disso, houve deterioração indicadores sociais, como um aumento em violência domestica em alguns países e a piora nos quadros de saúde mental da população como um todo, devido ao confinamento social (World Health Organization and others (WHO), 2020).

2.2 SISTEMA ELÉTRICO BRASILEIRO E AS PERDAS NÃO TÉCNICAS

Os sistemas elétricos são estruturados em três segmentos principais: Geração, Transmissão e Distribuição, responsáveis pela produção, transporte e entrega de energia. Ao longo desse percurso, parte da energia injetada não é comercializada, configurando as perdas de energia, que podem ser técnicas ou não técnicas (Agência Nacional de Energia Elétrica, 2024; Pooya et al., 2025; Jené-Vinuesa; Aragüés-Peñalba; Sumper, 2024)

No âmbito regulatório brasileiro, as perdas técnicas são definidas como aquelas decorrentes do fluxo de energia na rede, associadas à resistência elétrica, limitações dos equipamentos, transformação de energia, condições climáticas e o próprio transporte em si (Agência Nacional de Energia Elétrica, 2024; Pooya et al., 2025). Elas são previsíveis e calculadas pela diferença entre os pontos de injeção (energia gerada) e entregue nos pontos de carga, que são as unidades consumidoras diretamente conectadas no sistema de transmissão e as redes de distribuição. Essas perdas são apuradas mensalmente pela CCEE e o seu custo é rateado em 50% para geração e 50% para os consumidores (Agência Nacional de Energia Elétrica, 2024; Pooya et al., 2025).

A ultima etapa do sistema elétrico, a distribuição, é dividida de acordo com os segmentos de rede (alta, média e baixa tensão), subestações, redes, transformadores, ramais de ligação e medidores. Utilizando-se de modelos e de informações simplificadas das redes e equipamentos existentes, como por exemplo, comprimento e bitola dos condutores, potência dos transformadores e energia fornecida às unidades consumidoras, apura-se o percentual de perdas técnicas eficientes sobre a energia injetada, que é a energia elétrica inserida na rede de distribuição para atender aos consumidores, incluindo as perdas (Agência Nacional de Energia Elétrica, 2024).

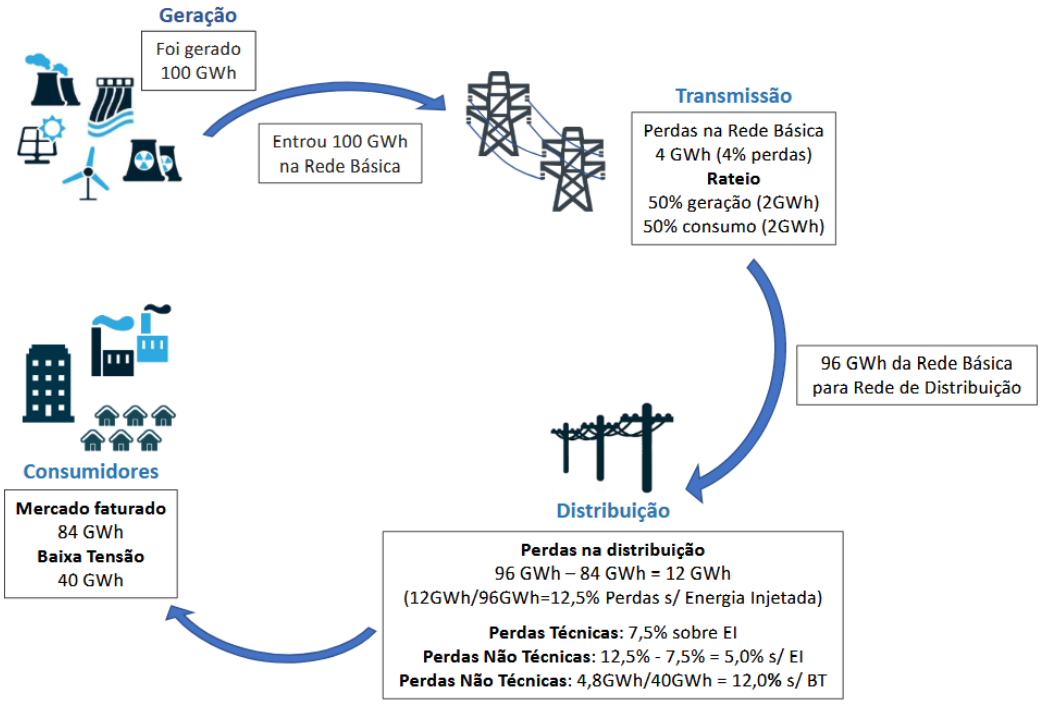
A AIE estima que uma redução mundial de 1% nas perdas técnicas podem vir a salvar até 200 TWh anualmente, no entanto, isso implica em um desafio de engenharia elétrica e de materiais, além de avanços tecnológicos e científicos (Agência Nacional de Energia Elétrica, 2024; Pooya et al., 2025).

As PNTs, por sua vez, correspondem à energia que deixa de ser contabilizada por motivos não relacionados ao desempenho físico da rede. Incluem furtos, irregularidades na medição, acessos não autorizados, erros de leitura e inconsistências cadastrais, apuradas pela diferença

entre as perdas totais e as perdas técnicas (Agência Nacional de Energia Elétrica, 2024; Savian et al., 2022a). Sua magnitude depende da eficácia dos sistemas de supervisão comercial e da capacidade de detecção e tratamento de irregularidades pelas distribuidoras.(Agência Nacional de Energia Elétrica, 2024; Savian et al., 2022a). As PNTs são mais frequentes no segmento de baixa tensão, onde se concentram consumidores residenciais e pequenos comércios, caracterizando-se como um dos principais desafios operacionais da distribuição (Agência Nacional de Energia Elétrica, 2024; Savian et al., 2022a).

A Figura 2 apresenta um esquema do sistema elétrico brasileiro e Figura 3 um exemplo de cálculo para energia gerada, injetada, consumida e as perdas. De forma geral, se 100GWh são gerados e enviados a rede de transmissão, já há uma perda de 4% sendo rateada os custos meio a meio, a energia restante (96 GWh) é enviada para a distribuição. Nesta etapa, as perdas de energia são a diferença entre a faturada e o total injetado (96 GWh - 84 GWh = 12 GWh). Há um valor pré definido de 7,5% do total da energia injetada sendo Perda Técnica, o excedente desse valor é definido como PNTs.

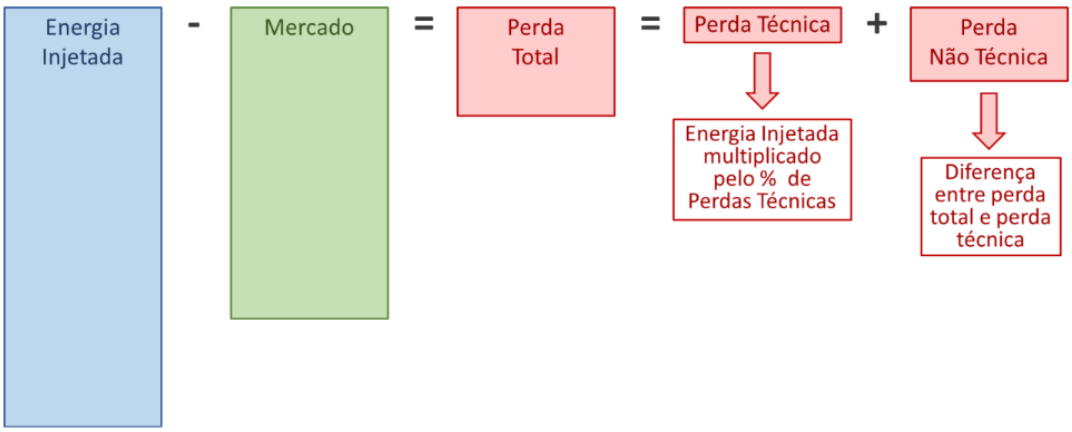
Figura 2 – Esquema do Sistema de Energia Elétrica



Fonte: Agência Nacional de Energia Elétrica (2024)

A série histórica apresentada no relatório de 2024 da ANEEL (Agência Nacional de Energia Elétrica, 2024) sobre Perdas de Energia Elétrica na Distribuição mostra que os números de PNTs vem se mantendo em um patamar bem similar desde 2008, variando entre valores de 6% até 7,5%, no qual o pico de 7,5% aconteceu em 2020, no ano de início da Pandemia do Covid-19. A Figura 4 apresenta os dados mencionados, no qual pode se verificar que os patamares das Perdas Técnicas se mantiveram, considerando que é uma característica física determinística do sistema

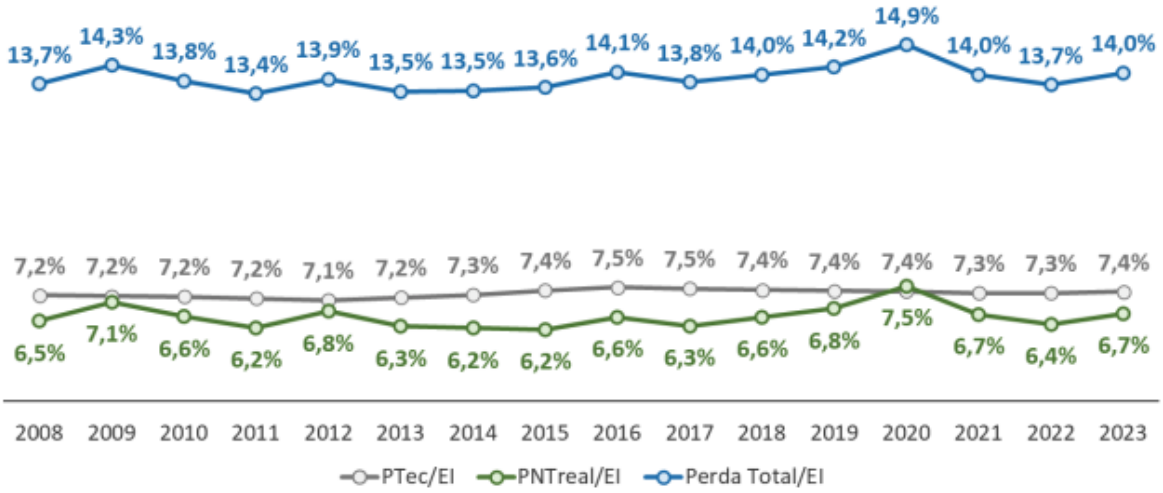
Figura 3 – Exemplo de Cálculo das Perdas de Energia



Fonte: Agência Nacional de Energia Elétrica (2024)

elétrico, porém as PNTs variaram ao longo do tempo.

Figura 4 – Série Temporal de 2008 até 2023 da Perdas do Sistema sobre Energia Injetada

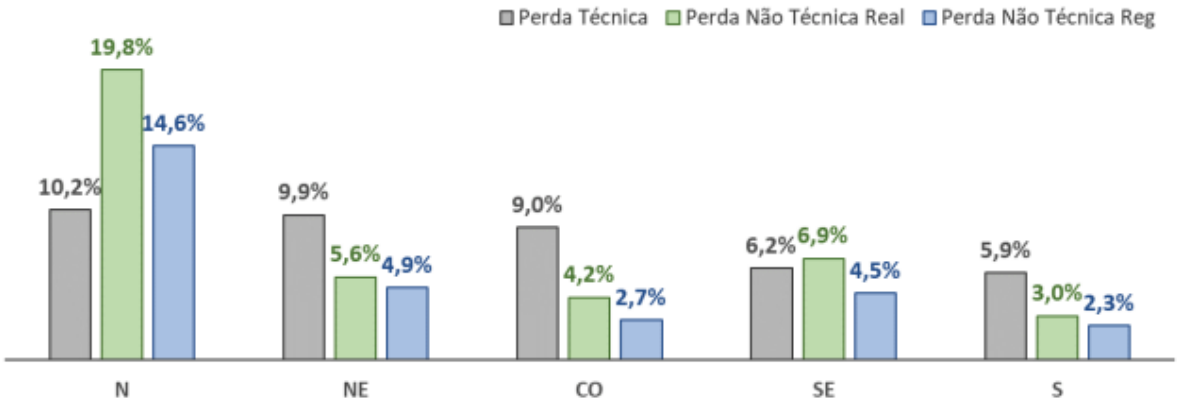


Fonte: Agência Nacional de Energia Elétrica (2024)

Também é disponibilizado no relatório os valores de 2023 das PNTs sobre a energia injetada separada por região geográfica no país. A Figura 5 apresenta os valores mencionados no qual a região que mais apresenta problemas com as PNTs é a região norte, o que está alinhado com o fato de que a concessionária de energia Amazonas Energia apresentou o segundo maior valor percentual sobre as PNTs totais de 2023, ficando atrás apenas da Light SESA, concessionária do Rio de Janeiro (Agência Nacional de Energia Elétrica, 2024).

As duas concessionárias mencionadas representaram 32% de todas as PNTs acumuladas de 2023, mesmo apresentando realidades socioeconômicas totalmente diferentes, como média salarial diferente, níveis de escolarização, desemprego e outros indicadores relacionados a realidade socioeconômica (Agência Nacional de Energia Elétrica, 2024; Brazilian Institute of Geography and Statistics, 2025).

Figura 5 – Perdas de Energia por Região Geográfica

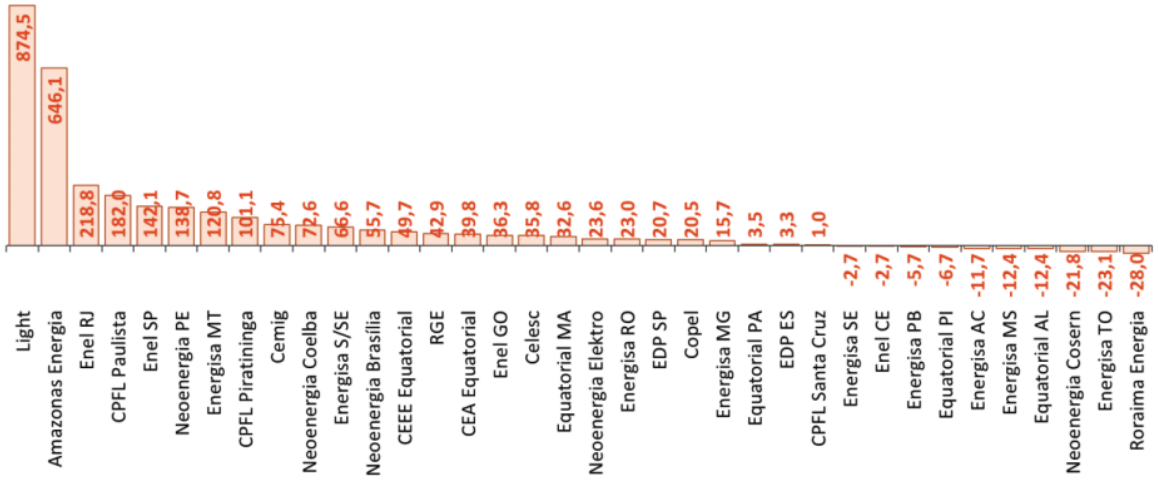


Fonte: Agência Nacional de Energia Elétrica (2024)

Em 2023, as Perdas Técnicas representaram R$ 10,3 Bilhões e as PNTs reais representaram R$ 9,9 Bilhões. Os valores são comparáveis, porém as Perdas Técnicas já consideram os processo eficiente das redes, não abrindo margem para maiores reduções (Agência Nacional de Energia Elétrica, 2024). No entanto, as PNTs representam uma perda que pode ser evitada na grande maioria dos casos e sua existência impacta diretamente na qualidade do fornecimento de energia e contribui para a fatura do consumidor final.

A Figura 6 apresenta os valores glosados pelas distribuidoras de grande porte em 2023 e as duas concessionárias com a maior representatividade nas PNTs também apresentaram os maiores valores glosados. (Agência Nacional de Energia Elétrica, 2024)

Figura 6 – Valores em milhões de reais glosados das perdas não técnicas por distribuidora (2023)



Fonte: Agência Nacional de Energia Elétrica (2024)

Estudos recentes reforçam que as PNTs estão fortemente associadas a variáveis socioeconômicas como educação, desemprego, urbanização e criminalidade (Yépez-García; Mori, 2024; Babar; Jamil; Haq, 2022; Savian et al., 2022b; Briseño; Rojas, 2020; Santiago et al., 2021).

Carr e Thomson (2022) mostram que as PNTs têm causas predominantemente socioeconômicas, relacionadas à informalidade, desigualdade, baixa escolaridade, criminalidade e tarifas elevadas. Os autores destacam que a qualidade da relação entre consumidor e concessionária, incluindo a confiança e percepção de justiça tarifária, é um dos principais determinantes das PNTs, reforçando que sua redução exige não apenas soluções técnicas, mas também ações sociais e institucionais.

A caracterização apresentada evidencia que as PNTs constituem um componente relevante do funcionamento do segmento de distribuição, com repercussões operacionais, regulatórias e financeiras. Dado que sua magnitude depende da efetividade dos mecanismos de medição, fiscalização, controle, bem estar social e justiça tarifária, compreender como eventos inesperados alteram essas condições é essencial para o setor elétrico, especialmente se esses eventos abalam pilares sociais, gerando um efeito composto e em cadeia sobre as PNTs.

3 MÉTODOS E DADOS

Para a realização das análises, aplicações de métodos e investigação da magnitude do impacto, esse estudo utilizou os dados abertos do balanço mensal da ANEEL para avaliar as PNTs durante a pandemia e a linguagem de programação Python, com pacotes de computação numérica e estatística. Esses dados contém informações de diversos setores e categorias, sempre referenciando a data de origem, o CPF/CNPJ a qual o dado está atrelado, o que inclui distribuição de energia, geração, energia injetada, vendida e diversas outras categorias de informações podem ser extraídas.

3.1 DADOS

As informações podem ser encontradas no portal de dados abertos do governo brasileiro (National Electric Energy Agency, 2025b) e apresentam um intervalo de registro de 2003 até o ultimo mês de fechamento. No presente trabalho, os dados foram obtidos em setembro de 2024, logo os dados vão até agosto de 2024. A Tabela 1 apresenta a descrição do conjunto de dados.

Além dos dados do balanço energético, foram coletados outros dados que podem estar associados às PNTs, como é observado em diversas pesquisas no qual urbanização, criminalidade, literacia, desemprego, justiça tarifária e etc (Carr; Thomson, 2022; Yépez-García; Mori, 2024; Viegas et al., 2017; Garcia et al., 2023; Babar; Jamil; Haq, 2022; Savian et al., 2022b; Briseño; Rojas, 2020; Santiago et al., 2021; Agência Nacional de Energia Elétrica, 2024; Savian et al., 2022a). Então, foram coletados dados sobre quantidade de homicídios anuais por estado brasileiro fornecidos pelo atlas da violência, do IPEA, percentual de desemprego, percentual de analfabetismo, literacia (percentual da população maior de 25 anos com colegial, ou equivalente, completo) e os valores da tarifa de energia para consumidor final (códigos B1 e B2 da tarifa de energia), através do IBGE e do conjunto de dados da ANEEL (Brazilian Institute of Geography and Statistics, 2025; National Electric Energy Agency, 2025b).

Cada uma dessas variáveis consideradas tem como objetivo controlar por diferentes fontes de variação que podem influenciar as PNTs. Por exemplo, a quantidade de homicídios representa o nível de crime de um determinado estado brasileiro, indicando que estados com maiores níveis de crime podem ter maiores níveis de PNTs do que estados com menores taxas de crime.

As tarifas de energia também são selecionadas por apresentar alta relação com as PNTs, uma vez que elas são os valores pagos pela a população em relação ao uso da energia elétrica no Brasil, controlando assim a noção de justiça tarifária (Savian et al., 2022a; Agência Nacional de Energia Elétrica, 2024; National Electric Energy Agency, 2025a). Há diferentes grupos de tarifas, para baixa e alta tensão, e outros casos específicos. Neste caso é usado as tarifas de energia do grupo B por representa os valores pagos pela população de baixa tensão, pois como foi mostrado, as PNTs se concentram na baixa tensão, levando a maiores relações com as tarifas

Tabela 1 – Descrição dos Campos do Conjunto de Dados

Nome do Campo	**Tipo do dado**	**Tamanho do Campo**	**Descrição**
DatGeracaoConjuntoDados	Data Simples	-	Data do processamento de carga automática no momento da geração para publicação do conjunto de dados abertos.
NumCPFCNPJ	Cadeia de caracteres	14	Código de identificação do agente (órgão).
NomAgente	Cadeia de caracteres	255	Nome completo do agente.
AnmCompetenciaBalanco	Cadeia de caracteres	6	Ano e mês de competência do balanço.
DscModalidadeBalanco	Cadeia de caracteres	80	Descrição da modalidade do balanço de energia.
DscFluxoEnergia	Cadeia de caracteres	80	Descrição do fluxo de energia.
DscCctBalanco	Cadeia de caracteres	80	Descrição da característica do balanço.
DscClassificacaoAgente	Cadeia de caracteres	80	Descrição da classificação do agente.
AnoReferenciaBalanco	Cadeia de caracteres	4	Ano de referência do balanço.
MesReferenciaBalanco	Cadeia de caracteres	2	Mês de referência do balanço.
DscDetalheBalanco	Cadeia de caracteres	80	Descrição do detalhamento do balanço de energia.
VlrEnergia	Numérico	21,2	Valor da energia em kWh.

Fonte: Próprio autor

pagas por essa parcela (Agência Nacional de Energia Elétrica, 2024; National Electric Energy Agency, 2025a).

Tomando-se os dados coletados, foi feita a manipulação de várias etapas para obter os valores anuais - para cada estado brasileiro, valores que representavam energia, ou contagem, foram somados - de fornecimento cativo de energia, PNTs, total de homicídios, população total, percentual de analfabetismo, percentual de desemprego, literacia e as tarifas de energia B1 e B2 separadas em TUSD (Tarifa de Uso do Sistema de Distribuição) e TE (Tarifa de Energia).

Essa manipulação foi realizada em 3 etapas. A primeira é agrupar os dados da ANEEL anualmente, por estado e categoria, usando um dataset que relaciona o CNPJ de cada distribuidora/concessionária com o estado, e agregando a coluna de valor de energia somando os valores, permitindo assim construir um conjunto de dados em painel, isto é, cada linha representa um ano específico, para uma categoria específica e o respectivo valor total somado.

A segunda etapa é pivotar a coluna de categorias de forma que os valores se tornem uma coluna, ou seja, se antes havia uma tabela de dados variando de 2011 a 2020 com 20 linhas de dados e 3 colunas para ano, categoria e valor, com valores das categorias "Energia Injetada"e "Perdas de Energia", a nova tabela terá 10 linhas com 3 colunas, sendo elas: Ano, Valor de Energia Injetada, Valor de Perdas de Energia; dispensando a necessidade da coluna de categoria. A terceira e ultima etapa é unir conjuntos de dados paralelos do IBGE, IPEA e outros, com a tabela resultante do pivotamento.

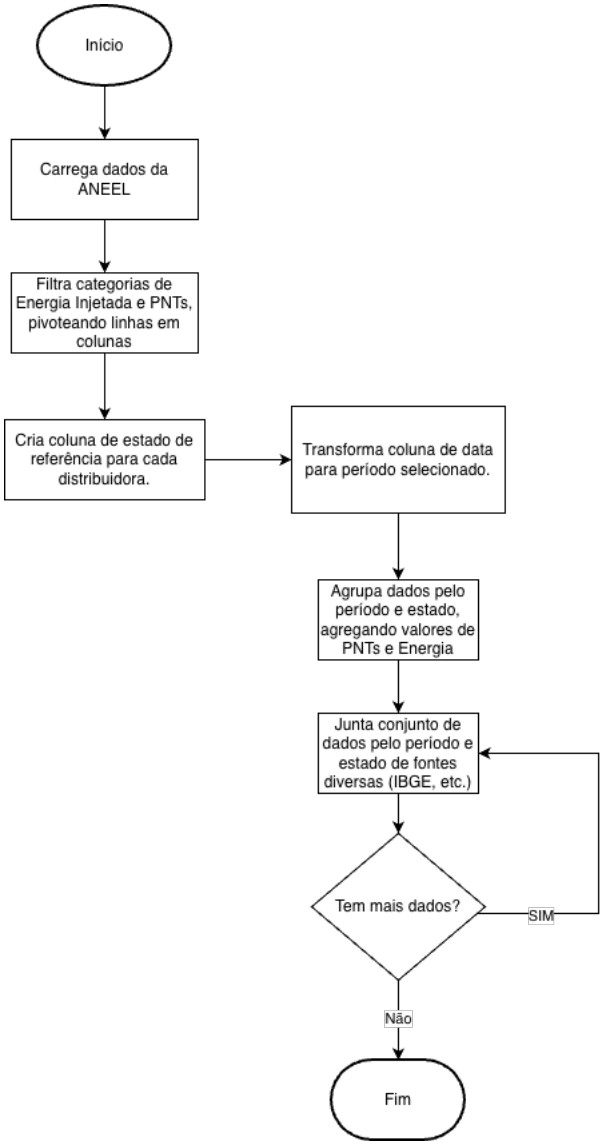
Na sequência, os dados foram transformados ou usados para criar novas variáveis explicativas, sendo estas consumo de energia cativa (Consumo de energia elétrica exclusivamente da concessionária de distribuição que atende sua região) per capita em logaritmo, total de homicídios a cada mil pessoas, e os valores das tarifas de energia em reais por MWh normalizadas pelo salário mínimo do ano de referência. A Figura 7 apresenta um fluxograma dos passos para transformação e seleção dos dados para utilização nos diferentes métodos aplicados, e a 2 apresenta um compilado das colunas, origens e suas unidades.

Os dados organizados de forma anual, considerando os anos de 2016 até 2019, foram utilizados para realizar uma regressão de dados em painel com efeitos fixos para identificar os estimadores independentes que variam no tempo mais relacionados com as PNTs (Huntington-Klein, N., 2025), mitigando efeitos fixos controlando pelo tempo e pelos estados. Finalmente, foram utilizada as variáveis mais explicativas para realizar um pareamento dos estados que não passaram por *lockdowns* com os que passaram, para então aplicar a metodologia de diferenças-em-diferenças para mensurar o impacto dos *lockdowns* (Baker et al., 2025; Callaway; Sant'Anna, 2021; Huntington-Klein, N., 2025).

3.2 REGRESSÃO COM EFEITOS FIXOS

Quando um conjunto de dados é organizado com amostras sendo coletada em diferentes pontos do tempo, formando um conjunto no qual há um mesmo grupo sendo observado no passar

Figura 7 – Fluxograma de criação do Dataset



Fonte: Próprio Autor

do tempo, esse tipo de conjunto é definido como um conjunto de dados em painel, ou dados longitudinais, e é constantemente usado por diferentes áreas para mensurar relações intra-grupos e causalidade (Huntington-Klein, N., 2025; Wooldridge, 2010; Wooldridge, 2025).

Nesse tipo de conjunto, variáveis variam através de i (dimensão de corte transversal) e t (dimensão de corte temporal), no qual N e T representam os tamanhos de quantos clusters transversais e quantas dimensões de tempo e N x T representa o tamanho do conjunto de dados final (Huntington-Klein, N., 2025; Wooldridge, 2010; Wooldridge, 2025).

Com esse tipo de conjunto disponível, é possível incluir efeitos fixos de unidade e tempo em uma equação de Mínimos Quadrados Ordinários, resultando em diferentes tipos de estimador (Wooldridge, 2010; Sommet; Lipps, 2025). Para casos no qual é interessante apenas considerar os efeitos fixos de um grupo específico, ou seja, trabalhando apenas com dados transversais que se tem N unidades para cada J grupos, é possível realizar uma regressão linear com variáveis

Tabela 2 – Descrição dos Campos Finais Transformados

Nome do Campo	**Unidade**	**Campo de Origem**
crime_per_capita	Homicídios/1000 pessoas	População total e Total de homicídios
log_cons_energy_per_capita	log de MWh/total de Pessoas	Energia Cativa Total e População
log_pop	log da População	População total
logTEB1	log MWh/R$	Coluna de tarifa de energia B1 em MWh
logTEB2	log MWh/R$	Coluna de tarifa de energia B2 em MWh
logTUSDB1	log MWh/R$	Coluna de tarifa de uso do sistema de distribuição B1 em MWh
logTUSDB2	log MWh/R$	Coluna de tarifa de uso do sistema de distribuição B2 em MWh
perc_analf	percentual da população	coluna original
perc_desemp	percentual da população trabalhadora	coluna original
perc_lit	percentual da população com 25 anos ou mais	coluna original
SalarTEB1	R$ / MWh/R$	Salário mínimo do Ano e TE B1
SalarTEB2	R$ / MWh/R$	Salário mínimo do Ano e TE B2
SalarTUSDB1	R$ / MWh/R$	Salário mínimo do Ano e TUSD B1
SalarTUSDB2	R$ / MWh/R$	Salário mínimo do Ano e TUSD B2

Fonte: Próprio autor

dummy de grupo ou uma centralização das variáveis. Esse tipo de modelagem expurga efeitos fixos médios de cada cluster, mantendo apenas efeitos intra-grupos (Wooldridge, 2010; Sommet; Lipps, 2025). As Equações 1 e 2, representam respectivamente a modelagem com dummies e com centralização de variáveis, no qual y_{ij} é a variável dependente de , e $\overline{y_j}$ média por grupo j; β o vetor de pesos das variáveis independentes (covariáveis preditoras); x_{ij} representa o vetor de covariáveis preditoras, e $\overline{x_j}$ a média por grupo j; c_j representa variáveis binárias de grupo multiplicadas por um peso, ou seja, são os efeitos médios de cada grupo; e u_{ij} o termo de erro da equação:

$$y_{ij} = \beta \times x_{ij} + c_j + u_{ij}, i = 1, ..., N; j = 1, ..., J \tag{1}$$

$$(y_{ij} - \overline{y_j}) = \beta \times (x_{ij} - \overline{x_j}) + (u_{ij} - \overline{u_j}), i = 1, ..., N; j = 1, ..., J \tag{2}$$

Sommet e Lipps (2025) apresentam de forma intuitiva a aplicação desse modelo, em que

para medir os efeitos de pratica de exercício físico e esportes sobre os quadros de depressão, eles geraram um conjunto de dados sintético e exploram as diferentes modelagens e aplicações que podem ser feitas. Junto disso eles dão instruções de como garantir que a modelagem está correta.

Além da modelagem transversal, pode ser feita modelagens que usam dados longitudinais, controlando pelos efeitos fixos de unidade N (como sexo, região, motivação intrínseca e outras variáveis invariantes no tempo) e tempo T (choques temporais, como chuva em um dia específico que afeta todos igualmente), resultando em um estimador conhecido como Efeitos Fixos de Duas Vias (No inglês, Two Way Fixed Effects) (Huntington-Klein, N., 2025; Wooldridge, 2010; Wooldridge, 2025). Essa modelagem é utilizada para inferir conclusão a cerca das relações intra-grupos e, por vezes, tirar conclusões causais.

No entanto, é importante levar em consideração que essa modelagem sofre com problemas de variáveis de confusão variáveis no tempo e de causalidade reversa (i.e., na simulação de esportes, participantes que aumentam seus quadros de depressão ao longo do tempo, praticam menos esporte, afetando menos a depressão) (Huntington-Klein, N., 2025; Wooldridge, 2010; Wooldridge, 2025; Sommet; Lipps, 2025).

A Equação 3 apresentam a modelagem usada no caso de Efeitos Fixos de Duas Vias, no qual f_t representa os efeitos fixos de tempo, que representam variáveis binárias que assumem valor igual a 1 quando fazem referência ao período específico.

$$y_{it} = x_{it}\beta + c_i + f_t + u_{it}, i = 1, ..., N; t = 1, ..., N \tag{3}$$

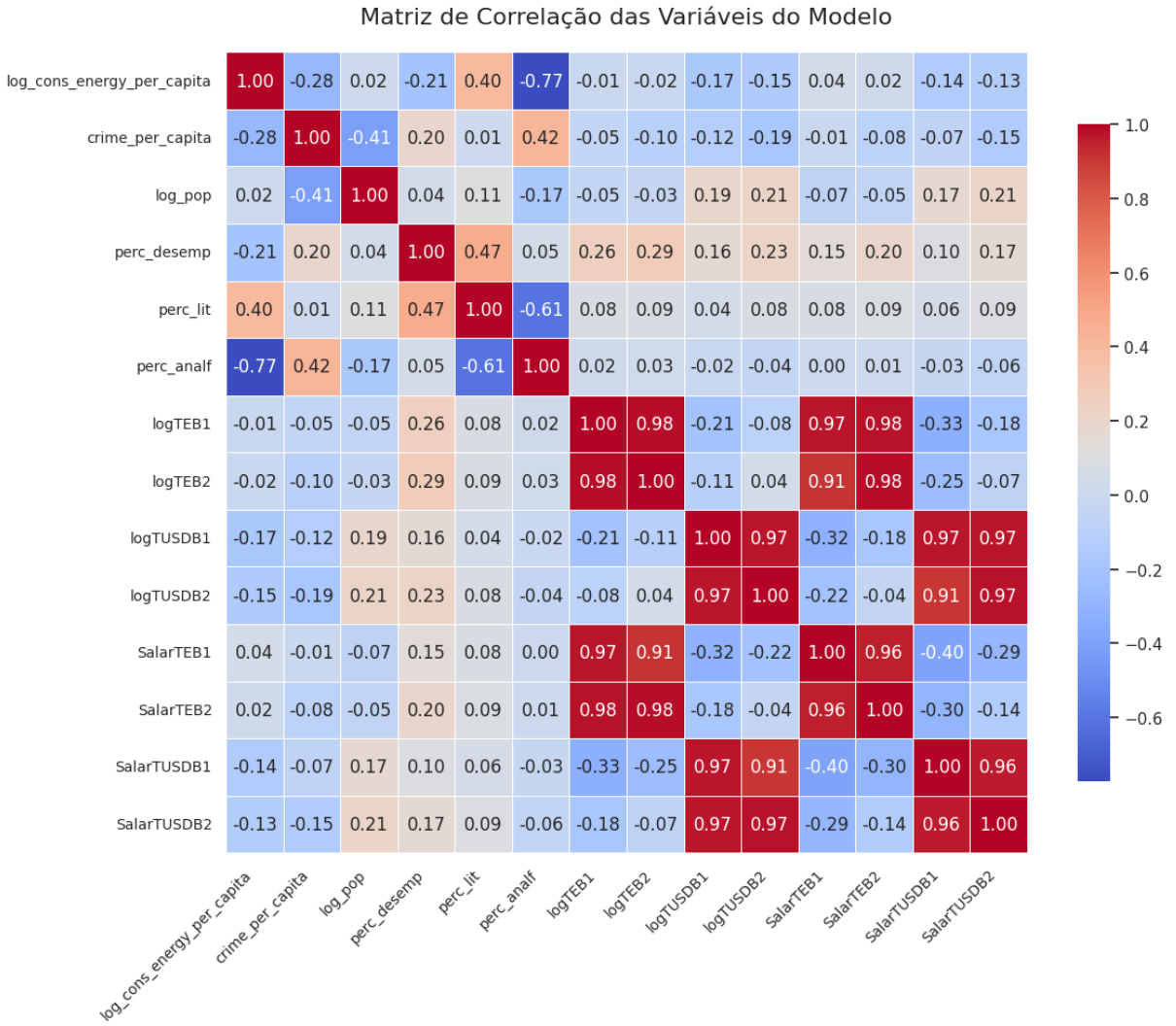
Para incluir variáveis que são variantes no tempo e reduzir possíveis erros com variáveis de confusão, Sommet e Lipps (2025) recomendam o uso de uma centralização dupla, para evitar efeitos de transbordo indesejados das variáveis.

O primeiro passo para realizar a regressão com efeitos fixos (Huntington-Klein, N., 2025; Wooldridge, 2025) foi identificar possíveis fontes de multicolinearidade, considerando que esse é uma potencial fonte de instabilidade e imprecisão na inferência econométrica (Huntington-Klein, N., 2025; Wooldridge, 2010). A literatura recomenda o uso do Fator de Inflação de Variância (VIF) como ferramenta para diagnosticar multicolinearidade entre as variáveis explicativas (Gujarati, 2009; Wooldridge, 2010). Assim, calculou-se o VIF para todas as variáveis do modelo, identificando aquelas que, ao serem removidas, reduziam substancialmente o indicador. Por meio dessa etapa inicial, foram descartadas as variáveis logaritmo da população, tarifas de energia B2, todas as tarifas de energia em log e o percentual de analfabetismo.

Para confirmação do VIF, também foi calculado a matriz de correlação de pearson das variáveis, identificando quais estão mais correlacionadas para eliminação e encontrar conjunto utilizável na inferência (Black; Babin, 2019). A Figura 8 apresenta os valores de correlação. Esse passo colabora na identificação de potenciais fontes de variação similar, ou seja, variáveis com alta correlação de pearson (negativa ou positiva) introduzem a mesma variação no modelo, incluindo muito pouca ou nenhuma nova informação.

De forma geral, deve-se ler os dados de correlação considerando alta correlação valores mais próximos de 1 ou -1. Não há um consenso geral na literatura sobre quais os valores exatos para se considerar correlação relevantes ou não. Então, neste caso, qualquer valor acima de 0.6 ou menor que -0.6 é considerado uma correlação relevantes o suficiente.

Figura 8 – Matriz de Correlação



Fonte: Próprio Autor

Após o calculo de multicolinearidade, realiza-se uma sequência de testes para garantir que a aplicação de um estimador de efeitos fixos tem sentido. O primeiro é o Teste de Agrupabilidade (*F-test for Poolability*), que testa se os interceptos e inclinações são constantes ao longo do tempo e entre indivíduos. O teste baseia-se na comparação da Soma dos Quadrados dos Resíduos (SQR) do Modelo Restrito (Pooled OLS) com o Modelo Irrestrito (Efeitos Fixos ou modelo de coeficientes variáveis) (Huntington-Klein, N., 2025; Wooldridge, 2010; Wooldridge, 2025).

De forma geral, testa a hipótese nula H_0: "os dados podem ser agrupados. O modelo restrito é adequado". Em caso de rejeição da hipótese nula, conclui-se que o agrupamento é inválido.

Em sequência, realiza o Teste de Hausman, que avalia se os coeficientes de um modelo

de Efeitos Aleatórios e de um modelo de Efeitos Fixos são estatisticamente diferentes. A diferença central está na relação entre a heterogeneidade não observada individual (α_i) e as variáveis explicativas (X_{it}). Basicamente, o modelo de Efeitos Fixos: Assume que α_i pode ser correlacionado com X_{it}, e é sempre consistente, mas apaga variáveis que não mudam no tempo (Huntington-Klein, N., 2025; Wooldridge, 2010; Wooldridge, 2025).

A hipótese nula do Teste de Hausman avalia se não há correlação entre os efeitos individuais e os regressores. E consequentemente, ambos os estimadores são eficientes, mas há uma preferência pelo de Efeitos Aleatórios. Em caso de rejeição da hipótese nula, apenas o de Efeitos Fixos é consistente (Huntington-Klein, N., 2025; Wooldridge, 2010; Wooldridge, 2025).

O modelo desenvolvido a partir das variáveis segue a equação 3, no qual as variáveis incluídas são as restantes depois da eliminação. O beneficio de se usar esse modelo antes de realizar a implementação do diferença-em-diferença é a sua capacidade de controlar por variáveis de confusão não observáveis ou mensuráveis em diferentes grupos. Isto é, variáveis como questões culturais, geografia, choques temporais e outras variáveis, que são difíceis de mensurar, são controladas quando controla-se a variabilidade do grupo, permitindo avaliação de variáveis mensuradas (Huntington-Klein, N., 2025; Wooldridge, 2025).

3.3 DIFERENÇAS-EM-DIFERENÇAS

As primeiras menções do diferenças-em-diferenças datam da década de 1840, em que esse método quase-experimental foi usado no Hospital Geral de Viena por Semmelweis para provar que a lavagem das mãos com solução de cal clorada prevenia a febre puerperal (pós-parto). Foram separadas duas clinicas de maternidade, uma com médicos e estudantes de medicina obrigados a higienizar as mãos, e outra com parteiras que não lidavam com os trabalhos da medicina e cadáveres, apenas faziam parto (Baker et al., 2025; Semmelweis, 1983).

O objetivo da separação em duas clinicas era comparar o efeito da higienização antes do trabalho de parto e, foi usando a metodologia de diferenças-em-diferenças (de forma rudimentar comparado com as técnicas mais modernas) que Semmelweis atestou a eficácia da higienização, resultando na primeira clínica do estudo igualar os níveis de mortalidade e patamar histórico das parteiras, provando que a contaminação por partículas cadavéricas era a causa das mortes (Baker et al., 2025; Semmelweis, 1983).

No momento de redação deste trabalho, Diferenças-em-Diferenças é um dos métodos mais mencionados em rascunhos de trabalhos acadêmicos em repositórios abertos, e em revistas de economia e econometria (Baker et al., 2025).

O design básico é o estimador de diferenças-em-diferenças com dois grupos, um recebendo o tratamento e outro não, e dois períodos de tempo, um antes e outro depois do tratamento. A estimativa é dada pela diferença das diferenças entre os períodos de tempo, para cada grupo tratado, sobre uma variável medida (como no caso de Semmelweis, as taxas mortalidades). Isso assumi que as diferenças entre os períodos para cada grupo seria igual, assumindo que

o tratamento não teria ocorrido. Essa suposição é o que a literatura chama de "suposição de tendências paralelas"(Baker et al., 2025; Huntington-Klein, N., 2025).

3.3.1 Formalização do Diferenças-em-Diferenças e ATT

Para iniciar qualquer análise causal, é necessário a definição formal da variável de interesse. Nesse estudo, segue-se a definição de resultados potenciais (Robins, 1986; Rubin, 1974; Baker et al., 2025). Seja $Y_{i,t}(0,0)$ o resultado potencial da unidade i no período t, se a unidade permanece sem tratamento nos dois períodos; e $Y_{i,t}(0,1)$ o resultado potencial da unidade i no período t, se a unidade permanece sem tratamento primeiro período e tratada no segundo (Baker et al., 2025; Huntington-Klein, N., 2025).

No estudo de Semmelweis, $Y_{i,t}(0,0)$ representa um cenário no qual não houve o tratamento de higienização das mãos, e $Y_{i,t}(0,1)$ representa o cenário que houve a separação. Para fins de simplificação, as formalizações modernas consideram $Y_{i,t}(0,0) = Y_{i,t}(0)$ e $Y_{i,t}(0,1) = Y_{i,t}(1)$ (Baker et al., 2025; Huntington-Klein, N., 2025).

Na prática, nunca se é observado $Y_{i,t}(0)$ e $Y_{i,t}(1)$ para uma mesma unidade i. O que existe é uma variável D_i que representa 1 quando uma unidade passou pelo tratamento, e 0 caso contrário. Ou seja, a formalização oficial seria como na Equação 4 (Baker et al., 2025; Huntington-Klein, N., 2025).

$$Y_{it} = (1 - D_i)Y_{i,t}(0) + D_i Y_{i,t}(1) \tag{4}$$

A partir da formalização da variável de interesse, adiciona-se suposições sobre o processo de geração dos dados que são importantes para garantir que o efeito estimado a partir dos modelos é de fato causal (Baker et al., 2025; Huntington-Klein, N., 2025). A primeira suposição que deve-se levar em consideração na metodologia de resultados potenciais é a de "Não Antecipação", definida na Suposição 3.1 (Baker et al., 2025; Huntington-Klein, N., 2025).

Suposição 3.1 (Não Antecipação)**.** *Para toda a unidade tratada i e todos os períodos pré-tratamento t, $Y_{i,t}(1) = Y_{i,t}(0)$*

A Suposição de não antecipação garante que a variável de interesse não foi afetada pelo tratamento antes do próprio tratamento acontecer. De forma intuitiva, é o mesmo que dizer que as taxas de mortalidade no estudo de Semmelweis não foram afetadas por saberem que os médicos passariam a higienizar as mãos (Baker et al., 2025; Huntington-Klein, N., 2025; Semmelweis, 1983).

Essa suposição parece fazer pouco sentido no caso de Semmelweis, no entanto, em casos de políticas publicas de alta capilaridade, pode ser que a população mude seu comportamento esperado a partir do conhecimento do tratamento, "sujando"as estimativas feitas posteriormente (Baker et al., 2025; Huntington-Klein, N., 2025).

A metodologia de resultados potenciais funciona a partir da definição de um efeito causal para todas as unidades em todos os períodos de tempo, $Y_{i,t}(1) = Y_{i,t}(0)$ (Baker et al., 2025; Huntington-Klein, N., 2025). Basicamente representa o impacto do tratamento nas unidades tratadas ou o que teria representado no caso das unidades não tratadas. Porém, não é possível observar os dois resultados ao mesmo tempo para cada unidade. Então, defini-se a métrica de "Efeito Médio do Tratamento nos Tratados", definida na Equação 5 (Baker et al., 2025; Huntington-Klein, N., 2025).

$$ATT(t) = E[Y_{it}(1) - Y_{i,t}(0)|D_i = 1] = E[Y_{it}|D_i = 1] - E[Y_{i,t}(0)|D_i = 1] \tag{5}$$

Com a definição do ATT, é possível definir a suposição que é utilizada como ferramenta de identificação do efeito causal, a suposição de "Tendências Paralelas"(Baker et al., 2025; Huntington-Klein, N., 2025). Essa suposição define que na ausência de um tratamento, a evolução (tendência) média do resultado é a mesma entre os grupos tratados e não tratados (Baker et al., 2025; Huntington-Klein, N., 2025).

Suposição 3.2 (Tendências Paralelas 2x2)**.** *A mudanças média para $Y_{i,t=2}(0)$ de $Y_{i,t=1}(0)$ é a mesma entre tratados e não tratados; isto é:*

$$E[Y_{i,t=2}(0)|D_i = 1] - E[Y_{i,t=1}(0)|D_i = 1] = E[Y_{i,t=2}(0)|D_i = 0] - E[Y_{i,t=1}(0)|D_i = 0] \tag{6}$$

Se a suposição de tendências paralelas for respeitada, então deriva-se de forma simples $E[Y_{i,t=2}(0)|D_i = 1]$ das quantidades observadas, ou seja, seguindo a Equação 7 (Baker et al., 2025; Huntington-Klein, N., 2025).

$$E[Y_{i,t=2}(0)|D_i = 1] = E[Y_{i,t=1}|D_i = 1] + (E[Y_{i,t=2}|D_i = 0] - E[Y_{i,t=1}|D_i = 0]) \tag{7}$$

A partir das suposições e equações derivadas até aqui é possível construir a estimativa ATT(2) (Efeito médio do tratamento nos tratados no pós tratamento), que representa o estimador do Diferença-em-Diferenças, o que segue a Equação 8 (Baker et al., 2025; Huntington-Klein, N., 2025).

$$\begin{aligned} ATT(2) &= \overbrace{E[Y_{i,t=2}|D_i = 1]}^{E[Y_{i,t=2}(1)|D_i=1]} - \overbrace{(E[Y_{i,t=1}|D_i = 1] + (E[Y_{i,t=2}|D_i = 0] - E[Y_{i,t=1}|D_i = 0]))}^{E[Y_{i,t=2}(0)|D_i=1]} \\ &= \underbrace{(E[Y_{i,t=2}|D_i = 1] - E[Y_{i,t=1}|D_i = 1])}_{\Delta E[Y|D_i=1]} - \underbrace{(E[Y_{i,t=2}|D_i = 0] - E[Y_{i,t=1}|D_i = 0])}_{\Delta E[Y|D_i=0]} \end{aligned} \tag{8}$$

A partir da Equação 8, confirma-se que a metodologia do Diferença-em-Diferenças é a diferença das diferenças entre os períodos de tempo, para cada grupo, sobre uma variável medida (Baker et al., 2025; Huntington-Klein, N., 2025). E essa simplicidade é o que o torna um modelo moderno amplamente utilizado, como mencionado anteriormente.

O mesmo resultado pode ser obtido através da especificação de uma regressão linear com mínimos quadrados, no qual ATT é medido sobre β^{2x2} na Equação 9 (Utilizando dados de dois períodos e dois grupos)(Baker et al., 2025; Huntington-Klein, N., 2025).

$$Y_{i,t} = \beta_0 + \beta_1\{D_i = 1\} + \beta_2\{t = 2\} + \beta^{2x2}(D_i = 1 \times t = 2) + e_{i,t} \tag{9}$$

Naturalmente, como o modelo DiD pode ser estimado por uma regressão linear com mínimos quadrados ordinários, este também pode ser estimado com efeitos fixos (Baker et al., 2025; Huntington-Klein, N., 2025). Essa é uma extensão natural no controle de variáveis de confusão do modelo, permitindo que variáveis invariantes no tempo e por agrupamento sejam controladas (Baker et al., 2025; Huntington-Klein, N., 2025).

A Equação 10 apresenta a modelagem para estimar o DiD considerando efeitos fixos de unidade (η_i) e tempo (θ_t) (Baker et al., 2025; Huntington-Klein, N., 2025).

$$Y_{i,t} = \theta_t + \eta_i + \beta_{treat}D_{i,t} + \beta_{cov}X_{i,t} + e_{i,t} \tag{10}$$

Mesmo com essas diferentes formas de modelagem, e muitas outras não mencionadas neste trabalho, um simples design de 2×2 é irreal em muitos cenários. Partindo das definições feitas do design básico, pode-se extender o DiD para um cenário de multiplos períodos de tempo, ou seja, $2xT$ (Baker et al., 2025; Huntington-Klein, N., 2025).

A modelagem de multiplos períodos é referencia como "Estudo de Evento", e considera que para cada $t \geq g$, no qual g é o período que ocorreu o tratamento, a suposição de tendências paralelas deveria ser respeitada (Baker et al., 2025; Huntington-Klein, N., 2025).

Baseado nessa suposição mais forte, é possível usar as estimativas para cada período pre e pós tratamento como formas de placebos, ou seja, estima-se o impacto de cada período de tempo entre os grupos. Espera-se que os períodos pré tratamento apresentem estimativas centradas em zero, concluindo que havia uma tendência paralela no pré tratamento (Baker et al., 2025; Huntington-Klein, N., 2025).

No entanto, mesmo que os períodos pré tratamentos não apresentem ajuste perfeito centrado no zero em todos os períodos, isso não indica que o modelo não respeita a suposição de tendência paralelas, apenas indica que não foi possível observar esta nos períodos do pré-tratamento (Baker et al., 2025; Huntington-Klein, N., 2025). Para que o DiD seja valido, a suposição deve ser respeitada apenas no pós-tratamento, porém argumenta-se que é possível que a precisão dos modelos sejam bem maiores em caso de bom ajuste no pré-tratamento (Baker et al., 2025; Huntington-Klein, N., 2025).

A Equação 11 apresenta a modelagem feita considerando multiplos períodos de tempo.

$$Y_{i,t} = \theta_t + \eta_i + \sum_{k=1}^{g-2}(\{G_i = g\}.\{t = k\}) + \sum_{k=g}^{T}(\{G_i = g\}.\{t = k\}) + e_{i,t} \quad (11)$$

O modelo de diferença-em-diferença é ideal no contexto de avaliação de políticas e choques exógenos no setor de energia, considerando que muito dos problemas de regulação, políticas e até mesmos choques socioeconômicos são externos aos sistemas de elétricos, permitindo que a modelagem estime valores confiáveis sobre suposições fortes (Huntington-Klein, N., 2025; Baker et al., 2025).

Alguns dos primeiros estudos com ideias similares ao diferença-em-diferença aplicados ao setor de energia mediram: comportamento de preços de energia em mercados desregulados (Puller, 2007); impacto de programas de eficiência energética sobre demanda de energia (Horowitz, 2007); impacto de programas de emissão de certificados de emissão de dióxido de enxofre em concessionárias de energia (Burtraw et al., 2000), e diversos outros.

3.3.2 Matching

Para garantir que a suposição de tendências paralelas seja respeitada, especialmente por ser a base da modelagem, foi feita aplicação da técnica de propensity matching score usando distancia de Mahalanobis (Huntington-Klein, N., 2025; Guo; Fraser; Chen, 2020). Essa metodologia foi inicialmente proposta em 1997 (Heckman; Ichimura; Todd, 1997) e aperfeiçoada com diversos métodos (Guo; Fraser; Chen, 2020; Lee et al., 2025; Roth et al., 2023) para garantir que a suposição de tendências paralelas fosse respeitada, uma vez que é uma suposição forte para garantir o contrafactual o mais próximo do ideal possível.

A análise estatística usando o Propensity Score se encaixa em um grupo de métodos estatísticos para estimar os efeitos de um tratamento através de dados observacionais, quando métodos randomizados são inviáveis ou antiéticos (Guo; Fraser; Chen, 2020; Lee et al., 2025; Roth et al., 2023).

Também pode ser caso de pesquisadores que precisam avaliar o tratamento ou efeitos causais a partir de dados de levantamento, dados censitários, dados administrativos, dados de registros médicos, ou outros tipos de observações onde um contrafatual deve ser construído (Guo; Fraser; Chen, 2020; Lee et al., 2025; Roth et al., 2023).

A premissa para essa metodologia é que pesquisadores quando usam desenhos quase-experimentais (ou seja, sem randomização), normalmente atribuem participantes do estudo em condições de tratamento e comparação baseadas em critérios não aleatórios (Guo; Fraser; Chen, 2020; Lee et al., 2025; Roth et al., 2023).

A presença de vários efeitos de seleção pode resultar na correlação do termo residual com as variáveis independentes. Considerando que análises de regressão de qualquer tipo assume que as variáveis independentes usadas na regressão não estão correlacionadas com o termo residual/erro, a presença de vies de seleção resulta em estimativas ineficientes e potencialmente enganosas

de efeitos de tratamento. Isso é conhecido como o problema da endogeneidade (Guo; Fraser; Chen, 2020; Lee et al., 2025; Roth et al., 2023). Em 1983, Rosenbaum e Rubin publicaram um artigo seminal que propôs usar uma pontuação de propensão para resolver o problema, e assim estabelecendo teorias e princípios de aplicação desenvolvidos a partir de seu trabalho, fundando as bases para toda a classe de modelos de pontuação de propensão (ROSENBAUM; RUBIN, 1983; Guo; Fraser; Chen, 2020).

A partir de um conjunto de dados, Rosenbaum e Rubin (1983) definiram o score de propensão para o participante $i(i = 1, ...N)$ como a probabilidade condicional de atribuição a um determinado tratamento $(W_i = 1)$ versus não tratamento $(W_i = 0)$ dado um vetor de covariáveis observadas, $x_i : e(x_i) = pr(W_i = 1)$. A pontuação de propensão $e(x_i)$ é uma medida de equilíbrio que resume a informação do vetor x_i em que cada x covariável.

Há diversos métodos que permitem o calculo do score de propensão, e para fins dessa pesquisa utiliza-se uma regressão logística para realizar o calculo do score e na sequência o pareamento com um método de distancia. Esse tipo de método é chamado na literatura de "Nearest Neighbor Within Caliper Matching"(ROSENBAUM; RUBIN, 1983; Guo; Fraser; Chen, 2020).

De forma geral, normalizam-se os valores das variáveis e da variável resposta em relação ao primeiro valor da série temporal de cada uma. Filtrou-se os dados apenas até o primeiro trimestre de 2020 - evitando assim dados pós choque exógeno - e foi feito o treinamento de um modelo de regressão logística para calcular a probabilidade de cada distribuidora passar pelo tratamento.

Com a probabilidade calculada, garantiu-se que havia suporte comum entre a probabilidade do grupo tratado e não tratado. Foram identificadas 3 distribuidoras que não passaram por lockdown, com comportamento mais similar a uma distribuidora a qual passou, usando a distância de Mahalanobis (Huntington-Klein, N., 2025; Guo; Fraser; Chen, 2020).

Essa metodologia garante que as séries temporais de PNTs das distribuidoras sejam comparadas com contrafactuais semelhantes antes do choque do lockdown, assumindo que o comportamento continuaria o mesmo caso o choque não acontecesse (Huntington-Klein, N., 2025; Guo; Fraser; Chen, 2020). Mesmo usando essa metodologia, foi aplicada a técnica de estudo de evento para observar estimativas de cada trimestre antes e depois do choque dos *lockdowns*, assim evidenciando que no pré choque havia semelhança entre as distribuidoras (Huntington-Klein, N., 2025; Guo; Fraser; Chen, 2020; Marcus; Sant'Anna, 2021).

Mediante a essas metodologias, buscou-se responder se a partir do segundo trimestre de 2020 até o final de 2020 houve um efeito indicando um aumento ou diminuição nas PNTs. O segundo trimestre foi selecionado como o inicio do tratamento por representar o trimestre com os primeiros *lockdowns* do Brasil e já alcançar mais de mil óbitos semanais (Brazilian Ministry of Health, 2025).

4 RESULTADOS E DISCUSSÕES

A primeira etapa para começar a análise foi a coleta dos dados e avaliação dos mesmos, isto é, primeiro identifica-se onde poderia ser obtido os dados. O presente trabalho obteve os dados do balanço mensal da ANEEL, que apresenta diversas informações sobre os instituições presentes no setor elétrico, seja elas distribuidoras, geradoras, compradoras e diversas outras categorias. Esse estudo focou apenas em analisar os dados de distribuidoras, então a partir da coleta do balanço realiza-se uma filtragem dos dados.

Formalmente, seja D o conjunto de dados do balanço da ANEEL. A operação de seleção que extrai os registros de D que satisfazem a uma condição conceitual C é formalizada por $\sigma_C(D)$. Para delimitar o escopo desta pesquisa, o conjunto de dados filtrado foi obtido por meio da conjunção de duas condições específicas, denotadas por C_{mod} e C_{cat}, que representam condições aplicadas sobre as colunas *DscModalidadeBalanco* (DMB) e *DscCctBalanco* (DCB), definidas conforme as Equações 12 e 13:

$$\begin{aligned} C_{\text{mod}} = (\text{DMB} = \text{“Perdas na Distribuição (valor faturado)”}) \\ \vee (\text{DMB} = \text{“Energia Vendida”}) \end{aligned} \tag{12}$$

$$C_{\text{cat}} = (\text{DCB} = \text{“Perdas Não-Técnicas”}) \vee (\text{DCB} = \text{“Fornecimento - Cativo TOTAL”}) \tag{13}$$

Desse modo, o filtro resultante aplicado à base de dados é expresso pela composição lógica das propriedades apresentadas, resultando no subconjunto D_{filtrado}, conforme apresentado na Equação (14), no qual, por meio da aplicação do operador de seleção (σ) sobre a matriz de dados basilar D, as unidades de observação são preservadas na relação resultante exclusivamente mediante a satisfação simultânea dos critérios preestabelecidos: a condição de restrição de modalidade (C_{mod}) e o critério de categorização (C_{cat}).

$$D_{\text{filtrado}} = \sigma_{C_{\text{mod}} \wedge C_{\text{cat}}}(D) \tag{14}$$

Esse conjunto resultante representa a composição no qual pode-se construir basicamente qualquer combinação de dados necessária, uma vez que tem-se uma tabela que cada linha representa os dados de PNTs ou de Energia Injetada cativa para uma distribuidora de energia em um mês específico de um determinado ano.

A segunda etapa da estruturação dos dados consistiu na identificação do estado brasileiro de atuação de cada distribuidora de energia elétrica, visando viabilizar os agrupamentos geográficos posteriores. Essa informação foi coletada diretamente no endereço eletrônico da ANEEL, me-

diante consulta individual por distribuidora. A partir desses dados, construiu-se um conjunto referencial secundário, mapeando o Cadastro Nacional da Pessoa Jurídica (CNPJ) ao seu respectivo estado. Por fim, realizou-se a junção (merge) desse conjunto secundário ao conjunto de dados basal por meio de uma operação de junção natural, conforme formalizado na Equação 15

$$D_{\text{enriquecido}} = D_{\text{basal}} \bowtie_{\text{CNPJ}} D_{\text{estados}} \tag{15}$$

Onde D_{estados} representa o conjunto referencial obtido junto à ANEEL e $D_{\text{enriquecido}}$ corresponde à base de dados integrada, pronta para as etapas de agregação.

A partir deste ponto, já é possível construir grande parte de todos os dados usados nessa pesquisa, desde os dados de agrupamento por estado brasileiro anual, até os dados posteriormente usados para gerar os impactos do DiD.

Na etapa subsequente, o conjunto de dados $D_{\text{enriquecido}}$ foi submetido a um processo de agregação e pivotamento para a reestruturação de suas dimensões. Os dados foram agrupados pelas variáveis de estado, ano e mês. Simultaneamente, aplicou-se uma operação de pivotamento sobre a coluna *DCB*, convertendo seus registros nominais em duas novas colunas de atributos distintos. Os valores correspondentes a essas novas colunas foram preenchidos pelo somatório da variável *VlrEnergia*. Esta transformação gerou o conjunto de dados consolidado $D_{\text{consolidado}}$, formalizado por meio da Equação 16:

$$D_{\text{consolidado}} =_{\{\text{Estado,Ano,Mês}\}} \gamma_{\sum(\text{VlrEnergia})}[D_{\text{enriquecido}}] \rightarrow M_{i,j} \tag{16}$$

Onde cada elemento $m_{i,j}$ da matriz resultante corresponde ao somatório do valor de energia associado ao grupo indexado pelas variáveis temporais e geográficas i, e segregado pelas duas novas categorias de balanço j, em que:

$$j \in \{\text{“Perdas Não-Técnicas”}, \text{“Fornecimento - Cativo TOTAL”}\} \tag{17}$$

A estrutura resultante deste tratamento matemático permite tanto o acoplamento futuro de novas observações quanto a posterior agregação em escala anual, caso seja necessário para o desenvolvimento do modelo.

Previamente ao início da análise de regressão com efeitos fixos, cujo propósito foi determinar as variáveis com maior correlação perante as PNTs, realizou-se uma nova agregação temporal da base de dados. O conjunto $D_{\text{consolidado}}$ foi consolidado no nível anual, gerando o subconjunto D_{anual}.

Para tanto, as observações foram agrupadas pelas variáveis de controle espacial e temporal, correspondentes ao estado e ao ano, aplicando-se a função de somatório sobre os atributos de PNTs e de Energia Injetada. Esta operação de sumarização é descrita matematicamente pela Equação (18):

$$D_{\text{anual}} =_{\{\text{Estado,Ano}\}} \gamma_{\{\sum(\text{PNT}),\sum(\text{Energia Injetada})\}}[D_{\text{consolidado}}] \tag{18}$$

A estrutura resultante em formato de painel balanceado viabiliza a correta aplicação do estimador de efeitos fixos, controlando as heterogeneidades inobserváveis constantes no tempo para cada unidade federativa.

Os dados das PNTs sobre a Energia Injetada, sem balanceamento, apenas as médias por estado, apresentam uma distribuição com variações ano a ano, especialmente um crescimento no desvio padrão, que é detalhado na tabela 3. Esse crescimento dos desvios padrões sugere um crescimento na dispersão dos dados, particularmente nos extremos, indicando uma tendência natural de crescimento dos valores de PNTs.

Tabela 3 – Resumo estatístico da porcentagem de PNTs por ano de referência

	Ano de Referência			
	2016	**2017**	**2018**	**2019**
Contagem	27.0	27.0	27.0	27.0
Média	0.0769	0.0782	0.0804	0.0792
Desvio padrão	0.0886	0.0929	0.0957	0.0976
Mínimo	0.0065	0.0096	0.0104	0.0088
25%	0.0223	0.0233	0.0260	0.0240
Mediana	0.0400	0.0412	0.0368	0.0403
75%	0.1097	0.1026	0.1093	0.0995
Máximo	0.3540	0.4055	0.4149	0.4371

Fonte: Próprio autor

A partir desse estágio, realizou-se a integração do conjunto de dados D_{tarifas}, o qual compreende os valores tarifários repassados pelas distribuidoras de energia elétrica aos consumidores finais. Essa base de dados, coletada por meio do endereço eletrônico da ANEEL, contém informações referentes aos valores praticados, às datas de início e término de vigência, bem como ao enquadramento da modalidade tarifária.

Para o escopo desta pesquisa, delimitaram-se as tarifas padrão de energia consumida e as tarifas de uso do sistema de distribuição, restritas às classes B1 e B2. Essa filtragem conceitual é definida pela condição C_{tar} na Equação (19):

$$C_{\text{tar}} = (\text{Classe} \in \{\text{“B1”}, \text{“B2”}\}) \wedge (\text{Tipo} \in \{\text{“TE”}, \text{“TUSD”}\}) \tag{19}$$

A nomenclatura de classes B1 e B2 faz referência aos consumidores finais atendidos em redes de baixa tensão, englobando unidades residenciais, pequenos empreendimentos e estabelecimentos comerciais, assim como o ambiente rural. Desse modo, o conjunto de dados tarifários filtrado, denotado por $D_{\text{tar_filtrado}}$, foi incorporado à base consolidada anterior por meio de uma operação de junção, conforme expresso na Equação (20):

$$D_{\text{anual_tarifas}} = D_{\text{anual}} \bowtie_{\{\text{Estado,Tempo}\}} \sigma_{C_{\text{tar}}}(D_{\text{tarifas}}) \tag{20}$$

Como etapa final antes de realizar as transformações e aplicação da regressão com efeitos fixos, realizou-se a integração dos conjuntos de dados D_{crime}, $D_{\text{desemprego}}$, $D_{\text{literacia}}$, $D_{\text{analfabetismo}}$ e $D_{\text{população}}$, os quais compreende dados de quantidade de homicídios, taxa de desemprego, taxa de literacia, taxa de analfabetismo e população total, para ano e estado. Esses dados foram coletados dos endereços eletrônicos do Atlas da Violência (para quantidade de homicídios anual) e api SIDRA do IBGE.

Considerando que todos os indicadores foram harmonizados espacial e temporalmente para o nível de estado e ano, a unificação com a base de dados anual foi operacionalizada por meio de uma junção natural, conforme expressa na Equação (25):

$$D_{\text{anual_socio}} = D_{\text{anual}} \bowtie_{\{\text{Estado,Ano}\}} D_{\text{socio}} \tag{21}$$

Onde o vetor de atributos de D_{socio} é definido de forma simplificada pela Equação (22)

$$D_{\text{socio}} = \{\text{Homicídios}, \text{Desemprego}, \text{Literacia}, \text{Analfabetismo}, \text{População}\} \tag{22}$$

Com os dados anuais unidos aos socioeconômicos, criou-se um série de colunas transformadas e derivadas dos dados base, sendo elas homicídios a cada mil habitantes, consumo de energia per capita, log do consumo de energia per capita, log das taxas de energia, log da população e taxas de energia normalizadas pelo salario mínio do ano.

A Regressão Linear de forma geral pode sofrer com a multicolinearidade, ou seja, quando uma das variáveis independentes pode ser construída pela combinação de outras variáveis independentes. Baseado nisso, antes de realizar a regressão todas as variáveis foram submetidas ao processo de VIF (Variance Inflation Factor), que mede o quanto a variância de um coeficiente de regressão estimado é inflada devido à multicolinearidade (correlação) entre variáveis independentes em um modelo. Foi eliminado todas as variáveis com VIF acima de 5, no qual identificou-se que a maior parte da multicolinearidade estava nas colunas de tarifa de energia. Para validar o método, também foi calculada a matrix de correlação de Pearson, representada na Figura 8.

Eliminou-se todas as variáveis de transformação logarítmica exceto a variável de log do consumo de energia per capita, por apresentarem alto VIF, assim como as tarifas B2 normalizadas pelo salário. Assim, seguiu-se para a aplicação de uma regressão com efeitos fixos de estado e ano, para encontrar as variáveis mais correlacionadas com as PNTs.

Para realizar a regressão, construiu-se um loop no qual o mesmo testou todas as combinações possíveis de 1 até 6 variáveis independentes, sempre checando se ao mesmo a maioria das variáveis independentes tinham significância estatística e o teste de Hausman apresentava significância estatística.

Apenas os resultados que estavam dentro dos padrões impostos foram checados, dando prioridade para os resultados com maior numero de variáveis independentes utilizadas. O

resultado com maior número de variáveis utilizadas, dentro dos padrões e alinhado com a literatura é apresentado na Seção 4.1.

4.1 ESPECIFICAÇÃO DA REGRESSÃO DE EFEITOS FIXOS E RESULTADOS

Para analisar os determinantes das Perdas Não-Técnicas (PNT) no nível das unidades federativas, estimou-se um modelo de regressão linear com dados em painel utilizando o estimador de Efeitos Fixos (Two-way Fixed Effects). Essa abordagem permite controlar tanto as heterogeneidades inobserváveis e invariantes no tempo de cada estado quanto os choques comuns a todas as entidades ao longo dos anos. Este é formalizada na Equação (23):

$$\begin{aligned}\text{PercPNT}_{it} = \beta_0 + \beta_1 \ln(\text{ConsEnergia}_{it}) + \beta_2 \text{Crime}_{it} + \beta_3 \text{Desemp}_{it} + \beta_4 \text{SalarTEB1}_{it} \\ + \beta_5 \text{SalarTUSDB1}_{it} + \alpha_i + \lambda_t + \varepsilon_{it}\end{aligned} \tag{23}$$

Em que:

- i e t indexam, respectivamente, o estado e o ano da observação;
- PercPNT_{it} é a variável dependente que representa o percentual de perdas não-técnicas;
- $\ln(\text{ConsEnergia}_{it})$ indica o logaritmo do consumo de energia per capita;
- Crime_{it} representa a taxa de criminalidade per capita;
- Desemp_{it} denota a taxa de desemprego;
- SalarTEB1_{it} e SalarTUSDB1_{it} correspondem às variáveis de tarifas de energia e uso do sistema de distribuição da classe B1;
- α_i representa o efeito fixo da entidade (estado);
- λ_t representa o efeito fixo temporal (ano);
- ε_{it} é o termo de erro idiossincrático, estimado com erros-padrão robustos agrupados por cluster (*Clustered Standard Errors*).

Os resultados da estimação do modelo de Efeitos Fixos estão consolidados na Tabela (4). A análise revela nuances fundamentais sobre o comportamento das PNTs. O primeiro ponto de destaque analítico reside nos testes de especificação do painel. O Teste de Agrupabilidade (*F-test for Poolability*) rejeitou fortemente a hipótese nula de homogeneidade dos interceptos ($F = 82,402$; $p < 0,0001$), evidenciando que modelos de dados empilhados (Pooled OLS) são inadequados devido à presença de marcantes características regionais inobserváveis.

Ademais, o Teste de Hausman indicou uma estatística $\chi^2 = 4,3652$ com p-valor de $0,4981$. Sob o limiar clássico de significância de 5%, não se rejeita a hipótese nula de que os efeitos

Tabela 4 – Resultados da Estimação por Efeitos Fixos (PanelOLS)

Variável Independente	**Coeficiente (β)**	**Erro-Padrão**	**Estatística t**	**p-valor**
Constante (β_0)	0,3152	0,1826	1,7262	$0,0885^{*}$
$\ln$(ConsEnergia per capita)	-0,0449	0,0244	-1,8378	$0,0702^{*}$
Crime per capita	-0,0015	0,0564	-0,0271	0,9785
Percentual Desemprego	0,0014	0,0031	0,4585	0,6480
Tarifa SalarTEB1	0,1778	0,1575	1,1287	0,2627
Tarifa SalarTUSDB1	0,2058	0,0861	2,3917	$0,0193^{**}$
Diagnósticos do Modelo	**Estatística**	**Distribuição**	**p-valor**	
R^2 (Within)	-0,0323	–	–	
R^2 (Overall)	0,0594	–	–	
Teste F de Agrupabilidade (*Poolability*)	82,402	$F(29, 73)$	$< 0,0001^{***}$	
Teste de Hausman (χ^2)	4,3652	$df = 5$	0,4981	
Estatística F Robusta (Erros Agrupados)	1,7639	$F(5, 73)$	0,1310	
Observações / Entidades (Períodos)	108 / 27 (4)	–	–	

Nota: *** Significativo a 1%, ** Significativo a 5%, * Significativo a 10%. Erros-padrão corrigidos por *cluster*

individuais não são correlacionados com os regressores. Contudo, dado o caráter do resultado (significativo a 10%), o tamanho da amostra e a premissa teórica de que as características político-estruturais dos estados possuem correlação intrínseca com o nível socioeconômico e tarifário regional, a manutenção do estimador de Efeitos Fixos preserva a consistência dos parâmetros contra o viés de variável omitida.

No tocante às variáveis econômicas estruturais, a variável $\ln$(ConsEnergia per capita) apresentou um impacto negativo e estatisticamente significante ao nível de 10% ($\beta = -0,0449$; $p = 0,0702$). Sendo um modelo log-lin, a interpretação indica que um aumento de 10% no consumo de energia *per capita* está associado a uma redução absoluta aproximada de 0,45 pontos percentuais no indicador de PNTs. Esse comportamento sugere que o incremento na intensidade de consumo regulamentado — muitas vezes associado ao desenvolvimento econômico regional — tende a diluir ou desincentivar proporcionalmente a incidência de irregularidades comerciais. Em contrapartida, as variáveis estruturais de criminalidade (Crime per capita) e mercado de trabalho (Percentual Desemprego) não demonstraram significância estatística dentro do modelo, indicando que, no horizonte temporal avaliado sob efeitos fixos, variações dessas métricas não exerceram influência direta na flutuação das perdas.

Ao avaliar a política tarifária, observa-se que a variável SalarTUSDB1 se consolidou como o principal determinante estatístico positivo do modelo, sendo significante ao nível de 5% ($\beta = 0,2058$; $p = 0,0193$). O coeficiente indica que um aumento unitário no valor da Tarifa de Uso do Sistema de Distribuição (TUSD) para a classe residencial/comercial de baixa tensão está atrelado a uma elevação de 20,58 pontos percentuais na taxa de PNT. Esse achado corrobora fortemente a literatura de incentivos econômicos ao furto de energia: à medida que os encargos e custos puramente sistêmicos de distribuição se elevam e oneram a fatura final do consumidor, há

um estímulo financeiro direto para a migração rumo à informalidade e à realização de fraudes ou "gatos".

Embora a tarifa de energia pura (SalarTEB1) tenha exibido um coeficiente positivo expressivo ($\beta = 0,1778$), ela não alcançou significância estatística padrão ($p = 0,2627$). No entanto, a significância estatística não é um determinante binário, mas sim um espectro de confiança, o que indica que não se pode descartar esse resultado uma vez que o mesmo indica que aumentos unitários no valor da Tarifa de Energia frente ao salário mínimo estão associados a uma elevação de 17,78 pontos percentuais na taxa de PNTs.

Por fim, é crucial notar um paradoxo de diagnóstico comum em painéis curtos com matrizes de covariância robustas: a estatística F robusta ajustada por agrupamento em *cluster* não obteve significância ($F_{\text{robust}} = 1,7639$; $p = 0,1310$). Esse fenômeno, aliado a um R^2 *within* negativo, decorre da forte autocorrelação serial intra-grupo e da baixa variação temporal das variáveis dentro dos estados no curto período de 4 anos avaliado. Isso indica que, embora os efeitos fixos capturem eficientemente a variação estrutural (*between*), o modelo possui limitações para explicar desvios de curto prazo em relação às médias históricas de cada estado.

4.2 DIFERENÇAS EM DIFERENÇAS PAREADO E RESULTADOS

A partir dessas considerações, para realizar-se a mensuração do efeito que os lockdowns e a pandemia em si tiveram sobre as PNTs, desenvolve-se uma técnica de *Propensity Score Matching*, no qual se calcula um score de propensão para cada estado e realiza um pareamento de estados que tem tendências e comportamento de variáveis similares.

Para realizar esse pareamento, dois grupos devem existir, um grupo de controle e um grupo de unidades tratadas. A definição do grupo controle e grupo tratado, especialmente em um cenário de choque global como a pandemia da Covid-19, é o que determina o design quase-experimental e o mecanismo de estimação.

No caso desta pesquisa, duas abordagens foram usada. A primeira é a separação de estados que abertamente declararam *lockdown* contra um grupo controle de estados que não declararam. Esse design testa se estados que foram declaram o distanciamento social tiveram um efeito mais expresivo nas PNTs do que estados que não declararam.

O primeiro passo foi então realizar a separação dos estados, no qual a lista de estados que declararam os lockdowns no segundo trimestre de 2020 foram: AM, AP, CE, MA, MG, PA, PB, PE e RJ. Todos os demais estados foram considerados controles.

Na fase subsequente do tratamento de dados, utilizou-se a relação $D_{\text{consolidada}}$ como estrutura inicial para a construção de um novo arranjo matricial orientado à dimensão trimestral. Inicialmente, realizou-se a agregação temporal dos vetores de perdas não técnicas (*PNT*) e de energia injetada (E_{injetada}) para o horizonte delimitado entre os anos de 2018 e 2020.

Formalmente, essa primeira etapa de agregação é modelada por meio do operador de agrupamento da álgebra relacional (γ), indexado pelas chaves espaciais de estado (i) e temporais

de ano (ano) e trimestre ($trim$). As funções de partição e acumulação linear geram a tabela intermediária $D_{\text{trimestral}}$, conforme expresso na Equação 24:

$$D_{\text{trimestral}} = {}_{\{i,\,\text{ano},\,\text{trim}\}}\gamma_{\{\sum(\text{PNT}),\,\sum(E_{\text{injetada}})\}}\left(\sigma_{\text{ano}\geq 2018\,\wedge\,\text{ano}\leq 2020}\left(D_{\text{consolidada}}\right)\right) \tag{24}$$

Na sequência, operacionalizou-se a integração das matrizes de atributos socioeconômicos ($D_{\text{socioeconômico}}$) à relação $D_{\text{trimestral}}$. Esse acoplamento estrutural deu-se por meio de uma operação de junção natural ($\bowtie$), cujos critérios de pareamento lógico-espacial e de homogeneização temporal basearam-se na concordância estrita dos atributos de identificação regional e de período, mapeados unificadamente na Equação 25:

$$D_{\text{trim_socio}} = D_{\text{trimestral}} \bowtie_{\{i,\,\text{ano},\,\text{trim}\}} D_{\text{socioeconômico}} \tag{25}$$

Vale destacar que parcela expressiva dos indicadores socioeconômicos coletados junto às fontes primárias de informação já se encontrava originalmente indexada em periodicidade trimestral. Tal consistência minimizou a necessidade de interpolações temporais artificiais, mitigando potenciais vieses de especificação e preservando o comportamento estocástico original das séries temporais integradas ao modelo.

A partir desse conjunto de dados, foi feita um pivotamento da base de dados, filtrando os dados até o primeiro trimestre de 2020, transformando as variáveis de dados em combinação com o trimestre. Cada variável se torna T variáveis, no qual T é o número de trimestres usado na base. O objetivo consistiu na conversão dos dados do formato longo para o formato largo, em que cada unidade de observação espacial corresponde a uma única linha da matriz resultante.

Formalmente, delimitou-se o conjunto de entidades como o conjunto de estados I, e o horizonte temporal como o conjunto de períodos T. A operação de reestruturação geométrica da tabela original $D_{\text{trim_socio}}$ para a tabela em formato largo $D_{\text{trim_largo}}$ pode ser expressa conforme a Equação 26, indexada pelos atributos de estado ($i \in I$), variável (v) e tempo ($t \in T$).

$$D_{\text{trim_largo}} = \pi_{i,\,\ldots,\,\text{v}_t,\,\ldots}\left(\sigma_{\text{condições}}\left(D_{\text{trim_socio}}\right)\right) \tag{26}$$

Cada coluna resultante v_t representa a realização de uma variável quantitativa específica em um período de tempo determinado, tais como a proporção de perdas não técnicas (PercPNT_{2018Q1}) ou o logaritmo natural do consumo de energia *per capita* ($\ln(\text{ConsEnergia per capita})_{2018Q1}$). Desse modo, assegura-se que a estrutura matricial final alinhe as cross-sections estaduais dispostas horizontalmente ao longo de suas respectivas séries temporais, viabilizando as etapas subsequentes de análise multivariada.

A partir desse ponto, um modelo de regressão logística foi estimado usando esse conjunto de dados, para prever a probabilidade de um determinado estado sofrer lockdown ou não baseado puramente nos preditores socioeconômicos listados na regressão com efeitos fixos até o primeiro trimestre de 2020, não usando dados de 2020 pós choque. O calculo pode ser expresso conforme a Equação 27, no qual X_i representa o vetor de variáveis ao i-ésimo estado.

$$P(D_i = 1 \mid X_i) = \sigma(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki}) \tag{27}$$

O score de propensão então foi atribuído ao conjunto de dados largo e um algoritmo de pareamento por distância de Mahalanobis encontrou os 3 estados controle com score de propensão mais similar a cada estado tratado.

Sabendo quais estados mais são mais próximos uns dos outros quando controla-se por covariáveis e pela própria tendência das PNTs antes do choque exógeno, é possível estimar um modelo de DiD (2x2), com boa garantia de que a suposição de tendências paralelas está sendo respeitada.

Além disso, os resultados empíricos obtidos por meio da estimação do modelo, combinados ao pré-processamento de pareamento por distância de Mahalanobis, fornecem evidências robustas a respeito do impacto causal das restrições de mobilidade sobre as Perdas Não-Técnicas (PNT) no setor elétrico brasileiro.

A estimativa pontual para o efeito do tratamento ($\beta_{\text{EffectTreat}} = 0,0287$) revela que o choque pandêmico provocou um incremento médio de $2,87$ pontos percentuais médios na taxa de PNT das distribuidoras classificadas no grupo de tratamento em relação aos seus pares de controle contrafactual. Este coeficiente apresenta significância estatística ao nível de 5% (p-valor $= 0,0368$), com um intervalo de confiança de 95% estritamente positivo e delimitado entre $0,0018$ e $0,0556$.

A estratégia de pareamento multivariado via distância de Mahalanobis mostrou-se crucial para garantir a validade interna do modelo. Ao restringir a amostra de regressão, o algoritmo mitigou o problema de falta de suporte comum (*lack of common support*). O descarte de unidades de controle que não pareavam permitiu a construção de uma trajetória contrafactual homogênea, validada pelo teste de hipótese para a incorporação de efeitos fixos (*F-test for Poolability* com p-valor $= 0,0000$).

A robustez da inferência é adicionalmente assegurada pela estrutura de modelagem de Efeitos Fixos em Duas Vias (*Two-Way Fixed Effects* - TWFE). A inclusão de efeitos fixos de entidade absorveu todas as características estruturais invariantes no tempo de cada unidade federativa, tais como o perfil socioeconômico histórico da concessão, a densidade demográfica regional e os arranjos regulatórios contratuais específicos.

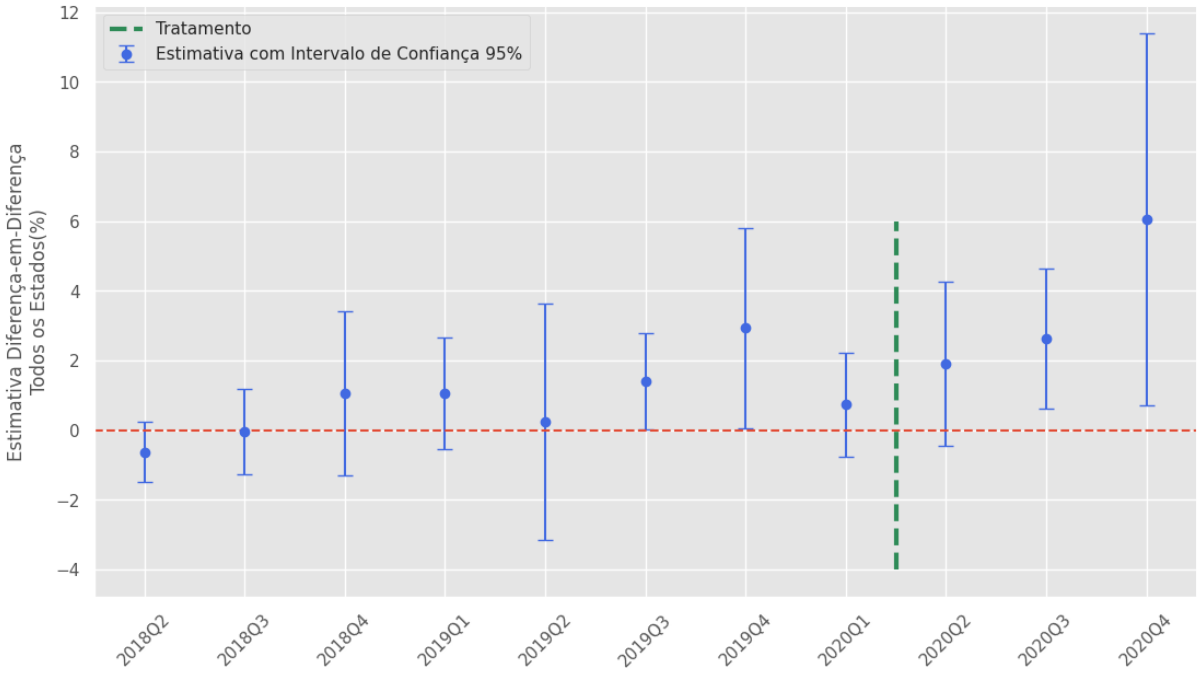
Paralelamente, os efeitos fixos de tempo expurgaram os choques sistêmicos contemporâneos que afetaram simultaneamente todas as distribuidoras do país, a exemplo da introdução do Auxílio Emergencial e de oscilações na tarifa básica de energia. Ademais, a utilização de uma matriz de covariância robusta a *clusters* no nível de UF corrige eventuais vieses decorrentes de autocorrelação serial e heterocedasticidade nos termos de erro (ε_{it}), blindando o erro padrão estimado (E.P. $= 0,0137$).

Do ponto de vista dos mecanismos econômicos subjacentes, o sinal positivo e a significância do estimador ATT (*Average Treatment Effect on the Treated*) validam a hipótese central desta

pesquisa: "Os lockdowns da pandemia do Covid-19 afetaram as PNTs e por consequência os roubos de energia?". A separação por restrição de mobilidade isola o fenômeno de restrição física ao policiamento e à fiscalização de fraudes cometidas na baixa tensão residencial.

Junto da estimativa pontual, foi feita a estimativa por estudo de evento, estimando o efeito de cada um dos períodos pre e pós choque. As estimativas podem ser vistas de forma visual na Figura 9. Através das estimativas nos períodos pré-choque é possível observar que as tendências paralelas foram respeitadas e o comportamento mudou apenas no pós choque, especialmente no terceiro trimestre de 2020, com uma piora significativa no quarto e ultimo trimestre.

Figura 9 – Estudo de Evento - lockdowns



Fonte: Próprio Autor

Um segundo teste de separação foi feito, no qual os grupos controles e tratados foram definidos não mais apenas pela declaração de lockdown, mas pelo Índice de medidas legais de distanciamento social (IMLDS) do IPEA. Através dos dados de 2020, coletou-se o índice de abril de 2020 e em dezembro de 2020, realizando uma média entre os dois valores e normalizando os valores através do máximo valor da média.

Isso permitiu definir estados que estavam acima do 3º quartil da média do IMLDS e abaixo do 1º quartil, separando em tratados e controles respectivamente. Assim os estados: MT,PB,ES,BA,PR,MS e TO, foram definidos como controle; e os estados: CE, GO, AM, PE, AC, RS e SP, como tratados.

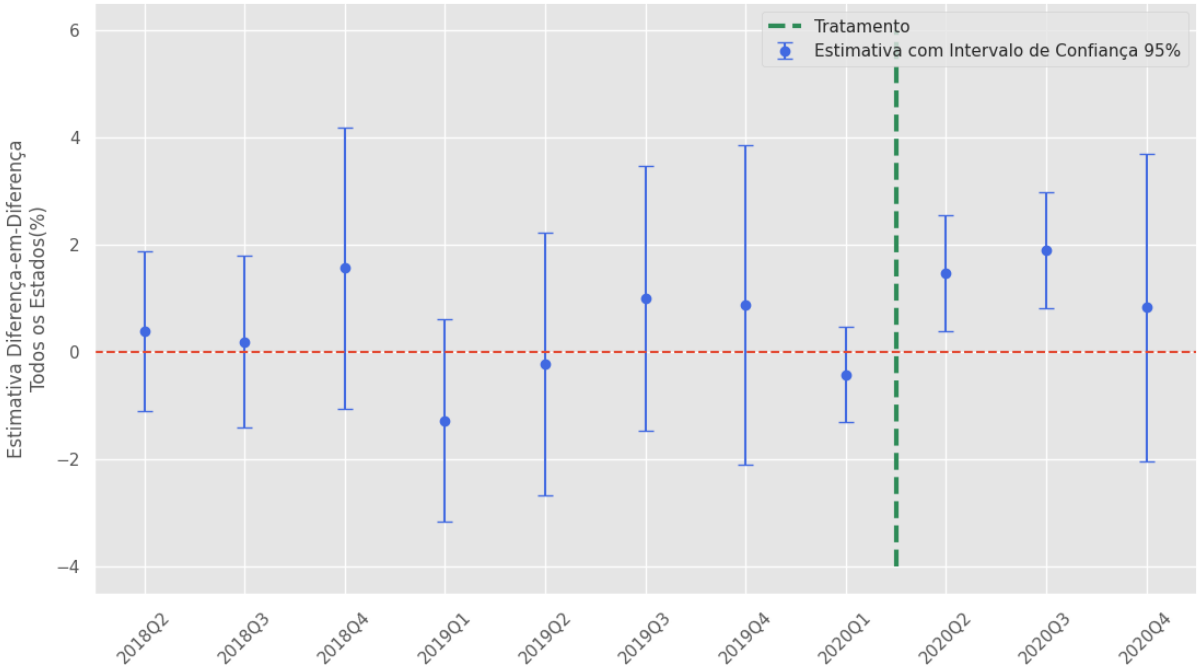
É possível observar que com essa nova definição, alguns estados mantiveram o perfil tratado, como CE, AM e PE, no entanto, PB se tornou um estado controle. Novos estados agora são considerados tratados e alguns nem são considerados em nenhum dos dois grupos. Essa diferença acontece baseado exclusivamente na intensidade das medidas legais de distanciamento.

Esse cenário também permite observa que mesmo que alguns estados tenham declarado *lock-down*, não necessariamente foi uma medida que gerou maior medidas legais de distanciamento social, como exemplo o estado de PB, que é considerado um controle por estar entre os menores indices, mas anteriormente foi tratado. Essa nova dinâmica permite isolar um novo cenário que é se os níveis de intensidade das medidas legais de distanciamento impactaram de alguma forma as PNTs.

Um novo processo de score de propensão foi realizando, e uma nova estimativa de estudo de evento foi calculado, agora olhando para estados que tiveram a sua mobilidade social pouco prejudicada em relação aos outros, contra estados que foram bastante prejudicados. As estimativas podem ser vistas de forma visual na Figura 10.

Através das estimativas nos períodos pré-choque é possível observar que as tendências paralelas foram respeitadas, até mais do que a separação por lockdowns, não tendo estimativas mais distantes de zero do que 1,5 pontos percentuais. O comportamento mudou apenas no pós choque, especialmente no segundo e terceiro trimestre de 2020, no entanto o quarto trimestre teve uma estimativa não tão significativa ao padrão de 5%.

Figura 10 – Estudo de Evento - separação por intensidade IMLDS



Fonte: Próprio Autor

Durante o ápice do período pandêmico (notadamente no segundo e terceiro trimestres de 2020), as medidas sanitárias e os decretos de isolamento social limitaram severamente a capacidade operacional das equipes de inspeção comercial das concessionárias. A redução drástica no volume de inspeções e de cortes por inadimplência gerou um cenário de tolerância forçada, mitigando o risco de detecção por parte dos agentes fraudadores. Dada a natureza do furto de energia como uma decisão econômica sob incerteza, a redução exógena da probabilidade

de auditoria reduziu o custo esperado do delito, estimulando o comportamento oportunista.

A diferença entre as estimativas no quarto trimestre de 2020, separando por lockdown e por severidade das medidas de distanciamento social, talvez sejam explicadas pela dinâmica do comportamento social. Estados que foram mais prejudicados na mobilidade podem ter cenários que estimularam a população a cometer mais furtos, uma vez que até a capacidade operacional das equipes de inspeção foi prejudica, mas parte disso se normalizou no quarto trimestre.

Em contra partida, quando é realizado a separação por *lockdowns*, observa-se um lento e progressiva choque que aumenta a medida que os trimestres passam, podendo ser explicado pela degradação dos indicadores socioeconômicos. Isto é, o choque inicial foi absorvido pela população, no entanto, conforme a incerteza econômica e a situação sanitária pioravam, um comportamento por sobrevivência e garantia dos recursos básicos foi instaurado, estimulando os furtos.

Além disso, a diferença entre os grupos tratados e controle em cada um dos cenários indica que mesmo com a declaração dos *lockdowns*, um estado pode ter medidas legais mais brandas do que outro, não representando de fato as medidas de distanciamento social.

Em suma, o avanço médio de $2,87$ pontos percentuais na PNT reflete um choque de ineficiência operacional repassado ao equilíbrio econômico-financeiro das distribuidoras. Estes resultados trazem implicações relevantes para a literatura de regulação tarifária e avaliação de impactos de choques de força maior, demonstrando que a governança regulatória deve desenhar mecanismos de incentivo que considerem a vulnerabilidade da fiscalização comercial perante restrições exógenas de mobilidade urbana.

Ademais, a importância das políticas públicas voltadas para proteção social e segurança energética se fazem mais importantes nesse cenário de crise sanitária, dado que o aumento nos níveis de PNTs compromete a qualidade e confiabilidade do fornecimento de energia, que impacta a saúde e os serviços estruturais ao longo do período de crises.

4.3 DISCUSSÕES

A literatura evidencia que as Perdas Não Técnicas (PNTs) estão intimamente vinculadas a condicionantes socioeconômicos, destacando-se a desigualdade social, os índices de criminalidade e a percepção de justiça tarifária no relacionamento com as concessionárias (Carr; Thomson, 2022; Yépez-García; Mori, 2024; Babar; Jamil; Haq, 2022; Savian et al., 2022b; Briseño; Rojas, 2020; Santiago et al., 2021). Regiões caracterizadas por maior vulnerabilidade socioeconômica registram níveis mais elevados de perdas, indicando origens compartilhadas entre o furto de energia e a dinâmica criminal geral. Ademais, embora a qualidade do fornecimento e tarifas percebidas como justas atenuem tais práticas, a restrição de renda da população atua como fator determinante, reforçando que as PNTs derivam primariamente de vulnerabilidades estruturais.

Nesse cenário, a pandemia de COVID-19 deflagrou uma profunda crise sanitária e socioeconômica global, exacerbando desigualdades, desacelerando economias e sobrecarregando

infraestruturas essenciais, especialmente em nações em desenvolvimento (Krarti; Aldubyan, 2021; Soava et al., 2021; Rouleau; Gosselin, 2021; Sánchez-Úbeda et al., 2022; Bahmanyar; Estebsari; Ernst, 2020; Pooya et al., 2025; Jené-Vinuesa; Aragüés-Peñalba; Sumper, 2024; World Bank, 2022). Dado que as PNTs respondem diretamente à deterioração das condições de vida e de renda, estabelece-se a hipótese de que o período pandêmico induziu um agravamento sistemático nos índices de perdas não técnicas.

Para investigar o cenário brasileiro das PNTs e da Covid-19, foi feito um processo extenso de coleta e manipulação de dados. Mesmo que esse não tenha sido um dos objetivos do estudo, o processo descrito aqui é uma contribuição para a literatura brasileira de Engenharia Elétrica e Energia. Esse estudo traz uma descrição extensa de onde encontrar os conjuntos de dados, como esses dados são organizados e como manipulá-los para obter diferentes visões para análise de dados.

No que tange este trabalho, O Sistema de Acompanhamento de Informações de Mercado para Regulação Econômica (SAMP) da ANEEL foi essencial na construção das visões de dados. No entanto, o mesmo também apresenta limitações. Até o momento do presente trabalho, não existe uma forma simples de determinar geograficamente as regiões e cidades atendidas por cada Concessionaria de energia, o que permitira uma granularidade rica na análise de dados de energia. Além disso, dados socioeconômicos também apenas são disponibilizados a nível estadual, o que limita a granularidade de estudos que investigam efeitos heterogêneos.

A partir dos dados do SAMP, construiu-se um conjunto de dados anual de PNTs e Energia Injetada por estado, com dados socioeconômicos coletados, e a aplicação de uma regressão com efeitos fixos, relacionando os indicadores socioeconomicos com os níveis de PNT em relação a Energia Injetada.

Observou-se que as correlações reportadas em diferentes estudos e trabalho na literatura de PNTs se confirma. No Brasil, as PNTs estão relacionadas também com indicadores socioeconômicos e justiça tarifária, e pode-se enfatizar o ultimo. A principal covariável do modelo foi a tarifa de uso do sistema de distribuição B1 normalizada pelo salário mínimo. Isso indica que o poder do salário mínimo frente as tarifas da concessionária é um dos principais determinantes da flutuação das PNTs, no qual o aumento unitário desta covariável está associado a um aumento médio de 20,6% das PNTs em relação a energia injetada.

Neste resultado reside a segunda contribuição deste trabalho. Mesmo que outros indicadores não tenham apresentado relevância estatística, provavelmente pela estrutura da regressão que absorveu os efeitos fixos de unidade e tempo, o qual está atrelado com os indicadores socioeconômicos; o resultado associado a justiça tarifária mostra que as concessionária tem oportunidade para melhorar, buscando relacionamentos mais saudáveis e justo com os consumidores finais. Esse resultado também abre oportunidades de políticas publicas, uma vez que incentivos federais para reduzir as tarifas de energia para a população mais pobre pode incentivar a redução dos furtos de energia.

Após as análises com a regressão com efeitos fixos, as covariáveis da melhor regressão foram

usadas para realizar pareamento entre grupos controle e tratados, definidos pela declaração de lockdown em 2020 e então realizar a estimativa de um modelo DiD 2x2 e um estudo de evento, produzindo um gráfico da dinâmica.

Este ultimo resultado apresenta a contribuição final deste trabalho. O modelo DiD 2x2 apresentou uma estimativa positiva, estatisticamente significativa, de 2,87 pontos percentuais sobre as PNTs. Isto significa que os estados que passaram pelos *lockdowns* tiveram um aumento médio de aproximadamente 2,9 pontos percentuais sobre o percentual de PNTs em relação a Energia Injetada.

Junto da figura da estimativa de estudo de evento, observa-se que estados que declararam o distanciamento social tiveram um aumento expressivo das PNTs. Isso pode estar atrelado a duas possíveis explicações, a primeira é que a diminuição do consumo de energia industrial e comercial cairam, diminuindo significativamente a injestão de energia, mas os níveis residências se mantiveram ou aumentaram, e estes representam a maioria das PNTs, ou seja, as PNTs se mantiveram; a segunda é que as PNTs tiveram um aumento significativo por conta da situação socioeconômica.

Evidentemente, através dos relatórios da ANEEL observa-se que as PNTs aumentaram em 2020. Além disso, observa-se uma diminuição significativa do consumo de energia industrial e comercial. Ou seja, tanto houve aumento no valor das PNTs, quanto diminuição da energia injetada, levando a um efeito mais acentuado no aumento relativo das PNTs sobre a Energia Injetada.

Esse resultado reforça a necessidade de planos de contingência em cenários de crises socioeconômicas. Mesmo com incentivos federais durante os primeiros dois trimestres da pandemia as PNTs cresceram, indicando que apenas incetivos não bastam. E, tendo em vista que PNTs apresentam um risco crítico ao sistema, aumentando os custos de manutenção e riscos de estabilidade, em momentos que a energia é um recurso critico operacional, é preciso garantia por parte da concessionária e dos orgãos responsáveis que o fornecimento de energia será estável e confiável para o funcionamento de outros recursos, como no caso da pandemia em que hospitais eram os recursos mais valiosos.

Assim sendo, as principais contribuições deste estudo são: um direcionamento de como usar os dados do SAMP, um dos mais ricos conjuntos de dados de energia elétrica brasileira; evidências brasileiras de que as PNTs estão associadas a indicadores socioeconômicos e principalmente justiça tarifaria; e por fim, a necessidade de ações e políticas por parte das concessionárias de energia e orgãos responsáveis em momentos de crise para garantir que as PNTs não apresentem um risco ao sistema e a outras infraestruturas básicas.

5 CONCLUSÃO

As Perdas Não-Técnicas de energia apresentam um desafio complexo e multivariado para orgãos reguladores e concessionárias de energia. Sua natureza pouco deterministica gera a necessidade de entender o que causa o seu surgimento, quais são os principais motivadores quando observa-se aumento, quedas e o que permite melhores controles da sua existência. Baseado nessa premissa e na literatura extensa indicado a origem socioeconomica das PNTs, esse estudo buscou entender o impacto da pandemia do Covid-19 sobre as PNTs.

Para isso, o estudo realizou uma extensa coleta de dados, combinando diferentes fontes para formar um conjunto de dados que relaciona indicadores socioeconômicos a Energia Injetada e PNTs das concessionárias de energia brasileira. Assim aplicou-se duas modelagens: a primeira, uma regressão de efeitos fixos para determinar os indicadores mais relacionados, no qual tarifas de energia se mostraram os mais relevantes; na sequência usou-se os indicadores para realizar o pareamento dos estados e aplicar um modelo de Diferença-em-Diferenças, avaliando estados que passaram pelos lockdowns, estimando um efeito médio de 2,9% e impacto persistente através de estudo de evento.

A partir desse estudo, busca-se um enriquecimento da literatura sobre as PNTs, especialmente a brasileira. Essa é a primeira aplicação de métodos quase-experimentais dentro do contexto das PNTs no Brasil, motivando possiveis trabalhos futuros a buscarem embasamento causal maior nas suas análises. Além disso, o trabalho enriquece os resultados brasileiros sobre as PNTs e seu comportamento frente a indicadores socioeconomicos. Por fim, evidência importancia desse tipo de análise na tomada de decisão de políticas públicas, buscando contribuir para o debate.

Esse estudo apresentou algumas limitações, dado a natureza do estudo, não foi feita uma análise mais ampla sobre as concessionárias em si, visto que um estado pode apresentar diferentes concessionárias de energia. Além disso, as granularidades temporais e estaduais limitaram os modelos de regressão de apresentarem amostras mais representativas, gerando intervalos de confiança mais consistentes.

Futuras pesquisas podem se debruçar em analisar granularidades menores, aplicando os mesmos modelos apresentados, com amostrar mais consistentes e ricas; realizar transformações diferentes sobre as PNTs, realizando uma média móvel de 3, 6 e 12 meses no nível mensal, para gerar uma série temporal mais estável e resiliente a ruídos; verificar diferentes crises socioeconômicas e seus efeitos sobre o sistema de Energia elétrico e as PNTs, trazendo mais evidências a literatura.

REFERÊNCIAS

Agência Nacional de Energia Elétrica. **Relatório de Análise de Impacto Regulatório sobre Perdas de Energia Elétrica em Distribuição**. [S.l.], 2024. Accessed on April 30, 2025. Disponível em: https://git.aneel.gov.br/publico/centralconteudo/-/raw/main/relatorioseindicadores/tarifaeconomico/Relatorio_Perdas_Energia.pdf.

BABAR, Z.; JAMIL, F.; HAQ, W. Consumer's perception towards electricity theft: A case study of islamabad and rawalpindi using a path analysis. **Energy policy**, Elsevier, v. 169, p. 113189, 2022.

BAHMANYAR, A.; ESTEBSARI, A.; ERNST, D. The impact of different covid-19 containment measures on electricity consumption in europe. **Energy Research & Social Science**, Elsevier, v. 68, p. 101683, 2020.

BAKER, A. et al. **Difference-in-Differences Designs: A Practitioner's Guide**. 2025. Disponível em: https://arxiv.org/abs/2503.13323.

BLACK, W.; BABIN, B. J. Multivariate data analysis: Its approach, evolution, and impact. In: **The great facilitator: Reflections on the contributions of Joseph F. Hair, Jr. to marketing and business research**. [S.l.]: Springer, 2019. p. 121–130.

Brazilian Institute of Geography and Statistics. **National Survey by Continuous Quarterly Household Sample**. 2025. https://rb.gy/3utp5g. Accessed on June 5, 2025.

Brazilian Ministry of Health. **Coronavirus COVID-19 Panel**. 2025. https://infoms.saude.gov.br/extensions/covid-19_html/covid-19_html.html. Accessed on June 5, 2025.

BRISEÑO, H.; ROJAS, O. Factors associated with electricity losses: A panel data perspective. **International Journal of Energy Economics and Policy**, v. 10, n. 5, p. 281–286, 2020.

BRISEñO, H.; ROJAS, O. Factors associated with electricity losses: A panel data perspective. **International Journal of Energy Economics and Policy**, v. 10, n. 5, p. 281–286, Aug. 2020. Disponível em: https://econjournals.com/index.php/ijeep/article/view/9599.

BURTRAW, D. H. et al. Sulfur dioxide control by electric utilities: What are the gains from trade? **Journal of Political Economy**, The University of Chicago Press, v. 108, n. 6, p. 1292–1326, 2000. ISSN 00223808, 1537534X. Disponível em: http://www.jstor.org/stable/10.1086/317681.

CALLAWAY, B.; SANT'ANNA, P. H. Difference-in-differences with multiple time periods. **Journal of Econometrics**, v. 225, n. 2, p. 200–230, 2021. ISSN 0304-4076. Themed Issue: Treatment Effect 1. Disponível em: https://www.sciencedirect.com/science/article/pii/S0304407620303948.

CARR, D.; THOMSON, M. Non-technical electricity losses. **Energies**, MDPI, v. 15, n. 6, p. 2218, 2022.

GARCIA, V. J. et al. **On the Impact of Socioeconomic Variables on Non-Technical Losses in Low Voltage Distribution Systems**. 2023. 1–6 p.

GUJARATI, D. N. **Basic econometrics**. [S.l.]: McGraw-Hill, 2009.

GUO, S.; FRASER, M.; CHEN, Q. Propensity score analysis: Recent debate and discussion. **Journal of the Society for Social Work and Research**, v. 11, n. 3, p. 463–482, 2020. Disponível em: https://doi.org/10.1086/711393.

HECKMAN, J. J.; ICHIMURA, H.; TODD, P. E. Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. **The Review of Economic Studies**, v. 64, n. 4, p. 605–654, 10 1997. ISSN 0034-6527. Disponível em: https://doi.org/10.2307/2971733.

HOROWITZ, M. J. Changes in electricity demand in the united states from the 1970s to 2003. **The Energy Journal**, International Association for Energy Economics, v. 28, n. 3, p. 93–119, 2007. ISSN 01956574, 19449089. Disponível em: http://www.jstor.org/stable/41323111.

Huntington-Klein, N. **The Effect: An Introduction to Research Design and Causality (2st ed.)**. [S.l.]: Chapman and Hall/CRC, 2025. v. 2.

JENÉ-VINUESA, M.; ARAGÜÉS-PEÑALBA, M.; SUMPER, A. Comprehensive data-driven framework for detecting and classifying non-technical distribution losses. **IEEE access**, IEEE, 2024.

KRARTI, M.; ALDUBYAN, M. Review analysis of covid-19 impact on electricity demand for residential buildings. **Renewable and Sustainable Energy Reviews**, v. 143, p. 110888, 2021. ISSN 1364-0321. Disponível em: https://www.sciencedirect.com/science/article/pii/S1364032121001829.

LEE, K. et al. Difference-in-differences with matching methods in leadership studies: A review and practical guide. **The Leadership Quarterly**, v. 36, n. 2, p. 101813, 2025. ISSN 1048-9843. Disponível em: https://www.sciencedirect.com/science/article/pii/S1048984324000420.

MARCUS, M.; SANT'ANNA, P. H. The role of parallel trends in event study settings: An application to environmental economics. **Journal of the Association of Environmental and Resource Economists**, The University of Chicago Press Chicago, IL, v. 8, n. 2, p. 235–275, 2021.

MOFIJUR, M. et al. Impact of covid-19 on the social, economic, environmental and energy domains: Lessons learnt from a global pandemic. **Sustain Prod Consumpt**, v. 26, p. 343–359, abr. 2021.

National Electric Energy Agency. **Electric Power Distribution Procedures in the National Electric System - PRODIST: Module 11 - Electrical Energy Invoice and Supplementary Information**. 2025. https://www2.aneel.gov.br/cedoc/aren2021956_2_10.pdf. Disponível em: https://www2.aneel.gov.br/cedoc/aren2021956_2_10.pdf.

National Electric Energy Agency. **SAMP - Energy Balance of Distributors**. 2025. https://dadosabertos.aneel.gov.br/dataset/samp-balanco. Accessed on May 26, 2025.

Pan-American Health Organization (PAHO). **History of the international emergency of COVID-19**. 2023. Disponível em: https://www.paho.org/pt/historico-da-emergencia-internacional-covid-19.

POOYA, P. et al. Technical losses in power networks: Mechanisms, mitigation strategies, and future directions. **Electronics**, MDPI AG, v. 14, n. 17, p. 3442, 2025.

PULLER, S. L. Pricing and firm conduct in california's deregulated electricity market. **The Review of Economics and Statistics**, The MIT Press, v. 89, n. 1, p. 75–87, 2007. ISSN 00346535, 15309142. Disponível em: http://www.jstor.org/stable/40043075.

ROBINS, J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. **Mathematical Modelling**, Pergamon, v. 7, n. 9-12, p. 1393–1512, 1986.

ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. **Biometrika**, v. 70, n. 1, p. 41–55, 04 1983. ISSN 0006-3444. Disponível em: https://doi.org/10.1093/biomet/70.1.41.

ROTH, J. et al. What's trending in difference-in-differences? a synthesis of the recent econometrics literature. **Journal of Econometrics**, v. 235, n. 2, p. 2218–2244, 2023. ISSN 0304-4076. Disponível em: https://www.sciencedirect.com/science/article/pii/S0304407623001318.

ROULEAU, J.; GOSSELIN, L. Impacts of the covid-19 lockdown on energy consumption in a canadian social housing building. **Applied Energy**, Elsevier, v. 287, p. 116565, 2021.

RUBIN, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. **Journal of Educational Psychology**, American Psychological Association, v. 66, n. 5, p. 688–701, 1974.

SÁNCHEZ-ÚBEDA, E. et al. Impact of covid-19 on electricity demand of latin america and the caribbean countries. **Sustainable Energy, Grids and Networks**, v. 30, p. 100610, 2022.

SANTIAGO, I. et al. Electricity demand during pandemic times: The case of the covid-19 in spain. **Energy policy**, Elsevier, v. 148, p. 111964, 2021.

SAVIAN, F. d. S. et al. Non-technical losses in brazil: Overview, challenges, and directions for identification and mitigation. **International Journal of Energy Economics and Policy**, v. 12, n. 3, p. 93–107, May 2022. Disponível em: https://econjournals.com/index.php/ijeep/article/view/12614.

SAVIAN, F. S. d. et al. Non-technical losses in brazil: Overview, challenges, and directions for identification and mitigation. **International Journal of Energy Economics and Policy**, v. 12, n. 3, p. 93–107, 2022.

SEMMELWEIS, I. **Etiology, Concept, and Prophylaxis of Childbed Fever**. Madison: The University of Wisconsin Press, 1983.

SILVA, L. N. F. D. et al. **Innovative Strategy for the Socioeconomic Variables Impact Evaluation on Non-Technical Losses**. [S.l.], 2024. 1–5 p.

SOAVA, G. et al. The impact of the covid-19 pandemic on electricity consumption and economic growth in romania. **Energies**, MDPI, v. 14, n. 9, p. 2394, 2021.

SOMMET, N.; LIPPS, O. A primer on fixed effects and fixed-effects panel modeling using r, stata, and spss. **Advances in Methods and Practices in Psychological Science**, v. 8, n. 4, p. 25152459251392843, 2025. Disponível em: https://doi.org/10.1177/25152459251392843.

VIEGAS, J. L. et al. Solutions for detection of non-technical losses in the electricity grid: A review. **Renewable and Sustainable Energy Reviews**, Elsevier, v. 80, p. 1256–1268, 2017.

WOOLDRIDGE, J. M. **Econometric analysis of cross section and panel data**. [S.l.]: MIT press, 2010.

WOOLDRIDGE, J. M. Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. **Empirical Economics**, 2025. ISSN 1435-8921.

World Bank. **World Development Report 2022: Finance for an Equitable Recovery**. Washington, D.C.: World Bank Group, 2022. v. 1. (World Development Report, v. 1).

World Health Organization and others (WHO). **Addressing violence against children, women and older people during the COVID-19 pandemic: key actions, 17 June 2020**. [S.l.], 2020.

YéPEZ-GARCíA, A.; MORI, R. A. J. **Economics of Electricity Losses in Latin America and the Caribbean**. 2024. Disponível em: https://doi.org/10.18235/0012971.