

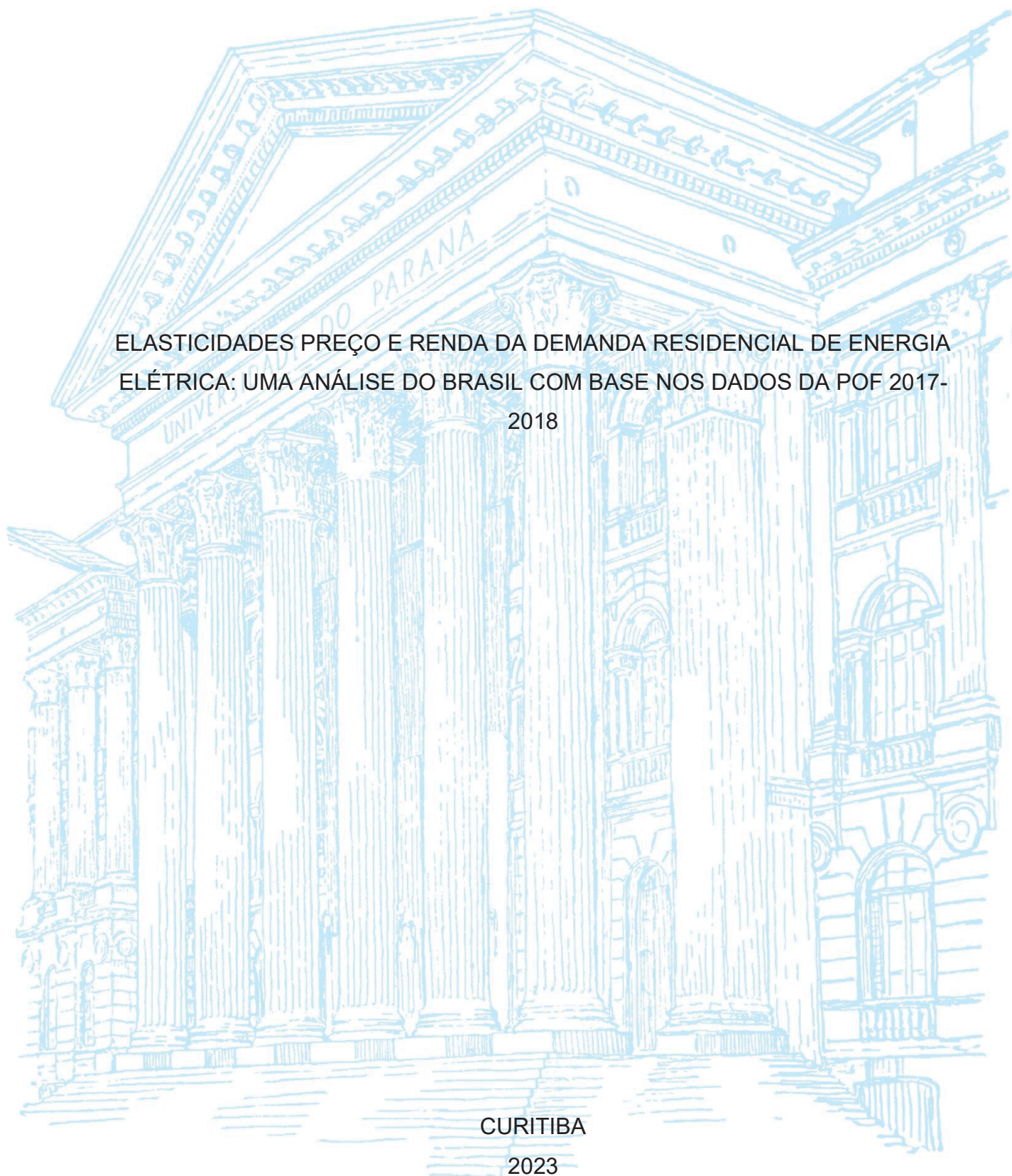
UNIVERSIDADE FEDERAL DO PARANÁ

GUILHERME BODINI DA CRUZ

ELASTICIDADES PREÇO E RENDA DA DEMANDA RESIDENCIAL DE ENERGIA
ELÉTRICA: UMA ANÁLISE DO BRASIL COM BASE NOS DADOS DA POF 2017-
2018

CURITIBA

2023



GUILHERME BODINI DA CRUZ

ELASTICIDADES PREÇO E RENDA DA DEMANDA RESIDENCIAL DE ENERGIA
ELÉTRICA: UMA ANÁLISE DO BRASIL COM BASE NOS DADOS DA POF 2017-
2018

Monografia apresentada ao curso de Especialização em Economia, Setor de Ciências Sociais, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Especialista em Análise de Conjuntura Econômica.

Orientadora: Prof^a. Dra^a. Kênia Barreiro de Souza

CURITIBA

2023

AGRADECIMENTOS

Agradeço aos professores e professoras do curso de especialização em Análise de Conjuntura Econômica pelo conhecimento que me foi passado, em especial a Profª Drª Kênia, pela ajuda e pelo apoio na elaboração deste trabalho.

RESUMO

O presente estudo apresenta a estimação das elasticidades preço e renda da demanda por energia elétrica em domicílios brasileiros. Para alcançar esse objetivo, foram utilizados dados da Pesquisa de Orçamentos Familiares 2017/2018. A metodologia empregada envolveu um procedimento iterativo de regressão linear múltipla, executado em quatro especificações distintas, de forma que a cada iteração, uma nova variável independente foi adicionada ao modelo. As variáveis utilizadas foram o preço da energia, o número de eletrodomésticos no domicílio, a renda familiar e a despesa total do domicílio. Este método iterativo e progressivo permitiu uma análise gradual do impacto de cada variável no desempenho do modelo. Os resultados obtidos indicam que a demanda por energia elétrica nos domicílios brasileiros é inelástica em relação ao preço, com uma correlação negativa. Em relação à elasticidade-renda, os dados sugerem que o consumo de energia elétrica pode ser considerado um bem normal, apresentando uma correlação positiva com a renda.

Palavras-chave: Elasticidade-preço. Elasticidade-renda. Consumo de eletricidade. Regressão Linear Múltipla. Pesquisa de Orçamentos Familiares.

ABSTRACT

This study presents the estimation of price and income elasticities of demand for electricity in Brazilian households. To achieve this objective, data from the 2017/2018 Brazilian Household Budget Survey were used. The methodology employed involved an iterative multiple linear regression procedure, executed in four distinct rounds, so that at each round, a new independent variable was added to the model. The variables used were the price of energy, the number of household appliances, household income and total household expenditure. This iterative and progressive method allowed for a gradual analysis of the impact of each variable on the model's performance. The results obtained indicate that the demand for electricity in Brazilian households is price inelastic, with a negative correlation. Regarding income elasticity, the data suggest that electricity consumption can be considered a normal good, presenting a positive correlation with income.

Keywords: Price elasticity. Income elasticity. Electricity consumption. Multiple Linear Regression. Brazilian Household Budget Survey.

LISTA DE FIGURAS

FIGURA 1 – CURVA DE DEMANDA POR APARTAMENTOS	13
FIGURA 2 – CURVA DE DEMANDA POR APARTAMENTOS COM MUITOS DEMANDANTES.....	13
FIGURA 3 – CURVA DE DEMANDA DE MERCADO	14

LISTA DE QUADROS

QUADRO 1 – MICRODADOS – LISTA DE ARQUIVOS DA BASE DA POF	
2017/2018	20
QUADRO 2 – RELAÇÃO DE VARIÁVEIS INDEPENDENTES E DEPENDENTES ..	23

LISTA DE TABELAS

TABELA 1 – ELETRODOMÉSTICOS NOS DOMICÍLIOS EM CORRESPONDÊNCIA COM O PROCEL	21
TABELA 2 – TARIFA MÉDIA POR ESTADO ENTRE OS ANOS DE 2017 E 2018...	22

SUMÁRIO

1 INTRODUÇÃO	10
2 REVISÃO DE LITERATURA	12
2.1 TEORIA DO CONSUMIDOR.....	12
2.2 ESTIMAÇÃO DE ELASTICIDADE-RENDA E PREÇO DO CONSUMO DE ENERGIA ELÉTRICA NO BRASIL.....	16
3 METODOLOGIA E DADOS	18
3.1 METODOLOGIA.....	18
3.2 DADOS.....	20
4 APRESENTAÇÃO DOS RESULTADOS	24
5 CONSIDERAÇÕES FINAIS	27
REFERÊNCIAS.....	28

1 INTRODUÇÃO

A energia elétrica exerce papel fundamental na sociedade moderna, segundo a *International Energy Agency (IEA, 2022)*, em 2021, o consumo de energia elétrica era responsável por 20% do consumo mundial de energia. Os três setores com maior representatividade foram o industrial, residencial e de serviços, com 41,9%, 26,6% e 21,2%, respectivamente. Além disso, é central em muitos aspectos da vida diária e tende a se tornar mais intensa devido a novas destinações, como a utilização de veículos elétricos em maior escala.

Segundo a Agência Nacional de Energia Elétrica (ANEEL, 2008), o nível de consumo de energia é um indicador fundamental do desenvolvimento econômico e da qualidade de vida de sua sociedade. Ele reflete a intensidade das atividades industrial, comercial e de serviços, bem como pode ser indicativo da capacidade da população em acessar bens e serviços de tecnologia superior, como carros e aparelhos domésticos e eletrônicos, que requerem conexão à eletricidade e aumentam a demanda de energia elétrica.

Conforme *International Energy Agency (IEA, 2022)*, a demanda por energia mundial, que em 2021 era de 24.700TWh, tem projeção para crescer 3,5% ao ano até atingir 33.733TWh em 2030. Essa previsão salienta a necessidade iminente dos países quanto ao planejamento da expansão do setor elétrico para atendimento da carga.

De acordo com os dados da Empresa de Pesquisa Energética (EPE, 2023a), o consumo de energia elétrica no Brasil em 2022 totalizou 509TWh, sendo as parcelas mais significativas os setores industrial, residencial e comercial, com 36,22%, 29,99% e 18,16%, respectivamente.

No Brasil, o planejamento do setor elétrico ocorre de forma integrada e coordenada entre diferentes entidades governamentais e agentes do setor, visando garantir o suprimento adequado da demanda eletroenergética do país, considerando aspectos técnicos, econômicos, ambientais e sociais. (EPE, 2023b). Desta forma, a compreensão do comportamento da demanda por energia elétrica é essencial para que se possa antecipar necessidades futuras e planejar de forma eficiente e sustentável a expansão do setor.

No âmbito nacional, diversos trabalhos foram desenvolvidos na tentativa de entendimento do tema. Dentre eles, destaca-se Modiano (1984), que estimou as

elasticidades preço e renda da demanda por energia elétrica no segmento residencial, industrial e comercial entre 1963 e 1981, utilizando o método de mínimos quadrados com correção para correlação serial pelo método de Corchrane – Orcutt. Andrade e Lobão (2007) analisaram a evolução do consumo residencial de energia elétrica no Brasil no período de 1963 a 1995 estimando a elasticidade renda e preço da sua demanda agregada através do modelo de vetor autoregressivo (VAR) sob a representação de um modelo de correção de erro (VEC). Utilizando a base de dados da POF 2008/2009, Dias (2015) investigou os padrões de consumo de energia elétrica para diferentes regiões e classes socioeconômicas do Brasil, através do método de mínimos quadrados. Dantas et al. (2016) analisaram as mudanças nas elasticidades preço e renda da demanda por energia elétrica entre 2000 e 2015. Os resultados foram obtidos através da metodologia de dados em painel, estimado pelo Método dos Momentos Generalizados (GMM), em uma versão conhecida como *System-GMM*.

Neste sentido, este trabalho tem por objetivo estimar as elasticidades preço e renda do consumo de energia elétrica residencial no Brasil, utilizando o método de regressão linear múltipla com base nos dados da Pesquisa de Orçamentos Familiares (POF) 2017-2018 realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE), fornecendo uma atualização dos parâmetros de elasticidade renda e preço para o consumo de energia elétrica domiciliar.

O presente trabalho está dividido em 6 seções. Na seção 2 é feita uma revisão da literatura, expondo os estudos já realizados a respeito das elasticidades preço e renda no Brasil. Na seção 3, são apresentadas a metodologia empregada para estimação das variáveis e as bases de dados. Na seção 4, são apresentadas estatísticas descritivas sobre as variáveis e os resultados obtidos. Por fim, a seção 5 expõe a conclusão do trabalho.

2 REVISÃO DE LITERATURA

Este capítulo apresenta uma breve revisão bibliográfica da teoria utilizada para determinação da equação de elasticidade e uma revisão dos principais estudos realizados no Brasil a respeito do tema elasticidade preço e renda da demanda por energia elétrica.

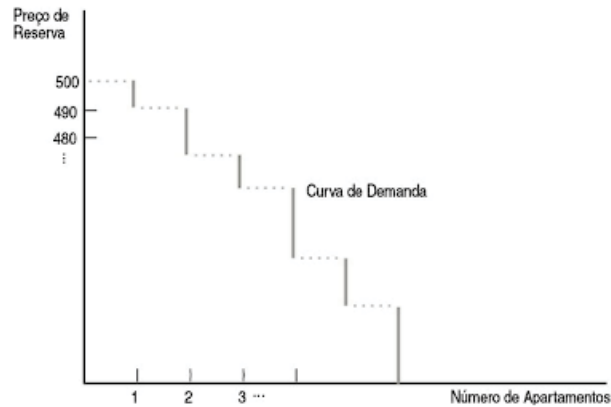
2.1 TEORIA DO CONSUMIDOR

A estimação das elasticidades preço e renda de um bem estão amparadas na teoria do consumidor, se fazendo necessário assim, a exposição de conceitos mais básicos até o entendimento das elasticidades. A teoria do consumidor é uma teoria econômica que busca compreender o comportamento dos agentes econômicos em relação às suas escolhas de consumo. A teoria se baseia na premissa de que os consumidores são racionais, buscando maximizar sua satisfação ou utilidade, dadas as restrições orçamentárias e de que os preços se ajustam até que a quantidade demandada seja igual a oferta (VARIAN, 2015). Segundo Varian (2015), o princípio de equilíbrio é quase tautológico, se as pessoas são livres para escolher, é razoável supor que elas tentem escolher as coisas que desejam, ao invés das que não desejam.

Quanto ao equilíbrio, Varian também afirma que é ao menos imaginável que, em algum momento, as demandas e as ofertas das pessoas não sejam compatíveis, indicando que algo está mudando. "Essas mudanças podem levar um longo tempo para se concretizarem e, pior ainda, podem induzir outras mudanças, capazes de 'desestabilizar' todo o sistema" (Varian, 2015, p. 41).

Uma vez exposta a estrutura baseada nos princípios de otimização e equilíbrio, é conveniente introduzir o conceito sobre curva de demanda. Conforme Varian (2015), a curva de demanda representa a relação entre a quantidade demandada de um bem e seu preço. Para melhor entendimento, o autor utiliza o exemplo da demanda por apartamentos, em que define o chamado preço de reserva, em que uma pessoa é indiferente entre comprar ou não comprar o bem e, só comprará, caso o preço de mercado atinja o preço em que está disposta a pagar. A **FIGURA 1** ilustra o exposto.

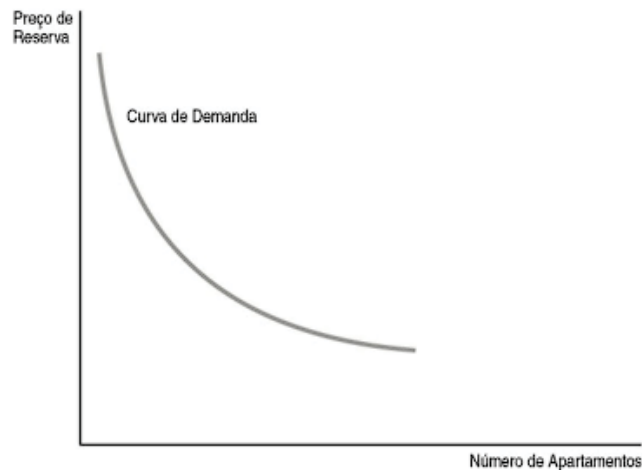
FIGURA 1 – CURVA DE DEMANDA POR APARTAMENTOS



FONTE: Varian (2015, p. 43).

De maneira semelhante, pode-se colocar a curva de demanda para o caso em que existem inúmeros demandantes, corroborando na suavização da curva, como observa-se na **FIGURA 2**.

FIGURA 2 – CURVA DE DEMANDA POR APARTAMENTOS COM MUITOS DEMANDANTES



FONTE: Varian (2015, p. 44).

Assim, a Lei geral da demanda estabelece que, *ceteris paribus*, mantendo-se constantes os outros fatores, a quantidade demandada de um bem ou serviço diminui

à medida que seu preço aumenta. Esta relação inversa entre preço e demanda pode ser verificada na **FIGURA 2**.

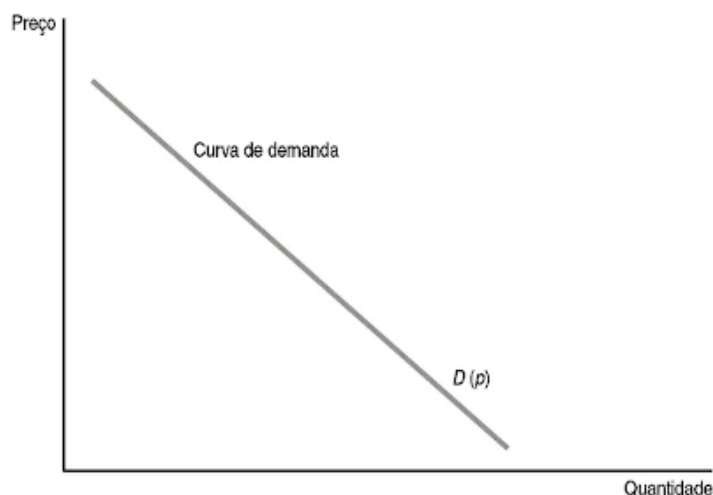
Na teoria do consumidor, os indivíduos escolhem suas cestas de consumo a partir de suas preferências e outras duas informações: os preços do mercado e sua própria renda. Assim, cada indivíduo i , dadas as suas preferências, escolhe a quantidade x_a a ser demandada de um bem qualquer a com base no preço desse bem p_a e no preço de todos os outros bens do mercado p_b . O gasto total desse consumidor fica limitada a sua renda m_i . Desta forma, no caso hipotético de dois bens, é possível representar a demanda do consumidor i para o bem 1, como $x_i^1(p_1, p_2, m_i)$ (VARIAN, 2015).

Para obter a demanda de mercado agregada, Varian (2015) utiliza a hipótese do consumidor representativo. Nesse caso, a função de demanda agregada do bem 1 terá a forma $X_1(p_1, p_2, M)$, em que M é a soma das rendas dos consumidores individuais.

$$X^1(p_1, p_2, M) = \sum_{i=1}^n x_i^1(p_1, p_2, m_i) \quad (1)$$

Fixando as rendas monetárias e o preço do bem 2, é possível observar a relação entre a demanda agregada do bem 1 e seu preço na FIGURA 3.

FIGURA 3 – CURVA DE DEMANDA DE MERCADO



FONTE: Varian (2015, p. 383).

Por fim, é possível apresentar quão sensível é a demanda em relação às variações no preço e na renda, através da medida chamada elasticidade. Para Varian (2015, p.389), “a elasticidade-preço da demanda é definida como a variação percentual na quantidade demandada dividida pela variação percentual no preço”. Assim, matematicamente:

$$\varepsilon = \frac{\Delta q/q}{\Delta p/p} \quad (2)$$

E rearranjada, obtêm-se a expressão mais comum:

$$\varepsilon = \frac{p}{q} \frac{\Delta q}{\Delta p} \quad (3)$$

Desta forma, a elasticidade de um bem pode ser inelástica, elástica ou unitária, considerando sempre o coeficiente em valores absolutos. No primeiro caso, um bem é considerado inelástico quando $|\varepsilon| < 1$, sendo insensível ou pouco sensível a mudanças nos preços. Quando $|\varepsilon| = 1$ a variação na demanda ocorre na mesma medida que a variação no preço. Se $|\varepsilon| > 1$ o bem é considerado elástico, isso significa que a variação na demanda é maior do que a variação dos preços. (VARIAN, 2015).

De forma semelhante, a elasticidade-renda da demanda é utilizada para medir a sensibilidade da demanda em relação a renda.

$$\varepsilon = \frac{r}{q} \frac{\Delta q}{\Delta r} \quad (4)$$

Quando $\varepsilon < 0$, o bem é chamado inferior e aumentos na renda reduzem a demanda do bem. Se $0 < \varepsilon < 1$, o bem é dado como normal e aumentos na renda provocam aumentos menores na demanda. Neste caso, a elasticidade-renda indica uma resposta relativamente baixa da demanda às mudanças na renda. Por fim, se $\varepsilon > 1$, aumentos na renda levam a aumentos maiores na demanda. (VARIAN, 2015).

2.2 ESTIMAÇÃO DE ELASTICIDADE-RENDA E PREÇO DO CONSUMO DE ENERGIA ELÉTRICA NO BRASIL

Para elaboração deste trabalho foi feita uma pesquisa sobre os diversos estudos publicados a respeito do tema elasticidade da demanda de energia elétrica, com objetivo compreender as diversas perspectivas e metodologias empregadas na estimação das elasticidades.

Modiano (1984) fez uma análise do comportamento da demanda de energia elétrica no Brasil, utilizando o Método de Mínimos Quadrados Ordinários (MQO), para estimar as elasticidades preço e renda dos consumidores residenciais, comerciais e industriais no período de 1963 a 1981. De acordo com o estudo realizado pelo autor, a demanda de energia elétrica demonstrou sensibilidade às variações da renda real da economia. Os valores encontrados de elasticidade-renda para as classes residencial, comercial e industrial no curto prazo foram, respectivamente, 0,332, 0,362 e 0,502. No longo prazo os valores encontrados foram 1,13, 1,068 e 1,360. Para elasticidade-preço de curto e longo prazo os resultados sugerem que o consumo de energia elétrica é inelástico.

Andrade e Lobão (2007) apresentaram uma atualização do estudo realizado por Modiano (1984) em que analisam a evolução do consumo residencial de energia elétrica no Brasil no período de 1963 a 1995 afim de estimar a elasticidade renda e preço da sua demanda agregada. Os modelos econométricos utilizados neste estudo foram o Método de Mínimo Quadrados Ordinários (MQO), estimação de Variáveis Instrumentais (VI) e o Modelo de Vetor Autorregressivo (VAR) sob a representação de um Modelo de Correção de Erro (VEC). Os valores obtidos para elasticidade-preço e renda entre os modelos foram bastante similares, na ordem de -0,05 e 0,21, respectivamente. Concluiu-se que a demanda de energia elétrica no Brasil é sensível aos preços como esperava-se, no entanto, é bastante inelástica às variáveis explicativas preço e renda.

Schmidt e Lima (2004) estimaram as elasticidades-preço e renda da demanda por energia elétrica, nas três classes de consumo, residencial, comercial e industrial para o período de 1969 a 1999. Os autores utilizaram o modelo de Vetor Autorregressivo (VAR), associado a um Modelo de Correção de Erro Vetorial (MCEV) para estimar as elasticidades, a partir do qual foi possível concluir que, em específico no caso da classe residencial, os valores obtidos para elasticidade-preço e renda

foram, respectivamente, -0,085 e 0,539. Segundo os autores, os resultados obtidos estão de acordo com o esperado pela teoria.

Dias (2015) realizou um estudo para estimar as elasticidades preço e renda com base na Pesquisa de Orçamentos Familiares (POF). Com base nos microdados disponibilizados pelo IBGE a autora utilizou o método de mínimos quadrados ponderados (MMQ), estabelecendo uma divisão de 20 grupos da população brasileira, de forma a garantir uma representação adequada das diferentes regiões geográficas e níveis socioeconômicos do Brasil na amostra da pesquisa. Os resultados obtidos indicam que as regiões do Brasil menos sensíveis às variações no preço seriam Centro-oeste e Sul, onde o consumo seria pouco afetado pelo aumento do valor da tarifa. O estudo ainda sugere que as regiões que sofreriam maior impacto devido a incrementos na renda seriam Norte, Nordeste e Sudeste.

Dantas, Costa e Silva (2016) estimaram as elasticidades preço e renda do consumo residencial, industrial e comercial de energia elétrica das regiões brasileiras para o período de 2000 a 2015. Neste estudo os autores utilizaram a metodologia de dados em painel, estimado pelo Método de Momentos Generalizados (GMM), em que, para o caso residencial, a região Norte apresentou a maior elasticidade-preço e a Região Nordeste o menor, respectivamente, -0,275 e -0,1527. Para elasticidade-renda a Região Centro-Oeste apresentou o maior valor de elasticidade e a Região Sul o menor, respectivamente, 0,3681 e 0,0432.

3 METODOLOGIA E DADOS

Neste capítulo são apresentadas a metodologia empregada para determinação do modelo, as bases de dados utilizadas e os tratamentos que foram aplicados às bases para que fosse possível encontrar às variáveis necessárias à obtenção das elasticidades.

3.1 METODOLOGIA

Conforme Wooldridge (2006), dados de corte transversal são amplamente utilizados na economia e estão fortemente relacionados com seus diversos campos, tais como microeconomia aplicada, economia do trabalho, demografia, economia urbana, saúde entre outros. O Autor também destaca a importância dos dados em determinado ponto do tempo para avaliar políticas governamentais e testar hipóteses microeconômicas. Este estudo, portanto, toma como base a Pesquisa de Orçamentos Familiares 2017/2018, usando um método econométrico ao estimar uma regressão com dados de corte transversal.

No caso de aplicações econométricas, Wooldridge (2006) coloca que a problemática inicial se dá, primeiramente, na tentativa de estudar como o comportamento de uma variável x pode explicar alterações y , considerando que x e y são variáveis que representam alguma população. A equação (5) define o modelo de regressão linear simples.

$$y = \beta_0 + \beta_1 x + u \quad (5)$$

Uma vez relacionadas as variáveis y e x , existem diversas terminologias para denominá-las, sendo “variável dependente” e “variável independente”, respectivamente, as denominações mais utilizadas na economia. A variável u , por sua vez, representa outros fatores não observados, além de x , que afetam y . Da equação (5) também se observa que, fixando outros fatores em u , a variação em u é zero, então x tem um efeito linear sobre y . Assim, tem-se que a variação em y pode ser encontrada através da multiplicação de β_1 pela variação em x , em que β_1 é o

parâmetro de inclinação da equação da relação entre y e x . β_0 é chamado de parâmetro de intercepto. (WOOLDRIDGE, 2006).

A utilização da regressão simples apresenta uma importante consideração relacionada à dificuldade de isolar outros efeitos não observáveis. Para contornar essa questão, Wooldridge (2006) destaca a vantagem da regressão múltipla em permitir controlar muitos fatores que afetam a variável dependente, justamente pela possibilidade de incluir outras variáveis explicativas que podem estar correlacionadas. Assim, o modelo de regressão linear múltipla pode ser escrito, na população, como:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k + u \quad (6)$$

Em que β_0 é o intercepto, β_1 é o coeficiente de x_1 , β_2 é o coeficiente de x_2 e assim por diante.

Diante disso, Gujarati e Porter (2008) destacam para o caso específico da determinação de elasticidade, que de maneira frequente, tem-se a utilização de modelos logarítmicos, como o modelo duplo log, ou simplesmente log-log. Para tanto, basta a aplicação do logaritmo natural em cada variável, conforme abaixo:

$$\ln y = \beta_0 + \beta_1 \ln x_1 + \beta_2 \ln x_2 + \beta_3 \ln x_3 + \dots + \beta_k \ln x_k + u \quad (7)$$

Isto se deve pela simplicidade na interpretação dos parâmetros β_k , que representam a variação percentual de y correspondente à uma variação percentual de x_k . (GUJARATI; PORTER, 2008).

Outro aspecto importante a ser considerado em análises de regressão, além das variáveis independentes, é justamente a utilização de variáveis de natureza essencialmente qualitativas, como gênero, religião, região geográfica, cor e afiliação partidária. Para tanto, se faz necessário a formulação de variáveis artificiais binárias, que assumem valor 1 quando possuem o atributo e 0, quando não possuem. Estas variáveis são comumente chamadas na econometria por variáveis *dummy*, e podem ser incorporadas ao modelo, conforme Gujarati e Porter (2008) pela seguinte expressão:

$$\ln y = \beta_0 + \beta_1 \ln x_1 + \dots + \beta_k \ln x_k + \gamma_1 D_1 + \dots + \gamma_k D_k + u \quad (8)$$

Em que, D_k é a variável qualitativa que pode assumir dois valores, $D_k = (1,0)$ e γ_k é o valor de seu coeficiente. Gujarati e Porter (2008), advertem ainda quanto ao uso correto do número de variáveis qualitativas no modelo para evitar a colinearidade perfeita, isto é, a relação linear exata entre variáveis. De modo que, na hipótese de uma variável qualitativa ter m classificações, recomenda-se a utilização de $(m - 1)$ variáveis binárias.

3.2 DADOS

A principal base utilizada para elaboração deste trabalho foi a Pesquisa de Orçamentos Familiares 2017/2018, conduzida pelo IBGE e publicada no ano de 2021. A pesquisa dispõe seus resultados em diferentes arquivos de despesas individuais e coletivas, condições de vida, características do domicílio, do morador, entre outros, como pode-se observar no QUADRO – 1. Pelo fato de o IBGE disponibilizar a base através de microdados, foi necessária a utilização da biblioteca *lerMicrodados*, em linguagem *python* para que fosse possível a leitura destes dados (o código completo encontra-se no Apêndice 1, ao final do trabalho).

QUADRO 1 – MICRODADOS – LISTA DE ARQUIVOS DA BASE DA POF 2017/2018

Arquivo
ALUGUEL_ESTIMADO.txt
CADERNETA_COLETIVA.txt
CARACTERÍSTICAS_DIETA.txt
CONDICOES_VIDA.txt
CONSUMO_ALIMENTAR.txt
DESPESA_COLETIVA.txt
DESPESA_INDIVIDUAL.txt
DOMICILIO.txt
INVENTARIO.txt
MORADOR.txt
OUTROS_RENDIMENTOS.txt
RENDIMENTO_TRABALHO.txt
RESTRICAO_PRODUTOS_SERVICOS_SAUDE.txt
SERVICO_NAO_MONETARIO_POF2.txt
SERVICO_NAO_MONETARIO_POF4.txt

FONTE: Elaboração própria (2023).

Para obtenção das variáveis necessárias à realização da estimação das elasticidades, com auxílio dos arquivos dicionário de variáveis e Cadastro de Produtos, ambos disponíveis na documentação da pesquisa, primeiramente, foram

criadas chaves de identificação combinando as variáveis 'UF', 'ESTRATO_POF', 'TIPO_SITUACAO_REG', 'COD_UPA' e 'NUM_DOM', com objetivo de obter um identificador exclusivo para cada domicílio.

Desta forma, utilizando o filtro do produto energia elétrica (600101) no arquivo DESPESA_COLETIVA, observando a quantidade do produto na variável V9005, obteve-se o consumo de energia elétrica em quilowatt-hora (kWh) por domicílio, para o qual adotou-se a sigla *Cde*. Definido o consumo, utilizando o mesmo arquivo de DESPESA_COLETIVA, foi possível obter a renda total por domicílio pela variável RENDA_TOTAL (*Y*).

Na sequência combinou-se os arquivos de DESPESA_COLETIVA, CADERNETA_COLETIVA, e DESPESA_INDIVIDUAL a fim de apurar a despesa total por domicílio, obtida através da variável V8000_DEFLA (*D*). Dispondo destes dados, optou-se pela determinação do número de eletrodomésticos no domicílio e pela relação entre o tipo do eletrodoméstico e seu consumo mensal em kWh, conforme estabelece o Programa Nacional de Conservação de Energia Elétrica – PROCEL.

Desta forma, da relação entre a lista do PROCEL e a lista dos eletrodomésticos presentes na POF, foi possível obter uma lista de 14 aparelhos, conforme TABELA 1.

TABELA 1 – ELETRODOMÉSTICOS NOS DOMICÍLIOS EM CORRESPONDÊNCIA COM O PROCEL

Código do produto	Descrição do produto	Consumo médio mensal (kWh)	Representatividade no domicílio (%)
1402101	Ar-condicionado	181,60	35,39
1400501	Chuveiro elétrico	80,00	15,59
1400401	Geladeira de 2 portas	48,24	9,40
1400201	Freezer	47,55	9,27
1400801	Máquina de lavar louças	30,86	6,01
1400301	Geladeira de 1 porta	25,20	4,91
1401401	Televisão de LED, plasma ou LCD	19,05	3,71
1402201	Ventilador ou circulador de ar	17,52	3,41
1401301	Televisão em cores (tubo)	15,15	2,95
1401001	Forno elétrico	15,00	2,92
1400901	Forno de micro-ondas	13,98	2,72
1401901	Microcomputador ou notebook	9,96	1,94
1401101	Ferro elétrico	7,20	1,4
1401201	Máquina de lavar roupas	1,76	0,34

FONTE: Elaboração própria com base em IBGE (2021) e PROCEL (2022).

A partir da representatividade de cada aparelho, exposta na TABELA 1, multiplicou-se seu valor pelo número de eletrodomésticos em quantidade, para cada produto, obtendo o número de eletrodomésticos ponderado (nEp).

Para controle, foram utilizadas *dummies* de região e classe socioeconômica obtidas através das colunas de região e classe, transformadas em vetores binários, de modo que, seus valores foram adicionados somente à fórmula de regressão, conforme pode-se observar na página 34 (APÊNDICE 1 – Código).

Por fim, foi necessário realizar o cálculo do preço da energia elétrica, ou tarifa de energia elétrica. Para tanto, utilizou-se da despesa com energia elétrica, obtida a partir do filtro do produto energia elétrica, observando a variável V8000_DEFLA, dividida pelo consumo de energia elétrica Cde . Entretanto, as tarifas calculadas apresentaram valores bastante incoerentes com os valores esperados para tarifa de energia elétrica no Brasil. Desta forma, optou-se por utilizar a base das tarifas médias por estado, para os anos de 2017 e 2018, obtida diretamente no site da Agência Nacional de Energia Elétrica – ANEEL, conforme **TABELA 2**. A partir dos valores da **TABELA 2**, temos o preço do kWh, qualificado neste trabalho como P .

TABELA 2 – TARIFA MÉDIA POR ESTADO ENTRE OS ANOS DE 2017 E 2018

Estado	Tarifa média (R\$/kWh)
Acre	0,561360
Alagoas	0,533790
Amapá	0,551380
Amazonas	0,655220
Bahia	0,481490
Ceará	0,483270
Distrito Federal	0,520290
Espírito Santo	0,527220
Goiás	0,525040
Maranhão	0,608640
Mato Grosso	0,533650
Mato Grosso do Sul	0,518930
Minas Gerais	0,539890
Pará	0,634340
Paraíba	0,524720
Paraná	0,475460
Pernambuco	0,501120
Piauí	0,589490
Rio de Janeiro	0,556120
Rio Grande do Norte	0,449710
Rio Grande do Sul	0,523520
Rondônia	0,556380
Roraima	0,546300
Santa Catarina	0,488850
São Paulo	0,474860
Sergipe	0,502990
Tocantis	0,577610

FONTE: Elaboração própria com base em ANEEL (2023).

Assim, exibe-se no **QUADRO 2** a relação das variáveis independentes, dependentes e as variáveis *dummy* que foram utilizadas no modelo de estimação.

QUADRO 2 – RELAÇÃO DE VARIÁVEIS INDEPENDENTES E DEPENDENTES

Variáveis	Característica	Descrição
<i>Cde</i>	Dependente	Consumo de Energia Elétrica
<i>P</i>	Independente	Tarifa de Energia Elétrica
<i>nEp</i>	Independente	Número de Eletrodomésticos ponderado
<i>Y</i>	Independente	Renda Total
<i>D</i>	Independente	Despesa Total
<i>DR</i>	<i>Dummy</i>	<i>Dummy</i> de Região
<i>DC</i>	<i>Dummy</i>	<i>Dummy</i> de Classe Socioeconômica

FONTE: Elaboração própria (2023).

4 APRESENTAÇÃO DOS RESULTADOS

A base resultante, produto do relacionamento exposto no capítulo 3, possui 51.858 observações. Na **TABELA 3** seguem algumas estatísticas descritivas da amostra:

TABELA 3 – CARACTERÍSTICAS DESCRITIVAS DAS VARIÁVEIS

Variáveis	Cde (kWh)	P (R\$/kWh)	nE	Y (R\$)	D (R\$)
Média	173,859	0,526	3,712	4902,798	10853,213
Desvio-padrão	138,096	0,005	1,950	8220,836	29386,803
Min	13,000	0,450	1,000	9,870	44,710
Máx.	2918,000	0,655	24,000	675212,200	2919917,750

FONTE: Elaboração própria (2023).

Observando a **TABELA 3**, nota-se que o consumo de energia elétrica dos domicílios, representado pela variável Cde, apresenta uma variação significativa de domicílio para domicílio, com valor médio igual a 173,859 kWh, desvio-padrão de 138,096 kWh, mínimo de 13,000 kWh e máximo de 2918,000 kWh. Para o valor da tarifa de energia elétrica (P), nota-se uma distribuição mais uniforme dos valores, com valor médio de 0,526 R\$/kWh, desvio-padrão de 0,005 R\$/kWh, mínimo de 0,450 R\$/kWh e máximo de 0,655 R\$/kWh, indicando que existe certa homogeneidade nos preços praticados pelas Distribuidoras de energia elétrica. Para o número de eletrodomésticos presentes no domicílio, a amostra apresenta uma dispersão moderada, com valor médio de 3,712, desvio-padrão de 1,950, mínimo de 1,000 e máximo de 24,000. Isso sugere que a maioria dos domicílios tem um número semelhante de eletrodomésticos. Quanto a renda domiciliar total, a amostra apresenta um conteúdo bastante heterogêneo, com valor médio de R\$ 4.902,798, desvio-padrão de R\$ 8.220,836, mínimo de R\$ 9,870 e máximo R\$ 675.212,200. Isso é consistente com o que se espera para a distribuição de renda, com um pequeno número de indivíduos ganhando muito mais do que a média. De maneira semelhante a renda, a despesa total dos domicílios indica alta dispersão entre os dados, com valor médio de R\$ 10.853,213, desvio-padrão de R\$ 29.386,803, mínimo de R\$ 44,710 e máximo de R\$ 2.919.917,750.

A **TABELA 4** exhibe os valores encontrados para estimação das elasticidades em 4 especificações. Primeiro, considerando somente a variável independente P, na

sequência P e nEp, e assim, sucessivamente até a última rodada com a utilização de todas as variáveis independentes (P, nEp, Y e D).

TABELA 4 – RESULTADOS DE ESTIMAÇÃO DE ELASTICIDADES

Modelos	ln P	nEp	ln Y	D	Constante	R ²
Modelo 1	-0,492669*	-	-	-	5,595017	0,250188
Modelo 2	-0,591190*	0,636771*	-	-	4,777612	0,322248
Modelo 3	-0,576758*	0,624967*	0,144211*	-	3,289323	0,326065
Modelo 4	-0,577183*	0,625397*	0,144863*	-7,48x10 ^{-8**}	3,288174	0,326074

FONTE: Elaboração própria (2023).

NOTA: * p < 0,001, ** p > 0,05.

Analisando os valores obtidos para elasticidade preço, notou-se que no Modelo 1, o valor encontrado de elasticidade preço foi de -0,492669, indicando que uma variação percentual de 1% na tarifa, teria uma contrapartida negativa de 0,492669%. Nos modelos 2, 3 e 4, houve um aumento em magnitude, 0,591190, 0,576758 e 0,577183, respectivamente. Para o coeficiente estimado do número de eletrodomésticos ponderado, o primeiro modelo apresentou valor igual a 0,636771, indicando relação positiva com o consumo de energia elétrica. Nos modelos seguintes os valores obtidos foram inferiores, indicando que o incremento no número de eletrodomésticos provocaria menores acréscimos no consumo de energia elétrica. Para elasticidade renda, os valores obtidos para os modelos 3 e 4 foram, respectivamente, 0,144211 e 0,144863, indicando, assim como para o coeficiente do número de eletrodomésticos ponderado, acréscimos no consumo de energia à medida que a renda aumenta, precisamente 0,144211% e 0,144863% de aumento no consumo, a cada 1% de aumento na renda. A Despesa total por domicílio(D), foi inserida apenas no modelo 4, em que foi possível obter o valor de $-7,48 \times 10^{-8}$ para o coeficiente da despesa, apresentando relação negativa com o consumo de energia elétrica. Cabe destacar que a variável D apresentou valor p superior a 0,05, o que indica que ela não é significativa neste modelo.

Também foi possível observar que do modelo 1 para o modelo 4, obteve-se um aumento significativo no valor de R², o que demonstra que a utilização de mais variáveis confere ao modelo uma capacidade aprimorada de explicar a variação dos dados, ainda que o ganho seja pequeno a partir da terceira especificação. Portanto, o modelo 4, que incorpora todas as variáveis independentes, aparenta maior robustez, em termos de poder explicativo, se comparado aos modelos anteriores.

Os valores estimados de elasticidade preço (-0,492669, -0,591190, -0,576758 e -0,577183) foram superiores aos valores obtidos em outros estudos, como por exemplo em relação a Dantas, Costa e Silva (2016) para classe de consumo residencial, que variou entre -0,1527 e -0,2751. Cabe destacar que Dias (2015), que realizou a estimação das elasticidades com base nos dados da POF 2008/2009, encontrou valores de elasticidade preço entre -3,892 e 0,747.

Já para elasticidade renda, em comparação ao que encontrou Dias (2015), os valores obtidos neste trabalho (0,144211 e 0,144863) corroboram o obtido pela autora na maioria das classes, para todas as regiões, com valores entre 0,0794 e 2,525, com exceção da classe AB na região Sul, com valor igual a -2,198. Em comparação à Dantas, Costa e Silva (2016), este estudo verificou valores estimados de elasticidade renda mais próximos aos valores obtidos pelos autores nas regiões Sudeste e Nordeste, 0,1954 e 0,1930, respectivamente.

5 CONSIDERAÇÕES FINAIS

Este estudo procurou contribuir para as estimações das elasticidades preço e renda do consumo de energia elétrica residencial no Brasil, utilizando os dados da Pesquisa de Orçamentos Familiares de 2017/2018. O estudo fornece base para entendimento do comportamento do consumo de energia elétrica residencial, em um cenário de demanda crescente por energia elétrica e desafios relacionados ao planejamento da expansão do setor.

Os resultados da análise empírica revelaram uma elasticidade-preço negativa, o que sugere que o consumo de energia elétrica residencial é pouco sensível a variações nos preços das tarifas. Para elasticidade-renda os valores obtidos foram positivos, indicando que o consumo de energia elétrica é um bem normal.

Cabe ressaltar que este estudo se limitou a utilizar algumas das principais variáveis utilizadas em outros estudos, mas que ainda seria possível a inserção de variáveis que poderiam influenciar o consumo de energia elétrica como Temperatura e número de cômodos do domicílio. Além disso, a utilização de modelos mais sofisticados poderia auxiliar na utilização das regiões e das classes socioeconômicas para entendimento das elasticidades para subgrupos mais específicos.

REFERÊNCIAS

ANDRADE, T. A.; LOBÃO, W. J. A. **Elasticidade renda e preço da demanda residencial de energia elétrica no Brasil**. 1997.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA (ANEEL). **Atlas de energia elétrica do Brasil – Fatores de Conversão**. 3. Ed. – Brasília: ANEEL, 2008. Disponível em: https://www.fisica.net/energia/atlas_de_energia_eletrica_do_brasil_3a-ed.pdf. Acesso em: 5 ago. 2023.

DANTAS, F. D. C.; COSTA, E. M.; DA SILVA, J. L. M. Elasticidade preço e renda da demanda por energia elétrica nas regiões brasileiras: uma abordagem através de painel dinâmico. **Revista de Economia**, v. 42, n. 3, 2016.

DIAS, T. F. **Elasticidades-preço e renda da demanda domiciliar de eletricidade: Estimação econométrica com dados da POF 2008/2009**. 2015.

EMPRESA DE PESQUISA ENERGÉTICA (EPE). **Anuário Estatístico de Energia Elétrica 2023: Ano base 2022**. Empresa de Pesquisa Energética. Rio de Janeiro: EPE, 2023a. Disponível em: <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-160/topico-168/Fact%20Sheet%202023%20-20Anu%C3%A1rio%20Estat%C3%ADstico%20de%20Energia%20El%C3%A9trica.pdf>. Acesso em: 3 set. 2023.

EMPRESA DE PESQUISA ENERGÉTICA (EPE). **Planejamento energético**. Empresa de Pesquisa Energética. Rio de Janeiro: EPE, 2023b. Disponível em: <https://www.epe.gov.br/pt/areas-de-atuacao/planejamento-energetico>. Acesso em: 5 ago. 2023.

GUJARATI, D. N.; PORTER, D. C. **Econometria básica-5**. São Paulo: Amgh Editora, 2011.

INTERNATIONAL ENERGY AGENCY (IEA). **World Energy Outlook 2022**. Paris, France, 2022. Disponível em: <https://iea.blob.core.windows.net/assets/830fe099-5530-48f2-a7c1-11f35d510983/WorldEnergyOutlook2022.pdf>. Acesso em: 5 ago. 2023.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Pesquisa de orçamentos familiares: 2017-2018: perfil das despesas no Brasil: indicadores de qualidade de vida / IBGE, Coordenação de Trabalho e Rendimento**. Rio de Janeiro: IBGE, 2021.

MODIANO, E. M. **Elasticidade-renda e preços da demanda de energia elétrica no Brasil**. Texto para discussão, 1984.ia.

SCHMIDT, C. A. J.; LIMA, M. A. M. A demanda por energia elétrica no Brasil. **Revista brasileira de economia**, v. 58, n. 1, p. 68-98, 2004.

VARIAN, H. R. **Microeconomia**: Uma Abordagem Moderna. São Paulo: Grupo GEN, 2015. E-book. 9788595155107. Disponível em: https://edisiplinas.usp.br/pluginfile.php/6909148/mod_resource/content/1/Microeconomia%20-%20Hal%20Varian%209a%20edic%CC%A7a%CC%83o.pdf. Acesso em: 10 jun. 2023.

WOOLDRIDGE, J. M. **Introdução à econometria** - uma abordagem moderna. São Paulo: Pioneira Thomson Learning, 2006.

APÊNDICE 1 – Código

```
# Import e instalação das bibliotecas
!pip install lerMicrodados
from lerMicrodados import lerMicrodados
import pandas as pd
import numpy as np
import statsmodels.api as sm
import matplotlib.pyplot as plt
import seaborn as sns

# Execução da função lermicrodados
caminho_arquivo = '/content/drive/MyDrive/Dados_20221226.zip'
lerMicrodados.ler_POF(caminho_arquivo, header=True)

# Carregar o arquivo Excel em um DataFrame
df = pd.read_excel('/content/drive/MyDrive/Dados-Tarifa-média-2017-2018.xlsx')

# Excluir a primeira coluna do DataFrame
df = df.drop(df.columns[0], axis=1)

# Definir a primeira linha como cabeçalhos
df.columns = df.iloc[0]
df = df[1:]

# Lendo o arquivo CSV e criando o DataFrame
df_despesa_coletiva = pd.read_csv('/content/DESPESA_COLETIVA.csv')

# Criação da chave
df_despesa_coletiva['CHAVE'] = df_despesa_coletiva['UF'].astype(str) +
 '_' + df_despesa_coletiva['ESTRATO_POF'].astype(str) + '_' +
df_despesa_coletiva['TIPO_SITUACAO_REG'].astype(str) + '_' +
df_despesa_coletiva['COD_UPA'].astype(str) + '_' +
df_despesa_coletiva['NUM_DOM'].astype(str)

# Seleção das colunas de interesse
colunas_desejadas = ['CHAVE', 'UF', 'V9001', 'V9005', 'RENDA_TOTAL']
df_despesa_coletiva_s = df_despesa_coletiva[colunas_desejadas]

# Agrupar pelos valores IGUAIS de 600101 e aplicar operações de
agregação
df_despesa_coletiva_600101 =
df_despesa_coletiva_s[df_despesa_coletiva_s['V9001'] ==
600101].groupby('CHAVE').agg({
    'V9001': 'first',
```

```

    'V9005': 'sum',
    'UF': 'first',
    'RENDA_TOTAL': 'sum'
}).reset_index()

# Defina as faixas de renda e as classes econômicas correspondentes
faixas_renda = [float('-inf'), 1908, 2862, 5724, 9540, 14310, 23850,
float('inf')]
classes_economicas = ['E', 'D', 'C2', 'C1', 'B2', 'B1', 'A']

# Crie a nova coluna com base nas faixas de renda
df_despesa_coletiva_600101['CLASSE_ECONOMICA'] =
pd.cut(df_despesa_coletiva_600101['RENDA_TOTAL'], bins=faixas_renda,
labels=classes_economicas, right=False)

# Criando dataframe resultante entre despesa coletiva com energia
elétrica e as tarifas por estado pelo id 'UF'
df_merge1 = pd.merge(df_despesa_coletiva_600101, df, on='UF')

# Renomeação de coluna
df_merge1.rename(columns={'Tarifa Média de Aplicação 2017-2018 (kWh)':
'Tarifa_kWh'}, inplace=True)
df_merge1['Tarifa_kWh'] = df_merge1['Tarifa_kWh'].astype(float)
df_merge1['log_quantidade'] = np.log(df_merge1['V9005'])
df_merge1['log_preco'] = np.log(df_merge1['Tarifa_kWh'])
df_merge1['log_renda'] = np.log(df_merge1['RENDA_TOTAL'])

# Lendo o arquivo CSV e criando o DataFrame
df_inventario = pd.read_csv('/content/INVENTARIO.csv')

# Criação da chave
df_inventario['CHAVE'] = df_inventario['UF'].astype(str) + '_' +
df_inventario['ESTRATO_POF'].astype(str) + '_' +
df_inventario['TIPO_SITUACAO_REG'].astype(str) + '_' +
df_inventario['COD_UPA'].astype(str) + '_' +
df_inventario['NUM_DOM'].astype(str)

# Seleção das colunas de interesse
colunas_desejadas12 = ['CHAVE', 'V9001', 'V9005']
df_inventario_s = df_inventario[colunas_desejadas12]
# Filtrando e somando os valores
inventario_filtro = df_inventario_s['V9001'].isin([1400301, 1400401,
1402101, 1400501, 1400201, 1401301, 1401401])
df_inventario_filtrado = df_inventario_s[inventario_filtro]
df_inventario_resultante = df_inventario_filtrado.groupby(['CHAVE',
'V9001'])['V9005'].sum().reset_index()
#df_inventario_resultante =
df_inventario_filtrado.groupby(['CHAVE'])['V9005'].sum().reset_index()

```

```

df_inventario_resultante.rename(columns={'V9005':
'n°_eletrodomesticos'}, inplace=True)

pesos = {'V9001': [1402101, 1400501, 1400401, 1400201, 1400801,
1400301, 1401401, 1402201, 1401301, 1401001, 1400901, 1401901, 1401101,
1401201],
        'PESOS': [0.353947804393163, 0.155924142904477,
0.0940222581713996, 0.0926774124388485, 0.060147738125402,
0.0491161050149102, 0.0371294365291286, 0.0341473872960805,
0.0295281345625353, 0.0292357767945894, 0.0272477439725573,
0.0194125557916074, 0.0140331728614029, 0.00343033114389849]}
df_eletro_pesos = pd.DataFrame(pesos)

df_inventario_resultante2 = pd.merge(df_inventario_resultante,
df_eletro_pesos, on='V9001', how='left')
df_inventario_resultante2['n°_eletro_ponderado'] =
df_inventario_resultante2['n°_eletrodomesticos'] *
df_inventario_resultante2['PESOS']

# Agrupar
df_inventario_resultante3 =
df_inventario_resultante2.groupby('CHAVE').agg({
    'n°_eletro_ponderado': 'sum',
    'n°_eletrodomesticos': 'sum'
}).reset_index()
df_inventario_resultante3

# Criando dataframe resultante entre despesa coletiva com energia
elétrica e as tarifas por estado pelo id 'UF'
df_merge2 = pd.merge(df_merge1, df_inventario_resultante3, on='CHAVE')
df_merge2['log_n_eletro'] = np.log(df_merge2['n°_eletro_ponderado'])

# Seleção das colunas de interesse
colunas_desejadas7 = ['CHAVE', 'V9001', 'V9005', 'V8000_DEFLA']
df_despesa_coletiva_s2 = df_despesa_coletiva[colunas_desejadas7]

# Agrupar pelos valores diferentes de 600101 e aplicar operações de
agregação
df_despesa_coletiva_outros =
df_despesa_coletiva_s2[df_despesa_coletiva_s2['V9001'] !=
1].groupby('CHAVE').agg({
    'V9001': 'last',
    'V9005': 'sum',
    'V8000_DEFLA': 'sum'
})

# Lendo o arquivo CSV e criando o DataFrame
df_caderneta_coletiva = pd.read_csv('/content/CADERNETA_COLETIVA.csv')

# Criação da chave

```

```

df_caderneta_coletiva['CHAVE'] =
df_caderneta_coletiva['UF'].astype(str) + '_' +
df_caderneta_coletiva['ESTRATO_POF'].astype(str) + '_' +
df_caderneta_coletiva['TIPO_SITUACAO_REG'].astype(str) + '_' +
df_caderneta_coletiva['COD_UPA'].astype(str) + '_' +
df_caderneta_coletiva['NUM_DOM'].astype(str)

# Seleção das colunas de interesse
colunas_desejadas2 = ['CHAVE', 'V9001', 'V9005', 'V8000_DEFLA']
df_caderneta_coletiva_s = df_caderneta_coletiva[colunas_desejadas2]

# Agrupar pelos valores diferentes de 600101 e aplicar operações de
agregação
df_caderneta_coletiva_SOMA =
df_caderneta_coletiva_s.groupby('CHAVE').agg({
    'V9001': 'last',
    'V9005': 'sum',
    'V8000_DEFLA': 'sum'
})

# Lendo o arquivo CSV e criando o DataFrame
df_despesa_individual = pd.read_csv('/content/DESPESA_INDIVIDUAL.csv')

# Criação da chave
df_despesa_individual['CHAVE'] =
df_despesa_individual['UF'].astype(str) + '_' +
df_despesa_individual['ESTRATO_POF'].astype(str) + '_' +
df_despesa_individual['TIPO_SITUACAO_REG'].astype(str) + '_' +
df_despesa_individual['COD_UPA'].astype(str) + '_' +
df_despesa_individual['NUM_DOM'].astype(str)

# Seleção das colunas de interesse
colunas_desejadas3 = ['CHAVE', 'V9001', 'V8000_DEFLA']
df_despesa_individual_s = df_despesa_individual[colunas_desejadas3]

# Agrupar pelos valores diferentes de 600101 e aplicar operações de
agregação
df_despesa_individual_SOMA =
df_despesa_individual_s.groupby('CHAVE').agg({
    'V9001': 'last',
    'V8000_DEFLA': 'sum'
})

#Concatenar os dois dataframes
df_concat1 = pd.concat([df_despesa_coletiva_outros,
df_caderneta_coletiva_SOMA])

# Agrupar pelos valores da coluna "CHAVE" e somar as colunas
"V8000_DEFLA"
df_resultante_concat1 =
df_concat1.groupby('CHAVE')[['V8000_DEFLA']].sum()

```

```

# Concatenar os dois dataframes
df_concat2 = pd.concat([df_resultante_concat1,
df_despesa_individual_SOMA])

# Agrupar pelo valores da coluna CHAVE" e somar as colunas
"V8000_DEFLA"
df_resultante_concat2 =
df_concat2.groupby('CHAVE')[['V8000_DEFLA']].sum().reset_index()

#Criando dataframe resultante entre despesa coletiva com energia
elétrica e as tarifas por estado pelo id 'UF'
df_merge3 = pd.merge(df_merge2, df_resultante_concat2, on='CHAVE')
df_merge3['log_despesa'] = np.log(df_merge3['V8000_DEFLA'])

#REGRESSÃO FINAL
# Criando variaveis dummy para as colunas 'Região' e 'CLASSE_ECONOMICA'
data_dummies = pd.get_dummies(df_merge3, columns=['Região',
'CLASSE_ECONOMICA'])

# Definindo as variáveis dummy e excluindo a primeira variável dummy de
cada categoria
dummy_vars = [col for col in data_dummies.columns if
col.startswith('Região_') or col.startswith('CLASSE_ECONOMICA_')]
dummy_vars_drop_first = [col for col in dummy_vars if not
col.endswith('_A') and not col.endswith('_N')]

# Definindo a variável dependente
Y = data_dummies['log_quantidade']

# Definindo uma função para realizar a regressão e retornar os
resultados
# A função adiciona uma constante aos dados.
def perform_regression(X, Y):
    X = sm.add_constant(X)
    model = sm.OLS(Y, X)
    results = model.fit()
    return results

# Inicializando um dicionario para armazenar os resultados de cada
rodada de regressão
results_rounds = {}

# Rodada 1: Apenas 'log_preco' e as variáveis dummy são incluídas como
variáveis independentes
X = data_dummies[['log_preco'] + dummy_vars_drop_first]
result = perform_regression(X, Y)
results_rounds['Round 1'] = {
    'log_preco': result.params['log_preco'],

```

```

'n°_eletro_ponderado': None,
'log_renda': None,
'V8000_DEFLA': None,
'const': result.params['const'],
'R-squared': result.rsquared,
'log_preco_pvalue': result.pvalues['log_preco'],
}

# Rodada 2: 'n°_eletro_ponderado' e adicionado ao modelo como uma
variável independente
X = data_dummies[['log_preco', 'n°_eletro_ponderado']] +
dummy_vars_drop_first
result = perform_regression(X, Y)
results_rounds['Round 2'] = {
    'log_preco': result.params['log_preco'],
    'n°_eletro_ponderado': result.params['n°_eletro_ponderado'],
    'log_renda': None,
    'V8000_DEFLA': None,
    'const': result.params['const'],
    'R-squared': result.rsquared,
    'log_preco_pvalue': result.pvalues['log_preco'],
    'n°_eletro_ponderado_pvalue':
result.pvalues['n°_eletro_ponderado'],
}

# Rodada 3: 'log_renda' é adicionado ao modelo como uma variável
independente
X = data_dummies[['log_preco', 'n°_eletro_ponderado', 'log_renda']] +
dummy_vars_drop_first
result = perform_regression(X, Y)
results_rounds['Round 3'] = {
    'log_preco': result.params['log_preco'],
    'n°_eletro_ponderado': result.params['n°_eletro_ponderado'],
    'log_renda': result.params['log_renda'],
    'V8000_DEFLA': None,
    'const': result.params['const'],
    'R-squared': result.rsquared,
    'log_preco_pvalue': result.pvalues['log_preco'],
    'n°_eletro_ponderado_pvalue':
result.pvalues['n°_eletro_ponderado'],
    'log_renda_pvalue': result.pvalues['log_renda'],
}

# Rodada 4: 'V8000_DEFLA' é adicionado ao modelo como uma variável
independente
X = data_dummies[['log_preco', 'n°_eletro_ponderado', 'log_renda',
'V8000_DEFLA']] + dummy_vars_drop_first
result = perform_regression(X, Y)
results_rounds['Round 4'] = {

```

```

    'log_preco': result.params['log_preco'],
    'n°_eletro_ponderado': result.params['n°_eletro_ponderado'],
    'log_renda': result.params['log_renda'],
    'V8000_DEFLA': result.params['V8000_DEFLA'],
    'const': result.params['const'],
    'R-squared': result.rsquared,
    'log_preco_pvalue': result.pvalues['log_preco'],
    'n°_eletro_ponderado_pvalue':
result.pvalues['n°_eletro_ponderado'],
    'log_renda_pvalue': result.pvalues['log_renda'],
    'V8000_DEFLA_pvalue': result.pvalues['V8000_DEFLA'],
}

# Convertendo o dicionário de resultados em um dataframe transposto
results_rounds_df = pd.DataFrame(results_rounds).T

```