

# ECONOMETRIA EM FOCO

## ENTRE JUROS, CONFIANÇA E ENERGIA NO BRASIL



**ORGANIZADORES:**  
**FERNANDO ANTONIO LUCENA AIUBE**  
**EDISON AMERICO HUARSAYA TITO**

# ECONOMETRIA EM FOCO

## ENTRE JUROS, CONFIANÇA E ENERGIA NO BRASIL



**ORGANIZADORES:**  
**FERNANDO ANTONIO LUCENA AIUBE**  
**EDISON AMERICO HUARSAYA TITO**

2025 by Atena Editora

Copyright © 2025 Atena Editora

Copyright do texto © 2025, o autor

Copyright da edição © 2025, Atena Editora

Os direitos desta edição foram cedidos à Atena Editora pelo autor.

*Open access publication by Atena Editora*

**Editora chefe**

Profª Drª Antonella Carvalho de Oliveira

**Editora executiva**

Natalia Oliveira Scheffer

**Imagens da capa**

iStock

**Edição de arte**

Yago Raphael Massuqueto Rocha



Todo o conteúdo deste livro está licenciado sob a Licença Creative Commons Atribuição 4.0 Internacional (CC BY 4.0).

O conteúdo desta obra, em sua forma, correção e confiabilidade, é de responsabilidade exclusiva dos autores. As opiniões e ideias aqui expressas não refletem, necessariamente, a posição da Atena Editora, que atua apenas como mediadora no processo de publicação. Dessa forma, a responsabilidade pelas informações apresentadas e pelas interpretações decorrentes de sua leitura cabe integralmente aos autores.

A Atena Editora atua com transparência, ética e responsabilidade em todas as etapas do processo editorial. Nosso objetivo é garantir a qualidade da produção e o respeito à autoria, assegurando que cada obra seja entregue ao público com cuidado e profissionalismo.

Para cumprir esse papel, adotamos práticas editoriais que visam assegurar a integridade das obras, prevenindo irregularidades e conduzindo o processo de forma justa e transparente. Nosso compromisso vai além da publicação, buscamos apoiar a difusão do conhecimento, da literatura e da cultura em suas diversas expressões, sempre preservando a autonomia intelectual dos autores e promovendo o acesso a diferentes formas de pensamento e criação.

# Econometria em Foco: Entre Juros, Confiança e Energia no Brasil

## | Organizadores:

Fernando Antonio Lucena Aiube  
Edison Americo Huarsaya Tito

## | Revisão:

Os autores

## | Diagramação:

Nataly Gayde

## | Capa:

Yago Raphael Massuqueto Rocha

### Dados Internacionais de Catalogação na Publicação (CIP)

E19    Econometria em foco: entre juros, confiança e energia no Brasil / Organizadores Fernando Antonio Lucena Aiube, Edison Americo Huarsaya Tito. – Ponta Grossa - PR: Atena, 2025.

Formato: PDF  
Requisitos de sistema: Adobe Acrobat Reader  
Modo de acesso: World Wide Web  
Inclui bibliografia  
ISBN 978-65-258-3683-6  
DOI: <https://doi.org/10.22533/at.ed.836252910>

1. Econometria. I. Aiube, Fernando Antonio Lucena (Organizador). II. Tito, Edison Americo Huarsaya (Organizador). III. Título.

CDD 330.015195

Elaborado por Bibliotecária Janaina Ramos – CRB-8/9166

## Atena Editora

📞 +55 (42) 3323-5493

📞 +55 (42) 99955-2866

🌐 [www.atenaeditora.com.br](http://www.atenaeditora.com.br)

✉ [contato@atenaeditora.com.br](mailto:contato@atenaeditora.com.br)

# CONSELHO EDITORIAL

## CONSELHO EDITORIAL

Prof. Dr. Alexandre Igor Azevedo Pereira – Instituto Federal Goiano  
Prof<sup>a</sup> Dr<sup>a</sup> Amanda Vasconcelos Guimarães – Universidade Federal de Lavras  
Prof. Dr. Antonio Pasqualetto – Pontifícia Universidade Católica de Goiás  
Prof<sup>a</sup> Dr<sup>a</sup> Ariadna Faria Vieira – Universidade Estadual do Piauí  
Prof. Dr. Arinaldo Pereira da Silva – Universidade Federal do Sul e Sudeste do Pará  
Prof. Dr. Benedito Rodrigues da Silva Neto – Universidade Federal de Goiás  
Prof. Dr. Cirênio de Almeida Barbosa – Universidade Federal de Ouro Preto  
Prof. Dr. Cláudio José de Souza – Universidade Federal Fluminense  
Prof<sup>a</sup> Dr<sup>a</sup> Daniela Reis Joaquim de Freitas – Universidade Federal do Piauí  
Prof<sup>a</sup> Dr<sup>a</sup>. Dayane de Melo Barros – Universidade Federal de Pernambuco  
Prof. Dr. Elio Rufato Junior – Universidade Tecnológica Federal do Paraná  
Prof<sup>a</sup> Dr<sup>a</sup> Érica de Melo Azevedo – Instituto Federal do Rio de Janeiro  
Prof. Dr. Fabrício Menezes Ramos – Instituto Federal do Pará  
Prof. Dr. Fabrício Moraes de Almeida – Universidade Federal de Rondônia  
Prof<sup>a</sup> Dr<sup>a</sup> Glécilla Colombelli de Souza Nunes – Universidade Estadual de Maringá  
Prof. Dr. Humberto Costa – Universidade Federal do Paraná  
Prof. Dr. Joachin de Melo Azevedo Sobrinho Neto – Universidade de Pernambuco  
Prof. Dr. João Paulo Roberti Junior – Universidade Federal de Santa Catarina  
Prof<sup>a</sup> Dr<sup>a</sup> Juliana Abonizio – Universidade Federal de Mato Grosso  
Prof. Dr. Julio Cândido de Meirelles Junior – Universidade Federal Fluminense  
Prof<sup>a</sup> Dr<sup>a</sup> Keyla Christina Almeida Portela – Instituto Federal de Educação, Ciência e Tecnologia do Paraná  
Prof<sup>a</sup> Dr<sup>a</sup> Miranilde Oliveira Neves – Instituto de Educação, Ciência e Tecnologia do Pará  
Prof. Dr. Sérgio Nunes de Jesus – Instituto Federal de Educação Ciência e Tecnologia  
Prof<sup>a</sup> Dr<sup>a</sup> Talita de Santos Matos – Universidade Federal Rural do Rio de Janeiro  
Prof. Dr. Tiago da Silva Teófilo – Universidade Federal Rural do Semi-Árido  
Prof. Dr. Valdemar Antonio Paffaro Junior – Universidade Federal de Alfenas

# INTRODUÇÃO

## INTRODUÇÃO

A coletânea *“Econometria em Foco: Entre Juros, Confiança e Energia no Brasil”* reúne cinco estudos aplicados que utilizam ferramentas econométricas para investigar temas centrais da economia brasileira recente, oferecendo ao leitor uma visão estruturada e integrada sobre os desafios e dinâmicas que estão presentes no cenário nacional.

O primeiro capítulo investiga a relação entre o Índice de Confiança do Consumidor (ICC) e a atividade econômica. Utilizando modelos VAR e VECM, os autores demonstram como as expectativas e o sentimento dos consumidores antecipam variações no PIB e outros indicadores, contribuindo para a compreensão dos canais não convencionais de transmissão de choques macroeconômicos.

No segundo capítulo, os autores analisam a taxa neutra de juros brasileira utilizando filtros estatísticos aplicados a séries temporais. A estimativa dessa taxa, que serve de referência para a condução da política monetária, revela-se fundamental para entender a orientação da Selic frente aos ciclos econômicos, sendo um ponto de partida analítico relevante para os agentes que lidam com decisões ligadas à macroeconomia.

O terceiro capítulo desloca o foco para o setor energético, com ênfase na geração eólica no Nordeste. A análise do *curtailment*, o desperdício de energia por restrições na transmissão, ilustra como a infraestrutura, regulação e eficiência impactam diretamente a expansão de fontes renováveis no Brasil. O estudo traz à tona a complexidade do planejamento energético diante de um sistema interligado e altamente dependente de variáveis climáticas e econômicas.

No quarto capítulo, os determinantes do spread bancário para pessoas físicas são explorados à luz de variáveis como Selic, inadimplência, inflação e concorrência bancária. A análise evidencia o peso da política monetária na formação do custo do crédito e suas implicações para o consumo e a intermediação financeira.

Por fim, o quinto capítulo foca na Formação Bruta de Capital Fixo (FBCF), investigando os fatores que condicionam os investimentos produtivos no país. O estudo identifica a importância do PIB, da confiança empresarial, da poupança e da capacidade instalada, trazendo à luz os elementos estruturais que influenciam as decisões de longo prazo no setor produtivo.

# INTRODUÇÃO

## INTRODUÇÃO

Ao reunir análises sobre juros, confiança e energia, três pilares interligados da economia brasileira, esta coletânea oferece uma contribuição valiosa ao debate econômico nacional, articulando teoria, dados e evidência empírica com precisão metodológica e relevância aplicada.

Os organizadores

# SUMÁRIO

## SUMÁRIO

### **CAPÍTULO 1 ..... 1**

UM ESTUDO SOBRE OS DETERMINANTES DA CONFIANÇA DO CONSUMIDOR BRASILEIRO

**Camila Lima Silva**

**Edison Americo Huarsaya Tito**

**Fernando Antonio Lucena Aiube**

 <https://doi.org/10.22533/at.ed.8362529101>

### **CAPÍTULO 2 ..... 19**

ESTIMAÇÃO DA TAXA NEUTRA: APLICAÇÃO DE FILTROS ESTATÍSTICOS

**Luiz Antônio Soares dos Santos Leopoldo**

**Fernando Antonio Lucena Aiube**

**Edison Americo Huarsaya Tito**

 <https://doi.org/10.22533/at.ed.8362529102>

### **CAPÍTULO 3 ..... 41**

ANÁLISE ECONOMÉTRICA DO CURTAILMENT NA GERAÇÃO EÓLICA NO BRASIL ENTRE 2022 E 2024

**Bruna Amaral Lanzillotti Barreto**

**Felipe Cardoso Wilhelm da Costa**

**Fernando Antonio Lucena Aiube**

**Edison Americo Huarsaya Tito**

**Daiane Rodrigues dos Santos**

 <https://doi.org/10.22533/at.ed.8362529103>

### **CAPÍTULO 4 ..... 63**

DETERMINANTES DO SPREAD BANCÁRIO PARA PESSOAS FÍSICAS NO BRASIL

**André Luis R. Vasconcellos**

**Marlon Cecílio de Souza**

**Edison Americo Huarsaya Tito**

**Fernando Antonio Lucena Aiube**

 <https://doi.org/10.22533/at.ed.8362529104>

# SUMÁRIO

## SUMÁRIO

### **CAPÍTULO 5.....86**

DETERMINANTES DA FORMAÇÃO BRUTA DE CAPITAL FIXO NO BRASIL (1999-2023): UMA ANÁLISE COM MODELO VECM

**Marianna Baia de Andrade**

**Matheus de Lucas Ribeiro Mota**

**Fernando Antonio Lucena Aiube**

**Edison Americo Huarsaya Tito**

 <https://doi.org/10.22533/at.ed.8362529105>

### **SOBRE OS AUTORES ..... 116**

# Um Estudo sobre os Determinantes da Confiança do Consumidor Brasileiro

Camila Lima Silva \*

Edison Americo Huarsaya Tito †

Fernando Antonio Lucena Aiube ‡

26 de junho de 2025

## Resumo

O estudo analisou os componentes do sentimento do consumidor brasileiro a partir do Índice de Confiança do Consumidor (ICC). As componentes analisadas foram incerteza econômica, expectativa de inflação, paridade do poder de compra, câmbio, renda disponível das famílias, taxa de juros e notícias de negócios obtidas a partir dos dados da FGV, BACEN e *Google Trends* para o período de março de 2012 a outubro de 2024. A estimativa foi realizada a partir do Modelo Autorregressivo de Defasagem Distribuída (ARDL), com *Bounds test* e Modelo de Correção de Erro (ECM) para capturar a relação de curto e longo prazo. Os resultados mostraram que o indicador de incerteza econômica, a expectativa de inflação, a paridade do poder de compra, as notícias de negócios e a renda disponível (apenas no ARDL) têm efeitos significativos sobre o ICC, enquanto taxa de câmbio e juros não apresentaram significância estatística.

**Palavras-chave:** sentimento do consumidor; Índice de Confiança do Consumidor; ARDL.

**Código JEL:** D12, C51

## 1 Introdução

Os índices de confiança visam refletir o sentimento e a percepção dos agentes sobre o presente e o futuro da economia, e agem como indicadores antecedentes de atividade econômica (Graminho, 2015) [11]. Sua maior vantagem em relação a outras variáveis macroeconômicas é a capacidade de gerar informações sobre o comportamento do consumo de forma mais rápida (Bentes, 2006 [1]). O Índice de Confiança do Consumidor (ICC) é um indicador econômico utilizado por diversos países para análise e tomada de decisões econômicas. No

---

\*Mestranda do Programa de Pós-Graduação em Ciências Econômicas da Universidade do Estado do Rio de Janeiro (PPGCE-UERJ)

†Professor Adjunto do Instituto de Matemática e Estatística - UERJ

‡Professor Associado da Faculdade de Ciências Econômicas - UERJ

Brasil, a Fundação Getulio Vargas desenvolveu a sondagem do consumidor na tentativa de sanar a necessidade de indicadores nacionais com características de monitoramento da situação corrente. Sendo assim, a sondagem do consumidor é aplicada mensalmente com o objetivo de gerar indicadores que antecipem tendências econômicas. Neste trabalho, o ICC será adotado como medida do sentimento do consumidor.

A confiança do consumidor está estritamente relacionada a seus níveis de consumo, e a literatura aponta o uso do ICC como preditor de consumo. De acordo com Carvalho et al. (2016) [4], no Brasil, o consumo das famílias é muito importante para o PIB, representando, nos últimos anos, cerca de 60% do total e sendo foco de políticas anticíclicas durante a crise de 2009. Sua análise é essencial tanto para compreender mudanças econômicas em diferentes cenários quanto para avaliar a qualidade de vida da população, os impactos de políticas públicas e questões relacionadas à pobreza. Assim, a disponibilidade de dados confiáveis sobre o consumo das famílias é indispensável para investigações econômicas amplas e relevantes.

O uso de dados do *Google Trends* é amplamente adotado em pesquisas econômicas (Penna e Huang (2009) [9]; Vosen e Schmidt (2011, 2012) [22, 23]; Woo e Owen (2018) [24]; Oust e Eidjord (2020) [16]; Choi e Varian (2012) [5]; Goel et al. (2010) [10]; Choi e Shin (2002) [6]). Nesse sentido, em consonância com os trabalhos anteriores, este trabalho utilizará a subcategoria notícias de negócios, a fim de incorporar como as notícias podem ter efeito sobre a confiança do consumidor, refletido no volume de pesquisas a respeito da economia para formar sua confiança ou a falta dela.

A investigação empírica neste trabalho é feita por meio do Modelo Autorregressivo de Defasagem Distribuída (ARDL). As variáveis utilizadas para a estimação econométrica são: Índice de Confiança do Consumidor com ajuste sazonal, Indicador de Incerteza Econômica, Expectativa Média de Inflação, Taxa de Paridade do Poder de Compra, Taxa de Câmbio Comercial Média, Renda Nacional Disponível Bruta das Famílias Restrita deflacionada, Taxa de Juros (Selic) e volume de pesquisas sobre notícias de negócios do *Google Trends*. Os dados foram analisados para o período de março de 2012 a outubro de 2024.

Este trabalho tem como objetivo analisar os componentes da confiança do consumidor brasileiro. A hipótese é que variáveis macroeconômicas tradicionais, como taxa de juros, taxa de câmbio e expectativa de inflação, estão relacionadas ao ICC, sendo assim o índice poderia resumir tais informações. Diante de incertezas econômicas, os indivíduos tomam decisões baseando-se em expectativas e confiança; dessa forma, cabe uma análise que busque entender o que compõe a confiança do consumidor. Este trabalho visa contribuir com a literatura já existente com a inclusão do uso de dados do *Google Trends* como proxy do volume de notícias para o período analisado e trazer o ICC como variável dependente, visto que a literatura se concentra em utilizá-lo como variável para prever o consumo.

O trabalho é composto de cinco seções, além desta introdução. Dessa forma, a seção 2 realiza uma breve análise sobre a literatura acerca do sentimento do consumidor e o uso do *Google Trends* na temática. A seção 3 apresenta os dados e a metodologia; a seção 4, os resultados; por fim, a seção 5 apresenta a conclusão.

## 2 Revisão da Literatura

### 2.1 Sentimento do Consumidor

Ao analisar as causas e os efeitos do sentimento do consumidor, Throop (1992) [20] observou que o sentimento do consumidor se move com as condições econômicas e normalmente possui uma relação estável com algumas variáveis, mas eventos como guerras podem provocar alterações abruptas no sentimento do consumidor, independentemente das condições econômicas habituais, influenciando ou rompendo os ciclos de negócios.

Existem diversos índices voltados a capturar os sentimentos dos consumidores no Brasil. Apesar de este trabalho focar no ICC fornecido pela FGV, podemos destacar também outros como o ICC da Federação do Comércio de São Paulo, o Índice de Expectativas do Consumidor (IEC) da Federação do Comércio do Estado do Rio de Janeiro e o Índice Nacional de Expectativas do Consumidor (Inec). A FGV ainda fornece outros dessa classe que denomina Sondagem do Consumidor.

Dada a importância do ICC, existe uma grande investigação acerca da capacidade de previsão dos índices de confiança e se eles fornecem informações adicionais que outras variáveis macroeconômicas são capazes. Bram e Ludvigson (1998) [2] analisaram se os principais índices utilizados nos Estados Unidos — o Índice de Confiança do Consumidor do *Conference Board* e o Índice de Sentimento do Consumidor da Universidade de Michigan — eram eficazes na previsão dos gastos dos consumidores. Os autores concluíram que o índice do *Conference Board* apresentou melhor desempenho, atribuído ao fato de conter um maior número de perguntas relacionadas ao mercado de trabalho. Isso sugere que os consumidores tendem a estar mais preocupados com suas perspectivas futuras de emprego do que com as condições favoráveis para os negócios. A dúvida é se os índices que capturam o sentimento do consumidor são capazes de fornecer informações adicionais e mais informativas comparadas a indicadores econômicos mais tradicionais (Ludvigson, 2004 [15]). Bruestle e Crain (2015) [3] argumentam que, se o ICC não adiciona nenhuma informação, mas apenas reflete as condições do mercado de trabalho, taxas de juros, preços de ações e assim por diante, serve apenas como uma proxy para essas outras dimensões econômicas. Isso significaria que as métricas de confiança do consumidor transmitem informações redundantes em modelos de previsão que levam em consideração essas outras variáveis econômicas.

Para o Brasil, Graminho (2015) [11] considera o sentimento do consumidor como uma variável de estado baseada na parcela dos índices de confiança não relacionada com variáveis macroeconômicas. O resultado encontrado apontou que o ICC exerce efeitos sobre o consumo e a produção industrial.

Bentes (2006) [1] utiliza vetores autorregressivos (VAR) para avaliar o poder preditivo do ICC (Fecomercio) em antecipar o consumo futuro e se transmitir informações adicionais. Concluiu-se que o Índice de Confiança do Consumidor atua como um indicador antecedente do consumo, trazendo mais eficiência quando se trata de bens duráveis, uma vez que estão associados às expectativas dos agentes no longo prazo. Além disso, observou-se que a renda e o consumo encontram-se negativamente correlacionados no longo prazo.

O autor destaca a fragilidade da ideia de que o ICC contém informações adicionais sobre o consumo futuro que não são capturadas pelas variáveis macroeconômicas tradicionais.

A literatura apresenta divergências quanto à formação da confiança do consumidor. De um lado, encontra-se o conceito do "espírito animal" de Keynes (1937) [13], caracterizado como um impulso espontâneo capaz de amplificar os ciclos econômicos. Por outro lado, destaca-se a "visão informativa", segundo a qual a confiança seria construída com base em informações relevantes sobre o futuro da economia, explicando a antecipação em relação à atividade econômica (Graminho, 2015 [11]). Tjandrasa e Dewi (2022) [21] destacam a relevância dos fatores psicológicos dos agentes para prever o estado da economia.

Curtin (2021) [8] analisa a tomada de decisão e o comportamento do consumidor como uma integração perfeita entre o raciocínio consciente e inconsciente, uma vez que o inconsciente processa, organiza e aprende automaticamente para tomar decisões mais informadas e precisas. Não há uma generalização do inconsciente como uma decisão puramente irracional ou permissão de uma maior aceitação da imprecisão ou irracionalidade. O consciente e o inconsciente não representam forças contrárias. Essa ideia está em consonância com o que Kahneman (2012) [12] chamaria de pensamento rápido, atrelado ao sistema 1, que tem como características ser impulsivo e intuitivo. Esse sistema usa heurísticas e atalhos mentais para tomar decisões com base em experiências anteriores.

Simões et al. (2016) [19] apontam o sentimento do consumidor como um indicador de grande relevância para a previsão da demanda agregada na economia, ressaltando que levantamentos destinados à sua mensuração constituem uma ferramenta eficaz para a análise de conjuntura econômica. Bruestle e Crain (2015) [3] destacam que o consenso na literatura é que os índices de confiança do consumidor aumentam as previsões econômicas dos gastos do consumidor, mas apenas modestamente.

A importância de uma medida que reflete o sentimento do consumidor está intimamente relacionada com a necessidade de entender a disposição de compra do consumidor. Tjandrasa e Dewi (2022) [21] relacionam os aumentos de inflação, taxa de juros e taxa de câmbio como fatores que podem afetar psicologicamente os agentes, principalmente considerando compras de bens duráveis, como imóveis e carros, que normalmente são feitas a prazo pela maioria da população. Ainda vale destacar como variáveis influenciadoras as notícias relacionadas à política, taxas de criminalidade e taxas de desemprego, que podem ser facilmente acessadas pela internet. O aprimoramento das fontes de confiança do consumidor é fundamental para recomendações de políticas monetárias e fiscais. Por isso, Choudhry e Wohar (2023) [7] investigam as variáveis importantes para a confiança do consumidor.

Apesar de diversos trabalhos trazerem o ICC como uma variável independente, pode-se destacar o trabalho de Choudhry e Wohar (2023) [7], que coloca o ICC como variável dependente. Os autores investigaram o efeito assimétrico da incerteza e outros determinantes da confiança do consumidor a partir do Método de Defasagem Distribuída Autoregressiva Assimétrica Não Linear (NARDL), uma vez que as variáveis macroeconômicas podem mudar assimetricamente de um estado expansionista para um estado recessivo (Keynes, 1937 [13]). Sendo assim, o modelo aplicado permitiu medir e distinguir entre os efeitos impostos

por uma diminuição ou um aumento nas variáveis. Como medida de confiança, foi utilizado o *Michigan Consumer Confidence Index* e o *US Consumer Confidence*. Os resultados fornecem evidências de efeitos assimétricos tanto no longo quanto no curto prazo. A taxa de desemprego, a inflação esperada e os preços das ações impõem o maior efeito assimétrico em ambas as versões dos índices de confiança. Os resultados de longo prazo fornecem mais evidências do efeito assimétrico da incerteza econômica em comparação com os resultados de curto prazo.

## 2.2 Google Trends

O *Google Trends* é uma ferramenta que fornece o interesse de pesquisa em tópicos específicos. Os dados não são refletidos em termos absolutos, mas são uma amostra do volume total de pesquisas e são normalizados para que os termos possam ser comparados. Cada ponto de dados é dividido pelo total de pesquisas da geografia e do intervalo de tempo que representa, para se comparar a popularidade relativa. Os resultados são apresentados entre 0 e 100, onde 100 indica o maior número de buscas e 0 indica que não há informações suficientes para a pesquisa. Além disso, a ferramenta garante a eliminação de pesquisas repetidas do mesmo usuário durante um curto período de tempo<sup>1</sup>.

A ferramenta permite as buscas por palavras-chave, mas também podem ser consultadas por categorias e subcategorias sem, necessariamente, se utilizar de uma palavra específica naquela busca. A combinação do uso da palavra-chave dentro de uma categoria também é possível. As buscas podem ser realizadas a nível mundial ou por país, e os resultados fornecidos podem ser vistos ao longo do tempo, por sub-região ou por cidade. Os dados do *Google Trends* utilizados para este trabalho podem ser baixados no site<sup>2</sup>.

Penna e Huang (2009) [9], Vosen e Schmidt (2011) [22] e Woo e Owen (2018) [24] utilizam dados do *Google Trends* para construir um novo ICC e compará-lo com os tradicionais já publicados nos EUA. Vosen e Schmidt (2012) [23] também criam um indicador de consumo mensal para a Alemanha. Choi e Varian (2011) [5] mostram como prever valores no curto prazo de indicadores econômicos, como vendas de automóveis, pedidos de seguro-desemprego, planejamento de viagem e confiança do consumidor. Goel et al. (2010) [10] usam o volume de consultas de pesquisa para prever a receita de bilheteria do fim de semana de abertura de longas-metragens, as vendas de videogames no primeiro mês e a classificação das músicas na parada *Billboard Hot 100*. Já Oust e Eidjord (2020) [16] investigaram se índices de volume de pesquisa do *Google* podem ser usados para prever os preços das casas e identificar bolhas no mercado imobiliário dos EUA. Choi e Shin (2002) [6] investigaram como melhorar a previsão do consumo de petróleo usando o *Google Trends*.

Penna e Huang (2009) [9], Vosen e Schmidt (2011) [22], Woo e Owen (2018) [24] e Oust e Eidjord (2020) [16] convergem ao encontrar uma capacidade preditiva superior nos índices formados por dados do *Google Trends*. Lolić, Matošec e Sorić (2024) [14] ressaltam a importância desse tipo de conjunto de dados, pois permite observar com maior profundidade os interesses de pesquisa das pessoas e também observaram que a pandemia de COVID-19

<sup>1</sup>[https://support.google.com/trends/answer/6248105?hl=pt&ref\\_topic=6248052](https://support.google.com/trends/answer/6248105?hl=pt&ref_topic=6248052)

<sup>2</sup><https://trends.google.com.br/home?hl=pt-BR>

gerou a necessidade ainda maior de fontes alternativas e de alta frequência para a tomada de decisões que capturassem o cenário macroeconômico.

Assim como Penna e Huang (2009) [9] destacam, não é possível encontrar uma correspondência perfeita, mas há uma tentativa de se aproximar ao utilizar categorias que possam refletir as informações correspondentes. Para esta pesquisa, não foi utilizada nenhuma palavra-chave específica, mas sim a busca por uma subcategoria — notícias de negócios — que faz parte da categoria de notícias.

## 3 Dados e Metodologia

Nesta seção são apresentados os dados utilizados neste trabalho e a metodologia. O Modelo Autorregressivo de Defasagem Distribuída (ARDL) foi utilizado para verificar a relação entre as variáveis investigadas como componentes importantes no sentimento do consumidor e o Índice de Confiança do Consumidor. Para isso, foram estimados o modelo ARDL, o *Bounds test* para verificar a cointegração e o Modelo de Correção de Erros (ECM).

### 3.1 Base de dados e estatísticas descritivas

Os dados utilizados neste trabalho foram obtidos a partir do Ipeadata do Instituto de Pesquisa Econômica Aplicada, Sistema Gerenciador de Séries Temporais (SGS) do Banco Central do Brasil e *Google Trends*. O período analisado vai de março de 2012 a outubro de 2024. A Tabela 1 apresenta o resumo das variáveis utilizadas no modelo e a Tabela 2, as estatísticas descritivas.

O Índice de Confiança do Consumidor é representado pela média aritmética dos cinco indicadores padronizados que compõem o Índice da Situação Atual (ISA) e o Índice de Expectativas (IE). A avaliação sobre o momento atual contém a situação econômica local e a situação financeira atual das famílias. As previsões em relação ao futuro próximo levam em consideração a situação econômica local futura, a situação financeira futura das famílias e os gastos previstos com a compra de bens duráveis.

O Indicador de Incerteza Econômica Brasil tem como objetivo mensurar a incerteza da economia brasileira a partir de informações dos principais jornais do país e das expectativas do mercado financeiro acerca das variáveis macroeconômicas.

A expectativa média de inflação trata-se da projeção do índice de inflação IPCA. A expectativa para a inflação dos próximos seis meses é calculada com base nas taxas de inflação mensais esperadas no último dia útil de cada mês para cada um dos seis meses subsequentes. A taxa é anualizada para os próximos seis meses.

A taxa de paridade do poder de compra (consumo das famílias) é um fator de conversão que representa, em cada mês, a razão entre os preços ao consumidor de uma mesma cesta de bens no Brasil (em R\$) e nos Estados Unidos (em US\$).

A taxa de câmbio é o preço de uma moeda estrangeira medido em unidades ou frações

(centavos) da moeda nacional. Neste caso, a moeda estrangeira é o dólar. O câmbio comercial ou livre corresponde à média das taxas efetivas de operações no mercado interbancário, ponderada pelo volume de transações de venda do dia. A taxa é calculada para transações de compra, utilizando-se as cotações diárias do período em referência.

A Renda Nacional Disponível Bruta das Famílias restrita é a soma dos rendimentos obtidos pelo uso de fatores de produção (trabalho e capital) e as transferências recebidas, descontadas das transferências pagas, como impostos e contribuições sociais. É apresentada em valores constantes (milhões de reais), deflacionada pelo IPCA, com ajuste sazonal.

A taxa de juros é o coeficiente que determina o valor dos juros, isto é, a remuneração do fator capital utilizado durante certo período de tempo. Refere-se à taxa de juros acumulada mensal, tendo como base a taxa Selic diária. A Selic é a taxa de juros média apurada no Sistema Especial de Liquidação e Custódia (Selic), que incide sobre operações de financiamento por um dia útil (*overnight*), lastreadas por títulos públicos registrados no sistema. É utilizada como instrumento primário de política monetária do Comitê de Política Monetária (Copom), visando ao cumprimento da meta para a inflação. O Copom estabelece a meta para a taxa Selic, e cabe à mesa de operações do mercado aberto do Banco Central do Brasil (BCB) manter a taxa Selic diária próxima à meta.

As notícias de negócios representam o volume de pesquisa no *Google* para o período analisado. Não foi utilizada nenhuma palavra-chave específica, apenas a subcategoria notícias de negócios, presente na categoria de notícias.

Tabela 1: Resumo dos Dados

Variável	Sigla	Unidade	Fonte
Índice de confiança do consumidor com ajuste sazonal	ICC	-	FGV
Indicador de Incerteza Econômica Brasil	IIE	-	FGV
Expectativa média de inflação	EIN	-	BACEN
Taxa de Paridade do Poder de Compra (consumo das famílias)	PPC	%	IPEA
Taxa de Câmbio Comercial média	TC	%	BACEN
Renda Nacional Disponível Bruta das Famílias restrita deflacionada	RD	R\$ milhões	BACEN
Taxa de juros - Over / Selic	TJ	%	BACEN
Notícias > Notícias de negócios	GTNE	Índice	<i>Google Trends</i>

Tabela 2: Estatísticas Descritivas

Variável	Mínimo	Máximo	Média	Desvio Padrão	CV (%)
ICC	58.2	113.2	86.38	11.39	13.19
IIE	85.1	210.5	113.27	18.22	16.09
EIN	1.61	8.66	4.94	1.44	29.06
PPC	1.68	2.50	2.20	0.25	11.29
TC	18.22	57.77	38.75	12.18	31.42
RD	4245	548317	407715.90	131003.60	32.13
TJ	1.79	5.65	3.86	1.21	31.37
GTNE	42	100	64.72	10.20	15.76

### 3.2 Especificação Econométrica

O modelo ARDL [17, 18], estimado usando mínimos quadrados ordinários (OLS), é um modelo linear que comprehende dois componentes principais: a parte autorregressiva(AR) e as defasagens distribuídas(DL) das variáveis independentes. Na parte AR, a variável dependente é considerada na forma defasada, enquanto na parte DL, as variáveis independentes são incluídas em seus níveis e formas defasadas. O modelo está interessado na relação de longo prazo independente das variáveis serem I(0), I(1) ou uma combinação de ambas. Isso ajuda a evitar os problemas de pré-teste associados à análise de cointegração padrão, que requer a classificação das variáveis em I(0) e I(1) como é o caso do modelo VAR. Um nível ótimo de defasagens pode ser determinado para cada uma das variáveis do modelo ARDL.

A equação genérica do modelo ARDL é dada por:

$$y_t = \alpha + \sum_{i=1}^p \beta_i y_{t-i} + \sum_{j=0}^q \gamma_j x_{t-j} + \sum_{k=0}^r \delta_k z_{t-k} + \epsilon_t \quad (1)$$

onde  $y_t$  é a variável dependente no tempo  $t$ ,  $x_t$  e  $z_t$  são as variáveis explicativas no tempo  $t$ ,  $\alpha$  é o intercepto,  $\beta_i$  são os coeficientes das defasagens da variável dependente  $y_t$ ,  $\gamma_j$  são os coeficientes das defasagens da variável explicativa  $x_t$ ,  $\delta_k$  são os coeficientes das defasagens da variável explicativa  $z_t$ ,  $\epsilon_t$  é o erro (resíduo),  $p$  é a ordem da defasagem de  $y_t$ ,  $q$  é a ordem da defasagem de  $x_t$ , e  $r$  é a ordem da defasagem de  $z_t$ .

A especificação do modelo ARDL para os dados utilizados nesse trabalho são:

$$\begin{aligned} ICC = & \alpha + \sum_{i=1}^p \beta_i \cdot ICC_{t-i} + \sum_{j=0}^q \gamma_j \cdot IEE_{t-j} + \sum_{k=0}^r \delta_k \cdot EIN_{t-k} \\ & + \sum_{l=0}^s \theta_l \cdot PPC_{t-l} + \sum_{m=0}^u \phi_m \cdot RD_{t-m} + \sum_{n=0}^v \lambda_n \cdot TC_{t-n} \\ & + \sum_{o=0}^w \psi_o \cdot TJ_{t-o} + \sum_{p=0}^x \xi_p \cdot GTNE_{t-p} + \epsilon_t \end{aligned} \quad (2)$$

No qual  $ICC$  é a variável dependente no tempo  $t$ ,  $IEE$ ,  $EIN$ ,  $PPC$ ,  $RD$ ,  $TC$ ,  $TJ$ , e  $GTNE$  são as variáveis explicativas no tempo  $t$ ;  $\alpha$  é o intercepto;  $\beta_i$  são os coeficientes das defasagens de  $ICC$ ;  $\gamma_j$  são os coeficientes das defasagens de  $IEE$ ;  $\delta_k$  são os coeficientes das defasagens de  $EIN$ ;  $\theta_l$  são os coeficientes das defasagens de  $PPC$ ;  $\phi_m$  são os coeficientes das defasagens de  $RD$ ;  $\lambda_n$  são os coeficientes das defasagens de  $TC$ ;  $\psi_o$  são os coeficientes das defasagens de  $TJ$ ;  $\xi_p$  são os coeficientes das defasagens de  $GTNE$ ;  $\epsilon_t$  é o erro (resíduo); e  $p, q, r, s, u, v, w, x$  são as ordens de defasagem de  $ICC$ ,  $IEE$ ,  $EIN$ ,  $PPC$ ,  $RD$ ,  $TC$ ,  $TJ$ , e  $GTNE$ , respectivamente.

A relação de cointegração é analisada para determinar se há associação entre as variáveis num equilíbrio de longo prazo. Sendo assim, para o teste de cointegração entre as variáveis regressoras para o modelo ARDL é aplicado o *Bound test*, que analisa essa relação independente dos regressores serem puramente I(0), I(1) ou ambos.

O *Bounds test* possui como hipótese nula a não cointegração. A análise do teste é baseada na comparação da estatística F com os valores críticos de I(0) e I(1). Se a estatística F for maior que o valor crítico de I(1), isso indica a presença de cointegração. Por outro lado, se a estatística F for inferior ao valor crítico de I(0), isso sugere a ausência de cointegração. Quando o valor de F está entre os valores críticos de I(0) e I(1), o teste resulta em uma resultado inconclusivo, o que pode exigir a utilização de outro teste ou a coleta de mais informações para uma análise mais precisa.

Quando há cointegração entre as variáveis é necessário utilizar o Modelo de Correção de Erros (ECM) [Pesaran et al (1996b)] derivado do modelo ARDL por meio de uma transformação linear simples, que integra ajustes de curto prazo com equilíbrio de longo prazo sem perder informações de longo prazo. Ou seja, após a confirmação da presença de vetores de longo prazo entre as variáveis analisadas, procede-se à estimativa dos coeficientes de longo e curto prazos, além da velocidade de ajuste em direção ao equilíbrio de longo prazo. O ECM utiliza a primeira diferença da variável dependente, regredida nas primeiras defasagens das variáveis dependentes e independentes. Os regressores restantes no modelo consistem nas defasagens das primeiras diferenças das variáveis dependentes e independentes.

A equação genérica do modelo ECM é dada por:

$$\Delta y_t = \alpha + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \sum_{j=0}^q \gamma_j \Delta x_{t-j} + \sum_{k=0}^r \delta_k \Delta z_{t-k} + \lambda \cdot ECM_{t-1} + \epsilon_t . \quad (3)$$

No qual  $\Delta y_t$ ,  $\Delta x_t$ , e  $\Delta z_t$  denotam as variações das variáveis utilizadas no presente modelo.  $ECM_{t-1}$  é o termo de erro de cointegração (obtido do modelo ARDL original), e  $\lambda$  é o coeficiente de ajuste que indica a velocidade de ajuste à relação de longo prazo.

O termo de erro de cointergação mede a relação entre curto e longo prazo, deve ser estatisticamente significativo, com sinal negativo confirmado a relação de equilíbrio no longo prazo. Caso contrário, o modelo é explosivo e não há convergência. A especificação do modelo ECM para os dados utilizados nesse trabalho é:

$$\begin{aligned} \Delta ICC = & \alpha + \sum_{i=1}^p \beta_i \Delta ICC_{t-i} + \sum_{j=0}^q \gamma_j \Delta IEE_{t-j} \\ & + \sum_{k=0}^r \delta_k \Delta EIN_{t-k} + \sum_{l=0}^s \theta_l \Delta PPC_{t-l} + \sum_{m=0}^u \phi_m \Delta RD_{t-m} \\ & + \sum_{n=0}^v \lambda_n \Delta TC_{t-n} + \sum_{o=0}^w \psi_o \Delta TJ_{t-o} + \sum_{p=0}^x \xi_p \Delta GTNE_{t-p} \\ & + \eta \cdot ECM_{t-1} + \epsilon_t \end{aligned} \quad (4)$$

onde  $\Delta ICC$ ,  $\Delta IEE$ ,  $\Delta EIN$ ,  $\Delta PPC$ ,  $\Delta RD$ ,  $\Delta TC$ ,  $\Delta TJ$ , e  $\Delta GTNE$  são as variações das variáveis,  $ECM_{t-1}$  é o termo de erro de cointegração (obtido do modelo ARDL original), e  $\eta$  é o coeficiente de ajuste que indica a velocidade de retorno ao equilíbrio de longo prazo.

## 4 Resultados

Essa seção apresenta os resultados do teste de raiz unitária utilizado para verificar se o modelo ARDL é adequado para os dados. Além disso, é apresentado o resultado da estimativa do modelo ARDL, teste de cointegração e o modelo ECM.

### 4.1 Teste de raiz unitária

O teste de raiz unitária é utilizado para verificar a estacionariedade de uma série afim de terminar sua ordem de integração, ou seja, verificar se a série é estacionária em nível ou em alguma diferença. No caso do modelo ARDL é necessário que as séries sejam integradas de ordem zero ou um.

Para o teste de raiz unitária foram utilizados: Augmented Dickey-Fuller (ADF) e Kwiatkowski-Phillips-Schmidt-Shin(KPSS). Os resultados dos testes estão apresentados na tabela 2 e 3, respectivamente.

Tabela 3: Teste ADF

Variável	p-valor Nível	p-valor 1 <sup>a</sup> Diferença	Ordem de Integração
ICC	0.666	0.01	I(1)
IIE	0.4804	0.01	I(1)
EIN	0.3156	0.01	I(1)
PPC	0.7844	0.01	I(1)
TC	0.4079	0.01	I(1)
RD	0.01	0.01	I(0)
TJ	0.3847	0.01	I(1)
GTNE	0.01	0.01	I(0)

Tabela 4: Teste KPSS

Variável	p-valor Nível	p-valor 1 <sup>a</sup> Diferença	Ordem de Integração
ICC	0.662	0.01	I(1)
IIE	0.08415	0.01	I(1)
EIN	0.01016	0.01	I(1)
PPC	0.9486	0.01	I(1)
TC	0.3722	0.01	I(1)
RD	0.01	0.01	I(0)
TJ	0.4765	0.01	I(1)
GTNE	0.01	0.01	I(0)

A hipótese nula de ambos os testes é que a série possui raiz unitária. Para as séries em nível ICC, IIE, EIN, PPC, TC, TJ a hipótese nula não é rejeitada. Enquanto, em primeira diferença as séries rejeitam a hipótese de raiz unitária ao nível de 5%, ou seja, as séries

apresentam estacionaridade em primeira diferença de forma que são consideradas I(1). Já as séries RD e GTNE em nível rejeitam a hipótese nula ao nível de 5%, logo, são I(0). Portanto, o modelo ARDL pode ser utilizado, uma vez que as variáveis apresentam I(0) ou I(1).

## 4.2 Modelo

O modelo ARDL selecionado com base no Critério de Informação de Akaike (AIC), foi o de defasagens de ordem (3, 4, 0, 2, 1, 0, 2, 4). A tabela<sup>3</sup> 5 apresenta os resultados do modelo ARDL, o qual apresentou um bom ajuste, com o coeficiente de determinação igual a 96.73% o que indica que o modelo explica bem o Índice de Confiança do Consumidor, a variável dependente. O p-valor <2.2e-16 também indica que o modelo é estatisticamente significativo.

Tabela 5: Resultados da Estimação do Modelo ARDL

Variável	Estimativa	Erro Padrão	t-valor	p-valor
Intercepto	29.18	8.812	3.312	0.001216 **
L(ICC, 1)	0.6735	0.08328	8.088	4.69e-13 ***
L(ICC, 2)	0.2147	0.0977	2.152	0.033345 *
L(ICC, 3)	-0.05862	0.0789	-0.734	0.464452
IEE	-0.2133	0.03133	-6.807	3.79e-10 ***
L(IEE, 1)	0.0191	0.04809	0.397	0.691954
L(IEE, 2)	0.09242	0.04830	1.913	0.057995 .
L(IEE, 3)	0.02934	0.04196	0.699	0.485808
L(IEE, 4)	-0.03887	0.02625	-1.480	0.141298
EIN	-0.5268	0.1849	-2.850	0.005128 **
PPC	-42.31	21.85	-1.937	0.055040 .
L(PPC, 1)	-7.282	35.54	-0.205	0.837976
L(PPC, 2)	42.07	22.70	1.854	0.066180 .
RD	-5.278e-07	1.308e-06	-0.403	0.687309
L(RD, 1)	3.760e-06	1.283e-06	2.930	0.004041 **
TC	-0.057	0.2016	-0.283	0.777911
TJ	-3.216	2.724	-1.181	0.240007
L(TJ, 1)	1.523	1.968	0.774	0.440615
L(TJ, 2)	2.163	1.402	1.542	0.125518
GTNE	0.0217	0.03679	0.602	0.547947
L(GTNE, 1)	0.01604	0.03715	0.432	0.666598
L(GTNE, 2)	0.1309	0.0352	3.709	0.000313 ***
L(GTNE, 3)	-0.01509	0.0382	-0.397	0.692036
L(GTNE, 4)	0.1067	0.03739	2.854	0.005066 **

Erro padrão residual: 1.949 com 124 graus de liberdade.  $R^2$ : 0.9724  $R^2$  ajustado: 0.9673. Estatística F: 190 com 23 e 124 graus de liberdade, p-valor: <2.2e-16

Os coeficientes das defasagens 1 e 2 para o índice de Confiança do Consumidor (ICC)

<sup>3</sup>Nível de significância: 0 \*\*\*, 0.001 \*\*, 0.01 \*, 0.05 ?, 0.1 ‘ ’ 1

são positivos e estatisticamente significativo ao nível de 1% e 5%, respectivamente. O coeficiente da defasagem 1 indica que 67,5% da variação do ICC de um período anterior é transmitida para o período atual. Em termos econômicos, há uma forte inércia ou persistência no sentimento do consumidor, o que é esperado em variáveis de confiança. O coeficiente da segunda defasagem, apesar de menor indica que os efeitos de choques na confiança se prolongam por pelo menos dois período, reforçando a ideia de que a confiança do consumidor responde de forma gradual a eventos passados, com memória de curto prazo.

O efeito contemporâneo do Indicador de Incerteza Econômica (IIE) é negativo e estatisticamente significativo ao nível de 1%, indicando que maiores níveis de incerteza reduzem a variável dependente imediatamente. Isso significa que quando a incerteza econômica aumenta, os consumidores rapidamente ajustam suas expectativas e percepções de forma negativa.

A Paridade do Poder de Compra apresenta sinal negativo, no entanto não é significativo ao nível de 5%, pois o p-valor é 0.055040, com esse valor o teste pode ser considerado inconclusivo, caso significativo poderia indicar que a valorização relativa do consumo das famílias pode sinalizar pressões inflacionárias, reduzindo a confiança. As defasagens da variável também não apresentaram significância.

As notícias de negócios (GTNE) possuem um efeito no curto prazo, as defasagens 2 e 4 são positivas e estatisticamente significativas ao nível de 1%, indicando que notícias positivas tem o poder de aumentar a confiança do consumidor, mas com um atraso, uma vez que não ocorre imediatamente.

Apesar de serem variáveis relevantes para o consumo, a Taxa de Câmbio Comercial Média (TC) e a Taxa de Juros (TJ) não são estatisticamente significativas.

Tabela 6: Análise de Resíduos

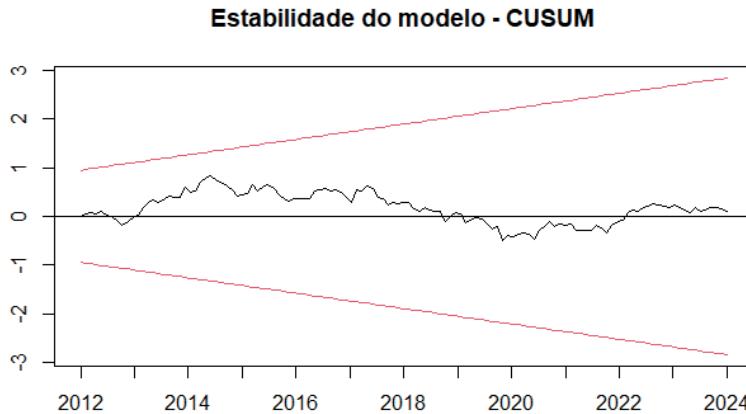
Teste	p-valor	Hipótese Nula
<b>BG</b>	0.5161	Não há autocorrelação
<b>BP</b>	0.1428	Homocedasticidade
<b>Skapiro-Wilk</b>	0.1789	Distribuição Normal

**Fonte:** Elaborado pela autora.

A tabela 6 apresenta o diagnóstico de resíduos. Para testar a autocorrelação serial foi utilizado o *Breusch-Godfrey test*(BG) que apresentou o p-valor maior que 0.05 dessa forma a hipótese nula não é rejeitada, ou seja, não há evidências estatísticas de correlação serial. Para testar o comportamento da variância dos resíduos foi utilizado o *studentized Breusch-Pagan test* (BP) que também apresentou um p-valor superior a 0.05, não rejeitando a hipótese nula, logo, não há evidências estatísticas para dizer que os resíduos não são homocedásticos. Por fim, foi testada a normalidade através do *Shapiro-Wilk normality test* que também não rejeitou a hipótese nula, dessa forma, os resíduos seguem uma distribuição normal. Portanto, os resultados são consistentes com um modelo robusto.

A estabilidade dos parâmetros do modelo é avaliada através do *cumulative sum of recursive (CUSUM) test*. Os testes são baseados nos resíduos das estimativas recursivas, a hipótese nula é que os parâmetros do modelo são estáveis ao longo do tempo, logo a distribuição está simetricamente em torno de zero. A hipótese nula não é rejeitada quando o gráfico está dentro do intervalo de confiança. A figura 1 apresenta o gráfico de estabilidade desse modelo. Portanto, como a curva está dentro do intervalo de confiança, há evidências estatísticas de estabilidade dos parâmetros do modelo.

Figura 1



O teste de cointegração foi utilizado para identificar se os modelos de longo e curto prazo devem ser especificados ou apenas o último. A tabela 6 apresenta os resultados do teste. A estatística F obtida é 4.203. Esse valor é superior aos valores críticos de I(1) em todos os níveis de significância considerados (10%, 5%, 2,5% e 1%), sendo assim, rejeita-se a hipótese nula, ou seja, há evidências da existência de uma relação de equilíbrio de longo prazo estatisticamente significativa entre as variáveis analisadas. Portanto, o Modelo de Correção de erros é estimado e seus resultados estão apresentados na tabela 7.

Tabela 7: ARDL Bound Test

Teste	Valor	Nível de Significância	I(0)	I(1)
<b>F-statistic</b>	4.203	10%	1.92	2.89
<b>k</b>	7	5%	2.17	3.21
		2,5%	2.43	3.51
		1%	2.73	3.9

**Fonte:** Elaborado pela autora.

A tabela 8<sup>4</sup> apresenta os resultados da estimação do Modelo ECM. O termo de correção de erro (ECT) representa o desvio da relação de longo prazo entre as variáveis dependentes

<sup>4</sup>Nível de significância: 0 \*\*\* 0.001 \*\* 0.01 \* 0.05 ? 0.1 ' ' 1

e independentes e neste caso é identificado como L(ICC,1). O ECT é estatisticamente significativo ao nível de 1% de significância, com sinal negativo, o que indica que 17.04% do desequilíbrio é corrigido a cada período, ou seja, o modelo converge para equilíbrio de longo prazo.

Tabela 8: Resultados da Estimação do Modelo ECM

Variável	Estimativa	Erro Padrão	t-valor	p-valor
<b>Intercepto</b>	29.18	8.812	3.312	0.001216 **
L(ICC, 1)	-0.1704	0.0333	-5.117	1.15e-06 ***
L(IEE, 1)	-0.1113	0.02797	-3.979	0.000117 ***
EIN	-0.5268	0.1849	-2.850	0.005128 **
L(PPC, 1)	-7.524	2.976	-2.528	0.012724 *
L(RD, 1)	3.232e-06	1.926e-06	1.678	0.095937 .
TC	-0.057	0.2016	-0.283	0.777911
L(TJ, 1)	0.4694	2.065	0.227	0.820585
L(GTNE, 1)	0.2607	0.0697	3.740	0.000280 ***
d(L(ICC, 1))	-0.1561	0.08162	-1.912	0.058161 .
d(L(ICC, 2))	0.05862	0.07989	0.734	0.464452
d(IIE))	0.21335	0.03133	-6.807	3.79e-10 ***
d(L(IIE, 1))	-0.08289	0.03710	-2.234	0.027243 *
d(L(IIE, 2))	9.532e-03	0.03643	0.262	0.794036
d(L(IIE, 3))	0.03887	0.02625	1.480	0.141298
d(PPC)	-42.31	21.85	-1.937	0.055040 .
d(L(PPC, 1))	-42.07	22.70	-1.854	0.066180 .
d(RD)	-5.278e-07	1.308e-06	-0.403	0.687309
d(TJ)	-3.216	2.724	-1.181	0.240007
d(L(TJ, 1))	-2.163	1.402	-1.542	0.125518
d(GTNE)	0.02217	0.03679	0.602	0.547947
d(L(GTNE, 1))	-0.2225	0.0160	-4.312	3.27e-05 ***
d(L(GTNE, 2))	-0.0916	0.04665	-1.964	0.051792 .
d(L(GTNE, 3))	-0.1067	0.03739	-2.854	0.005066 **

*Erro padrão residual: 1.949 com 124 graus de liberdade. R<sup>2</sup>: 0.7028 R<sup>2</sup> ajustado: 0.6477 Estatística F: 12.75 com 23 e 124 graus de liberdade, p-valor: <2.2e-16*

**Fonte:** Elaborado pela autora.

No longo prazo, o coeficiente de L(IIE,1) é negativo (-0.1704) e significativo ao nível de 1% indicando que as incertezas são prejudiciais no horizonte de longo prazo. O coeficiente de L(IIE,1) também é negativo (-0.1113) e estatisticamente significativo ao nível de 1%. L(PPC,1) também negativo e significante ao nível 5% indicando que desajustes no poder de compra geram um efeito negativo. L(GTNE,1) positivo ao nível 1% de significância, boas notícias promovem confiança aos agentes. Não há alteração no coeficiente de EIN em relação ao já apresentado no ARDL.

O efeito de curto prazo é identificado pelos coeficientes das diferenças das variáveis independentes sobre a variável dependente capturando o efeito transitório das mudanças. d(IIE) apresenta um coeficiente negativo e estatisticamente significativo ao nível de 1%.

Esse efeito é mais forte no curto prazo em comparação ao longo prazo, sugerindo um efeito imediato que se dissipa com o tempo.  $d(L(GTNE,1))$  e  $d(L(GTNE,3))$  também são negativas e estatisticamente significativas ao nível de 1% indicando que a confiança empresarial pode ter flutuações momentâneas antes de estabilizar no longo prazo.

Portanto, o ICC tem efeito significativo apenas no longo prazo, sugerindo que o otimismo dos consumidores afeta a economia mais como uma tendência de médio a longo prazo, não sendo um fator de ajuste imediato. O IEE tem efeito negativo tanto no curto quanto no longo prazo. O EIN e PCC possuem efeito apenas no longo prazo. Enquanto GTNE possui um efeito positivo no longo prazo, no curto prazo, mudanças abruptas nas notícias podem desestabilizar as expectativas, resultando em volatilidade e ajustes negativos.

## 5 Conclusão

O presente trabalho analisou os componentes do sentimento do consumidor brasileiro aqui observado a partir do Índice de Confiança do Consumidor, uma medida que busca capturar a percepção dos agentes sobre o presente e futuro da economia. A literatura faz uma ampla discussão se o ICC pode fornecer novas informações sobre o sentimento do consumidor que outras variáveis macroeconômicas não capturam. Além disso, investigam como em modelos de previsão a inserção do ICC pode prever melhor o consumo, utilizando o ICC como uma variável independente. No entanto, há uma investigação também acerca dos componentes da confiança do consumidor, tratando-o como uma variável dependente que é o caso desse trabalho.

A partir do modelo ARDL utilizando o ICC como variável dependente foi investigado se indicador de incerteza econômica, expectativa de inflação, paridade do poder de compra, câmbio, renda disponível das famílias, taxa de juros e notícias relacionadas a negócios poderiam explicar o sentimento do consumidor no curto e longo prazo. Foi realizado o teste de cointegração, que rejeitou a hipótese nula de ausência de cointegração. Assim, conclui-se que existe cointegração entre as variáveis, o que justifica a utilização do Modelo de Correção de Erro (ECM).

A partir dos modelos estimados conclui-se que a variáveis estatisticamente significativas são indicador de incerteza econômica, expectativa de inflação, paridade do poder de compra, notícias de negócios e renda nacional (significante apenas no ARDL). Em contrapartida, taxa de juros e taxa de câmbio não apresentaram significância estatística, dessa forma, o ICC não carrega informações dessas variáveis consigo, rejeitando a hipótese do trabalho de que o ICC resumiria de certa forma tais informações, de maneira que o ICC não apresenta informações redundantes.

A incerteza econômica desestimula investimentos e consumo, pois empresas e consumidores tendem a adotar comportamentos mais cautelosos. Esse resultado confirma que ambientes de alta incerteza são prejudiciais ao crescimento econômico no longo prazo. Portanto, choques de incerteza (IIE) têm efeito imediato e prolongado, prejudicando a estabilidade e o desempenho econômico. A inflação esperada corrói o poder de compra das famílias e aumenta a incerteza sobre preços futuros, prejudicando tanto o consumo quanto o investi-

mento. A inflação elevada desestabiliza a economia, afetando negativamente o crescimento.

A GTNE pode estar captando expectativas positivas que impulsionam o crescimento com algum atraso, indicando que as pessoas podem demorar a reagir a notícias positivas. Esse resultado sugere que GTNE possivelmente captando efeitos relacionados à renda, expectativas ou dinamismo econômico, tem um impacto positivo e significativo no crescimento econômico, mas com um atraso, sendo assim no curto prazo apresentasse sinal negativo enquanto no longo prazo positivo. Isso pode estar relacionado a reações graduais de consumo ou investimento diante de mudanças nos determinantes de GTNE.

A rejeição da hipótese inicial indica que o ICC não incorpora diretamente os movimentos de variáveis como juros e câmbio, sugerindo que o comportamento dos consumidores brasileiros é mais sensível a fatores de percepção direta, como inflação, renda e incerteza econômica, do que a variáveis financeiras mais abstratas. Esse achado está alinhado com a literatura internacional, como Ludvigson (2004), que destacam que consumidores tendem a reagir com maior intensidade a variáveis macroeconômicas com efeitos visíveis no seu cotidiano.

Adicionalmente, características estruturais da economia brasileira, como elevada informalidade, baixo nível de educação financeira e menor acesso ao crédito bancário formal, podem limitar a transmissão direta das variações de juros e câmbio para a percepção de confiança do consumidor médio. Dessa forma, os resultados sugerem que o ICC reflete dimensões comportamentais e expectativas que não são totalmente captadas pelas variáveis macroeconômicas tradicionais. Por fim, a rejeição da hipótese de que o ICC resumiria as informações de juros e câmbio reforça a importância de considerá-lo como um indicador independente, que acrescenta informações próprias ao modelo, evidenciando seu valor analítico para entender o comportamento dos consumidores.

O próximo passo seria investigar se a construção de um Índice de Confiança do Consumidor a partir de dados do *Google Trends* conseguiria ser explicado por variáveis como taxa de juros e taxas de câmbio que para o ICC da FGV não foi encontrado significância estatística. Isso traria mais robustez ao índice e usabilidade a medida que ele é estimado com maior rapidez do que outros dados macroeconômicos ficam disponíveis.

# Referências

- [1] BENTES, F. G. M. O poder preditivo do índice de confiança do consumidor no brasil: uma análise através de vetores autorregressivos. *Rio de Janeiro 83* (2006).
- [2] BRAM, J., AND LUDVIGSON, S. C. Does consumer confidence forecast household expenditure? a sentiment index horse race. *Economic Policy Review* 4, 2 (1998).
- [3] BRUESTLE, S., AND CRAIN, W. M. A mean-variance approach to forecasting with the consumer confidence index. *Applied Economics* 47, 23 (2015), 2430–2444.
- [4] CARVALHO, S. S. D., SANTOS, C. H. M. D., ALMEIDA, V. A. L. D., JOEL, Y. K. Z., PAIVA, K. C., AND CALDAS, L. F. O consumo das famílias no brasil entre 2000 e 2013: uma análise estrutural a partir de dados do sistema de contas nacionais e da pesquisa de orçamentos familiares.
- [5] CHOI, H., AND VARIAN, H. Predicting the present with google trends. *Economic record* 88 (2012), 2–9.
- [6] CHOI, J.-E., AND SHIN, D. W. How to improve oil consumption forecast using google trends from online big data?: the structured regularization methods for large vector autoregressive model. *Communications for Statistical Applications and Methods* 29, 1 (2022), 41–51.
- [7] CHOUDHRY, T., AND WOHar, M. What drives us consumer confidence? the asymmetric effects of economic uncertainty. *International Journal of Finance & Economics* 29, 4 (2024), 4268–4285.
- [8] CURTIN, R. Nonconscious cognitive reasoning: A neglected ability shaping economic behavior. *Journal of Behavioral Economics for Policy* 5, S3 (2021), 35–43.
- [9] DELLA PENNA, N., AND HUANG, H. Constructing consumer sentiment index for us using internet search patterns. *Department of Economics, WP* 26 (2009).
- [10] GOEL, S., HOFMAN, J. M., LAHAIE, S., PENNOCK, D. M., AND WATTS, D. J. Predicting consumer behavior with web search. *Proceedings of the National academy of sciences* 107, 41 (2010), 17486–17490.
- [11] GRAMINHO, F. M., ET AL. Sentimento e macroeconomia: uma análise dos índices de confiança no brasil. *Trabalhos para Discussão* 408 (2015).
- [12] KAHNEMAN, D. *Rápido e devagar: duas formas de pensar*. Objetiva, 2012.
- [13] KEYNES, J. M. The general theory of employment. *The quarterly journal of economics* 51, 2 (1937), 209–223.
- [14] LOLIĆ, I., MATOŠEC, M., AND SORIĆ, P. Diy google trends indicators in social sciences: A methodological note. *Technology in society* 77 (2024), 102477.
- [15] LUDVIGSON, S. C. Consumer confidence and consumer spending. *Journal of Economic perspectives* 18, 2 (2004), 29–50.

- [16] OUST, A., AND EIDJORD, O. M. Can google search data be used as a housing bubble indicator? *International Real Estate Review* 23, 2 (2020).
- [17] PESARAN, M. H., SHIN, Y., AND OTHERS. *An autoregressive distributed lag modelling approach to cointegration analysis*, vol. 9514. Department of Applied Economics, University of Cambridge Cambridge, UK, 1995.
- [18] PESARAN, M. H., SHIN, Y., AND SMITH, R. J. Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics* 16, 3 (2001), 289–326.
- [19] SIMÕES, J. E. M., DINIZ, M. J. T., DINIZ, M. B., AND DE OLIVEIRA JUNIOR, J. N. Índice de expectativas do consumidor no brasil: Uma análise do poder preditivo no período entre 2001 e 2014. *Planejamento e Políticas Públicas*, 47 (2016).
- [20] THROOP, A. W., ET AL. Consumer sentiment: Its causes and effects. *Federal Reserve Bank of San Francisco Economic Review* 1 (1992), 35–59.
- [21] TJANDRASA, B. B., AND DEWI, V. I. Determinants of consumer confidence index to predict the economy in indonesia. *Australasian Accounting, Business and Finance Journal* 16, 4 (2022), 3–13.
- [22] VOSEN, S., AND SCHMIDT, T. Forecasting private consumption: survey-based indicators vs. google trends. *Journal of forecasting* 30, 6 (2011), 565–578.
- [23] VOSEN, S., AND SCHMIDT, T. A monthly consumption indicator for germany based on internet search query data. *Applied Economics Letters* 19, 7 (2012), 683–687.
- [24] WOO, J., AND OWEN, A. L. Forecasting private consumption with google trends data. *Journal of Forecasting* 38, 2 (2019), 81–91.

# Estimação da Taxa Neutra: aplicação de filtros estatísticos

Luiz Antônio Soares dos Santos Leopoldo \*

Fernando Antonio Lucena Aiube †

Edison Americo Huarsaya Tito ‡

25 de julho de 2025

## Resumo

A taxa de juros neutra é um conceito fundamental para a formulação de políticas monetárias e estratégias de mercado financeiro, mas sua estimativa é desafiadora, pois se trata de uma variável não observável, sujeita a mudanças estruturais e incertezas. Diante desse desafio, este estudo propõe a utilização de filtros estatísticos, permitindo diferentes abordagens para a modelagem e interpretação da taxa neutra. Três métodos estatísticos amplamente utilizados em séries temporais: o Filtro de Kalman, que ajusta dinamicamente a taxa neutra a partir de um modelo estocástico; o Filtro de Hodrick-Prescott (HP), que separa tendências de longo prazo das flutuações cíclicas; e a Suavização Exponencial Holt-Winters, que captura padrões persistentes e responde rapidamente a mudanças recentes. A análise comparativa mostra que cada método oferece uma visão distinta da dinâmica da taxa neutra, permitindo que analistas e traders escolham a abordagem mais adequada conforme o horizonte de investimento e o grau de suavização desejado.

Dessa forma, argumenta-se que não há um único modelo ideal para estimar a taxa neutra, mas sim um conjunto de técnicas que podem ser combinadas estrategicamente para obter uma visão mais robusta. O estudo contribui para a literatura ao fornecer um arcabouço prático para a estimação da taxa neutra, ressaltando sua aplicabilidade em decisões de mercado e precificação de ativos.

**Palavras-chave:** Taxa de Juros; Taxa Neutra de Juros; Filtros estatísticos; Filtro de Kalman; Filtro de Hodrick-Prescott; Suavização Exponencial Holt-Winters.

**Código JEL:** E43, E52, C32, G11, G17.

---

\*Mestrando do Programa de Pós-Graduação em Ciências Econômicas da Universidade do Estado do Rio de Janeiro (PPGCE-UERJ)

†Professor Associado da Faculdade de Ciências Econômicas - UERJ

‡Professor Adjunto do Instituto de Matemática e Estatística - UERJ

# 1 Introdução

A taxa de juros neutra desempenha um papel crucial tanto na formulação da política monetária quanto na tomada de decisões no mercado financeiro. Em termos macroeconómicos, ela é definida como a taxa real de juros que equilibra a economia, mantendo a inflação dentro da meta sem gerar pressões inflacionárias ou deflacionárias, análise contida em Holston et al (2017)[6]. No entanto, para os agentes do mercado financeiro, a taxa neutra serve como um referencial dinâmico, ajudando na avaliação do nível apropriado das taxas de juros e suas implicações sobre ativos financeiros, expectativas de inflação e ciclos econômicos. Operadores de mercado, traders e gestores de fundos utilizam a taxa neutra para calibrar estratégias de investimento, estimar prêmios de risco e ajustar posições em resposta às mudanças nas condições monetárias e macroeconómicas.

Apesar de sua importância, a taxa neutra é uma variável não observável, o que impõe desafios significativos à sua estimação. Para lidar com essa incerteza, diferentes abordagens estatísticas e econométricas vêm sendo empregadas, permitindo que analistas de mercado extraiam estimativas robustas a partir de dados históricos. Métodos como o Filtro de Kalman, o Filtro de Hodrick-Prescott (HP) e a Suavização Exponencial Holt-Winters oferecem formas distintas de modelagem da taxa neutra, cada qual com vantagens e limitações. Enquanto o Filtro de Kalman permite capturar dinâmicas latentes e ajustar-se a mudanças estruturais, o Filtro HP é amplamente utilizado para separar componentes de tendência e ciclo em séries temporais macroeconómicas. Já a Suavização Exponencial Holt-Winters fornece um método responsivo para acompanhar variações ao longo do tempo sem exigir conhecimento detalhado da estrutura da série.

Diante desses desafios, este artigo propõe a adoção de diferentes métodos estatísticos, permitindo que analistas utilizem abordagens complementares para estimar a taxa neutra. Essa abordagem evita a busca por um modelo único e definitivo, reconhecendo que diferentes técnicas oferecem perspectivas distintas sobre a dinâmica da taxa neutra. O objetivo é avaliar a aplicabilidade de três filtros estatísticos amplamente utilizados em séries temporais: o Filtro de Kalman, o Filtro de Hodrick-Prescott (HP) e a Suavização Exponencial Holt-Winters, ressaltando como cada um pode ser empregado em diferentes contextos analíticos no mercado financeiro.

O artigo está organizado da seguinte forma: a Seção 2 apresenta a revisão da literatura sobre a taxa neutra e os métodos de estimação. A Seção 3 descreve a formulação matemática dos filtros estatísticos utilizados. A Seção 4 detalha os dados empregados na análise. Na Seção 5, são apresentados os resultados empíricos e a comparação entre os diferentes filtros. Por fim, a Seção 6 discute as implicações dos achados e sugere direções para futuras pesquisas.

## 2 Revisão da Literatura

A taxa de juros neutra é aquela que equilibra a economia, ou seja, não acelera nem desacelera a atividade econômica. Ela serve como referência fundamental para a política monetária, sendo definida como a taxa consistente, no médio prazo, com a inflação na

meta e o crescimento do produto igual ao seu crescimento potencial (Blinder, 1999)[2]. Na literatura macroeconômica, os termos taxa de juros real de equilíbrio, taxa de juros de equilíbrio, taxa de juros real neutra, taxa de juros neutra e taxa neutra são frequentemente utilizados como sinônimos, embora possam ter nuances metodológicas distintas. O conceito de taxa de juros neutra tem origem em Wicksell (1898;1936)[12], que identificou a "taxa natural de juros" como aquela que equilibra poupança e investimento sem pressão inflacionária.

A taxa de juros neutra é um conceito teórico, a dificuldade do seu uso se deve ao fato de ser uma variável não observável e ao longo do tempo, evolui em linha com outras variáveis, tal como o crescimento do produto, consumo, poupança, investimento, fundamentos do sistema financeiro e o prêmio de risco. Estimá-la com precisão é difícil, mas ela serve como um ponto de referência: se a expectativa de inflação, estiver acima da meta de inflação, deve se elevar a taxa de juros nominal, no caso brasileiro, a taxa Selic, superior à taxa neutra, afim de reduzir a inflação no futuro. Dessa forma, a taxa neutra funciona como um guia para as decisões do Banco Central na calibragem da política monetária, como pode ser visto em Clarida et al(1999)[3].

Embora a taxa neutra seja amplamente utilizada na formulação da política monetária, sua relevância vai além do âmbito dos bancos centrais. Para analistas financeiros, traders e gestores de fundos, compreender a taxa de juros neutra é essencial para a tomada de decisões estratégicas, no entanto, métodos tradicionais de estimação da taxa neutra, como modelos DSGE (Dynamic Stochastic General Equilibrium) ou VARs estruturais (Vector Autoregression), muitas vezes são inacessíveis para profissionais do mercado financeiro, pois exigem calibração frequente, longos períodos de tempo e suposições subjetivas sobre parâmetros econômicos.

Diante dessas dificuldades, o uso de filtros estatísticos, como o filtro de Kalman, usado no trabalho de Laubach e Williams(2003)[8], o filtro de Hodrick-Prescott (HP) e a suavização exponencial Holt-Winters, emerge como uma abordagem mais prática para estimar a taxa neutra em tempo real. Esses métodos permitem capturar tendências latentes e filtrar ruídos da série temporal, tornando-se ferramentas úteis para agentes do mercado financeiro que necessitam de estimativas rápidas e adaptáveis às condições de mercado.

Tendo em vista a explicação de Nielsen (2019)[9], o Filtro de Kalman é um método estatístico amplamente utilizado para estimar variáveis não observáveis em processos dinâmicos, combinando previsões baseadas em modelos com observações ruidosas. Esse filtro foi aplicado inicialmente em missões do projeto Apollo na exploração lunar, onde os engenheiros da NASA perceberam que os recursos computacionais a bordo eram limitados para técnicas mais intensivas em memória. Sua principal vantagem reside na capacidade de atualização recursiva das estimativas sem a necessidade de armazenar grandes quantidades de dados passados, tornando-se um método eficiente para modelagem de séries temporais.

O Filtro de Hodrick-Prescott (HP), por sua vez, foi originalmente proposto por Whittaker (1923)[11] e ganhou notoriedade após os trabalhos de Hodrick e Prescott (1997)[5]. Este filtro é um caso especial de um Spline de suavização (Smoothing Spline) e é amplamente

empregado em macroeconomia para separar tendências de longo prazo das flutuações de curto prazo em séries temporais econômicas. Ao minimizar uma função de penalização baseada na segunda diferença da série, o Filtro HP permite extrair um componente suavizado da série temporal, ajudando na identificação de ciclos econômicos e na análise de políticas monetárias.

Por fim, a Suavização Exponencial Holt-Winters teve suas estruturas fundamentais propostas por C. C. Holt (1957)[7] e seu aluno Peter Winters (1960)[4][10]. Essa técnica se diferencia da média móvel simples ao atribuir pesos exponencialmente decrescentes às observações passadas, permitindo uma maior adaptação a padrões de tendência e sazonalidade. Esse método é amplamente utilizado para previsão de séries temporais devido à sua capacidade de responder rapidamente a mudanças estruturais nos dados.

No entanto, a escolha do método adequado para a estimativa da taxa neutra depende das características específicas dos dados e dos objetivos da análise. Enquanto o Filtro de Kalman permite capturar mudanças dinâmicas na estrutura da taxa neutra, o Filtro HP é amplamente utilizado para extrair tendências de longo prazo, e a Suavização Exponencial fornece uma abordagem mais simples para modelar padrões persistentes. Assim, uma abordagem combinada pode oferecer uma visão mais robusta da taxa neutra, equilibrando rigor estatístico e aplicabilidade prática no contexto do mercado financeiro.

No mercado financeiro, o Filtro de Kalman é útil pois permite que traders e analistas monitorem a evolução da taxa neutra em tempo real, ajustando posições de investimento conforme mudanças estruturais ocorrem na economia. Já o filtro HP pode vir ser utilizado como uma referência rápida para identificar tendências de longo prazo nos juros, facilitando comparações com as condições atuais da política monetária. E a Suavização Exponencial é utilizada para extrair uma tendência subjacente da taxa real ex-ante, permitindo uma interpretação dinâmica do comportamento da taxa neutra ao longo do tempo.

Para traders e analistas financeiros, a abordagem mais interessante pode ser combinar esses métodos permitindo uma visão mais robusta da dinâmica da taxa neutra. Em um ambiente de mercado volátil, um trader pode utilizar o Filtro de Kalman para monitoramento em tempo real, o filtro HP para um referencial de longo prazo e a Suavização Exponencial para ajustes táticos de curto prazo.

Tabela 1: Comparação e Aplicabilidade dos Métodos no Mercado Financeiro

Método	Vantagens	Limitações
<b>Filtro de Kalman</b>	Adaptativo, permite modelagem de variáveis latentes	Exige parametrização e conhecimento estatístico avançado
<b>Filtro de Hodrick-Prescott</b>	Simples e amplamente utilizado	Problema de borda e sensível ao parâmetro $\lambda$
<b>Suavização Exponencial Holt-Winters</b>	Responde rapidamente a mudanças estruturais	Depende da calibração dos parâmetros $\alpha, \beta$

Portanto, a aplicação dessas metodologias estatísticas não apenas permite uma melhor compreensão da taxa de juros neutra, mas também viabiliza decisões estratégicas.

## 2.1 Relevância da Estimação da Taxa Neutra para o Mercado Financeiro

A estimativa da taxa de juros neutra não é apenas uma referência teórica para formuladores de política monetária, mas também uma ferramenta essencial para traders, gestores de fundos e analistas do mercado financeiro. O conhecimento da taxa neutra permite avaliar se a política monetária está em território restritivo ou expansionista, influenciando decisões sobre alocação de ativos, estratégias de hedge e precificação de derivativos.

Gestores de Fundos podem utilizar a taxa neutra para calibrar suas estratégias de alocação de portfólio e avaliar riscos de juros. Traders ajustam posições de renda fixa com base em desvios entre a taxa neutra estimada e a taxa real observada. Economistas, analistas de Macro monitoram a taxa neutra para antecipar mudanças nas diretrizes do Banco Central e reprecificar ativos.

Dessa forma, a escolha da metodologia estatística para estimar a taxa neutra pode impactar diretamente as estratégias do mercado financeiro. Métodos como o Filtro de Kalman são mais responsivos a mudanças na estrutura da economia, enquanto o Filtro de Hodrick-Prescott é útil para análises de longo prazo, e a Suavização Exponencial Holt-Winters pode ser adequada para ajustes táticos.

## 3 Formulação Matemática dos Filtros Estatísticos

### 3.1 Filtro de Kalman

O Filtro de Kalman pode ser representado por um modelo de espaço de estado, onde:

$$x_t = x_{t-1} + v_t, \quad v_t \sim N(0, Q) \quad (1)$$

$$y_t = x_t + w_t, \quad w_t \sim N(0, R) \quad (2)$$

onde  $x_t$  representa a taxa neutra latente e  $y_t$  é a taxa real ex-ante observável. O termo  $Q$  são variâncias do ruído e regula a variabilidade da tendência, enquanto  $R$  reflete a incerteza das observações.

### 3.2 Filtro de Hodrick-Prescott (HP)

O Filtro HP separa a série original  $y_t$  em uma componente de tendência  $g_t$  e um componente cíclico, resolvendo o seguinte problema de minimização:

$$\min_{g_t} \sum_{t=1}^T (y_t - g_t)^2 + \lambda \sum_{t=2}^{T-1} [(g_{t+1} - g_t) - (g_t - g_{t-1})]^2 \quad (3)$$

onde  $\lambda$  é o parâmetro de suavização. Para séries mensais, um valor comum é  $\lambda = 14.400$ .

### 3.3 Suavização Exponencial Holt-Winters

O modelo Holt-Winters é baseado nas seguintes equações:

$$l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1}) \quad (4)$$

$$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} \quad (5)$$

$$y_t = l_t + b_t + \epsilon_t \quad (6)$$

onde  $l_t$  representa o nível suavizado da série,  $b_t$  é a tendência estimada, e  $\alpha$ ,  $\beta$  são parâmetros de suavização.

## 4 Análise de Dados

Os dados mensais que compõe este artigo são séries temporais que compreendem o período de janeiro de 2014 até dezembro de 2024. Os dados utilizados no estudo foram obtidos do Banco Central do Brasil [1].

As principais variáveis utilizadas na análise incluem:

1. Taxa SELIC Mensal: Representa a taxa básica de juros da economia brasileira, definida pelo Comitê de Política Monetária (COPOM) do Banco Central do Brasil. É um dos principais instrumentos de política monetária e serve de referência para a precificação de ativos e operações de crédito.
2. Inflação Esperada (5 e 30 dias úteis): Mediana das expectativas de inflação medida a partir das projeções do mercado financeiro, com diferentes horizontes de previsão. Essas expectativas são fundamentais para estimar a taxa de juros real ex-ante.

3. Taxa Real Ex-Ante (5 e 30 dias úteis): Calculada como a diferença entre a taxa SELIC e a inflação esperada, reflete a taxa de juros ajustada pelas expectativas de mercado, sendo um indicativo relevante para decisões de investimento.

O gráfico abaixo representa a evolução da Taxa SELIC mensal ao longo do período analisado. Alguns pontos de destaque: a taxa permaneceu constante em 0,65 p.p. por um longo período. Entre 2014 e 2015, observa-se um ajuste gradual para 0,66 p.p. A partir de 2024, houve um novo aumento, estabilizando-se em 0,67 p.p..

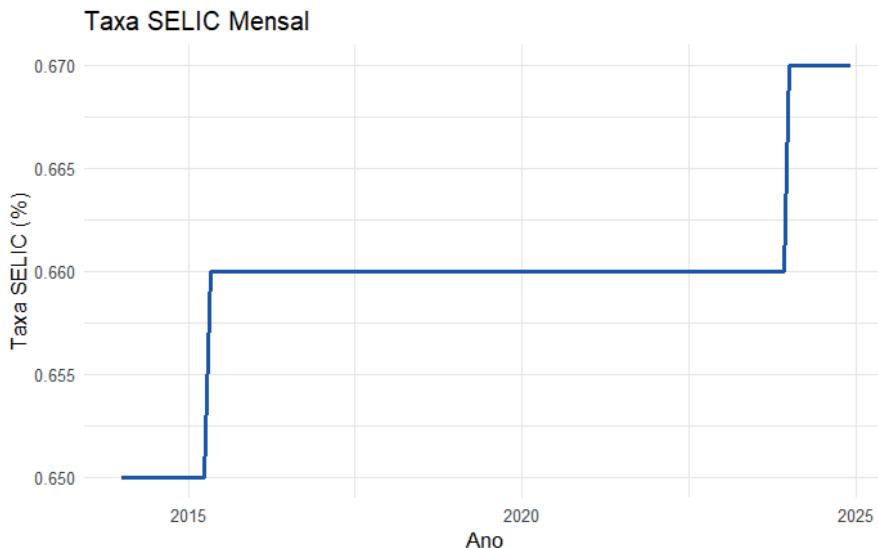


Figura 1: Evolução da Taxa SELIC Mensal

A Taxa SELIC é um referencial importante para o custo do dinheiro e serve de base para operações no mercado financeiro. Alterações na SELIC impactam diretamente a curva de juros, influenciando a precificação de ativos, decisões de investimento e expectativas inflacionárias.

As estatísticas descritivas (média, mediana, desvio padrão, valores máximos e mínimos, quartis e percentis) estão apresentadas na tabela abaixo.

Tabela 2: Estatísticas Descritivas das Variáveis

Variável	Média	Mediana	Desvio Padrão	Mínimo	Máximo	Q1	Q3	P5	P95
Taxa Real Ex-Ante (5d)	0.192	0.230	0.368	-0.730	1.300	-0.040	0.380	-0.444	0.750
Taxa Real Ex-Ante (30d)	0.004	0.000	0.027	-0.100	0.100	-0.010	0.020	-0.039	0.050
SELIC Mensal	0.660	0.660	0.0046	0.650	0.670	0.660	0.660	0.650	0.670
Inflação Esperada (5d)	0.468	0.430	0.367	-0.640	1.380	0.280	0.700	-0.090	1.100
Inflação Esperada (30d)	0.464	0.430	0.355	-0.650	1.400	0.275	0.685	-0.089	1.080

Fonte: Elaboração própria com base nos dados utilizados no estudo.

A taxa real ex-ante (5 dias úteis) apresenta uma média próxima de 0,19 p.p. ao mês, com um desvio padrão de aproximadamente 0,36 pontos percentuais, sugerindo variações expressivas ao longo do tempo. Seu primeiro quartil (Q1) é de -0,04 p.p., indicando que

em 1/4 das observações, a taxa esteve abaixo desse nível, enquanto o terceiro quartil (Q3) é de 0,38 p.p., refletindo uma concentração das observações dentro desse intervalo.

A taxa SELIC mensal exibe uma baixa dispersão, com um desvio padrão de 0,0046 pontos percentuais, indicando estabilidade na política monetária ao longo do período analisado. Em contrapartida, a inflação esperada (5 dias úteis) apresenta uma variação significativa, evidenciada por seu desvio padrão de 0,36 pontos percentuais, demonstrando oscilações nas expectativas inflacionárias do mercado.

Tabela 3: Estatísticas Descritivas das Variáveis Analisadas

Variável	Assimetria	Curtose
Taxa Real Ex-Ante (5d)	-0.077	3.440
Taxa Real Ex-Ante (30d)	0.138	5.610
SELIC Mensal	-0.114	4.699
Inflação Esperada (5d)	0.070	3.445
Inflação Esperada (30d)	0.031	3.495

Sobre a assimetria, um valor positivo indica uma cauda mais longa à direita (assimetria à direita), enquanto um valor negativo sugere uma cauda mais longa à esquerda (assimetria à esquerda). No caso da Taxa Real Ex-Ante (5d), a assimetria é ligeiramente negativa (-0.077), o que sugere uma distribuição um pouco inclinada para a esquerda. Por outro lado, a Taxa Real Ex-Ante (30d) apresenta uma assimetria positiva (0.138), indicando que a cauda direita da distribuição é mais pronunciada.

Já para a curtose, os resultados mostram que a Taxa Real Ex-Ante (30d) possui curtose de 5.61, enquanto a Taxa Real Ex-Ante (5d) apresenta um valor menor (3.44), mas ainda acima de 3, sugerindo que ambas as variáveis possuem caudas relativamente pesadas e podem ser influenciadas por eventos extremos.

A Taxa SELIC Mensal, apresenta assimetria ligeiramente negativa (-0.114) e curtose de 4.699. Já as variáveis de Inflação Esperada (5d e 30d) mostram um comportamento mais simétrico, com assimetrias próximas de zero (0.070 e 0.031, respectivamente) e curtozes em torno de 3.4, indicando que seus valores seguem uma distribuição relativamente concentrada, porém com caudas levemente pesadas.

Esses resultados são importantes porque distribuições com caudas mais pesadas podem indicar maior propensão a choques inesperados, o que é relevante para a modelagem da taxa neutra e para a formulação de estratégias de hedge e tomada de decisão no mercado financeiro. A presença de assimetria e alta curtose sugere que modelos estatísticos robustos devem ser empregados para capturar corretamente as dinâmicas das variáveis analisadas, justificando a aplicação dos filtros estatísticos utilizados neste estudo.

A Taxa Real Ex-Ante de 5 dias úteis exibe uma série temporal altamente volátil, com períodos de forte oscilação e picos expressivos. Podemos observar que, ao longo dos anos, há momentos de elevação e quedas acentuadas, o que pode estar relacionado a choques econômicos, decisões de política monetária ou mudanças na percepção dos agentes de mer-

cado em relação às expectativas inflacionárias. A volatilidade parece se intensificar em determinados períodos, indicando possíveis episódios de instabilidade macroeconômica.

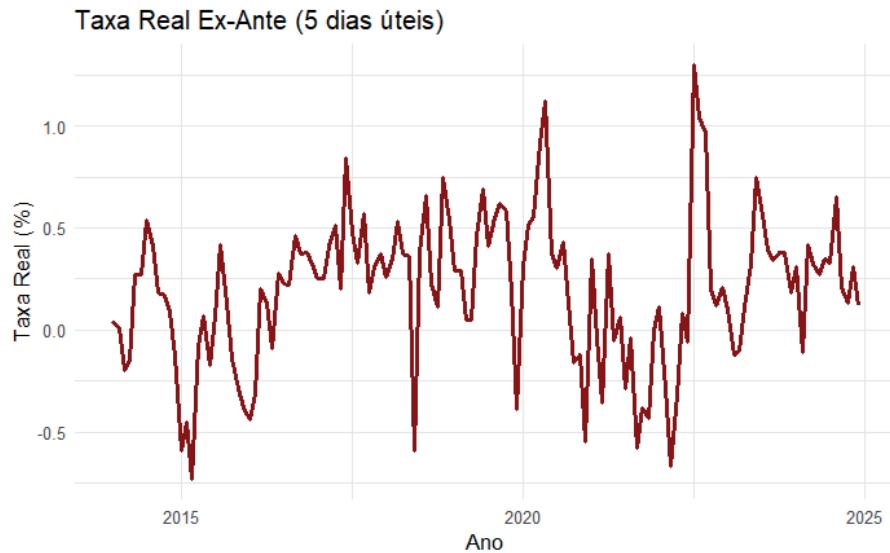


Figura 2: Evolução da Taxa Real Ex-Ante (5 dias úteis)

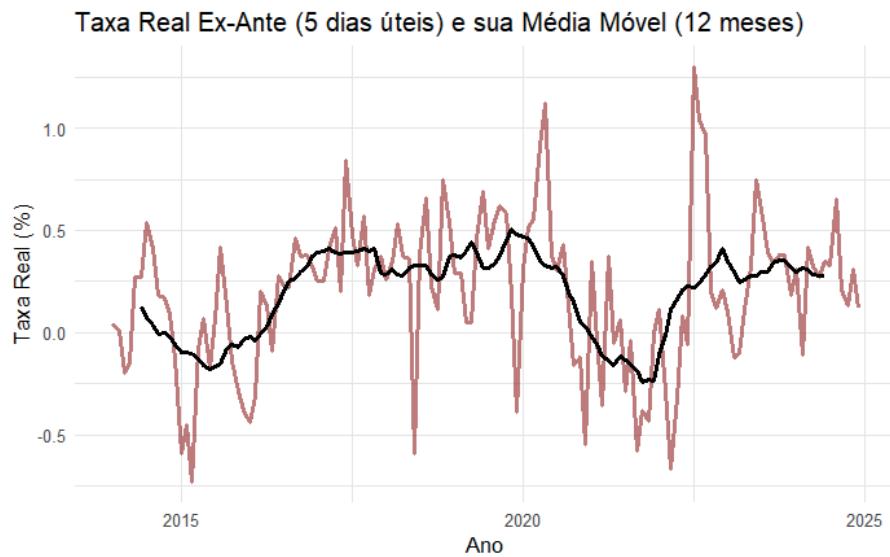


Figura 3: Evolução da Taxa Real Ex-Ante (5 dias úteis) e sua Média Móvel (12 meses)

A figura acima, exibe a evolução da Taxa Real Ex-Ante (5 dias úteis) ao longo do tempo, acompanhada de uma média móvel de 12 meses (linha preta) que suaviza as flutuações de curto prazo.

Observa-se que a taxa real ex-ante apresenta uma elevada volatilidade, com oscilações

expressivas ao longo dos anos.

A média móvel permite identificar tendências subjacentes na série, evidenciando períodos de alta e baixa persistentes. Destacam-se os seguintes padrões: entre 2015 e 2019, a taxa real ex-ante apresentou momentos de estabilização e crescimento moderado. A partir de 2020, houve uma intensificação da volatilidade, refletindo possivelmente os efeitos da pandemia e das respostas da política monetária. Em períodos mais recentes, a taxa parece convergir para níveis próximos da neutralidade, mas ainda apresenta oscilações.

Esse tipo de análise é essencial para traders e analistas de mercado, pois a média móvel auxilia na identificação de tendências estruturais e pode servir como um indicativo do comportamento da taxa neutra ao longo do tempo.

### Taxa Real Ex-Ante vs. Inflação Esperada (5 dias úteis)

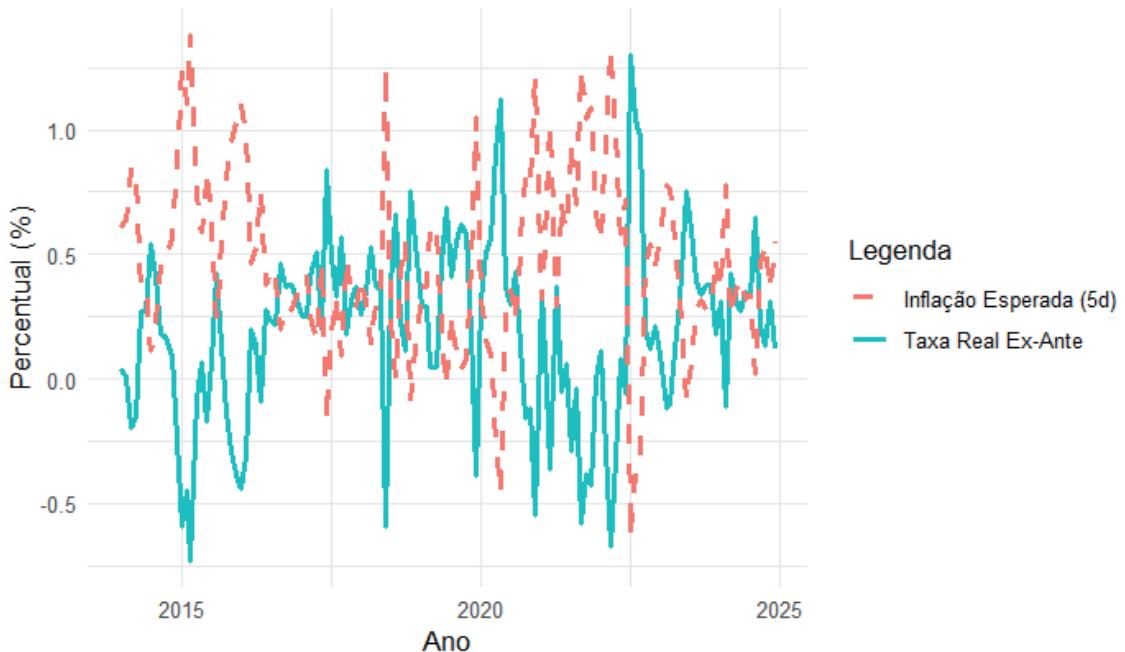


Figura 4: Comparação entre a Taxa Real Ex-Ante (5 dias úteis) e a Inflação Esperada (5 dias úteis)

O gráfico acima ilustra a relação entre a Taxa Real Ex-Ante (5 dias úteis) (linha azul) e a Inflação Esperada para 5 dias úteis (linha vermelha tracejada). Esse comparativo é fundamental para avaliar como as expectativas de inflação afetam a percepção da taxa real ao longo do tempo.

Em diversos períodos, nota-se que a inflação esperada oscila com certa independência da taxa real ex-ante, indicando momentos de desalinhamento entre expectativas inflacionárias e política monetária. Durante 2020-2022, há um aumento na volatilidade de ambas as séries, refletindo incertezas macroeconômicas e prováveis choques econômicos. A taxa real ex-ante tende a reagir às variações da inflação esperada, reforçando sua relação com as

expectativas de mercado.

O teste de Dickey-Fuller aplicado à taxa real ex-ante revelou que a série é estacionária reforçando a adequação das metodologias de filtragem utilizadas na análise. Isso significa que os filtros (Filtro de Kalman, Filtro HP e Suavização Exponencial) podem ser aplicadas sem a necessidade de diferenciação prévia da série.

Tabela 4: Teste de Dickey-Fuller para Estacionariedade da Taxa Real Ex-Ante (5d)

Estatística	Valor
Estatística Dickey-Fuller	-3.8395
Ordem do Lag	5
p-valor	0.01927
Hipótese alternativa	Série estacionária

Essas estatísticas iniciais e os testes estatísticos fornecem um panorama da estrutura dos dados e subsidiam a aplicação dos filtros estatísticos para estimação da taxa neutra. Nos próximos passos, analisaremos a dinâmica temporal das variáveis e a comparação entre os diferentes métodos de estimação da taxa de juros neutra.

## 5 Metodologia

A seguir, detalhamos a implementação de cada um desses métodos e os resultados obtidos com sua aplicação à Taxa Real Ex-Ante de 5 dias úteis.

### 5.1 Justificativa para o Horizonte de 5 Dias

A escolha da taxa real ex-ante com horizonte de 5 dias úteis fundamenta-se na necessidade de capturar as expectativas de curto prazo do mercado financeiro. Enquanto a taxa de 30 dias reflete projeções mais amplas, a taxa de 5 dias está mais alinhada com o comportamento dos traders e gestores de ativos, que frequentemente ajustam suas posições em função de choques macroeconômicos e anúncios de política monetária.

Além disso, em períodos de maior volatilidade, como durante crises financeiras, decisões de investimento precisam ser tomadas com base em projeções que consideram mudanças rápidas nos fundamentos econômicos. Assim, a taxa real ex-ante de 5 dias se mostra mais responsável e aplicável para uma análise dinâmica da taxa neutra no contexto de mercado.

## 5.2 Filtro de Kalman: Modelagem Estocástica da Taxa Neutra

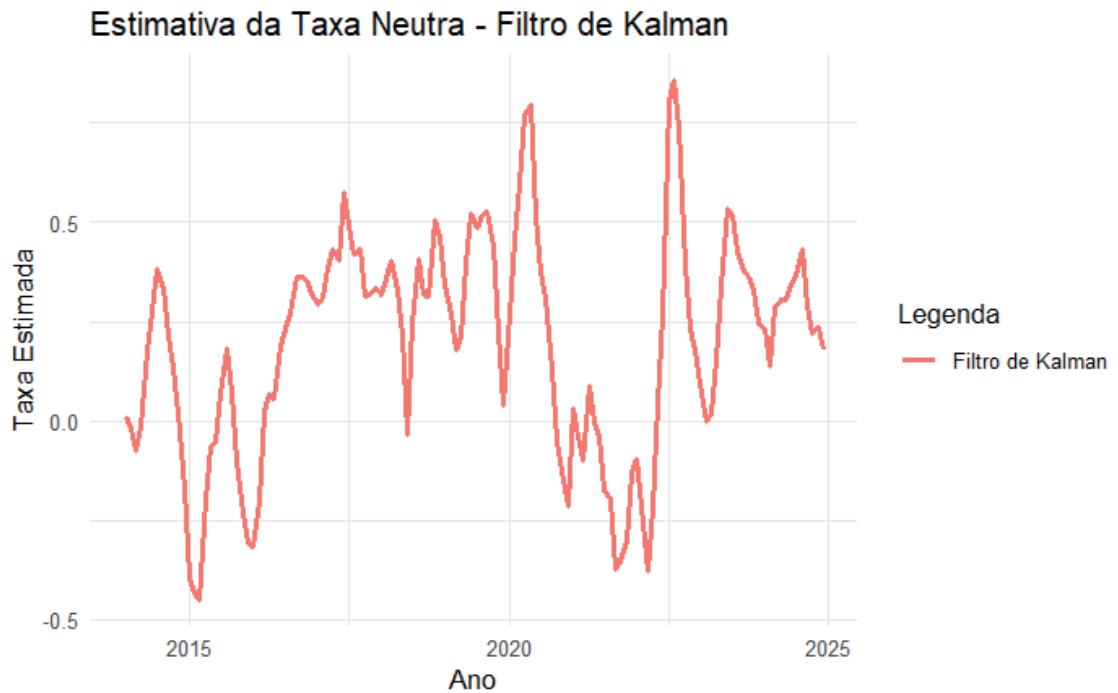


Figura 5: Estimativa da Taxa Neutra - Filtro de Kalman

A figura acima apresenta a estimativa da taxa neutra utilizando o Filtro de Kalman. Esse método é amplamente reconhecido por sua capacidade de adaptação a mudanças estruturais e por permitir uma modelagem dinâmica da taxa de juros neutra ao longo do tempo.

Os resultados mostraram que o Filtro de Kalman apresentou a maior suavização entre os métodos, reduzindo oscilações bruscas da Taxa Real Ex-Ante e capturando tendências mais persistentes. Além disso, o coeficiente de autocorrelação da série filtrada foi alto ( $\rho \approx 0.88$ ), indicando uma evolução relativamente estável da taxa estimada ao longo do tempo.

- **Dinâmica da Taxa Neutra:** O gráfico demonstra uma estimativa que se ajusta rapidamente a variações na economia, refletindo tanto períodos de estabilidade quanto momentos de maior volatilidade.
- **Reação a Choques Econômicos:** Observa-se uma resposta imediata do Filtro de Kalman a eventos como as crises de 2015-2016 e 2020-2022, onde há picos e quedas abruptas. Isso indica que esse método é eficiente para capturar mudanças estruturais e choques na economia.
- **Suavização e Ruído:** Diferentemente de outros métodos, o Filtro de Kalman não impõe uma suavização excessiva à série, permitindo que as flutuações sejam mantidas. Isso pode ser vantajoso para traders e analistas que buscam capturar mudanças rápidas na taxa neutra para tomar decisões ágeis no mercado financeiro.

- **Aplicabilidade:** No contexto do mercado financeiro, a estimativa da taxa neutra via Filtro de Kalman pode ser utilizada para antecipar movimentos da política monetária e ajustar estratégias de investimento conforme as mudanças de ciclo econômico.

Dessa forma, o Filtro de Kalman se destaca como um método flexível e dinâmico, sendo uma ferramenta útil para modelar a taxa neutra em ambientes de incerteza e volatilidade.

### 5.3 Filtro de Hodrick-Prescott: Extração da Tendência Suavizada

Estimativa da Taxa Neutra - Filtro HP

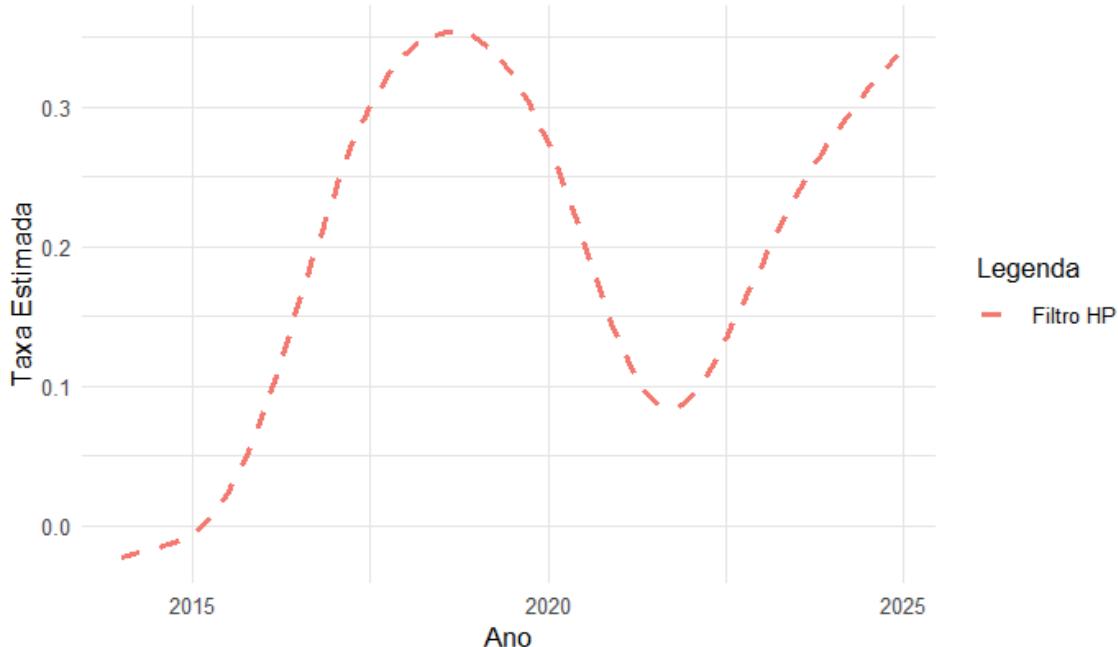


Figura 6: Estimativa da Taxa Neutra - Filtro HP

A figura acima exibe a estimativa da taxa neutra utilizando o Filtro de Hodrick-Prescott (HP). Esse método é amplamente empregado para decompor séries temporais em seus componentes de tendência e ciclo.

A aplicação do filtro HP apresentou uma taxa neutra menos volátil do que a série original, mas mais sensível a oscilações do que o Filtro de Kalman. A autocorrelação da série filtrada foi a mais alta ( $\rho \approx 0.98$ ), indicando um ajuste extremamente suave. Esse resultado sugere que o HP preserva melhor as tendências de longo prazo, mas pode subestimar choques de mercado.

- **Suavização da Série:** O Filtro HP produz uma estimativa muito mais suave da taxa neutra em comparação com o Filtro de Kalman. Isso significa que ele é eficiente para captar tendências de longo prazo, minimizando oscilações de curto prazo.

- **Sensibilidade ao Parâmetro  $\lambda$ :** A suavidade da estimativa é altamente influenciada pelo valor de  $\lambda$ . Para séries mensais, utilizamos  $\lambda = 14.400$ , que equilibra a resposta cíclica e a suavização da tendência.
- **Tendência de Longo Prazo:** O gráfico sugere que a taxa neutra apresentou um crescimento sustentado até meados de 2018, seguido por uma redução até 2022, com posterior retomada da alta. Essa dinâmica pode refletir mudanças estruturais na economia, como alterações na política monetária e na percepção do prêmio de risco.
- **Problema de Borda:** Um dos principais desafios do Filtro HP é a sua sensibilidade às extremidades da amostra. Como observado no gráfico, os valores mais recentes podem sofrer distorções devido à ausência de observações futuras para cálculo da tendência.
- **Aplicabilidade:** No contexto do mercado financeiro, essa estimativa pode ser útil para identificar a trajetória da taxa neutra ao longo do tempo, servindo como referência para estratégias de investimento e precificação de ativos.

Embora o Filtro HP ofereça uma interpretação clara da tendência da taxa neutra, sua principal limitação é a falta de resposta a choques econômicos de curto prazo. Dessa forma, ele pode ser complementado por métodos mais dinâmicos, como o Filtro de Kalman.

## 5.4 Suavização Exponencial Holt-Winters: Captura de Movimentos Recentes

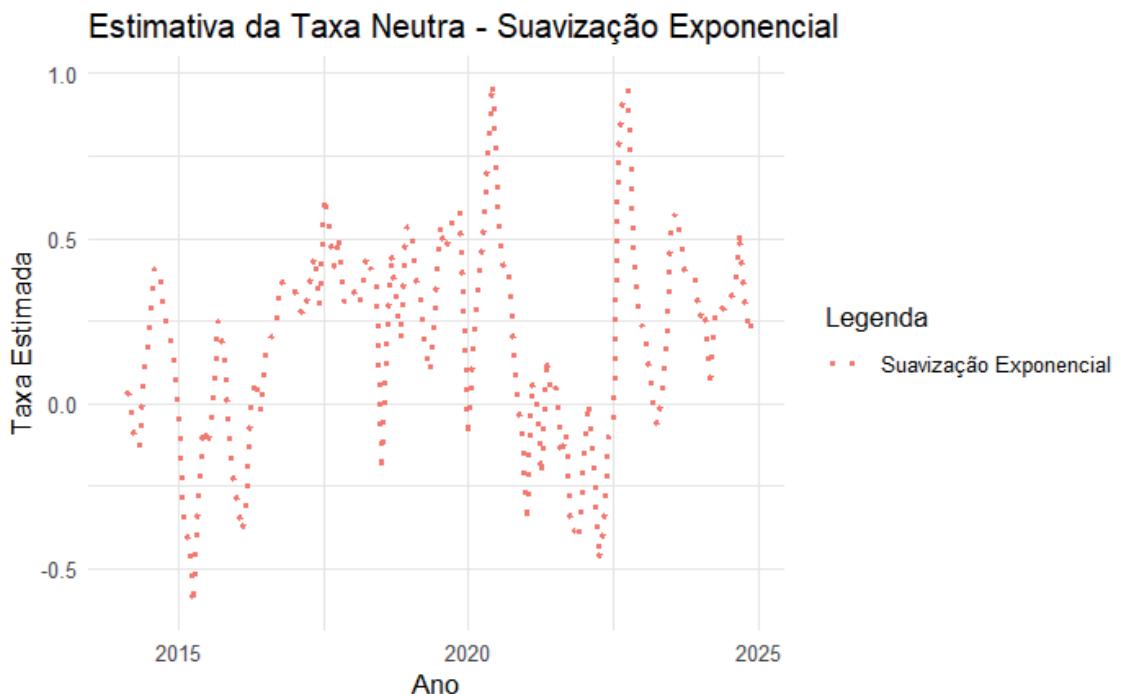


Figura 7: Estimativa da Taxa Neutra - Suavização Exponencial

A figura acima apresenta a estimativa da taxa neutra utilizando a Suavização Exponencial Holt-Winters. Este método é amplamente aplicado para capturar padrões de tendência e sazonalidade em séries temporais.

A autocorrelação da série suavizada foi a menor ( $\rho \approx 0.79$ ), reforçando a ideia de que esse método capta movimentos mais voláteis. Isso pode ser útil para traders que buscam ajustes rápidos nas expectativas de mercado, mas pode resultar em um comportamento mais instável da estimativa da taxa neutra.

- **Alta Sensibilidade a Mudanças:** A suavização exponencial confere pesos mais elevados às observações mais recentes, permitindo que a taxa neutra reaja rapidamente a novas informações do mercado.
- **Ruído e Oscilações:** Diferentemente dos outros métodos, a suavização exponencial apresenta uma maior volatilidade na estimativa, o que pode indicar uma resposta rápida a choques econômicos de curto prazo, mas também pode introduzir ruído excessivo.
- **Impacto da Escolha dos Parâmetros  $\alpha$  e  $\beta$ :** Os parâmetros de suavização  $\alpha$  e  $\beta$  determinam a influência das observações passadas e a suavidade da tendência. Ajustes inadequados podem resultar em um comportamento altamente errático da estimativa.
- **Aplicabilidade no Mercado Financeiro:** No contexto de trading e gestão de ativos, essa abordagem pode ser útil para identificar \*\*mudanças abruptas\*\* na taxa neutra e avaliar a dinâmica de curto prazo das taxas de juros.

Embora a Suavização Exponencial ofereça a vantagem de captar mudanças estruturais rapidamente, sua instabilidade e sensibilidade aos parâmetros podem comprometer sua confiabilidade para análises de longo prazo. Portanto, pode ser mais adequada quando combinada com métodos mais robustos, como o Filtro de Kalman.

## 5.5 Comparação dos Métodos

Os três filtros apresentaram características distintas, e não há um único método ideal, mas sim diferentes formas de interpretar a taxa neutra:

- O Filtro de Kalman gerou uma estimativa estável e robusta, sendo útil para interpretações macroeconômicas e formulação de política monetária.
- O Filtro HP apresentou a estimativa mais suave e de longo prazo, útil para análises estruturais da taxa neutra.
- A Suavização Exponencial Holt-Winters capturou movimentos rápidos da taxa real ex-ante, sendo um método adequado para análises mais dinâmicas.

A tabela abaixo mostra os resultados dos testes: Shapiro-Wilk e o Teste t pareado e Wilcoxon

Tabela 5: Testes Estatísticos de Normalidade e Diferença Entre Métodos

Teste	Estatística	p-valor
Shapiro-Wilk (Kalman)	0.9759	0.0191**
Shapiro-Wilk (HP Filter)	0.9235	1.45e-6
Teste t Pareado (Kalman vs HP)	-3.64e-16	1.0000
Teste de Wilcoxon (Kalman vs HP)	4359.0	0.9466

## 5.6 Interpretação dos Testes Estatísticos

Os testes estatísticos foram aplicados para avaliar as propriedades das séries e garantir a adequação dos métodos empregados. O teste de Shapiro-Wilk revelou que as distribuições das taxas neutras estimadas não seguem uma distribuição normal, o que indica que métodos tradicionais baseados na normalidade podem não ser os mais apropriados para modelar a série.

O teste t pareado não encontrou diferenças estatisticamente significativas entre as estimativas do Filtro de Kalman e do Filtro HP, sugerindo que, em termos médios, ambos fornecem resultados próximos. No entanto, a análise da variabilidade das séries mostra que o Filtro HP gera uma estimativa mais suave, enquanto o Filtro de Kalman captura melhor oscilações de curto prazo.

Já o teste de Wilcoxon, ao avaliar a distribuição das diferenças entre os métodos, corroborou os achados do teste t, indicando que não há um método significativamente superior. Isso reforça a necessidade de combinar os filtros estatísticos dentro de uma abordagem de "caixa de ferramentas" para uma estimativa mais robusta da taxa neutra.

Embora o p-valor do teste de normalidade (Shapiro-Wilk) sobre a série filtrada por Kalman seja numericamente semelhante ao p-valor do teste de Dickey-Fuller aplicado à série original, os testes são distintos e essa semelhança se deve a uma coincidência estatística, não sendo indicativo de relação direta. O que ocorreu foi uma coincidência numérica estatística, em que o p-valor do teste de normalidade sobre a série filtrada por Kalman acabou se aproximando bastante do p-valor do teste ADF aplicado à série original. E isso pode acontecer — principalmente quando as séries são correlacionadas ou derivadas entre si, como é o caso aqui (a série filtrada é obtida a partir da original).

## Volatilidade Móvel da Taxa Real Ex-Ante (12 meses)



Figura 8: Volatilidade Móvel da Taxa Real Ex-Ante (12 meses)

A Figura 8 apresenta a volatilidade móvel da Taxa Real Ex-Ante calculada com uma janela de 12 meses, capturando a variação do desvio-padrão ao longo do tempo. Alguns pontos de destaque:

- **2014-2016:** A volatilidade se mantém relativamente estável, mas com algumas oscilações que podem estar associadas a mudanças no ambiente macroeconômico ou expectativas de política monetária.
- **2018-2020:** Observa-se uma queda na volatilidade, indicando um período de maior previsibilidade da taxa real ex-ante.
- **2020-2022:** Há picos acentuados de volatilidade, sugerindo momentos de forte incerteza no mercado, possivelmente relacionados à pandemia da COVID-19 e suas implicações na política monetária.
- **2023-2024:** A volatilidade volta a cair, mas ainda mantém certa flutuação, o que pode refletir ajustes no mercado financeiro conforme novas expectativas inflacionárias se consolidam.

O comportamento da volatilidade ao longo do tempo pode fornecer insights valiosos para traders e analistas financeiros, auxiliando na calibração de modelos de risco e na definição de estratégias de investimento.

## Comparação da Taxa Real Ex-Ante com as Estimativas da Taxa Neutra

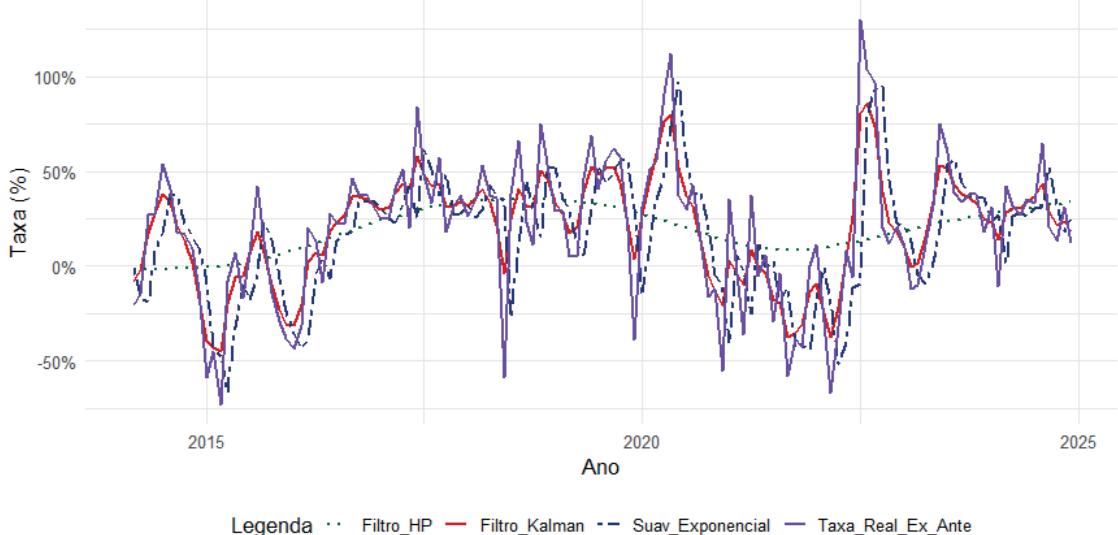


Figura 9: Comparação das Estimativas da Taxa Neutra

A Figura 9 apresenta as diferentes estimativas da taxa neutra, obtidas a partir dos três métodos de filtragem estatística: Filtro de Kalman, Filtro de Hodrick-Prescott (HP) e Suavização Exponencial Holt-Winters. Essa análise permite avaliar as características individuais de cada método e sua aplicabilidade na identificação da dinâmica subjacente da taxa neutra.

- **Comportamento Geral:** As três estimativas seguem padrões semelhantes ao longo do tempo, indicando que os métodos são consistentes na identificação das tendências de longo prazo da taxa neutra. No entanto, algumas diferenças sutis aparecem, sobretudo nos períodos de maior volatilidade.
- **Respostas às Mudanças:**
  - O Filtro de Kalman (linha vermelha) apresenta uma resposta mais dinâmica e ajustável às variações da série, refletindo mudanças estruturais da economia com menor defasagem.
  - O Filtro HP (linha verde pontilhada) gera uma suavização mais acentuada da série, destacando a tendência de longo prazo e eliminando parte da variabilidade de curto prazo.
  - A Suavização Exponencial (linha azul pontilhada) equilibra as duas abordagens, apresentando uma resposta que captura tanto flutuações de curto prazo quanto a tendência de longo prazo.
- **Divergências Entre os Métodos:** Em momentos de alta volatilidade, como nos anos de 2015-2016 e 2020-2022, as diferenças entre as abordagens tornam-se mais evidentes:
  - O Filtro de Kalman ajusta-se rapidamente às variações extremas.

- O Filtro HP apresenta um padrão mais estável, demorando a refletir mudanças abruptas.
- A Suavização Exponencial responde de maneira intermediária, captando parte das oscilações, mas sem absorver ruídos excessivos.

Tabela 6: Estatísticas das Estimativas da Taxa Neutra

Método	Média	Desvio-Padrão	Autocorrelação (lag 1)
Filtro de Kalman	0.1917	0.2770	0.885
Filtro HP	0.1917	0.1213	0.980
Suavização Exp.	0.1911	0.3082	0.793

A Tabela 6 apresenta as estatísticas descritivas das estimativas da taxa neutra para os três métodos utilizados: Filtro de Kalman, Filtro de Hodrick-Prescott (HP) e Suavização Exponencial.

**Média:** Observa-se que as médias das três estimativas são muito próximas, situando-se em aproximadamente 0,19. Isso sugere que, independentemente da metodologia empregada, a taxa neutra estimada converge para valores semelhantes ao longo do tempo.

**Desvio-Padrão:** As diferenças mais significativas aparecem na dispersão das estimativas. O Filtro HP apresenta o menor desvio-padrão (0,1213), indicando uma estimativa mais estável e menos volátil ao longo do tempo. Já a Suavização Exponencial tem o maior desvio-padrão (0,3082), refletindo uma maior sensibilidade a mudanças recentes na série. O Filtro de Kalman, com um desvio-padrão intermediário (0,2770), ajusta-se bem às variações sem apresentar extrema volatilidade.

**Autocorrelação de Primeira Ordem:** A análise da autocorrelação de primeira ordem mostra que o Filtro HP tem a maior persistência na série (0,980), sugerindo uma estimativa muito suavizada e com inércia elevada. Isso pode tornar o método mais adequado para análises de longo prazo, mas menos responsivo a mudanças rápidas na economia. O Filtro de Kalman apresenta uma autocorrelação ligeiramente menor (0,885), indicando um ajuste dinâmico mais equilibrado. Por outro lado, a Suavização Exponencial tem a menor autocorrelação (0,793), evidenciando que a série gerada por esse método responde rapidamente às oscilações, o que pode introduzir mais ruído na estimativa.

**Resumo:** Diante dessas estatísticas, nota-se que cada metodologia oferece um balanço diferente entre suavização e resposta a choques econômicos. O Filtro HP proporciona uma estimativa muito estável, mas pode não captar mudanças estruturais em tempo hábil. O Filtro de Kalman equilibra bem suavização e reatividade, enquanto a Suavização Exponencial é mais responsiva, porém mais volátil.

## 5.7 Interpretação do Erro Quadrático Médio (MSE)

A Tabela 7 apresenta os valores do Erro Quadrático Médio (MSE) para cada uma das metodologias utilizadas na estimativa da taxa neutra. O MSE é uma métrica que mede a média dos erros quadráticos entre a taxa real ex-ante observada e as estimativas fornecidas por cada filtro. Valores menores indicam melhor desempenho preditivo, ou seja, menor discrepância entre o valor observado e o valor suavizado.

Tabela 7: Erro Quadrático Médio (MSE) das Estimativas da Taxa Neutra

Método de Estimação	MSE
Filtro de Kalman	0,02295
Filtro HP	0,11099
Suavização Exponencial (Holt-Winters)	0,10808

Observa-se que o Filtro de Kalman apresentou o menor MSE (0,02295), indicando maior capacidade de capturar a tendência subjacente da taxa real ex-ante em relação aos outros métodos. Os filtros de Hodrick-Prescott (HP) e a Suavização Exponencial de Holt-Winters tiveram MSEs significativamente maiores (em torno de 0,11 e 0,108, respectivamente), o que sugere menor aderência às oscilações da série original ou possível superajuste.

Esse resultado reforça a robustez do filtro de Kalman em ambientes com maior volatilidade ou ruído estocástico, como o caso da taxa real ex-ante brasileira no período analisado (2014–2024).

## 5.8 Limitações dos Métodos Aplicados

Embora a abordagem adotada neste artigo forneça flexibilidade na estimativa da taxa neutra, cada método apresenta limitações inerentes. O Filtro de Kalman depende fortemente da especificação dos parâmetros e da modelagem da equação de estado, o que pode introduzir viés caso as premissas do modelo não sejam adequadas. O Filtro de Hodrick-Prescott, por sua vez, sofre com o problema de borda, tornando a estimativa dos períodos mais recentes menos confiável e sensível à escolha do parâmetro  $\lambda$ . Já a Suavização Exponencial Holt-Winters apresenta um comportamento mais reativo às variações do mercado, o que pode resultar em maior volatilidade da estimativa e menor robustez em cenários de longo prazo.

Diante dessas limitações, a escolha do método mais apropriado depende do horizonte de análise e do objetivo do analista. Em aplicações práticas, recomenda-se a utilização combinada dessas metodologias, explorando as complementaridades entre suavização e sensibilidade às oscilações do mercado.

## 6 Conclusão

Este estudo propôs a estimativa da taxa neutra utilizando diferentes abordagens estatísticas, enfatizando a ideia de que não há um único método ideal. Ao invés de buscar um

modelo definitivo, esta pesquisa analisou a complementaridade entre três metodologias: o Filtro de Kalman, o Filtro de Hodrick-Prescott (HP) e a Suavização Exponencial Holt-Winters, ressaltando como cada uma dessas abordagens pode ser aplicada para diferentes propósitos no mercado financeiro.

Os resultados empíricos mostraram que os três métodos convergem para valores médios similares da taxa neutra, mas diferem em termos de volatilidade e persistência. O Filtro de Kalman destacou-se pela sua capacidade de adaptação dinâmica, tornando-se uma ferramenta útil para capturar mudanças estruturais ao longo do tempo. O Filtro HP, por sua vez, apresentou a estimativa mais suave, sendo adequado para análises de longo prazo e tendências macroeconômicas. Já a Suavização Exponencial Holt-Winters mostrou-se mais responsiva às variações recentes da série, o que pode ser vantajoso para traders que necessitam de ajustes rápidos em suas estratégias.

A principal contribuição deste artigo é demonstrar que a combinação dessas metodologias pode fornecer uma visão mais robusta da taxa neutra, unindo estabilidade, flexibilidade e sensibilidade às mudanças do mercado. Para analistas financeiros e formuladores de política monetária, essa abordagem permite calibrar estimativas conforme diferentes cenários e objetivos. Em termos práticos, um operador de mercado pode utilizar o Filtro de Kalman para capturar mudanças estruturais em tempo real, o Filtro HP como referência de longo prazo e a Suavização Exponencial para ajustes táticos.

Como proposta para trabalhos futuros, sugere-se a integração dessas técnicas com modelos de aprendizado de máquina, visando aumentar a adaptabilidade e precisão das estimativas. Além disso, a aplicação dessas metodologias em mercados emergentes pode fornecer insights adicionais sobre a dinâmica da taxa neutra em diferentes contextos econômicos.

# Referências

- [1] BANCO CENTRAL DO BRASIL. SGS - Sistema Gerenciador de Séries Temporais, 2024. Acessado em: 21 jan. 2025.
- [2] BLINDER, A. S. *Bancos Centrais: teoria e prática*. Ed. 34, 1999.
- [3] CLARIDA, R., GALI, J., AND GERTLER, M. The science of monetary policy: a new keynesian perspective. *Journal of economic literature* 37, 4 (1999), 1661–1707.
- [4] GOODWIN, P., ET AL. The holt-winters approach to exponential smoothing: 50 years old and going strong. *Foresight* 19, 19 (2010), 30–33.
- [5] HODRICK, R. J., AND PRESCOTT, E. C. Postwar us business cycles: an empirical investigation. *Journal of Money, credit, and Banking* (1997), 1–16.
- [6] HOLSTON, K., LAUBACH, T., AND WILLIAMS, J. C. Measuring the natural rate of interest: International trends and determinants. *Journal of International Economics* 108 (2017), S59–S75.
- [7] HOLT, C. Forecasting trends and seasonals by exponentially weighted averages, vol. 52. *Carnegie Institute of Technology, Pittsburgh, Pa, USA* (1957).
- [8] LAUBACH, T., AND WILLIAMS, J. C. Measuring the natural rate of interest. *Review of Economics and Statistics* 85, 4 (2003), 1063–1070.
- [9] NIELSEN, A. *Practical time series analysis: Prediction with statistics and machine learning*. O'Reilly Media, 2019.
- [10] PAIGE, R. L., AND TRINDADE, A. A. The hodrick-prescott filter: A special case of penalized spline smoothing.
- [11] WHITTAKER, E. T. On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society* 41 (1922), 63–75.
- [12] WICKSELL, K. *Interest and prices*. Ludwig von Mises Institute, 1936.

# Análise econométrica do *curtailment* na geração eólica no Brasil entre 2022 e 2024

Bruna Amaral Lanzillotti Barreto \*

Felipe Cardoso Wilhelm da Costa †

Fernando Antonio Lucena Aiube ‡

Edison Americo Huarsaya Tito §

Daiane Rodrigues dos Santos ¶

25 de julho de 2025

## Resumo

O mercado de geração eólica e solar brasileiro tem sido impactado por um significativo crescimento do *curtailment*<sup>1</sup> da produção. Apesar de esse ser um fenômeno esperado no processo de integração das Energias Renováveis Variáveis (VRE) nos sistemas, fatores específicos do caso brasileiro como o perfil do crescimento da capacidade instalada, as dinâmicas da demanda e a gestão do sistema de transmissão parecem ter formado nos últimos anos a "tempestade perfeita" da restrição de geração. O presente trabalho realizou uma decomposição de séries temporais associadas ao fenômeno no período de janeiro de 2022 a dezembro de 2024 pelo método STL (*Seasonal and Trend decomposition using LOESS*), seguida de uma modelagem ARIMA com verificação dos efeitos das variáveis explicativas e elaboração de previsão de 12 semanas, considerando os princípios de modelagem Box-Jenkins. A sazonalidade foi identificada como elemento determinante do *curtailment* e o crescimento da Geração Distribuída (GD) como fator de maior impacto dentre as variáveis explicativas da parcela do fenômeno não atrelada ao seu caráter sazonal.

**Palavras-chave:** Restrição de Geração; Energia eólica, Energia Elétrica, Economia da Energia, *Curtailment*

**Código JEL:** C22, C53, N7.

---

\*Mestranda do Programa de Pós-Graduação em Ciências Econômicas da Universidade do Estado do Rio de Janeiro (PPGCE-UERJ)

†Mestrando do Programa de Pós-Graduação em Ciências Econômicas da Universidade do Estado do Rio de Janeiro (PPGCE-UERJ)

‡Professor Associado da Faculdade de Ciências Econômicas - UERJ

§Professor Adjunto do Instituto de Matemática e Estatística - UERJ

¶Professora Adjunta da Faculdade de Ciências Econômicas - UERJ

<sup>1</sup>Também conhecido como *constrained-off*, geração frustrada ou restrição de geração, o *curtailment* é definido pela Resolução Normativa Agência Nacional de Energia Elétrica (ANEEL) nº 1.030/2022 como "[...] a redução da produção de energia por usinas eólioelétricas despachadas centralizadamente ou usinas/conjuntos de usinas eólioelétricas considerados na programação, decorrente de comando do Operador Nacional do Sistema (ONS), que tenham sido originados externamente às instalações das respectivas usinas." [3]

# 1 Introdução

O crescimento das Energias Renováveis Variáveis<sup>2</sup> vem sendo uma tendência global dos sistemas de energia nas últimas décadas, com conhecidas vantagens econômicas e ambientais. Essas formas de geração são responsáveis por mudanças na gestão dos sistemas de eletricidade, até então dominados por fontes mais controláveis, como a hidro e a termoelétrica, as quais se pode ajustar mais facilmente conforme a demanda.

O sistema de energia elétrica brasileiro, hoje considerado um sistema hidro-termo-eólico, é dominado majoritariamente pela geração hidroelétrica. Como forma de redução dessa dependência do regime hídrico e enfrentamento aos racionamentos e apagões vivenciados no final do século XX, sua expansão tem sido focada na diversidade tecnológica, com o crescimento de usinas eólicas, solares e hidroelétricas a fio d'água, além de alguns projetos de grandes usinas (Goldemberg, 2015; Mercedes, Rico e Pozzo, 2015) [15] [29].

No caso do Brasil, a partir de 2006, as tecnologias eólica e solar saltaram de 0,2% da capacidade de geração instalada no Sistema Integrado Nacional (SIN) para 21% em 2024. Um crescimento acelerado e concentrado na região Nordeste que oferece recursos de alta qualidade para a produção de energia através dessas tecnologias. A fonte eólica, que aparece pela primeira vez na Matriz Elétrica brasileira com 49 MW instalados alcançou 29.185 MW em 2024, menos de 20 anos depois: um crescimento médio de 32,7% ao ano, passando a representar 14% da matriz elétrica nacional (Operador Nacional do Sistema (ONS), 2025; Empresa de Pesquisa Energética (EPE), 2024) [32] [13].

Atualmente no Brasil um dos temas chaves do processo de integração dessas tecnologias no sistema é a gestão do excedente de oferta nos períodos que, pela disponibilidade climática de recursos, há uma alta de produção. O relatório *Integrating Solar and Wind: Global Experience and Emerging Challenges* da IEA (2024) [23] sobre o processo de integração das VRE nos sistemas nacionais aponta que, para maximizar o benefício dessas fontes, é necessária uma integração eficaz dessas tecnologias aos sistemas de energia, de modo a maximizar o volume de energia fornecido de maneira segura e com menor custo, que reduzam a dependência de combustíveis fósseis e evitando medidas de estabilização que sejam caras ou mesmo gerem prejuízo.

Em relação ao tema do excedente de oferta, por exemplo, o relatório indica a necessidade de uma maior flexibilidade dos sistemas e ferramentas como o uso de baterias, mecanismos de resposta à demanda<sup>3</sup>, uma gestão eficiente das tecnologias despacháveis e o próprio *curtailment* das VRE. Esse último, que é a limitação da produção de energia pelo operador do sistema, teve um crescimento expressivo no Brasil nos últimos anos e é o alvo da presente análise, uma vez que, adotado de forma crescente e quase como ferramenta exclusiva de gestão desse excedente, se comporta como um "cobertor curto", sendo origem de prejuízo para as empresas de geração eólica e solar e do crescimento do nível de incerteza nesse

<sup>2</sup>Acrônimo para a sigla em inglês *Variable Renewable Energy*, adotada pela Agência Internacional de Energia (IEA) para se referir as fontes de energia cuja produção é dependente da disponibilidade de recursos ambientais flutuantes, como energia gerada pela força dos ventos e correntezas marítimas ou de rios e irradiação solar.

<sup>3</sup>Por exemplo políticas de incentivo à redução do consumo de energia em horários de pico.

mercado.

Os principais marcos desse crescimento são bem conhecidos pelo setor. Foi observada uma desaceleração do crescimento da demanda de energia em relação às previsões nos últimos anos. Essa desaceleração está intimamente relacionada ao aumento exponencial da capacidade solar distribuída que tem recebido incentivos fiscais bastante expressivos. Por outro lado, em termos de infraestrutura do sistema, um evento de apagão em Agosto de 2023 teve como uma de suas consequências uma gestão mais conservadora da transmissão de energia entre os subsistemas, com a restrição da exportação da produção do Nordeste para as demais regiões do país.

Com o objetivo de investigar este fenômeno que tem estado no centro das discussões do setor de renováveis no Brasil, o presente trabalho conduziu uma análise de séries temporais associadas ao fenômeno do *curtailment* em usinas eólioelétricas no período de janeiro de 2022 a dezembro de 2024. Foi feita uma decomposição pelo método STL, seguida de uma modelagem ARIMA com verificação dos efeitos das variáveis explicativas e elaboração de previsão de 12 semanas, considerando os princípios de modelagem Box-Jenkins.

Na seção 2 é apresentada uma breve revisão da literatura sobre o tema. Na seção 3 é detalhada a metodologia, além de apresentados os dados trabalhados. A seção 4 discute os resultados encontrados e a seção 5 conclui o estudo apontando para algumas possibilidades de avanço da análise do tema.

## 2 Revisão da Literatura

O tema das energias renováveis é alvo de uma vasta literatura, mas o fenômeno do *curtailment* especificamente tem um volume mais restrito e mais recente de produções acadêmicas, grande parte delas falando do fenômeno na China e nos Estados Unidos da América (EUA), que são dois países com parque de geração eólica maior e mais antigo. Muitos dos trabalhos são centrados nos aspectos de Engenharia, também em aspectos econômicos em escala micro ou ainda, como este trabalho, análises dos sistemas de forma geral. Nessa seção apresentamos brevemente algumas dessas abordagens.

López Prol e Zilberman (2023) [26] examinam as dinâmicas do *curtailment* no estado da Califórnia, nos EUA, entre 2014 e 2022. Suas análises usando métodos OLS, FGLS Prais-Winsten, regressão Beta e modelos aditivos generalizados mostram que o crescimento da geração solar, eólica e nuclear aumentam as taxas de restrição eólica e solar, por outro lado, como esperado, o aumento da demanda o reduz.

Acum et al (2023)[1], também analisando sistemas dos EUA, apontam que uma fração significativa (5-15%) da energia renovável gerada atualmente é desperdiçada nas redes elétricas ao redor do mundo devido a problemas de excesso de oferta e restrições de transmissão. Ser capaz de prever quando e onde o *curtailment* de energia renovável ocorre poderia melhorar a utilização dessas fontes renováveis. Os autores exploram estratégias de mitigação a partir de dados compilados de preços marginais nodais de sistemas elétricos dos EUA, com o objetivo de ajudar na descarbonização das redes elétricas, desbloqueando

o potencial da energia renovável.

Já Hurtado et al (2023)[20] por meio de análise direta de dados (*dispatch down reports*) e categorização de eventos, analisam o *curtailment* da energia eólica nos sistemas de energia irlandeses em 2020 e 2021. Os autores observaram, como se espera, uma correlação positiva entre a capacidade instalada de energia eólica e os níveis de *curtailment*. Junto a isso também demonstraram que o principal fator para o fenômeno no sistema energético foi a restrição operacional que exigiu cerca de 80% da operação das fontes convencionais e menos de 20% de fontes não síncronas como a eólica. Outras limitações sistêmicas, como a taxa de variação da frequência e a inércia, mostraram ter impacto insignificante no *curtailment* da energia eólica.

Alguns trabalhos interessantes se dedicaram a avaliar possibilidades de aproveitamento da energia que seria objeto de *curtailment* para outros usos que canalizem seu potencial econômico. Li et al (2023)[24], por exemplo aplicaram uma metodologia de aprendizado por *reinforcement learning* usada para tomar decisões otimizadas sob incerteza e simulações com dados realistas de geração eólica na Austrália. Os autores propõem desacoplar a operação conjunta do parque eólico e do sistema de armazenamento de energia por baterias que atuam como fontes de backup local, criando uma estratégia de oferta conjunta de energia que armazena o excedente de produção e realiza de arbitragem energética posterior.

Niaz, Liu e You (2022) [30] investigam a possibilidade de mitigação do *curtailment* utilizando a energia que seria restrita para alimentar a baixo custo uma fazenda de mineração de *bitcoin* no Texas, EUA, atividade que consome grandes volumes de energia. Segundo os autores, de forma melhor que baterias, por exemplo, uma fazenda de mineração de cripto moeda não apenas mitigaria a perda de energia, como a empregaria em uma atividade de alto potencial lucrativo, sendo uma opção mais vantajosa aos investidores do setor.

No Brasil, alguns trabalhos interessantes consideram também o tema em escala micro e aplicaram modelos para auxiliar as empresas a lidarem com a restrição de sua produção e terem maior previsibilidade de seus resultados. No trabalho de Silva (2022) [34], por exemplo, é feita uma tentativa de predição do *curtailment* na unidade com algoritmo de *Random Forest* usando dados anemométricos coletados *in loco*. Já Silva (2023) [33] calcula a produtividade de empresas tendo o *cutailment* como parâmetro. Os trabalhos de Menezes (2022) [28] e de Lima (2021) [12] fazem análises da implementação de sistemas de baterias em usinas para gestão do *curtailment* e da variação diária do preço da energia.

### 3 Metodologia

Chatfield (1996)[10] cita alguns métodos de previsão de séries temporais. Entre os principais estão: extrapolação de Têndencia (Harrison e Pearce, 1972[18], Meade, 1984[27]), Sua-viização Exponencial (Holt, 1958)[19], e Metodologia Box-Jenkins (Box e Jenkins, 1970)[8]. Mais recentemente Hyndman (2021)[21] considera outros métodos de previsão mais avançados e robustos que se baseiam nesses modelos.

Há ainda metodologias como o Modelo Profeta (*Prophet Model*) (SJ Taylor e Letham,

2018)[35], e o modelo de Vetores Autorregressivos, a partir de Hamilton (1994)[17] que insere uma relação bidirecional entre a variável dependente e as explicativas, diferente dos modelos citados até agora que impõem uma relação unidirecional. Ainda citam-se os Modelos de Redes Neurais e Previsões Bootstrapping (Bergmeir, C., Hyndman, RJ, e Benítez, JM, 2016)[6].

No presente artigo iremos aplicar a abordagem Box-Jenkins para realizar a previsão do *curtailment* na geração de energia eólica no Brasil. Antes disso, iremos decompor a série temporal para compreender melhor os padrões presentes nos dados e facilitar as análises e as respectivas previsões.

### 3.1 Decomposição de Séries Temporais

Conforme Hyndman (2021)[21] os dados de séries temporais podem exibir uma variedade de padrões, e geralmente é útil dividir uma série temporal em 3 componentes: tendência, sazonalidade e resíduos ou resto (*remainder*). A tendência ou tendência-ciclo captura mudanças de longo prazo ou movimentos de baixa frequência na série, a sazonalidade representa padrões que se repetem regularmente ao longo do tempo e o componente de resíduos representa os valores que não podem ser explicados pelas componentes sazonais e de tendência, incluindo a aleatoriedade.

A decomposição da série pode ser realizada de diferentes maneiras. As mais comuns são a aditiva, multiplicativa e a decomposição STL. A decomposição aditiva pode ser observada pela equação (1).

$$y_t = S_t + T_t + R_t \quad (1)$$

em que  $y_t$  é o dado observado;  $S_t$  é o componente sazonal;  $T_t$  é o componente de tendência-ciclo; e  $R_t$  é o componente de resíduo, todos no período  $t$ .

Já a decomposição multiplicativa pode ser representada pela equação (2).

$$y_t = S_t \times T_t \times R_t. \quad (2)$$

O autor ressalta que a decomposição aditiva é mais apropriada se a magnitude das flutuações sazonais, ou a variação em torno da tendência, não variar com o nível da série temporal. Quando a variação no padrão sazonal, ou a variação em torno da tendência-ciclo, parecer proporcional ao nível da série temporal, a decomposição multiplicativa é mais apropriada.

A Decomposição STL é um método versátil e robusto para decompor séries temporais. Ele lida com inúmeros padrões sazonais, o que é útil para as séries em análise neste trabalho, que foram semanais. Consegue também aliar a estimativa de métodos não lineares com uma técnica de suavização não paramétrica<sup>4</sup>. O método também é mais vantajoso do que a decomposição clássica porque captura quando há variação do componente sazonal ao longo do tempo e essa taxa de mudança pode ser controlada de forma explícita a partir do conhecimento sobre as séries que estão sendo modeladas. Também é relevante a robustez

---

<sup>4</sup>Veja em Cleveland et al (1990)[11] e Bandara et al. (2022)[5]

frente a *outliers*, com as observações incomuns eventuais não afetando as estimativas de tendência e sazonalidade, apenas os resíduos.

### 3.2 Estratégia de Modelagem Box & Jenkins

A metodologia Box-Jenkins foi utilizada para mensurar o comportamento da série temporal, tendo sido realizadas as etapas de: identificação, estimação, diagnóstico e previsão. Conforme podemos observar na Figura 1 na primeira fase ocorre a identificação do processo aleatório que gerou os dados, em seguida a estimação dos parâmetros que o caracterizam, após isso ocorre o diagnóstico e a verificação se as hipóteses do modelo foram cumpridas. Caso negativo, uma nova fase de identificação deve ser considerada até que a verificação das hipóteses seja finalmente positiva, permitindo assim a última fase de previsão.

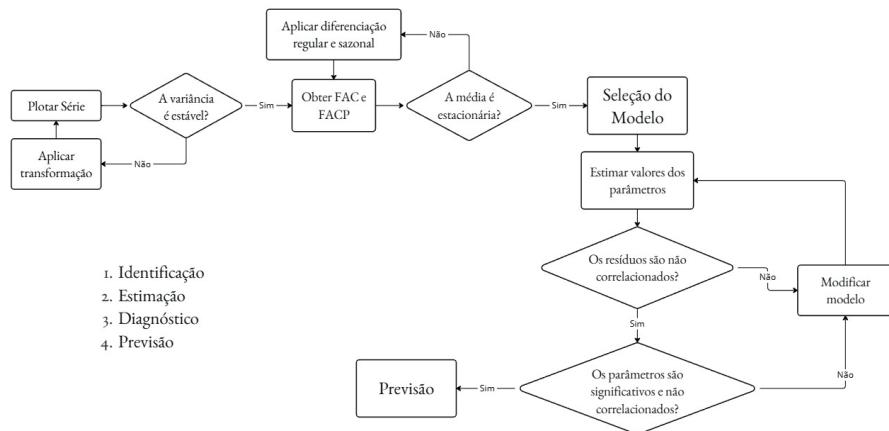


Figura 1: Abordagem de modelagem Box & Jenkins. Adaptado de Agarwal (2018) [2] pelos autores.

#### 3.2.1 Identificação

De acordo com Gujarati e Porter (2011)[16] no estágio da identificação, a princípio, investiga-se os valores apropriados do ARIMA ( $p, d, q$ ) a ser estimado. De modo que  $p$  é o número de termos autoregressivos (AR);  $d$  é o número de diferenças necessárias para tornar a série estacionária;  $q$  é o número de termos de média móvel (MA)<sup>5</sup>. O modelo ARIMA ( $p, d, q$ ) é dado pela equação (3)

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) (1 - L)^d y_t = c + \left(1 + \sum_{j=1}^q \theta_j L^j\right) \varepsilon_t, \quad (3)$$

onde  $y_t$  é o valor da série temporal no instante  $t$ ;  $L$  representa o operador de defasagem ( $L^k y_t = y_{t-k}$ );  $\phi_i$  os coeficientes autorregressivos;  $\theta_j$  os coeficientes de média móvel;  $c$  a

<sup>5</sup>Uma análise inicial dos modelos AR( $p$ ), MA( $q$ ), ARMA( $p, q$ ), pode ser vista em Wooldridge (2013) [36]

constante do modelo e  $\varepsilon_t$  o termo de erro aleatório (ruído branco).

Vale ressaltar que se a série for sazonal, é necessário utilizar um modelo SARIMA, que é uma extensão do ARIMA para capturar padrões sazonais. Para isso, além dos parâmetros  $p$ ,  $d$  e  $q$ , será necessário ajustar os parâmetros sazonais  $P$ ,  $D$  e  $Q$ , juntamente com o período sazonal. O modelo SARIMA(p,d,q)(P,D,Q)s é dado pela equação (4):

$$\left(1 - \sum_{i=1}^P \Phi_i L^{i \cdot s}\right) (1 - L^s)^D \left(1 - \sum_{i=1}^p \phi_i L^i\right) (1 - L)^d y_t = c + \left(1 + \sum_{j=1}^Q \Theta_j L^{j \cdot s}\right) \left(1 + \sum_{j=1}^q \theta_j L^j\right) \varepsilon_t, \quad (4)$$

onde  $y_t$  é o valor da série temporal no tempo  $t$ ;  $L$  o operador de defasagem ( $L^k y_t = y_{t-k}$ );  $s$  o período sazonal (ex.:  $s = 7$  para dados semanais);  $\Phi_i$  os coeficientes autorregressivos sazonais;  $\phi_i$  os coeficientes autorregressivos não sazonais;  $\Theta_j$  os coeficientes de média móvel sazonais;  $\theta_j$  os coeficientes de média móvel não sazonais;  $c$  a constante do modelo (se aplicável) e  $\varepsilon_t$  o termo de erro aleatório (ruído branco).

As principais ferramentas na etapa de identificação são a função de correlação amostral (ACF), a função de correlação amostral parcial (PACF) e os correlogramas resultantes, que são simplesmente as representações de ACF e PACF contra a extensão da defasagem. Assim, como tarefa inicial é preciso determinar  $p$  e  $q$  para a identificação de modelos que serão testados. Para isso, procede-se ao exame dos coeficientes de autocorrelação e dos coeficientes de autocorrelação parcial, que permitem medir a força relativa de interação entre as variáveis  $y_t$  defasadas.

Numa segunda parte desse processo é necessário efetuar testes de estacionariedade no modelo. Esse processo é de suma importância ao melhorar a capacidade de precisão e previsão dos modelos estatísticos. A estacionariedade de segunda ordem implica que as propriedades estatísticas da série, como média, variância e autocorrelação, permaneçam constantes ao longo do tempo. Isso permite que o modelo capture padrões e dependências temporais de forma mais eficaz. De acordo com Hyndman (2021)[21] uma série temporal estacionária é aquela cujas propriedades estatísticas não dependem do tempo em que a série é observada. Assim, a tendência e a sazonalidade afetarão o valor da série temporal em momentos diferentes. Por outro lado, uma série de ruído branco é estacionária — não importa quando você a observa, ela deve parecer a mesma em qualquer ponto do tempo.

Dessa forma, são realizados testes na série como o *Augmented Dickey-Fuller* (ADF) que ajuda a verificar se a mesma é estacionária ou não. A análise de seu p-valor abaixo de 0,05 ao rejeitar a hipótese nula, confirma a estacionariedade da série. Se ela não for estacionária, ou seja, suas propriedades estatísticas mudam ao longo do tempo, deve-se efetuar um processo de diferenciação na mesma para torná-la estacionária. Após a transformação é possível ajustar modelos como ARIMA ou SARIMA de forma mais eficaz, garantindo um modelo de previsão mais robusto.

### 3.2.2 Estimação

Após identificar as ordens adequadas para  $p$ ,  $d$  e  $q$  (e  $P$ ,  $D$  e  $Q$ , no caso de sazonalidade), o modelo ARIMA é ajustado aos dados, estimando os parâmetros do modelo. Nesse estágio é

necessário estimar os parâmetros dos termos autorregressivos e dos termos de média móvel incluídos no modelo. Esse cálculo pode ser feito de forma mais simples por mínimos quadrados, ou, se necessário, de maneira mais complexa a partir de métodos de estimação não linear dos parâmetros. Ao estimar o modelo ARIMA por máxima verossimilhança (MLE), a técnica encontra os valores dos parâmetros que maximizam a probabilidade de observar os dados fornecidos.

Para modelos ARIMA, os estimadores de MLE são idênticos aos de mínimos quadrados que seriam obtidos ao minimizar:

$$\sum_{t=1}^T \varepsilon_t^2. \quad (5)$$

Ainda seguindo Hyndman (2021)[21], vale ressaltar o critério de informação de Akaike (AIC), que é útil para determinar a ordem de um modelo ARIMA. Ele pode ser escrito como:

$$AIC = -2 \log(L) + 2(p + q + k + 1), \quad (6)$$

onde  $L$  é a verossimilhança dos dados,  $k = 1$  se  $c \neq 0$  e  $k = 0$  se  $c = 0$ . Observe que o último termo entre parênteses é o número de parâmetros no modelo (incluindo  $\sigma^2$ , a variância dos resíduos).

Para modelos ARIMA, o AIC corrigido (AICc) pode ser escrito como

$$AICc = AIC + \frac{2(p + q + k + 1)(p + q + k + 2)}{T - p - q - k - 2}, \quad (7)$$

e o Critério de Informação Bayesiano (BIC) pode ser escrito como:

$$BIC = AIC + [\log(T) - 2](p + q + k + 1). \quad (8)$$

Os modelos mais adequados são obtidos ao minimizar o AIC, AICc ou BIC. Entretanto, vale ressaltar que esses critérios de informação tendem a não ser bons guias para selecionar a ordem apropriada de diferenciação ( $d$ ) de um modelo, mas apenas para selecionar os valores de  $p$  e  $q$ . Isso ocorre porque a diferenciação altera os dados nos quais a verossimilhança é computada, tornando os valores de AIC entre modelos com diferentes ordens de diferenciação não comparáveis.

### 3.2.3 Diagnóstico

Após escolhido um modelo ARIMA específico, e tendo estimado seus parâmetros, ocorre o processo de verificação do modelo selecionado e se o mesmo ajusta-se aos dados razoavelmente bem. Um teste simples do modelo selecionado é verificar se os resíduos estimados com base nesse modelo são ruídos brancos; se forem, pode-se aceitar o ajuste específico; do contrário, deve-se recomeçar e voltar para a etapa de identificação.

Ou seja, ao ajustarmos um modelo ARIMA, estamos buscando capturar toda a estrutura de dependência serial presente na série. Logo, os resíduos do modelo não devem apresentar qualquer tipo de dependência no tempo. Num modelo ARIMA bem ajustado, os erros de previsão devem ser uma sequência de ruídos brancos, ou seja, não devem exibir

autocorrelações (e autocorrelações parciais) significativas.

Nessa etapa é necessário obter a partir dos resíduos do modelo estimado a função de autocorrelação e a função de autocorrelação parcial até uma certa defasagem. A partir de testes de autocorrelação como teste Ljung-Box<sup>6</sup> é possível auferir se os resíduos estimados são puramente aleatórios, se forem não há mais necessidade de procurar outro modelo ARIMA, e pode-se seguir em frente.

Junto a isso, outros procedimentos para análise de resíduos são importantes, como a análise da normalidade e a presença de heterocedasticidade. De acordo com Chatfield (1996) [10], a normalidade dos resíduos é desejável para que as inferências estatísticas como intervalos de confiança e testes de hipóteses sejam válidos. Quanto à heterocedasticidade, a variabilidade não constante pode comprometer as estimativas do modelo. Assim, ao identificar heterocedasticidade, deve ser feita uma modificação ou escolha de um modelo adequado que trate esse comportamento, como o uso de modelos GARCH<sup>7</sup> ou outras abordagens robustas.

### 3.2.4 Previsão

Por fim, chega-se à última etapa da metodologia que, depois de identificar, estimar e diagnosticar um modelo ARIMA(p,d,q), pode usá-lo para gerar previsões da série de forma pontual ou por intervalo. Na previsão pontual é necessário definir o previsor ótimo da série, de modo que a partir de um valor único, o modelo estima como o resultado mais provável para um período futuro da série temporal, condicional à informação disponível até o tempo  $t$ . Já a previsão por intervalo fornece um intervalo de confiança, que representa a incerteza associada à previsão, de modo que é necessário conhecer a distribuição do erro da previsão.<sup>8</sup>

## 3.3 Dados

Foram considerados na análise dados disponibilizados em portais de dados abertos do ONS (2024) [31] e da ANEEL (2024)[4]. Abaixo, além da descrição dos dados através de suas estatísticas principais na Tabela 1, cada uma das variáveis consideradas na elaboração dos modelos é apresentada. Todos os dados foram coletados em sua versão diária e posteriormente trabalhados como médias semanais para evitar a influência da sazonalidade ligada a mudança nos padrões de geração e consumo de energia nos dias úteis e fins de semana, que não é o objetivo deste estudo.

---

<sup>6</sup>Veja em On a Measure of a Lack of Fit in Time Series Models. Ljung, G. M., & Box, G. E. P. (1978) [25].

<sup>7</sup>Veja em Bollerslev (1986) [7]

<sup>8</sup>O detalhamento da parte algébrica pode ser visto em Hyndman (2021)[21]

Tabela 1: Principais estatísticas

Variáveis	<i>Curtailment</i> (MWh)	Capacidade (MW)	Demanda (MWmed)	Transmissão (MWmed)	GD (MW)
Média	44,4	24,7	73 791,5	79 732,0	46,3
Desv Pad	37,7	23,2	5 717,2	31 654,8	12,0
Mediana	25,5	19,6	73 720,2	80 759,7	45,1
Mínimo	7,8	0,0	63 450,8	-11 675,9	25,7
Máximo	171,4	135,1	86 190,2	149 206,5	80,0

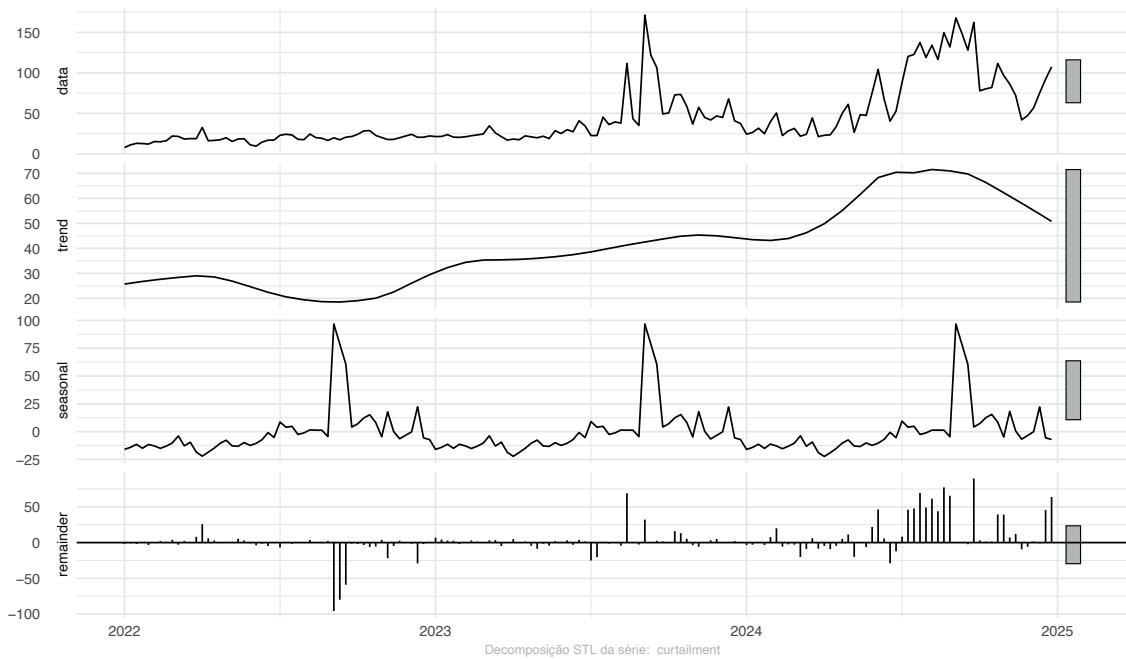
Nota: Os dados de Transmissão são referentes ao fluxo de energia da região Nordeste para as demais, por isso podem assumir valores negativos quando o volume de importação de energia para a região for maior que o de exportação.

### 3.3.1 *Curtailment*

O *curtailment* tem suas informações disponibilizadas na base de restrição de operação por *constrained-off* de usinas eólicas. Não foram incluídos dados do fenômeno na tecnologia solar em função do baixo volume de informações disponível. A unidade de medida é o MWh. Essa série foi a que definiu o recorte temporal da análise. Só há dados disponíveis a partir de outubro de 2021.

Na Figura 2 podemos observar a decomposição da série temporal. Na série original observamos uma clara mudança de padrão a partir de meados de 2023, quando o *curtailment* é expressivamente mais alto. A presença de uma sazonalidade anual também pode ser observada. Esse período de elevação sazonal do *curtailment* é justamente a "safra dos ventos", época do ano de maior disponibilidade de recurso para produção de energia eólica.

Figura 2: Decomposição STL da série *Curtailment* médio semanal - 2022-2024 (MWh)



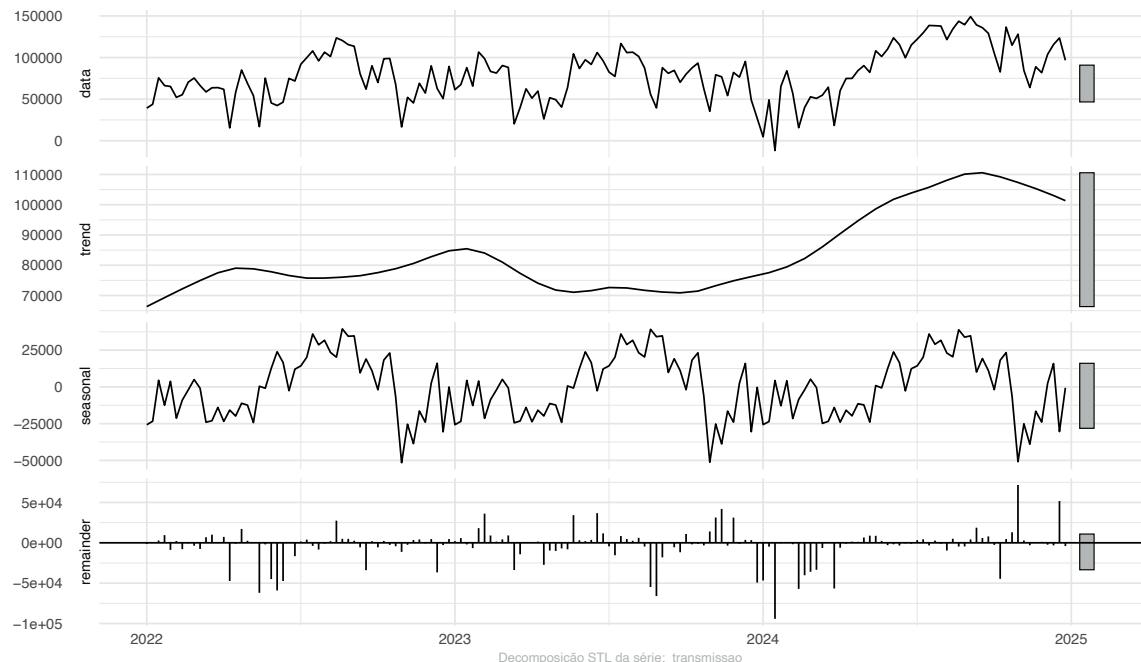
Fonte: Os autores

### 3.3.2 Transmissão de Energia

A transmissão de energia foi representada pelo intercâmbio de energia entre subsistemas, cujas grandezas são a soma das medidas de fluxo de potência ativa nas linhas de transmissão de fronteira entre os subsistemas. Foram selecionados apenas os valores dos fluxos de energia do subsistema Nordeste para as demais regiões. A unidade de medida é o MW médio.

Na decomposição da série na Figura 3 é interessante observar que a tendência capta bem o movimento realizado pelo ONS após o evento de apagão de agosto de 2021, quando a exportação de energia foi reduzida em cerca de 20% em prol de uma gestão mais conservadora das linhas. Alguns meses depois essa limitação seguiu presente mas em percentual menor, o que também aparece bem representado na série.

Figura 3: Decomposição STL da série Transmissão de Energia do subsistema Nordeste média semanal - 2022-2024 (MWmed)

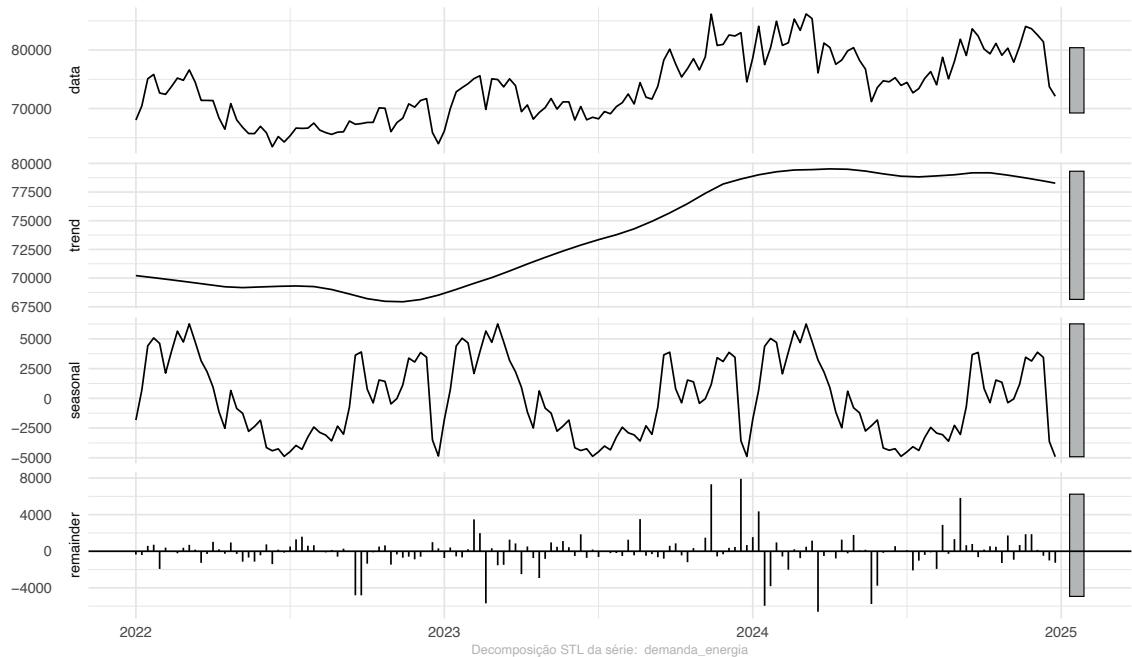


Fonte: Os autores

### 3.3.3 Demanda

A demanda de energia foi representada pela carga de energia diária, considerada em sua totalidade - para todo país. A unidade de medida também é o MW médio. Na decomposição da série na Figura 4 também observa-se a presença de uma sazonalidade anual e uma tendência de crescimento.

Figura 4: Decomposição STL da série Demanda de Energia nacional média semanal - 2022-2024 (MWmed)



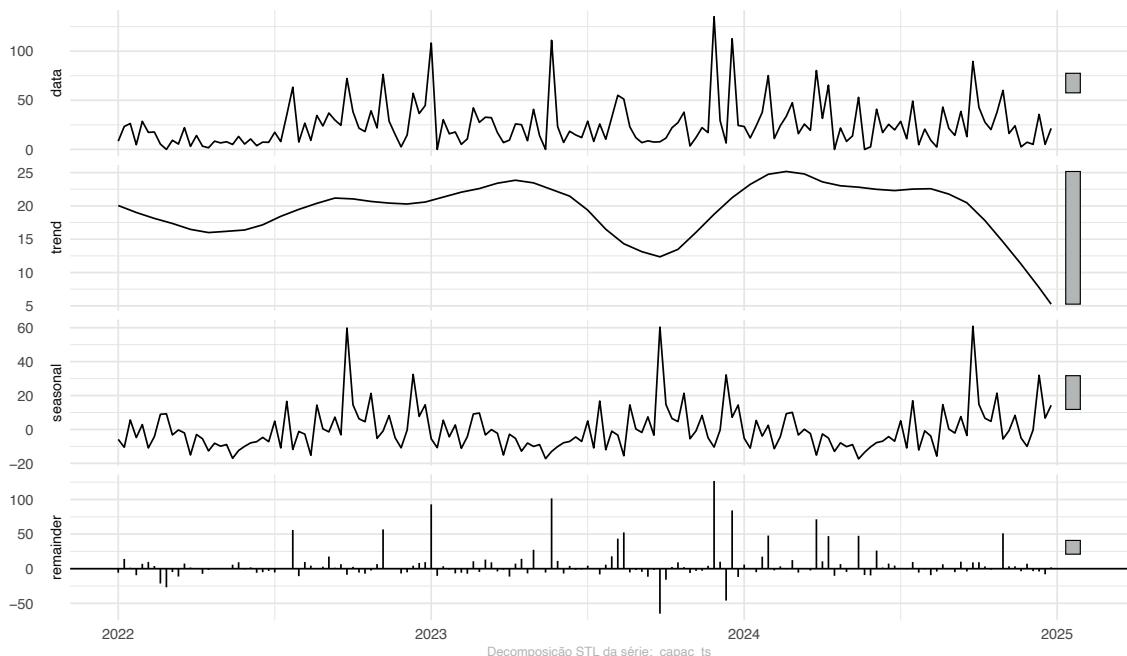
Fonte: O autores

### 3.3.4 Capacidade Instalada Centralizada

A capacidade instalada centralizada foi representada pela potência nominal de usinas despatchadas pelo ONS. É a maior parte da capacidade instalada no país, incluindo usinas das tecnologias eolioelétrica, fotovoltaica, hidroelétrica, nuclear e térmica. Os dados foram considerados como adição de nova capacidade em MW.

Na Figura 5 observamos uma queda da tendência no último ano, que pode estar associada a desaceleração dos investimentos em nova capacidade instalada eólicas e solar de projetos em escala industrial [22].

Figura 5: Decomposição STL da série Capacidade de Geração Centralizada adicionada média semanal - 2022-2024 (MW)

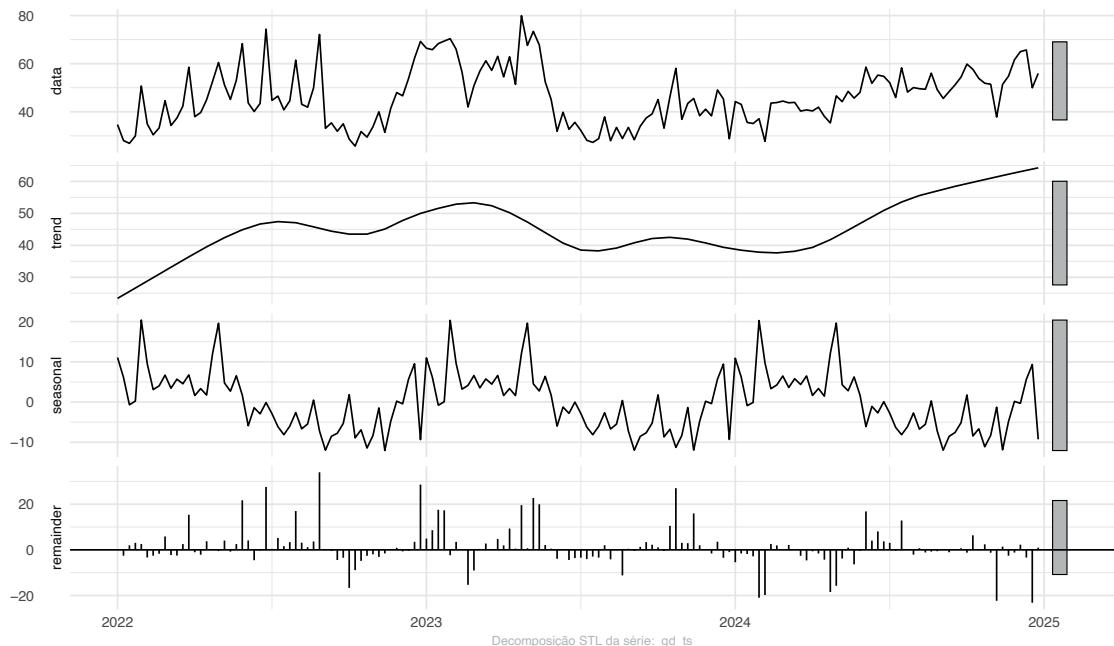


Fonte: Os autores

### 3.3.5 Geração Distribuída

A capacidade instalada em GD foi extraída da base de empreendimentos geração distribuída fotovoltaica. A GD é um modelo de produção de energia no qual ela é gerada no local ou próximo de onde será consumida. Abrange desde de consumidores residenciais até empresas e indústrias. Os dados foram considerados como adição de nova capacidade e estão expressos em MW. Na Figura 6 é possível ver a tendência de crescimento nos últimos anos e que não há uma sazonalidade muito definida.

Figura 6: Decomposição STL da série Capacidade de Geração Distribuída solar adicionada média semanal - 2022-2024 (MW)



Fonte: Os autores

## 4 Resultados

A aplicação da metodologia foi iniciada pela decomposição de todas as séries temporais e retirada do elemento de sazonalidade. A série sem sazonalidade foi tornada estacionária com a aplicação de uma diferenciação. O modelo identificado foi um ARIMA de ordem (1,0,3), conforme apresentado na Tabela 2. Os resíduos do modelo foram analisados<sup>9</sup> e ele foi considerado robusto e apto a realização de previsões.

<sup>9</sup>Veja no Apêndice a análise completa dos resíduos 5

Tabela 2: Informações do Modelo de Regressão ARIMA (1,0,3)

Variável	Estimativa	Erro padrão	Valor z	Pr(> z )	Signif.
AR1	0.982	0.016	59.848	<2.2e-16	***
MA1	-0.333	0.083	-4.020	0.0	***
MA2	-0.158	0.084	-1.887	0.059	.
MA3	-0.183	0.094	-1.953	0.051	.
Intercepto	147.840	56.693	2.608	0.009	**
Capac.Inst.	0.043	0.055	0.793	0.428	
GD	0.375	0.162	2.308	0.021	*
Demanda	-0.0018	0.001	-2.507	0.012	*
Transmissão	0.0002	0.0	2.214	0.027	*
$\sigma^2$	350.10				
Log likelihood	-674.510				
AIC	1369.20				
BIC	1399.510				

Notas. Códigos de significância: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1.

O parâmetro AR1, termo de autorregressão foi bastante significativo. Seu valor próximo de 1 sugeriu inércia no processo. É positivo, sugerindo que choques passados tendem a persistir e se somar positivamente à série. Dos parâmetros de média móvel, até o MA3 possui significância estatística a 10%, sendo esse escolhido para compor o ARIMA. O intercepto é positivo e significativo, indicando a existência de um termo constante e relevante.

Os coeficientes encontrados para as variáveis explicativas são coerentes com as dinâmicas já esperadas para os dados. A demanda é inversamente proporcional ao *curtailment*, uma vez que havendo maior demanda não há, ou há menos excedente de oferta. A capacidade centralizada adicionada não apresentou significância estatística.

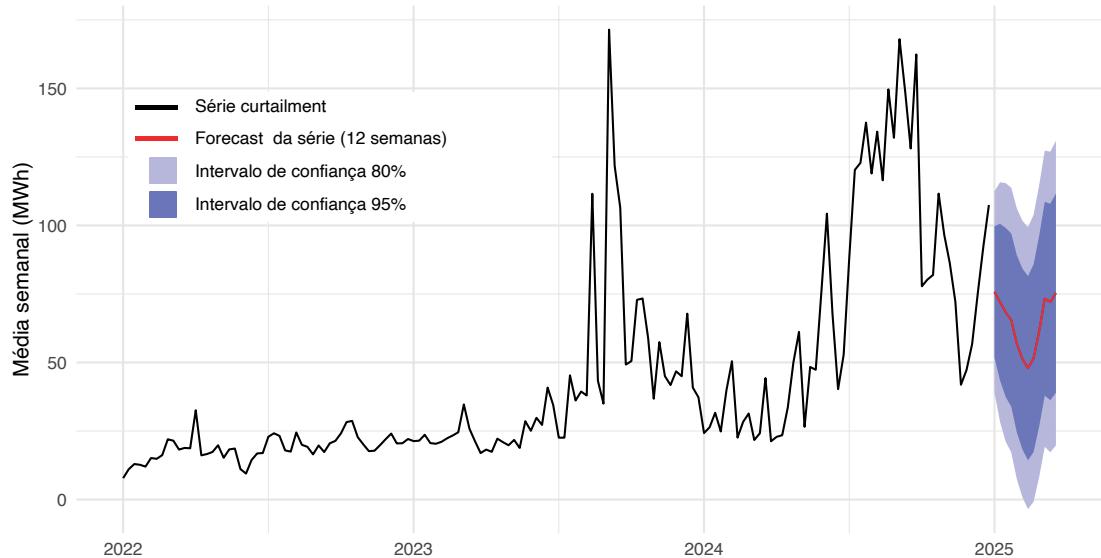
A geração solar distribuída, também conforme esperado, tem um impacto positivo e maior dentre as quatro. Com um crescimento exponencial nos últimos anos, tem sido a fonte de geração que mais cresce no Brasil e é hoje uma das mais representativas na matriz energética do país (EPE, 2024) [14]. Isso significa uma redução rápida da demanda - ou aumento da oferta, dependendo do ponto de vista. É, ao que parece pelo modelo e análise do contexto do setor, a provável maior responsável pela grande aceleração do crescimento do *curtailment*.

Já a transmissão apresenta um comportamento que à primeira vista causaria estranhanamento, já que está positivamente relacionada ao *curtailment*. Entretanto, além de a relação ter um valor muito baixo, entendemos que, em função do recorte temporal da análise, há uma aparência de crescimento que pode ter influência nesse resultado. Como após o apagão há uma redução na capacidade de escoamento da produção do Nordeste, seguida de uma flexibilização dessa restrição, a janela analisada apresenta um crescimento.

Na Figura 7 é apresentada uma previsão do *curtailment* para as 12 primeiras semanas

de janeiro. Ela foi elaborada a partir do modelo com a sazonalidade reintegrada a série. Em uma análise dos dados reais preliminares do *curtailment* de janeiro de 2025, com três semanas completas e uma parcial, a previsão realizada mostrou-se coerente, com os dados reais dentro de suas margens, confirmando uma tendência de continuidade dos patamares altos de restrição.

Figura 7: *Forecast Curtailment* - 12 semanas, média semanal (MWh)



Fonte: Os autores

## 5 Conclusão

O trabalho apresentado nesse artigo realizou uma investigação sobre o *curtailment* na geração eólica no Brasil entre 2022 e 2024. Foi feita uma análise de séries temporais associadas ao fenômeno incluindo a decomposição pelo método STL e modelagem de um ARIMA (1, 0, 3) para análise de variáveis explicativas e elaboração de uma previsão.

A decomposição dos dados mostrou de forma interessante como a dinâmica dessas variáveis têm-se comportado para além da sazonalidade, reforçando as tendências gerais percebidas sobre o tema. O impacto mais significativo no *curtailment*, para além da sazonalidade, é o crescimento da geração solar distribuída. A proporção da influência da demanda indica que mesmo diante de um aumento dela, para compensar os efeitos da Geração Distribuída, o mesmo teria que ser bastante expressivo.

Se considerarmos o processo de integração das VRE nos sistemas de energia sob a ótica da IEA, podemos afirmar que o cenário atual é de subaproveitamento da capacidade de geração através das VRE, o que tende a significar o desperdício de benefícios econômicos e ambientais proporcionados por essas tecnologias. No caso do sistema brasileiro, que é

predominantemente hídrico, essa afirmação pode parecer estranha à primeira vista, mas a necessidade de acionamento de térmicas em momentos de pico de demanda convivendo com o cenário de *curtailment* é um forte indicativo de que o uso de tecnologias como as baterias ou mesmo mecanismos de resposta à demanda, poderiam maximizar os resultados do setor de forma geral.

O setor hoje está envolto em uma série de discussões regulatórias ligadas ao tema. A provável entrada de baterias no sistema parece encomendada com as consultas públicas sobre o tema e a previsão de um leilão de capacidade para baterias ainda em 2025. Além disso, o futuro dos incentivos da geração distribuída são elementos importantes e em discussão. Do mesmo modo, novos usos da eletricidade como *data centers* com alto consumo e a eletrificação de usos finais na mobilidade urbana podem influenciar fortemente o fenômeno.

Com todos esses elementos, não há dúvidas de que as discussões acerca da gestão do *curtailment* seguirão por bastante tempo em pauta e a construção de bons modelos de *forecast* serão úteis. Investigações futuras podem lançar mão de dados climáticos na construção da previsão, uma vez que a disponibilidade de recurso eólico está cada vez mais sujeita a incertezas em função de mudanças nos padrões climáticos, como foi o longo *El Niño* de 2024.

Metodologicamente é possível avançar com modelos mais sofisticados, baseados em redes neurais, como *Long Short-Term Memory (LSTM)*, *Gated Recurrent Unit (GRU)* e *Convolutional Neural Network (CNN)*<sup>10</sup>, por exemplo. São metodologias capazes de captar e modelar não linearidades na série de dados e realizar previsões mais complexas, sendo alternativas apropriadas para análises futuras sobre o tema.

## Referências

- [1] ACUN, B., MORGAN, B., WU, C.-J., RICHARDSON, H., AND STEINSULTZ, N. Unlocking the potential of renewable energy through curtailment prediction. *Meta, Watt Time* (2023). Meta (1), WattTime (2).
- [2] AGARWAL, V. Box-jenkins method of forecasting, May 2018. Apresentação no Slideshare, publicada em 30 de maio de 2018.
- [3] ANEEL. Resolução normativa anel nº 1.030, de 26 de julho de 2022, 2022. Accessed on 26 Jan. 2025.
- [4] ANEEL. Portal de dados abertos anel. relação de empreendimentos de mini e micro geração distribuída, 2024. Período de 2022 a 2024. Acesso em: 23 jan. 2025.
- [5] BANDARA, K., HYNDMAN, R. J., AND BERGMEIR, C. Mstl: A seasonal-trend decomposition algorithm for time series with multiple seasonal patterns. *International Journal of Operational Research* (2022). In press.
- [6] BERGMEIR, C., HYNDMAN, R. J., AND BENÍTEZ, J. M. Bagging exponential smoothing methods using stl decomposition and box-cox transformation. *International Journal of Forecasting* 32, 2 (2016), 303–312.

---

<sup>10</sup>Ver mais em: Brownlee (2018) [9]

- [7] BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 3 (1986), 307–327.
- [8] Box, G. E., AND JENKINS, G. M. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, CA, 1970.
- [9] BROWNLEE, J. *Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python*. Machine Learning Mastery, Australia, 2018.
- [10] CHATFIELD, C. *The Analysis of Time Series: An Introduction*, 5th ed. Chapman and Hall/CRC, London, 1996.
- [11] CLEVELAND, R. B., CLEVELAND, W. S., MCRAE, J. E., AND TERPENNING, I. J. Stl: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics* 6, 1 (1990), 3–33.
- [12] DE LIMA, M. M. Sistemas de armazenamento em parques eólicos considerando pld horário e situações de constrained off. Dissertação de mestrado, Universidade Tecnológica Federal do Paraná, Curitiba, 2021.
- [13] EPE. Cadernos de estudo do pde 2034. caderno de consolidação de resultados. Tech. rep., Empresa de Pesquisa Energética / Ministério de Minas e Energia, 2024. Acesso em: 23 jan. 2025.
- [14] EPE. Cadernos de estudo do pde 2034. caderno de mmgd - baterias. Tech. rep., Empresa de Pesquisa Energética / Ministério de Minas e Energia, 2024. Acesso em: 23 jan. 2025.
- [15] GOLDEMBERG, J. O estado atual do setor elétrico brasileiro. *Revista USP - Dossiê Energia Elétrica*, 104 (2015).
- [16] GUJARATI, D. N., AND PORTER, D. C. *Basic Econometrics*, 5th ed. McGraw-Hill Education, New York, 2011.
- [17] HAMILTON, J. *Análise de séries temporais*. Princeton University Press, Princeton, 1994.
- [18] HARRISON, P., AND PEARCE, D. Extrapolation of time series and the determination of control limits. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 21, 1 (1972), 2–10.
- [19] HOLT, C. C. Forecasting trends and seasonals by exponentially weighted averages. *Carnegie Institute of Technology, ONR Memorandum* (1957). Reprinted in: International Journal of Forecasting, 2004, 20(1), pp. 5–10.
- [20] HURTADO, M., O'CONNOR, M., FLYNN, D., AND O'MALLEY, M. Analysis of wind energy curtailment in the ireland and northern ireland power systems. *arXiv preprint arXiv:2302.07143* (2023).
- [21] HYNDMAN, R. J. E ATHANASOPOULOS, G. *Forecasting: Principles and Practice*, 3rd ed. OTexts, Melbourne, Australia, 2021. Accessed on <current date>.

[22] IEA. World energy outlook 2023. Tech. rep., IEA, Paris, 2023. Licence: CC BY 4.0 (report); CC BY NC SA 4.0 (Annex A).

[23] IEA. Integrating solar and wind: Global experience and emerging challenges. Tech. rep., IEA, Paris, France, 2024. Acesso em: 23 Jan. 2025.

[24] LI, J., WANG, C., AND WANG, H. Optimal energy storage scheduling for wind curtailment reduction and energy arbitrage: A deep reinforcement learning approach. *arXiv preprint arXiv:2304.02239* (2023).

[25] LJUNG, G. M., AND BOX, G. E. P. On a measure of a lack of fit in time series models. *Biometrika* 65, 2 (1978), 297–303.

[26] LÓPEZ PROL, J., AND ZILBERMAN, D. No alarms and no surprises: Dynamics of renewable energy curtailment in california. *Energy Economics* 126 (2023), 106974.

[27] MEADE, N. *Forecasting with Exponential Smoothing Methods*. Gower Publishing Limited, Aldershot, England, 1984.

[28] MENEZES, R. F. A. *Programação Diária da Operação e Formação de Preços em Múltiplos Cenários Considerando Sistemas de Armazenamento de Energia por Bateria*. Tese de doutorado, Universidade Federal de Pernambuco, Recife, 2022.

[29] MERCEDES, S. S. P., RICO, J. A. P., AND POZZO, L. D. Y. Uma revisão histórica do planejamento do setor elétrico brasileiro. *Revista USP*, 104 (fev. 2015), 13–36.

[30] NIAZ, H., LIU, J. J., AND YOU, F. Can texas mitigate wind and solar curtailments by leveraging bitcoin mining? *Journal of Cleaner Production* 364 (2022), 132700.

[31] ONS. Portal de dados abertos ons. conjuntos de dados: Capacidade instalada de geração, carga de energia diária, intercâmbio entre subsistemas e restrição de operação por constrained-off de usinas eólicas., 2024. Período de 2022 a 2024. Acesso em: 23 jan. 2025.

[32] ONS. Capacidade instalada de geração - composição, 2025. Acesso em: 23 Jan. 2025.

[33] SILVA, V. E. M. D. *Influência do Tratamento de Dados no Desempenho de Modelos de Previsão da Geração Eólica Usando Informações de Restrição de Potência pelo Operador Nacional do Sistema Elétrico*. Trabalho de conclusão de curso, Universidade Federal do Ceará, Fortaleza-CE, 2023. Disponível em: <https://repositorio.ufc.br/handle/riufc/54035>.

[34] SILVA, W. C. O. D. *Estudo da Restrição de Geração Eólioelétrica por Constrained-Off e Aplicação da Metodologia de Obtenção de Curva de Produtividade Eólica do Operador Nacional do Sistema Elétrico*. Trabalho de conclusão de curso, Universidade Federal do Ceará, Fortaleza, 2022. Disponível em: <https://repositorio.ufc.br/handle/riufc/54034>.

[35] TAYLOR, S., AND LETHAM, B. Previsão em escala. *The American Statistician* 72, 1 (2018), 37–45.

[36] WOOLDRIDGE, J. M. *Introductory Econometrics: A Modern Approach*, 5 ed. South-Western, Cengage Learning, Mason, OH, 2013. Michigan State University.

# Apêndice - Análise dos Resíduos do Modelo

Foram realizados testes para avaliar a robustez dos resíduos do modelo cujos resultados são apresentados neste apêndice.

Como é possível observar na Tabela 3, para o teste de autocorrelação Ljung-Box não rejeitamos a hipótese nula ( $p\text{-valor} < 0.05$ ), indicando que os resíduos não apresentam autocorrelação significativa. O mesmo para o teste de heterocedasticidade: não rejeitamos a hipótese nula de que os resíduos não apresentam heterocedasticidade, indicando homocedasticidade. Para a verificação da normalidade, realizamos o teste de Shapiro-Wilk e o teste de Jarque-Bera. Ambos rejeitaram a hipótese nula de que os resíduos seguem uma distribuição normal.

Tabela 3: Testes aplicados aos resíduos

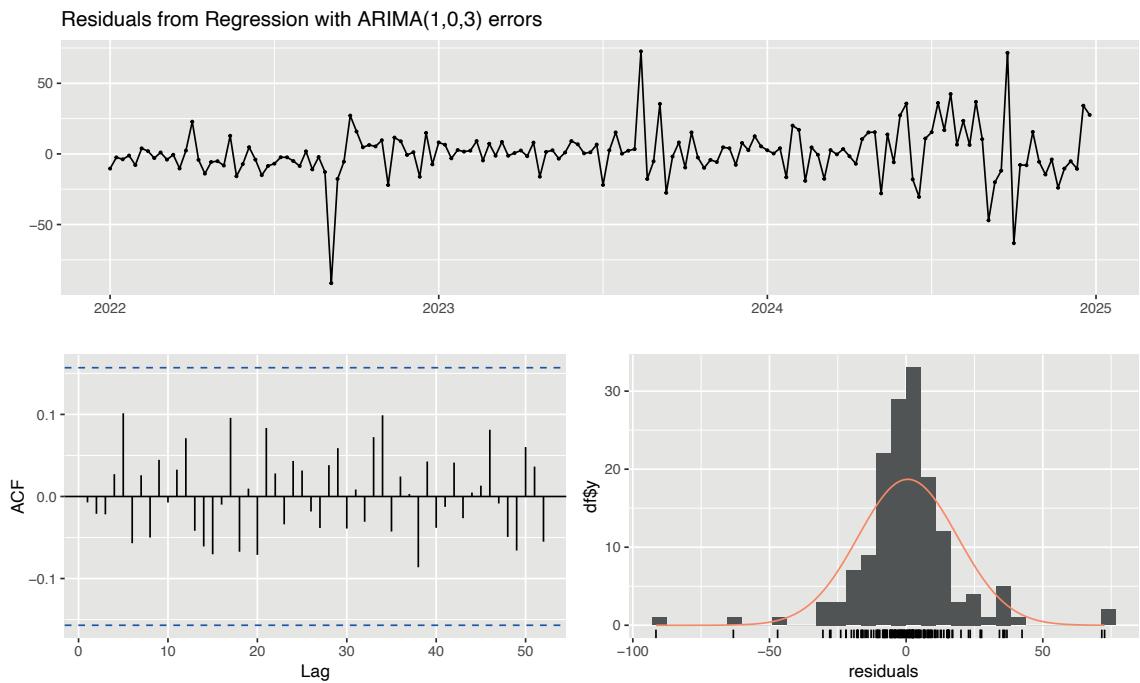
Teste	p-valor
Ljung-Box (Autocorrelação)	0.972
Shapiro-Wilk (Normalidade)	9.421521e-10
Jarque-Bera (Normalidade)	0
ARCH (Heterocedasticidade)	0.892

Apesar dos resíduos não serem distribuídos normalmente, o modelo captura bem a dependência temporal dos dados e não há autocorrelação residual. Um fator relevante nessa análise que nos ajuda a compreender o perfil dos resíduos são as mudanças abruptas características da série temporal do *curtailment* no período analisado, que podem ser os responsáveis por assimetrias nos dados e consequentemente afetaram a normalidade nos resíduos.

Além disso os resíduos apresentaram média 0.61, valor muito próximo de zero, considerando a escala dos dados. Isso associado ao resultado do teste ARCH nos indica que os ruídos tem um comportamento muito próximo ou igual a um passeio aleatório.

Levando esses fatores em consideração, consideramos o modelo robusto à análise dos resíduos e portanto apto para realizar previsões. No gráfico 1 é possível observar a série dos resíduos plotada, bem como a função de autocorrelação parcial (FACP) e o histograma da distribuição dos resíduos.

Figura 1: Resíduos da Regressão com modelo ARIMA (1,0,3)



Fonte: Os autores

# Determinantes do Spread Bancário para Pessoas Físicas no Brasil

André Luis R. Vasconcellos \*

Marlon Cecílio de Souza †

Edison Americo Huarsaya Tito ‡

Fernando Antonio Lucena Aiube §

25 de julho de 2025

## Resumo

Este estudo investiga os determinantes do spread bancário para pessoas físicas no Brasil por meio de dados de abril de 2011 a julho de 2024, utilizando um modelo de Vetores de Correção de Erros (VECM) e tendo por variáveis explicativas a Selic, inflação, PIB, inadimplência, total de instituições financeiras autorizadas e dívida bruta do governo geral. O modelo evidenciou que aproximadamente 87,51% dos desvios do equilíbrio são ajustados a cada período, confirmando a existência de relações de equilíbrio de longo prazo no sistema, nas quais a Selic destacou-se como a principal variável explicativa, apresentando impactos significativos e positivos sobre o spread em defasagens de um e dois períodos, ressaltando o papel central da política monetária na modulação dos custos de crédito. A inflação também mostrou influência relevante em defasagens específicas, indicando a transferência de custos financeiros decorrentes da instabilidade de preços. Por outro lado, variáveis macroeconômicas, como o PIB, apresentaram impacto significativo, mas de baixa magnitude, sugerindo que fatores estruturais têm maior peso no comportamento do spread bancário.

**Palavras-chave:** spread bancário, VECM, macroeconomia, mercado de crédito, Selic, análise econométrica

**Código JEL:** C32, G11, R39

## 1 Introdução

O sistema bancário brasileiro exerce papel fundamental na intermediação financeira, conectando poupadore e tomadores de recursos, promovendo a alocação eficiente de capital

---

\*Mestrando do Programa de Pós-Graduação em Ciências Econômicas da Universidade do Estado do Rio de Janeiro (PPGCE-UERJ)

†Mestrando do Programa de Pós-Graduação em Ciências Econômicas da Universidade do Estado do Rio de Janeiro (PPGCE-UERJ)

‡Professor Adjunto do Instituto de Matemática e Estatística - UERJ

§Professor Associado da Faculdade de Ciências Econômicas - UERJ

e influenciando a estabilidade macroeconômica. No entanto, uma característica persistente do mercado bancário nacional é a manutenção de spreads elevados, que encarecem o crédito e impõem barreiras ao consumo e à expansão de investimentos por parte das famílias. O spread bancário — definido como a diferença entre a taxa de juros ativa e a taxa de captação dos bancos — sintetiza os custos, os riscos e as ineficiências operacionais presentes no sistema, refletindo uma combinação de fatores macroeconômicos, institucionais e comportamentais Demirguc-Kunt (1998).

Nos últimos anos, o debate sobre os determinantes do spread bancário tem enfatizado o papel central da política monetária. A taxa Selic, como principal instrumento de controle da liquidez na economia, afeta diretamente o custo de captação das instituições financeiras, sendo rapidamente repassada às taxas de empréstimo. Em mercados concentrados e com baixa competição, como o brasileiro, esse repasse tende a ocorrer de forma quase automática, elevando o custo final do crédito Mishkin (1995). Além disso, o Brasil enfrenta um problema crônico de juros altos, que pode ser um obstáculo significativo para as condições de financiamento. De acordo com Garcia et al. (2019), a manutenção de juros elevados pode levar a uma redução da demanda por crédito, o que, por sua vez, pode afetar negativamente a atividade econômica.

Outro fator importante que influencia o spread bancário é a assimetria de informação. Como destacado por Stiglitz (1981), a assimetria de informação entre os bancos e os tomadores de crédito pode levar a uma precificação mais alta do risco, o que, por sua vez, aumenta o spread. Isso ocorre porque os bancos têm mais informações sobre a solvência dos tomadores do que os próprios tomadores, o que pode levar a uma seleção adversa e a uma precificação mais alta do risco.

Contudo, outros elementos também se mostram fundamentais para entender a persistência dos spreads. A inadimplência das famílias impacta diretamente o risco das operações de crédito, levando os bancos a elevarem os juros como forma de precificar a incerteza. Já a concorrência no sistema bancário, refletida no número de instituições financeiras atuantes, tem o potencial de reduzir spreads ao limitar o poder de mercado das grandes instituições. Nesse contexto, mercados mais abertos e menos concentrados tendem a oferecer melhores condições de financiamento.

Além disso, variáveis como a inflação e o Produto Interno Bruto (PIB) influenciam a dinâmica do spread. A inflação, ao reduzir o poder de compra e aumentar a incerteza, leva à adoção de prêmios de risco mais elevados. Por sua vez, o desempenho do PIB está associado à qualidade das carteiras de crédito e à capacidade de pagamento dos agentes econômicos, influenciando a percepção de risco sistêmico.

Diante disso, o presente estudo tem como objetivo investigar os principais determinantes do spread bancário para pessoas físicas no Brasil entre 2013 e 2024. Utilizando o modelo de Vetores de Correção de Erros (VECM), a análise busca identificar os fatores com maior influência no curto e no longo prazo, com ênfase na atuação da política monetária, nas condições macroeconômicas e nos elementos estruturais do sistema financeiro. Ao integrar essas dimensões, a pesquisa contribui para o debate sobre os custos do crédito e para a

formulação de políticas públicas voltadas à melhoria da eficiência e da inclusão financeira no país.

## 2 Revisão da Literatura

O spread bancário é frequentemente utilizado como indicador da eficiência e competitividade do sistema financeiro. Sua magnitude resulta da interação de múltiplos fatores, como política monetária, risco de crédito, inflação, atividade econômica e estrutura de mercado. Em economias emergentes como o Brasil, os spreads elevados refletem, em grande parte, a combinação entre riscos macroeconômicos e falhas estruturais de mercado Levine (2005).

A política monetária, especialmente por meio da taxa básica de juros (Selic), é amplamente reconhecida como um dos principais determinantes do spread. Conforme discutido por Mishkin (1995), alterações na taxa Selic são transmitidas rapidamente ao mercado de crédito, influenciando os custos de captação bancária e, por consequência, os juros cobrados aos tomadores. Em ambientes com baixa concorrência bancária, esse canal de transmissão tende a ser ainda mais eficaz, uma vez que os bancos possuem maior poder de precificação. Além disso, autores como De Paula (2020) de Paula (2020) destacam o impacto das altas taxas de juros no Brasil para o custo do crédito, ressaltando a importância da política monetária na determinação dos spreads.

Outro determinante relevante é a inadimplência. Níveis elevados de inadimplência aumentam o risco de crédito e exigem provisões maiores por parte das instituições financeiras, o que eleva o custo das operações e, portanto, o spread. Como demonstrado por Stiglitz (1981), os mercados de crédito com assimetrias informacionais tendem a responder com spreads mais altos para compensar o risco não observável dos tomadores.

Além disso, a estrutura do setor bancário exerce influência direta sobre o spread. Mercados com maior número de instituições financeiras tendem a apresentar maior concorrência e spreads mais baixos, conforme argumentado por Demirgüç-Kunt (1998). Em contrapartida, setores altamente concentrados dificultam a barganha dos tomadores e perpetuam custos elevados de crédito. No Brasil, a concentração bancária é apontada como um dos fatores que limitam a redução dos spreads, tornando essencial a ampliação da concorrência por meio da entrada de novos agentes e da digitalização do sistema financeiro. Desde 2016, o Banco Central do Brasil tem adotado medidas para facilitar a concorrência no setor bancário, como a criação de novas instituições financeiras e a promoção da inovação financeira Goldfajn (2017). Estudos como Oliveira (2020) discutem a importância dessas medidas para a redução dos spreads e a melhoria da eficiência do sistema financeiro.

No campo macroeconômico, a inflação atua como um importante condicionante do spread. Em cenários inflacionários, os bancos adotam prêmios de risco mais altos para proteger suas margens de rentabilidade frente à incerteza futura. Conforme exposto por Garcia et al. (2019), a inflação afeta a previsibilidade dos fluxos de caixa, aumentando a precificação do risco. Já o Produto Interno Bruto (PIB), ao refletir o nível de atividade econômica, impacta a capacidade de pagamento dos tomadores. Fases de crescimento reduzem o risco

de crédito e incentivam a oferta de financiamento, enquanto períodos recessivos produzem o efeito inverso, conforme Dantas Dantas (2016).

Portanto, o comportamento do spread bancário é multifacetado, sendo influenciado por determinantes monetários, macroeconômicos e estruturais. A literatura converge para a compreensão de que os elevados spreads no Brasil não podem ser explicados por um único fator, mas sim pela interação entre riscos, concorrência, política monetária e ambiente econômico. O presente estudo busca, com base nesse arcabouço teórico, avaliar empiricamente o peso relativo dessas variáveis na formação dos spreads para pessoas físicas, contribuindo para o avanço do conhecimento e para o aprimoramento das políticas de crédito no país.

## 3 Dados e Metodologia

### 3.1 Levantamento dos dados

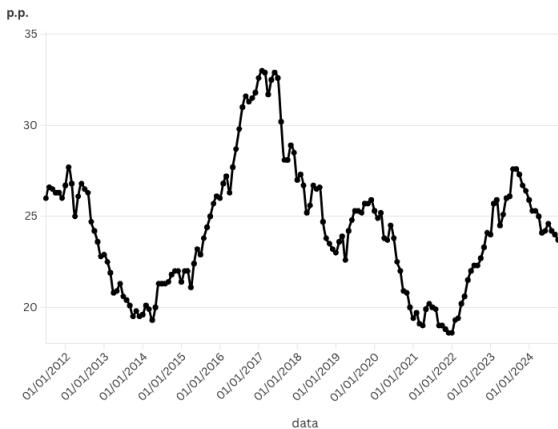
O levantamento de dados foi realizado com base na literatura vigente, considerando variáveis amplamente utilizadas em estudos relacionados aos determinantes do spread bancário para pessoas físicas no Brasil. Os dados foram coletados diretamente de fontes oficiais, do Banco Central do Brasil e do IBGE, acessados através de seus respectivos portais. Após a captura, as informações foram organizadas e tratadas no Excel, onde passaram por uma etapa de tratamento para garantir a consistência e a qualidade das informações. Em seguida, os dados organizados foram importados para o software RStudio, onde foram processados e analisados, possibilitando a execução das metodologias propostas na pesquisa.

Tabela 1: Variáveis, Descrição e Fontes de Dados

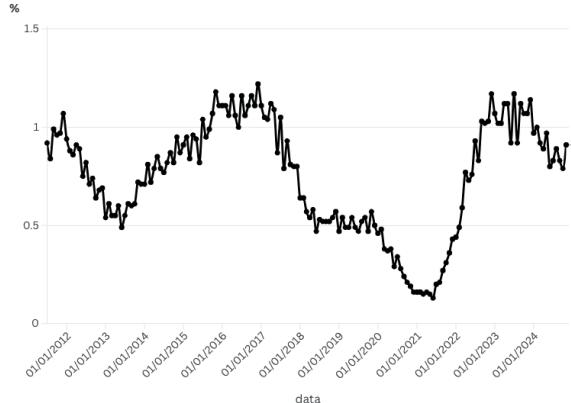
Variável	Descrição	Fonte	Origem
Spread Bancário – Pessoa Física (p.p.)	Diferença entre a taxa de juros cobrada pelos bancos e a taxa de captação, específica para pessoa física.	Bacen	Tabelas de Estatísticas Monetárias e de Crédito Tabela 14 SGS 20785
Taxa de juros – Selic acumulada no mês	Taxa básica de juros da economia brasileira, acumulada mensalmente.	Bacen	SGS - Sistema Gerenciador de Séries Temporais SGS: 4390
IPCA (Índice Nacional de Preços ao Consumidor Amplo)	Índice oficial de inflação utilizado no Brasil, calculado pelo IBGE.	IBGE	Índice Nacional de Preços ao Consumidor Amplo
PIB Mensal em unidade de R\$ (milhões)	Produto Interno Bruto mensal, em milhões de reais, indicando a atividade econômica do país.	Bacen	Boletim, Seção Atividade Econômica (Bacen / Boletim / Ativ. Ec.)
Inadimplência por modalidade – Pessoas Físicas (%)	Percentual de inadimplência de pessoas físicas em diferentes modalidades de crédito.	Bacen	Tabelas de Estatísticas Monetárias e de Crédito Tabela 20
Número total de Instituições Financeiras Autorizadas	Quantidade total de instituições financeiras autorizadas a operar no sistema financeiro nacional.	Bacen	Informações sobre evolução do SFN divulgação mensal Quadro 1
Dívida bruta do governo geral (% PIB)	Relação entre a dívida bruta do governo geral e o PIB, expressa em percentual.	Bacen	SGS - Sistema Gerenciador de Séries Temporais SGS: 13762

### 3.2 Análise Gráfica

Evolução do Spread PF



Evolução da Taxa Selic

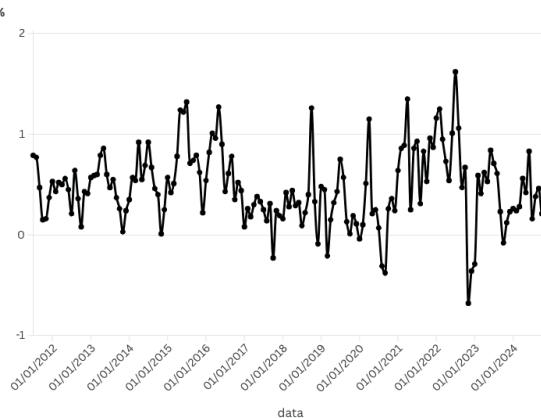


*Fonte: Os autores.*

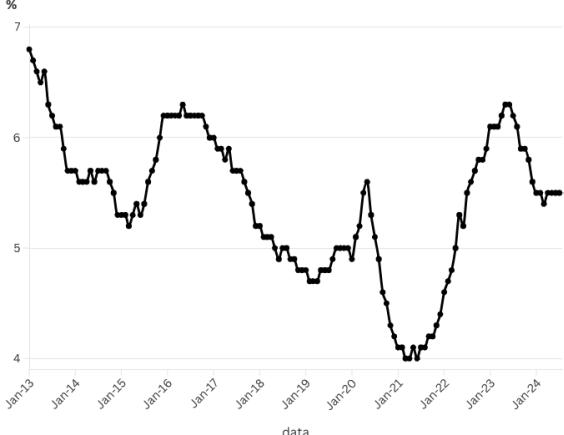
O gráfico intitulado 'Evolução do Spread Bancário para Pessoa Física' apresenta a variação do spread bancário para pessoas físicas ao longo do tempo. Nota-se uma queda acentuada em 2017, seguida por uma recuperação gradual, o que pode estar associado a alterações na política monetária ou a fatores macroeconômicos específicos. Além disso, observam-se períodos de relativa estabilidade nos últimos anos, especialmente após 2020, o que pode indicar maior previsibilidade nas condições econômicas ou uma acomodação dos spreads a patamares mais constantes. A análise visual também sugere que a série não é estacionária, dado que as oscilações e as tendências de alta e baixa parecem variar ao longo do tempo, possivelmente refletindo mudanças estruturais no mercado de crédito ou choques econômicos. Esses aspectos reforçam a importância de se considerar fatores conjunturais e estruturais ao analisar a dinâmica dos spreads bancários no Brasil.

O gráfico intitulado 'Evolução da Taxa Selic' apresenta a variação de uma variável associada à taxa básica de juros, provavelmente a Taxa Selic, ao longo do tempo. É evidente que a série apresenta oscilações significativas, com valores variando dentro de um intervalo considerável, destacando-se movimentos de queda mais acentuados em determinados períodos, como por volta de 2017, seguidos de estabilização em patamares mais baixos e subsequentes aumentos a partir de 2021. A análise visual sugere que a série não é estacionária, uma vez que há tendências distintas ao longo do tempo, caracterizadas por períodos de alta, baixa e estabilidade, refletindo as respostas da política monetária às condições econômicas prevalentes em cada momento. Esses comportamentos reforçam a necessidade de investigar os fatores conjunturais que influenciam a variável representada, especialmente no que tange à gestão macroeconômica e ao impacto sobre o mercado de crédito.

## Evolução da Inflação



## Evolução da Inadimplência

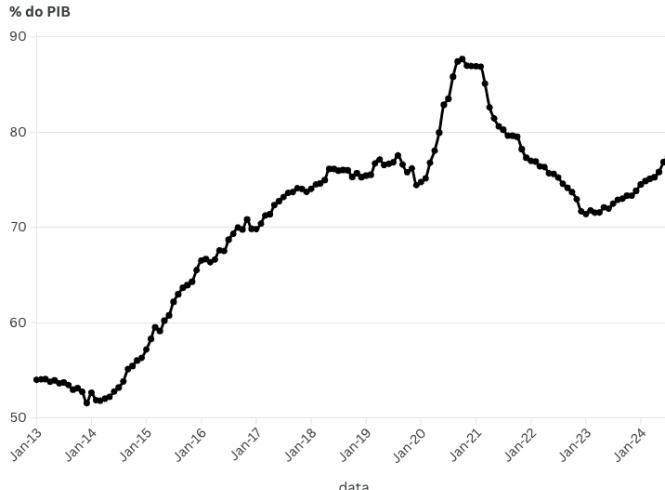


*Fonte: Os autores*

O gráfico intitulado 'Evolução da Taxa de Inflação' apresenta a evolução de uma série temporal entre 2012 e 2025. Visualmente, a série oscila em torno de dois níveis principais, 0 e 1, sugerindo um comportamento típico de uma variável categórica ou indicadora. Esse padrão de variação discreta indica que a série não apresenta tendência clara de longo prazo, e sua média permanece aproximadamente constante ao longo do tempo. Com base na análise visual, é possível inferir que a série é estacionária, dado que não há sinais de tendência, heterocedasticidade ou mudanças estruturais ao longo do período. No entanto, uma análise mais detalhada, incluindo testes estatísticos formais, seria necessária para confirmar a estacionariedade e avaliar a consistência desse comportamento em diferentes condições.

O gráfico intitulado 'Evolução da Taxa de Inadimplência' apresenta a evolução temporal de uma série relacionada à taxa de inadimplência no período de 2010 a 2025. A série mostra variações significativas ao longo do tempo, com flutuações que indicam períodos de aumento e redução na taxa de inadimplência, refletindo possivelmente alterações no ambiente econômico, como crises ou medidas governamentais. Visualmente, a série não exibe uma tendência clara de longo prazo, mas apresenta variações em torno de um intervalo relativamente estável, sugerindo que ela pode ser estacionária. No entanto, essas flutuações periódicas podem ser atribuídas a fatores sazonais ou conjunturais. Para confirmar a estacionariedade, seria necessário aplicar testes estatísticos formais, como o teste de Dickey-Fuller aumentado, mas, com base apenas na análise visual, não há sinais evidentes de tendência crescente ou decrescente.

## Evolução da Dívida Bruta do Governo Geral



*Fonte: Os autores*

O gráfico intitulado Evolução da Dívida Bruta do Governo Geral apresenta a evolução de uma série temporal que mede a dívida do governo entre 2010 e 2025. Observa-se que a série apresenta uma tendência de crescimento moderado até aproximadamente 2016, quando ocorre uma aceleração no aumento dos valores registrados. Após esse ponto, os dados mostram uma estabilização em níveis elevados, especialmente a partir de 2019, sugerindo a presença de uma possível quebra estrutural no comportamento da série.

Do ponto de vista da estacionariedade, não há evidências visuais de que a série seja estacionária, uma vez que apresenta uma clara tendência de crescimento, alternada por períodos de estabilidade prolongada — o que indica a ausência de um comportamento que reverta a valores constantes ao longo do tempo. Tal comportamento sugere que a série pode ser integrada de ordem 1 (não estacionária em níveis, mas possivelmente estacionária em diferenças), mas uma análise econometrífica mais rigorosa seria necessária para confirmar essa hipótese.

### 3.3 Tratamento estatístico de dados

1. **Análise de Correlação:** avaliação da dependência linear entre variáveis, reconhecendo que correlação não implica causalidade. A análise de correlação é uma ferramenta fundamental na estatística para identificar a força e a direção do relacionamento linear entre duas variáveis. Conforme destacado por Pearson (1895), a correlação permite um entendimento preliminar das possíveis associações entre variáveis e é frequentemente utilizada como ponto de partida em estudos exploratórios. No entanto, é crucial lembrar que correlação não implica causalidade, como enfatizado por Yule (1926), sendo necessário cautela ao interpretar esses resultados em termos de relações causais.

Tabela 2: Matriz de Correlação

	spread_pf	selic	inflação	PIB	inad	inst_fin	div_gov
spread_pf	1,0000	0,5695	-0,2607	-0,0814	0,3977	0,0890	-0,0262
selic	0,5695	1,0000	-0,0151	0,0447	0,5237	0,2950	-0,4283
inflação	-0,2607	-0,0151	1,0000	-0,0388	-0,1789	0,0520	-0,0687
PIB	-0,0814	0,0447	-0,0388	1,0000	-0,3952	-0,7095	0,6887
inad	0,3977	0,5237	-0,1789	-0,3952	1,0000	0,6705	-0,7191
inst_fin	0,0890	0,2950	0,0520	-0,7095	0,6705	1,0000	-0,8806
div_gov	-0,0262	-0,4283	-0,0687	0,6887	-0,7191	-0,8806	1,0000

Para este estudo, verificou-se que não há correlação perfeita ( $\text{cor} = 1$ ) entre as variáveis analisadas.

2. **Estatística Descritiva:** sumarização e caracterização das séries temporais envolvidas. A estatística descritiva é uma área da estatística que organiza, sumariza e apresenta dados de forma comprehensível. Seu propósito é descrever as características de um conjunto de dados, utilizando ferramentas como gráficos, tabelas e medidas de tendência central (média, mediana, moda) e dispersão (desvio padrão, variância).

Tabela 3: Estatísticas descritivas das variáveis

Variável	Mín	Máx	Mediana	Média	DP	Assimetria	Curtose	EP
spread_pf	18,6	33,0	24,2	24,34	3,49	0,48	-0,15	0,27
selic	0,13	1,22	0,8	0,75	0,29	-0,41	-0,83	0,02
inflacao	-0,68	1,62	0,44	0,48	0,37	0,19	0,60	0,03
pib	347879,6	987472,6	555644,6	596522,47	170354,75	0,72	-0,55	13425,83
inad	4,0	7,2	5,6	5,58	0,77	-0,01	-0,41	0,06
inst_fin	1625,0	2504,0	1742,0	1851,57	246,00	1,31	0,82	19,39
div_gov	51,27	87,68	72,1	68,05	10,85	-0,28	-1,26	0,85

Com base nos dados apresentados na tabela, observa-se que as médias e medianas das variáveis estão próximas, o que indica que não há grandes desvios causados por valores extremos para a maioria das variáveis. Um exemplo disso é **spread\_pf**, com média de 24,34 e mediana de 24,2. Algumas exceções, como **pib**, podem ter a média levemente influenciada por valores extremos devido à alta variância ( $DP = 170.354,75$ ). Em suma, os dados são relativamente estáveis em algumas variáveis, como inadimplência, enquanto outras, como spread e inst\_fin, mostram maior dispersão.

### 3.4 Teste de Estacionariedade e Teste de Cointegração

Os testes de estacionariedade e cointegração desempenham um papel fundamental na modelagem econométrica, especialmente quando se busca identificar relações de longo prazo

entre variáveis econômicas. Neste estudo, a aplicação desses testes teve como objetivo definir a abordagem de modelagem mais apropriada — seja um modelo VAR, no caso de séries estacionárias, ou um modelo VECM, caso haja evidências de cointegração entre séries não estacionárias.

Para essa finalidade, foi aplicado o teste de Dickey-Fuller Aumentado (ADF), que permite avaliar a presença de raiz unitária nas séries temporais. A classificação das séries quanto à estacionariedade baseou-se no resultado desse teste, que forneceu evidências sobre a presença ou ausência de raiz unitária. Embora o uso isolado de um único teste possa ser considerado limitado, o teste ADF é amplamente utilizado e aceito na literatura econômica para essa finalidade. Além disso, sua adoção pode ser justificada pela simplicidade e eficácia em detectar a presença de raiz unitária.

De acordo com os resultados do teste ADF, apenas as variáveis *inflação* (*inflacao*) e *total de instituições financeiras* (*inst\_fin*) se mostraram estacionárias em nível, enquanto as demais — *spread PF* (*spread\_pf*), *taxa Selic* (*selic*), *produto interno bruto* (*pib*), *inadimplência* (*inad*) e *dívida do governo* (*div\_gov*) — apresentaram evidências de raiz unitária, caracterizando-se como não estacionárias.

Tabela 4: Teste ADF para a variável *spread\_pf*

Sem dif. e sem tendência			Com dif. e sem tendência			Com dif. e com tendência		
Lag	ADF	p-valor	Lag	ADF	p-valor	Lag	ADF	p-valor
0	-0.1804	0.5915	0	-1.6690	0.4584	0	-3.3976	0.0261
1	-0.1956	0.5871	1	-1.6682	0.4812	1	-3.3413	0.0498
2	-0.4077	0.5284	2	-2.2218	0.4125	2	-2.0433	0.5542
3	-0.4810	0.5064	3	-2.3928	0.1759	3	-2.2884	0.4727
4	-0.6763	0.4366	4	-2.8955	0.0501	4	-2.7871	0.2474

Tabela 5: Teste ADF para a variável *selic (juros)*

Sem dif. e sem tendência			Com dif. e sem tendência			Com dif. e com tendência		
Lag	ADF	p-valor	Lag	ADF	p-valor	Lag	ADF	p-valor
0	-0.4160	0.5751	0	-2.7993	0.0616	0	-3.7962	0.0173
1	-0.4175	0.5744	1	-3.0416	0.0311	1	-4.0147	0.0099
2	-0.6848	0.4536	2	-2.9831	0.0367	2	-3.9212	0.0136
3	-0.8857	0.3932	3	-2.8135	0.0600	3	-3.9001	0.0145
4	-1.0082	0.3525	4	-2.8544	0.0545	4	-3.9433	0.0127

Tabela 6: Teste ADF para a variável *pib (crescimento)*

Sem dif. e sem tendência			Com dif. e sem tendência			Com dif. e com tendência		
Lag	ADF	p-valor	Lag	ADF	p-valor	Lag	ADF	p-valor
0	-1.4091	0.5610	0	-2.8857	0.0476	0	-4.5432	0.0043
1	-1.3647	0.5775	1	-3.2524	0.0201	1	-4.7935	0.0027
2	-1.2183	0.6220	2	-3.2407	0.0206	2	-4.7022	0.0032
3	-1.1753	0.6451	3	-3.2556	0.0199	3	-4.6310	0.0037
4	-1.1476	0.6418	4	-3.1553	0.0257	4	-4.4216	0.0057

Tabela 7: Teste ADF para a variável *inadimplência*

Sem dif. e sem tendência			Com dif. e sem tendência			Com dif. e com tendência		
Lag	ADF	p-valor	Lag	ADF	p-valor	Lag	ADF	p-valor
0	-0.3829	0.5387	0	-2.7109	0.0712	0	-3.8427	0.0155
1	-0.4173	0.5272	1	-3.1405	0.0312	1	-4.0276	0.0095
2	-0.5837	0.4851	2	-2.9381	0.0435	2	-3.9488	0.0126
3	-0.8026	0.4214	3	-2.8180	0.0585	3	-3.7610	0.0188
4	-0.9448	0.3806	4	-2.7601	0.0641	4	-3.7371	0.0198

Tabela 8: Teste ADF para a variável *inst\_fin*

Sem dif. e sem tendência			Com dif. e sem tendência			Com dif. e com tendência		
Lag	ADF	p-valor	Lag	ADF	p-valor	Lag	ADF	p-valor
0	-4.7881	0.0124	0	-3.7701	0.0088	0	-4.1710	0.0121
1	-3.7663	0.0295	1	-3.7810	0.0086	1	-4.2272	0.0108
2	-3.3516	0.0536	2	-3.8791	0.0070	2	-4.1587	0.0125
3	-3.0205	0.0849	3	-3.7798	0.0087	3	-4.2076	0.0113
4	-2.8617	0.0932	4	-3.8459	0.0076	4	-4.1130	0.0138

Tabela 9: Teste ADF para a variável *div\_gov*

Sem dif. e sem tendência			Com dif. e sem tendência			Com dif. e com tendência		
Lag	ADF	p-valor	Lag	ADF	p-valor	Lag	ADF	p-valor
0	-1.3963	0.5682	0	-2.7239	0.0690	0	-3.8141	0.0161
1	-1.2932	0.6150	1	-3.1985	0.0272	1	-3.9340	0.0130
2	-1.2783	0.6217	2	-3.0926	0.0334	2	-3.8165	0.0160
3	-1.2240	0.6448	3	-2.8643	0.0572	3	-3.8021	0.0164
4	-1.2429	0.6368	4	-2.8824	0.0554	4	-3.8213	0.0158

Tabela 10: Teste ADF para a variável *inflacao*

Sem dif. e sem tendência			Com dif. e sem tendência			Com dif. e com tendência		
Lag	ADF	p-valor	Lag	ADF	p-valor	Lag	ADF	p-valor
0	-4.2032	0.0067	0	-3.9184	0.0066	0	-4.2326	0.0107
1	-4.1267	0.0082	1	-3.9942	0.0056	1	-4.3195	0.0090
2	-3.9732	0.0110	2	-3.8734	0.0072	2	-4.2014	0.0117
3	-3.9016	0.0125	3	-3.9343	0.0064	3	-4.2715	0.0100
4	-3.8451	0.0139	4	-3.8897	0.0069	4	-4.2231	0.0110

Em seguida, foi realizado o teste de cointegração de Johansen com o objetivo de identificar a existência de relações de equilíbrio de longo prazo entre as variáveis econômicas selecionadas. Para assegurar a validade estatística do teste, procedeu-se à exclusão temporária das variáveis *inflação* (*inflacao*) e *total de instituições financeiras* (*inst\_fin*), por apresentarem estacionariedade em nível ( $I(0)$ ). Conforme destacado por Johansen (1995), a inclusão de séries com ordens distintas de integração pode comprometer a identificação adequada de vetores de cointegração, sendo recomendável trabalhar exclusivamente com variáveis integradas da mesma ordem ( $I(1)$ ).

A seleção do número ótimo de defasagens foi realizada por meio dos critérios de informação de Akaike (AIC), Hannan-Quinn (HQ), Schwarz (SC) e Erro de Previsão Final (FPE), sendo o AIC e o FPE os que indicaram a necessidade de utilizar sete defasagens para capturar adequadamente a dinâmica temporal das séries.

Tabela 11: Critérios de Seleção para o Número Ótimo de Defasagens

Critério	AIC(n)	HQ(n)	SC(n)	FPE(n)
Número de Lags	7	3	3	7

Dos quatro critérios utilizados, dois (AIC e FPE) apontaram o uso de sete defasagens como o mais adequado para a modelagem, enquanto HQ e SC sugeriram três defasagens. Considerando a maior sensibilidade do critério AIC à presença de autocorrelação residual e seu bom desempenho em amostras de tamanho moderado, optou-se por seguir com sete defasagens, como número ideal para aplicação do Teste de Johansen.

Tabela 12: Resultados do Teste de Cointegração (Johansen)

$H_0$	Est do Teste	Crítico (10%)	Crítico (5%)	Crítico (1%)	Rejeita $H_0$ ?
$r \leq 4$	5.63	6.50	8.18	11.65	Não
$r \leq 3$	17.10	15.66	17.95	23.52	Sim
$r \leq 2$	35.47	28.71	31.52	37.22	Sim
$r \leq 1$	59.90	45.23	48.28	55.43	Sim
$r = 0$	97.84	66.49	70.60	78.87	Sim

Conforme apresentado nos resultados do Teste de Cointegração de Johansen, existem quatro relações de cointegração ( $r = 4$ ) entre as cinco variáveis I(1) analisadas, indicando que quatro combinações lineares dessas variáveis são estacionárias.

Os resultados apresentados na Tabela 6 refletem a aplicação do teste de cointegração de Johansen, cuja finalidade é verificar a existência de relações de equilíbrio de longo prazo entre as variáveis do sistema. A hipótese nula ( $H_0$ ) estabelece que o número de relações de cointegração é menor ou igual a  $r$ . Com base nos valores críticos ao nível de 5%, rejeita-se  $H_0$  para  $r = 0$  até  $r = 3$ , mas não se rejeita  $H_0$  para  $r \leq 4$ , indicando que o número apropriado de vetores de cointegração é quatro.

A seleção do número de relações de cointegração é feita pela comparação entre a estatística de teste e os valores críticos para cada  $r$ . Quando a estatística excede o valor crítico, rejeita-se  $H_0$ , implicando a existência de pelo menos  $r + 1$  relações de cointegração. Assim, ao rejeitarmos  $H_0$  até  $r \leq 3$  e não rejeitarmos para  $r \leq 4$ , concluímos que  $r = 4$ .

Essa escolha está em linha com a literatura econométrica, segundo a qual deve-se selecionar o maior valor de  $r$  para o qual  $H_0$  ainda é rejeitada, evitando a superestimação do número de relações estacionárias no sistema. Como discutido por Johansen (1995), a determinação correta de  $r$  é essencial para garantir a identificação e a estabilidade do modelo VEC.

Dessa forma, os resultados confirmam a existência de quatro relações de cointegração no modelo, validando a aplicação de um modelo VEC (Vetores de Correção de Erros) para capturar adequadamente as dinâmicas de curto e longo prazo entre as variáveis.

A Tabela 13 apresenta os autovalores do Teste de Cointegração (Johansen), que indicam a força relativa de cada vetor de cointegração. O maior autovalor ( $\lambda = 0.2498$ ) está associado ao primeiro vetor de cointegração, refletindo a relação de longo prazo mais forte entre as variáveis. À medida que os autovalores diminuem, as relações de cointegração tornam-se progressivamente menos influentes. O quinto autovalor, de apenas 0.0417, está abaixo dos demais, sugerindo fraca contribuição adicional.

Tabela 13: Autovalores do Teste de Cointegração (Johansen)

Rank (r)	Autovalor (Eigenvalue)
$r = 0$	0.2498
$r = 1$	0.1690
$r = 2$	0.1299
$r = 3$	0.0833
$r = 4$	0.0417

A matriz de pesos ( $W$ ), também chamada de matriz dos coeficientes de ajuste ( $\alpha$ ), indica como cada variável do sistema responde a desvios das relações de cointegração estimadas. Cada entrada da matriz representa o grau de resposta de uma variável endógena a uma

determinada relação de equilíbrio de longo prazo.

Por exemplo, o valor  $W[1, 1] = -0.1531$  indica a contribuição da primeira relação de cointegração para o ajuste da variável `spread_pf`. De forma semelhante, o valor  $W[3, 1] = -25.6879$  mostra que a variável `pib` reage fortemente à primeira relação de cointegração, com um ajuste significativo em resposta a desvios do equilíbrio de longo prazo.

Valores baixos ou próximos de zero em determinadas entradas da matriz  $W$  indicam que certas relações de cointegração exercem influência marginal sobre algumas variáveis do sistema, refletindo a seletividade com que os equilíbrios de longo prazo afetam os ajustes de curto prazo.

Com base nos resultados obtidos, foram identificadas **quatro relações de cointegração** entre as variáveis, conforme evidenciado pelo teste de Johansen. Esse achado confirma a existência de vínculos de longo prazo robustos entre as séries  $I(1)$  analisadas, mesmo diante das flutuações observadas no curto prazo. Assim, a modelagem será conduzida por meio de um modelo de Vetores de Correção de Erros (VECM), capaz de capturar os ajustes de curto prazo condicionados pelas relações de equilíbrio de longo prazo.

Adicionalmente, a magnitude das cargas na matriz  $W$  revela a sensibilidade de cada variável a desvios de equilíbrio. Exemplos representativos incluem:

- $W[1, 3] = -0.0202$ : a terceira relação de cointegração tem impacto moderado sobre `spread_pf`.
- $W[3, 1] = -25.6879$ : a variável `pib` responde fortemente à primeira relação de cointegração.
- $W[4, 2] = -0.00008$ : o impacto da segunda relação de cointegração sobre a inadimplência é praticamente nulo.

Esses resultados reforçam a importância de incorporar as relações de cointegração ao modelar sistemas dinâmicos. O uso do VECM permite capturar não apenas a interdependência de longo prazo entre as variáveis, mas também a forma como essas variáveis reagem, de maneira diferenciada, aos desvios dessas relações de equilíbrio.

### 3.4.1 Verificação da significância dos resíduos

Para garantir a validade do modelo VECM estimado, os resíduos foram submetidos a uma bateria de testes estatísticos que avaliam pressupostos fundamentais como normalidade, ausência de autocorrelação e homocedasticidade. A adequação desses requisitos é essencial para que as inferências do modelo sejam confiáveis e robustas, conforme apontado por Lütkepohl (2005).

1. **Teste de normalidade:** O teste de Shapiro-Wilk foi aplicado aos resíduos do componente `spread_pf.d`, revelando um valor de  $W = 0.9829$  com  $p$ -valor = 0.0953. Como

o p-valor é superior ao nível de significância de 5%, não se rejeita a hipótese nula de normalidade. Embora o teste tenha sido exemplificado com apenas um componente, os resultados preliminares indicam que não há evidência estatística suficiente para afirmar que os resíduos desse componente se desviam significativamente de uma distribuição normal. Esse resultado fortalece a hipótese de que os erros do modelo seguem um padrão compatível com os pressupostos do VECM.

2. **Teste de autocorrelação:** O teste de Ljung-Box foi aplicado aos resíduos de cada equação do sistema VECM, com 7 defasagens. Os resultados estão apresentados na Tabela 14. Como todos os p-valores foram superiores a 0,05, conclui-se que não há evidência de autocorrelação serial nos resíduos. Isso significa que os resíduos se comportam como ruído branco, ou seja, não possuem estrutura temporal não explicada pelo modelo. A ausência de autocorrelação indica que o VECM é capaz de capturar adequadamente as dinâmicas das variáveis, assegurando que os erros não carregam informação remanescente. Esse resultado reforça a validade do modelo sob o critério de independência dos resíduos, o que é desejável em análises de séries temporais multivariadas.

Tabela 14: Resultados do Teste de Autocorrelação (Ljung-Box)

Componente	Estatística $X^2$	p-valor
spread_pf.d	5.0867	0.6494
selic.d	4.5257	0.7176
inflacao.d	1.6899	0.9750
pib.d	5.1684	0.6394
inad.d	8.9585	0.2556
inst_fin.d	4.6097	0.7075
div_gov.d	2.9241	0.8919

3. **Teste de homocedasticidade:** O teste de Breusch-Pagan foi utilizado para avaliar a constância da variância dos resíduos ao longo do tempo. Os resultados estão apresentados na Tabela 15. Os p-valores são todos superiores ao nível de 5%, indicando que não há evidência estatística para rejeitar a hipótese nula de homocedasticidade. Isso implica que os resíduos não apresentam variância não constante, ou seja, são homocedásticos, o que contribui positivamente para a confiabilidade dos estimadores.

Tabela 15: Resultados do Teste de Homocedasticidade (Breusch-Pagan)

Componente	Estatística BP	p-valor
spread_pf.d	0.2144	0.6434
selic.d	0.8240	0.3640
inflacao.d	2.3685	0.1238
pib.d	2.9549	0.0856
inad.d	2.5184	0.1125
inst_fin.d	0.2787	0.5976
div_gov.d	0.0009	0.9761

Dessa forma, os testes diagnósticos sugerem que os resíduos do modelo não apresentam violações importantes aos pressupostos de normalidade (ao menos no exemplo testado), autocorrelação ou heterocedasticidade. Em conjunto, esses achados sustentam a robustez do modelo VECM ajustado, conferindo maior confiabilidade às inferências derivadas da análise.

## 4 Resultados

Nesta seção, são apresentados os resultados da modelagem econométrica realizada com o intuito de examinar os determinantes do spread bancário para pessoas físicas no Brasil. Os modelos estimados utilizam a abordagem de Vetores de Correção de Erros (VECM) para capturar tanto os ajustes de longo prazo quanto os efeitos dinâmicos de curto prazo entre as variáveis. Os testes de estacionariedade, cointegração e critérios de seleção de defasagens fundamentaram a especificação do modelo, cujos coeficientes e métricas são analisados detalhadamente nas subseções a seguir.

### 4.1 Modelo Global

O modelo estimado, apresentado após os testes preliminares e as análises conduzidas, reflete a dinâmica geral do sistema. Este modelo global, também chamado de genérico, inclui todas as variáveis e seus respectivos coeficientes, independentemente da significância estatística, com o objetivo de capturar as interações entre as variáveis. A equação a seguir representa este modelo:

$$\begin{aligned}
spread\_pf.d = & -0.1050 \cdot ect1 + 1.0103 \cdot ect2 + 0.0040 \cdot ect3 - 0.000018 \cdot ect4 \\
& + 0.0946 \cdot spread\_pf.dl1 + 2.7284 \cdot selic.dl1 \\
& + 0.0018 \cdot inflacao.dl1 + 0.0002 \cdot pib.dl1 \\
& + 0.2809 \cdot inadimplencia.dl1 - 0.0216 \cdot inst\_fin.dl1 \\
& - 0.1525 \cdot div\_gov.dl1 - 0.2773 \cdot spread\_pf.dl2 \\
& + 3.9043 \cdot selic.dl2 + 0.0040 \cdot inflacao.dl2 \\
& - 0.0008 \cdot pib.dl2 - 0.0068 \cdot inadimplencia.dl2 \\
& - 0.0075 \cdot inst\_fin.dl2 - 0.0422 \cdot div\_gov.dl2 \\
& + 0.0329 \cdot spread\_pf.dl3 + 2.7666 \cdot selic.dl3 \\
& + 0.0033 \cdot inflacao.dl3 - 0.0001 \cdot pib.dl3 \\
& + 0.3765 \cdot inadimplencia.dl3 - 0.0518 \cdot inst\_fin.dl3 \\
& + 0.0417 \cdot div\_gov.dl3 - 0.1314 \cdot spread\_pf.dl4 \\
& + 1.4274 \cdot selic.dl4 + 0.0095 \cdot inflacao.dl4 \\
& - 0.0006 \cdot pib.dl4 - 1.3576 \cdot inadimplencia.dl4 \\
& - 0.0321 \cdot inst\_fin.dl4 + 0.1274 \cdot div\_gov.dl4 \\
& - 0.0194 \cdot spread\_pf.dl5 + 1.3373 \cdot selic.dl5 \\
& + 0.0094 \cdot inflacao.dl5 - 0.0007 \cdot pib.dl5 \\
& + 1.6178 \cdot inadimplencia.dl5 + 0.0019 \cdot inst\_fin.dl5 \\
& - 0.0880 \cdot div\_gov.dl5 + 0.0554 \cdot spread\_pf.dl6 \\
& + 1.1091 \cdot selic.dl6 + 0.0051 \cdot inflacao.dl6 \\
& - 0.0004 \cdot pib.dl6 + 0.8361 \cdot inadimplencia.dl6 \\
& + 0.0005 \cdot inst\_fin.dl6 + 0.1749 \cdot div\_gov.dl6
\end{aligned}$$

## 4.2 Modelo Ajustado

O modelo ajustado considera apenas as variáveis com significância estatística (*p*-valor menor que 0,05). Assim, a equação do modelo ajustado é dada por:

$$\begin{aligned}
spread\_pf.d = & -0.1050 \cdot ect1 + 2.7284 \cdot selic.dl1 \\
& - 0.2773 \cdot spread\_pf.dl2 + 3.9043 \cdot selic.dl2 \\
& - 0.0008 \cdot pib.dl2 - 0.0518 \cdot inst\_fin.dl3 \\
& + 1.6178 \cdot inadimplencia.dl5
\end{aligned} \tag{1}$$

Os resultados indicam que a relação de equilíbrio de longo prazo representada por *ect1* é estatisticamente significativa, o que reforça a presença de cointegração entre as variáveis do sistema e valida o uso do VECM. O coeficiente negativo de *ect1* sugere um mecanismo de correção de erros que reduz os desequilíbrios no spread bancário ao longo do tempo.

No curto prazo, os lags do spread (*spread\_pf.dl2*) apresentaram efeito negativo e estatisticamente significativo, o que indica a presença de certa inércia no comportamento dos spreads. A taxa Selic, por sua vez, demonstrou efeitos positivos tanto nos lags 1 quanto 2,

indicando que aumentos recentes na taxa básica de juros estão associados a elevações nos spreads bancários para pessoas físicas. A variável PIB, embora com coeficiente pequeno, apresentou significância no segundo lag, sugerindo que flutuações da atividade econômica também contribuem — ainda que marginalmente — para a variação do spread. Adicionalmente, os resultados mostram efeitos significativos para a inadimplência (lag 5), com impacto positivo, e para o total de instituições financeiras (lag 3), com efeito negativo, indicando que o aumento da concorrência tende a reduzir o spread.

De forma geral, o modelo ajustado apresentou um  $R^2$  ajustado de 34,52%, com erro padrão residual de 0,5995. Apesar de não explicar toda a variância do spread, os resultados indicam boa adequação e validade estatística, conforme também evidenciado pelo teste F global (p-valor = 0,00011).

### 4.3 Estrutura do Modelo

O termo de correção de erro (ECT1) apresentou coeficiente negativo e estatisticamente significativo ( $-0.1050$ ,  $p < 0.05$ ), indicando que aproximadamente 10,5% dos desvios em relação ao equilíbrio de longo prazo são corrigidos a cada período. Esse resultado evidencia a presença de mecanismos de ajuste dinâmico entre as variáveis cointegradas, o que confirma a adequação do modelo VECM para representar sistemas econômicos interdependentes que retornam ao equilíbrio após choques. Como destacado por Lütkepohl (2005), o termo de correção de erro é fundamental para capturar a força e a direção dos ajustes de curto prazo provocados por desequilíbrios persistentes entre variáveis integradas de ordem um.

No curto prazo, algumas variáveis exibiram efeitos significativos sobre a variação do spread de crédito. A taxa Selic demonstrou impacto positivo estatisticamente robusto, tanto na defasagem de um período ( $2.7284$ ,  $p < 0.05$ ) quanto na defasagem de dois períodos ( $3.9043$ ,  $p < 0.05$ ), sugerindo que elevações na taxa básica de juros tendem a aumentar o spread bancário em períodos subsequentes. Isso corrobora a sensibilidade do spread às mudanças de política monetária, especialmente no curto prazo.

Além disso, a variável *spread\_pf.d* também apresentou efeitos significativos nas defasagens, com coeficiente negativo na segunda defasagem ( $-0.2773$ ,  $p < 0.05$ ). Esse resultado indica que variações passadas no spread contribuem para uma reversão parcial na variação atual, caracterizando um comportamento de autorregulação da série. Esse efeito de memória, ou persistência, reforça o caráter dinâmico do spread, que tende a ajustar-se gradualmente em direção a uma trajetória de equilíbrio.

Outras variáveis também apresentaram significância estatística em defasagens específicas, como o Produto Interno Bruto (*pib.dl2*, coeficiente =  $-0.0008$ ,  $p < 0.05$ ), a inadimplência de pessoas físicas (*inadimplencia.dl5*, coeficiente =  $1.6178$ ,  $p < 0.05$ ) e o número de instituições financeiras (*inst\_fin.dl3*, coeficiente =  $-0.0518$ ,  $p < 0.01$ ). Esses achados sugerem que tanto a atividade econômica quanto o risco de crédito e a estrutura do mercado bancário exercem influência direta e defasada sobre o comportamento do spread.

De forma geral, os resultados destacam a relevância do VECM na modelagem conjunta

de variáveis macroeconômicas com vínculos de longo prazo. O modelo permite isolar os efeitos transientes, capturados pelas defasagens, ao mesmo tempo em que incorpora os ajustes estruturais via os termos de correção de erro, promovendo uma análise mais robusta das relações econômicas dinâmicas.

#### 4.3.1 Estatísticas do Modelo

Conforme apresentado na Tabela 16, o coeficiente de determinação  $R^2 = 0,5734$  (57,34%) indica que aproximadamente 57% da variação na variável dependente ( $\Delta spread_{-pf}$ ) é explicada pelas variáveis independentes do modelo. Esse valor sugere que o modelo captura uma parcela relevante da dinâmica entre as variáveis do sistema, ainda que não integralmente.

O  $R^2$  ajustado, por sua vez, é igual a 0,3452 (34,52%), refletindo a penalização pela inclusão de variáveis irrelevantes. A diferença entre os dois coeficientes sinaliza que, apesar de algumas variáveis contribuírem pouco para a explicação da variação do spread, o modelo ainda preserva coerência e capacidade explicativa.

Tabela 16: Tabela 16: Estatísticas do Modelo VECM

Estatística	Valor
Erro Padrão Residual	0,5995
$R^2$	0,5734
$R^2$ ajustado	0,3452
Estatística F	2,513
<i>p</i> -valor	0,0001136

O Erro Padrão Residual, igual a 0,5995, quantifica o desvio médio entre os valores observados e os valores previstos pelo modelo. Esse valor, relativamente baixo, sugere que o modelo apresenta uma precisão razoável na previsão da variável dependente, embora exista margem para melhorias, sobretudo por meio de possíveis ajustes na especificação.

Já a Estatística F, com valor de 2,513 e associada a um *p*-valor de 0,0001136, indica que o modelo é globalmente significativo ao nível de 1%. Isso significa que podemos rejeitar a hipótese nula de que todos os coeficientes das variáveis explicativas são simultaneamente nulos. Portanto, o conjunto das variáveis independentes exerce um impacto estatisticamente significativo sobre o spread de crédito.

De modo geral, os resultados sugerem que o modelo é estatisticamente válido para descrever a dinâmica das relações no sistema. A diferença entre o  $R^2$  e o  $R^2$  ajustado alerta para a presença de variáveis com baixo poder explicativo, mas a significância global do modelo valida a utilidade da abordagem VECM para fins analíticos.

#### 4.3.2 Principais Variáveis Significativas e Suas Implicações

Dado o conjunto de variáveis significativas identificadas no modelo VECM, é possível avançar para uma interpretação econômica mais profunda dos mecanismos que explicam a evo-

lução do *spread* bancário para pessoas físicas no Brasil. O modelo revelou um padrão de persistência no comportamento do *spread*, sinalizado pela significância das defasagens da própria variável, especialmente a de dois períodos. Tal resultado indica um caráter inercial relevante, em que aumentos passados influenciam diretamente o comportamento atual, o que pode ser explicado pela rigidez das taxas contratuais e pelas políticas internas de precificação das instituições financeiras.

A importância da taxa Selic como principal variável explicativa do *spread* se destaca tanto pela magnitude de seus coeficientes quanto pela significância estatística robusta nas defasagens de um e dois períodos. Esse resultado reforça a centralidade da política monetária na estruturação do custo do crédito, como discutido por Mishkin (1995). Variações na taxa básica são transmitidas com relativa rapidez ao mercado bancário, elevando os custos de captação e, consequentemente, os spreads repassados aos tomadores finais. Tais achados endossam a eficácia do canal de juros em ambientes com alta sensibilidade ao custo de oportunidade dos fundos emprestáveis.

Por outro lado, a inflação apresentou significância marginal em algumas defasagens, com impacto positivo sobre o *spread*. Essa relação pode ser interpretada como um reflexo da transferência de riscos inflacionários para os tomadores, exigindo spreads mais elevados como forma de proteger o poder de compra do retorno real. Além disso, variações de preços podem implicar incerteza sobre o comportamento futuro da taxa de juros, o que tende a ser precificado de forma conservadora pelos bancos.

Entre os determinantes estruturais, o número de instituições financeiras demonstrou efeito negativo e estatisticamente significativo sobre o *spread*, sugerindo que o aumento da concorrência no setor bancário contribui para a compressão dos custos de intermediação financeira. Esse resultado é coerente com o argumento de Demirgüç-Kunt (1998), segundo o qual ambientes regulatórios que favorecem a entrada de novos agentes no mercado reduzem o poder de precificação dos bancos incumbentes e promovem maior eficiência. No contexto brasileiro, marcado por elevada concentração bancária, esse achado evidencia a importância de políticas de fomento à concorrência e de redução de barreiras regulatórias.

A inadimplência das famílias, por sua vez, mostrou impacto positivo e significativo na defasagem de cinco períodos, o que sugere que os bancos respondem a aumentos no risco de crédito com elevações posteriores no *spread*. Essa defasagem temporal pode refletir tanto o atraso na identificação de deteriorações na qualidade da carteira quanto a rigidez institucional para mudanças imediatas nas taxas. Como argumentado por Stiglitz (1981), a presença de assimetrias informacionais e riscos ocultos tende a ser compensada por spreads mais altos, principalmente em contextos onde o sistema de garantias é inefficiente ou o histórico de crédito dos tomadores é limitado.

O PIB, ainda que com efeito estatisticamente significativo, apresentou coeficiente de pequena magnitude. Isso indica que, no curto prazo, a atividade econômica possui influência limitada sobre a determinação dos spreads, o que corrobora a visão de que, em economias emergentes, os fatores financeiros e institucionais têm maior peso na precificação do crédito do que variáveis agregadas Levine (2005).

Essas evidências permitem inferir que a formação do *spread* não é apenas uma função de variáveis macroeconômicas ou de política monetária isoladamente, mas resulta de um sistema interdependente de ajustes defasados, expectativas, riscos e estrutura de mercado. Políticas que visem à redução sustentada do *spread* devem, portanto, atacar múltiplos vetores simultaneamente: promover estabilidade macroeconômica, ampliar a competição bancária, melhorar os mecanismos de avaliação de risco e garantir a previsibilidade da política monetária. Apenas uma abordagem coordenada e multidimensional será capaz de produzir efeitos duradouros sobre os custos do crédito no Brasil.

## 5 Conclusão

O presente estudo investigou os determinantes do *spread* bancário para pessoas físicas no Brasil no período de janeiro de 2013 a julho de 2024, empregando um modelo VECM para capturar as relações dinâmicas de curto e longo prazo entre variáveis monetárias, estruturais e macroeconômicas. A metodologia adotada foi precedida por testes de raiz unitária e cointegração, que evidenciaram a presença de relações estáveis de equilíbrio de longo prazo no sistema.

A significância do termo de correção de erro reforçou a capacidade do modelo em representar ajustes graduais após choques temporários, validando a presença de vínculos estruturais entre as variáveis. Isso indica que o mercado de crédito brasileiro opera em torno de relações de equilíbrio, sendo afetado por perturbações conjunturais que tendem a se dissipar ao longo do tempo. Como argumenta Lütkepohl (2005), modelos VECM são particularmente úteis para descrever a interação de variáveis cointegradas, pois impõem restrições que refletem relações econômicas fundamentais.

Entre os principais achados empíricos, destaca-se o papel central da taxa Selic na formação do *spread*. Sua significância em múltiplas defasagens confirma a efetividade da política monetária como instrumento de gestão do custo do crédito. No entanto, os efeitos defasados e a presença de inércia no próprio *spread* indicam que alterações na Selic não são absorvidas de forma imediata e integral, o que exige prudência na condução das taxas básicas, especialmente em contextos de instabilidade. Estudos recentes, como o de Carvalho (2018), também destacam a importância da política monetária na determinação do *spread* bancário em economias emergentes.

A inadimplência, outro componente essencial, mostrou-se positivamente associada ao *spread*, refletindo a precificação do risco de crédito pelas instituições financeiras. Em contraste, a presença de mais instituições financeiras no mercado apresentou impacto negativo significativo, evidenciando que a ampliação da concorrência contribui de forma efetiva para a redução do custo do crédito. Esse resultado oferece subsídios para políticas que busquem estimular a entrada de novos agentes e reduzir a concentração bancária. De acordo com Beck (2018) Beck et al. (2018), a concorrência bancária é um fator importante na determinação do *spread* bancário, especialmente em economias emergentes.

Embora o PIB e a inflação tenham apresentado significância estatística em alguns lags,

seus efeitos foram relativamente modestos, indicando que, no curto prazo, variáveis de natureza financeira e institucional exercem papel mais decisivo na determinação do *spread*. Tais achados estão em sintonia com a literatura sobre economias emergentes, nas quais a fragilidade das instituições e a concentração bancária intensificam a dependência dos spreads em relação à política monetária e à estrutura do mercado, conforme apontado por Demirgütç-Kunt (1998) Demirgütç-Kunt (1998).

Dessa forma, as evidências empíricas sugerem que estratégias voltadas à redução do *spread* devem ir além da manipulação da taxa de juros, incorporando reformas estruturais no sistema bancário e instrumentos de controle de risco. Em particular, políticas que incentivem maior transparéncia na precificação do crédito, fortaleçam a gestão de risco e ampliem o acesso de novos entrantes ao mercado bancário poderão contribuir para um ambiente de crédito mais eficiente e competitivo.

# Referências

Beck, T., Demirgüç-Kunt, A., and Levine, R. (2018). Bank concentration and crises. *Journal of Financial Economics*, 130(2):267–285.

Carvalho, F. P. (2018). A política monetária e o spread bancário no brasil. *Revista de Economia Contemporânea*, 22(2):257–276.

Dantas, Medeiros, C. (2016). Determinantes do spread bancário ex post no mercado brasileiro. *Revista de Administração Mackenzie*, 17(3):102–133.

de Paula, L. F. (2020). O impacto das taxas de juros no crédito no brasil. *Revista de Economia Política*, 40(2):257–274.

Demirgüç-Kunt (1998). Law, finance, and firm growth. *Journal of Finance*, 53(6):2107–2137.

Garcia, M. N., de Oliveira Neto, O. J., and de Oliveira, R. M. (2019). Influência de variáveis socioeconômicas no spread bancário brasileiro. *Revista Global Manager*, 18(1):1–15.

Goldfajn, I. (2017). A política monetária e a concorrência no setor bancário. *Revista do Banco Central*, 25(1):1–15.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Levine, R. (2005). *Finance and growth: Theory and evidence*. Handbook of economic growth, Elsevier.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer-Verlag Berlin Heidelberg.

Mishkin, F. S. (1995). Symposium on the monetary transmission mechanism. *Journal of Economic Perspectives*, 9(4):3–10.

Oliveira, A. L. (2020). A concorrência no setor bancário e a redução dos spreads no brasil. *Revista de Economia e Finanças*, 15(3):341–356.

Pearson, K. (1895). Notes on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58:240–242.

Stiglitz, J. (1981). Credit rationing in markets with imperfect information. *The American Economic Review*, 71(3):393–410.

Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between time-series?—a study in sampling and the nature of time-series. *Journal of the Royal Statistical Society*, 89(1):1–63.

# Determinantes da Formação Bruta de Capital Fixo no Brasil (1999-2023): Uma Análise com Modelo VECM

Marianna Baia de Andrade\*

Matheus de Lucas Ribeiro Mota†

Fernando Antonio Lucena Aiube ‡

Edison Americo Huarsaya Tito §

16 de Junho de 2025

## Resumo

Este artigo investiga os determinantes da Formação Bruta de Capital Fixo (FBCF) no Brasil entre 1999 e 2023, por meio da aplicação do Modelo Vetorial de Correção de Erros (VECM). As variáveis explicativas analisadas incluem o Produto Interno Bruto (PIB), a Taxa Selic, os Depósitos de Poupança (DP), o Índice de Confiança do Empresário Industrial (ICEI) e a Utilização da Capacidade Instalada. Os resultados indicam que, no curto prazo, PIB, DP, ICEI e a própria FBCF exercem influência positiva sobre o investimento, enquanto a Capacidade Instalada apresenta efeito negativo. A Taxa Selic mostrou-se estatisticamente pouco relevante no período analisado. No longo prazo, foram identificadas relações de cointegração entre as variáveis, confirmado a existência de equilíbrios econômicos estáveis. As funções de impulso-resposta reforçam os resultados obtidos, evidenciando os efeitos dinâmicos dos choques nas variáveis sobre a FBCF. Os resultados contribuem para o entendimento da dinâmica dos investimentos produtivos no Brasil e fornecem subsídios para políticas voltadas à sua promoção.

**Palavras-chave:** Formação Bruta de Capital Fixo, Modelo Vetorial de Correção de Erros, Taxa Selic, PIB, Investimento, Economia Brasileira.

**Código JEL:** C32, E22, E27

## 1 Introdução

A Formação Bruta de Capital Fixo (FBCF) desempenha um papel central no crescimento econômico de longo prazo, ao representar os investimentos em ativos fixos necessários à

\*Mestranda do Programa de Pós-Graduação em Ciências Econômicas da Universidade do Estado do Rio de Janeiro (PPGCE-UERJ)

†Mestrando do Programa de Pós-Graduação em Ciências Econômicas da Universidade do Estado do Rio de Janeiro (PPGCE-UERJ)

‡Professor Associado da Faculdade de Ciências Econômicas - UERJ

§Professor Adjunto do Instituto de Matemática e Estatística - UERJ

expansão da capacidade produtiva de uma economia. No Brasil, os movimentos da FBCF refletem fortemente as condições macroeconômicas, os níveis de confiança empresarial e a disponibilidade de recursos financeiros, fatores que se tornam ainda mais relevantes em períodos de instabilidade.

A crise cambial enfrentada pelo Brasil recentemente, marcada pela saída significativa de Investimento Estrangeiro Direto (IED), intensificou questionamentos sobre o futuro do investimento doméstico e seus efeitos sobre a produtividade e o crescimento. Em um cenário de elevada volatilidade, compreender os determinantes da FBCF torna-se essencial para a formulação de políticas econômicas que favoreçam a retomada da atividade produtiva.

A literatura econômica destaca a importância do investimento para a geração de renda e emprego (Meurer, 2018) [9] e aponta sua sensibilidade às condições monetárias, fiscais e institucionais (Da Rocha Gonçalves et al., 2019) [3]. Experiências históricas, como a década de 1980 – caracterizada por hiperinflação e baixo crescimento – e a década de 1990 – marcada pela estabilização proporcionada pelo Plano Real – ilustram como mudanças nos fundamentos econômicos impactam diretamente os níveis de investimento.

Diante desse contexto, o presente estudo tem como objetivo analisar os principais fatores que influenciam a Formação Bruta de Capital Fixo no Brasil, utilizando o Modelo Vetorial de Correção de Erros (VECM). Busca-se mensurar o impacto de variáveis macroeconômicas, como o Produto Interno Bruto (PIB), a Taxa Selic, o Índice de Confiança do Empresário Industrial (ICEI), os Depósitos de Poupança e a Utilização da Capacidade Instalada, sobre a dinâmica da FBCF no curto e longo prazo.

Este artigo está organizado em sete seções: além desta introdução, a Seção 2 revisa a literatura relevante; a Seção 3 descreve a metodologia empregada; a Seção 4 apresenta os dados utilizados; a Seção 5 explora as relações lineares entre as variáveis; a Seção 6 discute os resultados econométricos; e, finalmente, a Seção 7 sintetiza as conclusões e propõe direções para futuras pesquisas.

## 2 Revisão da Literatura

Os modelos econômicos de investimento são fundamentais para debater o funcionamento da dinâmica econômica. Dentre diversos arcabouços que surgiram ao longo da história econômica, faz-se necessário evocar os conceitos propostos pelo modelo keynesiano, o modelo acelerador e o modelo neoclássico, cujo direcionamento teórico aborda a decisão de investir sob diferentes perspectivas. Neste sentido, esta seção fará um resgate sintético dessas teorias, relacionando com o desenvolvimento de estudos empíricos que basearam-se nesse referencial teórico, e que também serviram como objeto de inspiração na composição deste presente trabalho.

Primeiramente, apresenta-se o modelo keynesiano. Este foi pioneiro ao trazer o princípio do investimento como uma variável endógena, sendo esta a motivação para a criação de diversas teorias subsequentes. Para John Maynard Keynes, a decisão de investir está

intrinsecamente ligada às expectativas de retorno, especialmente à eficiência marginal do capital, definida como as previsões dos empresários sobre os fluxos futuros de rendimento de um ativo, descontados pelo seu custo de aquisição. Além disso, fatores como a taxa de juros, os preços de bens de capital e o custo de oportunidade desempenham papel fundamental nas decisões de investimento. Keynes enfatiza o caráter monetário da taxa de juros, que depende da preferência pela liquidez e da oferta de moeda regulamentada pelas autoridades monetárias. Outro ponto central de sua teoria é o caráter intertemporal do investimento, que implica que o capital seja comprometido antes que os retornos sejam realizados, o que torna a formação de expectativas um elemento crucial. Isso confere ao investimento uma natureza volátil e altamente sensível às condições econômicas esperadas, conforme Ribeiro (2021) [11].

Por outro lado, o modelo acelerador, proposto por Clark (1917) [2] e amplamente difundido nas décadas de 1950 e 1960, apresenta uma abordagem mais simples, postulando uma relação linear e proporcional entre investimento e produto. Nesse modelo, o nível de investimento varia diretamente com o nível de produção, mantendo uma relação constante entre ambos. Apesar de sua simplicidade, o modelo negligencia aspectos importantes, como a autocorrelação do investimento e a influência do período anterior sobre as decisões presentes. Para enfrentar essas limitações, surgiu o modelo acelerador flexível, que incorpora uma variável de defasagem para capturar a velocidade de ajuste do investimento entre períodos. Contudo, mesmo essa evolução deixa de lado fatores relevantes, como custos de capital, rentabilidade e as expectativas dos tomadores de decisão.

Já o modelo neoclássico, proposto por Jorgenson (1963) [6], combina elementos da teoria keynesiana e do modelo acelerador, mas com maior foco nos custos e incentivos associados ao investimento. Ele sugere que o capital investido depende tanto do nível de produção quanto do custo de utilização do capital, sendo influenciado por variáveis como a taxa de juros, a depreciação, os preços de bens de capital, os impostos e os subsídios ao investimento. Chirinko (1993) [?], formalizou esse modelo em uma função que relaciona o investimento total a esses fatores, incluindo parâmetros como elasticidade de substituição entre capital e trabalho, defasagens na entrega de equipamentos e erros aleatórios. Apesar de sua abrangência, o modelo neoclássico se distancia das expectativas subjetivas defendidas por Keynes, priorizando fatores objetivos que moldam as decisões de investimento.

Esses modelos, ao longo de suas evoluções, revelam a complexidade da dinâmica do investimento, integrando aspectos de expectativas, custos, incentivos e ajustes intertemporais. Juntos, oferecem uma base teórica robusta para analisar como os agentes econômicos respondem às variáveis macroeconômicas e institucionais que influenciam suas decisões de investimento. A partir desse arcabouço conceitual, foram desenvolvidos inúmeros trabalhos empíricos que buscaram validar essas teorias e explorar suas implicações práticas em diferentes contextos econômicos, contribuindo para um entendimento mais amplo e aplicado das dinâmicas de investimento nas economias modernas.

Gonzales et al. (2014)[4], utilizando-se do modelo vetorial de correção de erros (VECM), objetivou analisar as principais variáveis determinantes do investimento no Brasil, com dados trimestrais compreendidos entre o período de 1995 a 2013. Obtiveram a seguinte

constatação, há uma relação positiva e significativa entre PIB, taxa de câmbio, demanda agregada e infraestrutura com o investimento. Por outro lado, observou-se uma relação negativa entre o aumento da taxa de juros e o nível de investimento.

Da Rocha Gonçalves et al. (2019) [3], analisaram os determinantes macroeconômicos do investimento no Brasil entre os anos 1996 e 2014, utilizando a Formação Bruta de Capital Fixo como proxy para investimento, sendo utilizada a metodologia vetorial de correção de erros no modelo. Com os resultados e constatações do estudo, foram estruturadas funções de impulso-resposta para analisar no longo prazo, como um choque em determinadas variáveis impactaria o comportamento de outras séries dentro de um dado período de tempo. Além disso, foi empregado o teste de Causalidade de Granger para investigar a influência que uma variável pode ter sobre a outra. As variáveis analisadas foram poupança interna, poupança externa, crescimento do PIB, inflação e taxa de juros. Os resultados indicaram causalidade unidirecional entre o investimento e as variáveis PIB, poupança interna, inflação e taxa de juros. A formação bruta de capital fixo apresentou maior elasticidade em relação à poupança interna, mas também foi sensível aos choques na taxa de juros e inflação.

Lélis, Bredow e Cunha (2015) [8], aplicaram o modelo vetorial de correção de erros para analisar as variáveis macroeconômicas que determinaram o investimento privado no Brasil entre 1996 e 2012. Os resultados indicaram que variáveis como o consumo das famílias e a utilização da capacidade instalada influenciam diretamente a formação bruta de capital. No entanto, o crédito total da economia foi identificado como a variável mais significativa para o aumento do investimento. Por outro lado, variáveis como o preço de máquinas e equipamentos, taxa de juros e custo do crédito, a longo prazo, atuam como condicionantes do investimento.

Meurer (2018) [9], analisou os determinantes da Formação Bruta de Capital Fixo no Brasil de 1996 a 2016, utilizando séries temporais e mínimos quadrados ordinários. No curto prazo, apenas o PIB foi significativo, enquanto no longo prazo, o PIB, a Utilização da Capacidade Instalada e as Operações de Crédito impactaram positivamente. Ainda assim, a inflação e taxa de juros tiveram um impacto negativo, sendo apenas a inflação significativa. O modelo concluiu que o PIB, a Utilização da Capacidade Instalada e as Operações de Crédito impactam positivamente a Formação Bruta de Capital Fixo (FBCF), enquanto a inflação exerce um efeito negativo. A taxa de juros real não apresentou significância estatística.

Em Ribeiro (2021) [11], o autor buscou investigar a determinação da Formação Bruta de Capital Fixo (FBCF) no Brasil entre 1996 e 2020, utilizando um modelo vetorial de correção de erros (VECM). Os resultados destacaram a importância do PIB e seu efeito cíclico sobre a FBCF, além da relação negativa e estatisticamente significativa entre FBCF e as variáveis Juros e Câmbio, refletindo os custos de investimento. O modelo mostrou relevância estatística para o coeficiente do próprio FBCF, indicando autocorrelação positiva, enquanto os coeficientes do PIB e do IPCA não apresentaram significância estatística, sugerindo a necessidade de explorar lags maiores na análise do impulso-resposta.

Apesar de distintas em suas formulações, as abordagens keynesiana, aceleradora e neo-

clássica reconhecem a relevância de fatores macroeconômicos na decisão de investir. A modelagem com VECM permite incorporar simultaneamente elementos dessas correntes, captando tanto o ajuste dinâmico quanto os equilíbrios de longo prazo.

### 3 Metodologia

Este estudo emprega o modelo vetorial de correção de erros para analisar as relações entre a Formação Bruta de Capital Fixo e variáveis econômicas determinantes, incluindo a Taxa Selic, o volume de Depósitos de Poupança, o Produto Interno Bruto (PIB), o Índice de Confiança do Empresário Industrial (ICEI) e a Utilização da Capacidade Instalada. Os dados utilizados foram transformados em trimestres e logaritmados, abrangendo o período do primeiro trimestre de 1999 ao último trimestre de 2023, totalizando 100 observações. A delimitação do período de análise até 2023 deve-se à indisponibilidade de dados completos para algumas variáveis no ano de 2024, garantindo assim a consistência e a robustez das séries temporais analisadas. Dessa forma, nesta seção serão abordados os testes referentes às propriedades necessárias para séries temporais, a metodologia utilizada e o modelo econométrico proposto.

#### 3.1 Raiz unitária e Teste de Dickey-Fuller Aumentado

Em 1979, Dickey e Fuller desenvolveram o teste chamado Dickey-Fuller Aumentado (ADF). A construção do teste visa avaliar a estacionariedade de uma série temporal, ou seja, verificar se a série possui uma raiz unitária. Anteriormente a este teste, foi proposto o modelo original (apenas "Dickey-Fuller"), entretanto, após problemas devido a inclusão de defasagens da variável dependente, foi introduzido o modelo aumentado, permitindo assim que fossem corrigidos os problemas de correlação serial nos erros da regressão conforme Bueno (2018) [1].

Segundo Gujarati [5], o teste ADF é formulado como:

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^m \alpha_i \Delta Y_{t-i} + \epsilon_t \quad (1)$$

No qual  $\Delta Y_t$  representa a primeira diferença da série,  $\beta_1$  e  $\beta_2$  são coeficientes associados a um intercepto e uma tendência, respectivamente,  $\delta$  é o coeficiente a ser testado, e  $\epsilon_t$  é um termo de erro ruído branco.

As hipóteses do teste ADF são definidas como:

- **Hipótese Nula ( $H_0$ ):**  $\delta = 0$ , o que indica que a série possui uma raiz unitária e é não estacionária.
- **Hipótese Alternativa ( $H_1$ ):**  $\delta < 0$ , o que sugere que a série é estacionária.

Caso o valor do teste ADF indique que a hipótese nula deve ser rejeitada, isso significa que a série é estacionária. Se a hipótese nula não for rejeitada, será necessário realizar transformações adicionais na série, sendo comum a diferenciação, para garantir a estacionariedade e adequação do modelo.

### 3.2 Teste de Johansen

Conforme proposto por Bueno (2018) [1], o Teste de Cointegração de Johansen tem como finalidade verificar a estacionariedade de um conjunto de variáveis endógenas. A metodologia do teste se concentra em determinar o posto da matriz  $\Phi$  ( $n \times n$ ), onde a presença de cointegração é indicada quando  $r < n$ .

A concepção de  $\Phi = 0$  é análoga à ideia de raiz unitária em modelos univariados; no entanto, no contexto do Teste de Cointegração de Johansen, uma matriz  $\Phi = 0$  sugere uma raiz unitária multivariada. Na análise, o determinante da matriz é obtido através do produto de seus autovalores. Para que o posto de  $\Phi$  esteja dentro dos limites entre 0 e  $n$ , é necessário que existam  $r$  autovalores distintos de zero e  $n - r$  autovalores iguais a zero conforme Bueno (2018) [1].

O Teste de Cointegração de Johansen oferece dois formatos estatísticos para calcular os autovalores de  $\Phi$ . O primeiro é conhecido como Teste do Traço, expresso pela equação abaixo:

$$\lambda_{tr}(r) = -T \sum_{i=r+1}^n \ln(1 - \hat{\lambda}_i) \quad (2)$$

Neste caso, a hipótese nula postula a presença de  $\hat{r}$  vetores de cointegração ( $r = \hat{r}$ ), enquanto a hipótese alternativa sugere que  $r > \hat{r}$ . Na ausência de cointegração, os autovalores obtidos tendem a ser próximos de zero, sinalizando a não estacionariedade e a instabilidade da matriz  $\Phi$ , impossibilitando a rejeição da hipótese nula. Se a hipótese nula de que  $r = \hat{r}$  for rejeitada, isso indica a possibilidade de múltiplos vetores de cointegração. Sendo este teste o escolhido neste estudo.

A segunda abordagem é através do Teste de Máximo Autovalor, cuja estatística é dada pela equação:

$$LR(r) = -T \ln(1 - \hat{\lambda}_{r+1}) \quad (3)$$

Neste teste, a hipótese nula sugere a existência de  $r = \hat{r}$  vetores de cointegração, enquanto a hipótese alternativa indica que  $\hat{r} + 1 = r$  vetores de cointegração. Assim como no Teste do Traço, a rejeição da hipótese nula implica que há mais de um vetor de cointegração.

Entretanto, para aplicar o Teste de Cointegração, necessita-se não apenas verificar a autocorrelação dos resíduos, mas também analisar o menor valor obtido pelos critérios de informação para determinar o número adequado de defasagens em todas as variáveis endógenas conforme Bueno (2018) [1]. Os critérios utilizados incluem Akaike (AIC), Schwarz (SBIC) e Hannan-Quinn (HQIC), expostos através das equações a seguir:

$$AIC(m) = \ln |T_0(m)| + \frac{2Tm}{n^2} \quad (4)$$

$$SBIC(m) = \ln |T_0(m)| + \frac{\ln T}{T} mn^2 \quad (5)$$

$$HQIC(m) = \ln |T_0(m)| + \frac{\ln \ln T}{T} \frac{2m}{n^2} \quad (6)$$

onde  $mn^2$  representa a totalidade de parâmetros estimados nas equações.

A principal usabilidade do Teste de Cointegração seria para determinar a existência de um equilíbrio de longo prazo entre as variáveis em modelos de séries temporais, algo importante para a escolha do modelo econômétrico mais adequado para alcançar os objetivos do estudo.

Destaca-se que se tivermos  $n$  variáveis I(1) que são modeladas conjuntamente em um sistema dinâmico, pode haver até  $n - 1$  relações de cointegração ligando-as. Cada relação de cointegração pode ser vista como uma tendência comum que conecta algumas ou todas as séries no sistema. Nós consideraremos "relação de cointegração" e "tendência comum" como sinônimos. O posto de cointegração do sistema é o número de tais tendências comuns, ou o número de relações de cointegração conforme Judge (1991)[7].

Para determinar o posto de cointegração  $r$ , realizamos uma sequência de testes. Primeiro, testamos a hipótese nula de  $r = 0$  contra  $r \geq 1$  para determinar se existe pelo menos uma relação de cointegração. Se não rejeitarmos  $r = 0$ , concluímos que não existem relações de cointegração ou tendências comuns entre as séries. Neste caso, não precisamos de um modelo VEC e podemos simplesmente usar um VAR nas diferenças das séries.

Se rejeitarmos  $r = 0$  na etapa inicial, então pelo menos algumas das séries são cointegradas, e queremos determinar o número de relações de cointegração. Procedemos para uma segunda etapa para testar a hipótese nula de  $r \leq 1$  contra  $r \geq 2$ . Se não pudermos rejeitar a hipótese de que não existem mais do que uma tendência comum, então estimamos um sistema VEC com uma relação de cointegração.

Se rejeitarmos a hipótese de que  $r \leq 1$ , então avançamos para testar  $r \leq 2$  contra  $r \geq 3$ , e assim por diante. Escolhemos  $r$  como o menor valor para o qual não rejeitamos a hipótese nula de que não existem relações de cointegração adicionais.

### 3.3 Modelo Vetorial de Correção de Erros (VECM)

O Modelo Vetorial de Correção de Erros (VECM) é uma extensão do modelo VAR, que integra tanto a dinâmica de curto quanto de longo prazo entre as variáveis endógenas. De acordo com Bueno (2018) [1], o VECM é particularmente útil quando há cointegração entre as variáveis, pois captura as relações de equilíbrio de longo prazo, ajustando as flutuações.

A denominação "*Modelo Vetorial de Correção de Erros*" deriva da Teoria de Cointegração, que foca no teste de estacionariedade dos resíduos  $u_t$ . Quando os resíduos se mostram estacionários, a cointegração é confirmada. A equação do VECM, conforme Judge (1991)[7], inclui tanto os fatores de curto prazo, representados pela soma das diferenças defasadas  $\sum_{i=1}^{p-1} \Lambda_i \Delta X_{t-i}$ , quanto a relação de longo prazo, dada por  $\Phi X_{t-1}$ , que descreve o equilíbrio entre as variáveis.

Em um VECM (Modelo Vetorial de Correção de Erros) com  $n$  variáveis cointegradas, a estrutura geral das equações é expandida para acomodar múltiplas variáveis, mantendo a premissa central de que as variáveis são cointegradas, ou seja, apresentam uma relação de longo prazo que pode ser modelada por meio de um vetor de cointegração. Considerando que as séries temporais  $y_1, y_2, \dots, y_n$  são de ordem  $I(1)$  (ou seja, são não estacionárias, mas se tornam estacionárias após a diferenciação), o modelo assume que essas variáveis são cointegradas, ou seja, existe uma combinação linear dessas variáveis que é estacionária, ou seja,  $\alpha_0 + \sum_{i=1}^n \alpha_i y_i$  é uma série  $I(0)$ , indicando que há uma relação de equilíbrio de longo prazo entre elas, Judge (1991)[7].

O modelo para  $n$  variáveis pode ser representado pelas seguintes equações de correção de erro:

$$\Delta y_{1,t} = \beta_{10} + \sum_{i=1}^n \beta_{1i} \Delta y_{i,t-1} + \lambda_1 \left( y_{1,t-1} - \alpha_0 - \sum_{i=1}^n \alpha_i y_{i,t-1} \right) + \epsilon_{1,t} \quad (7)$$

$$\Delta y_{2,t} = \beta_{20} + \sum_{i=1}^n \beta_{2i} \Delta y_{i,t-1} + \lambda_2 \left( y_{2,t-1} - \alpha_0 - \sum_{i=1}^n \alpha_i y_{i,t-1} \right) + \epsilon_{2,t} \quad (8)$$

⋮

$$\Delta y_{n,t} = \beta_{n0} + \sum_{i=1}^n \beta_{ni} \Delta y_{i,t-1} + \lambda_n \left( y_{n,t-1} - \alpha_0 - \sum_{i=1}^n \alpha_i y_{i,t-1} \right) + \epsilon_{n,t} \quad (9)$$

Destaca-se que os coeficientes de correção de erro  $\lambda_1, \lambda_2, \dots, \lambda_n$  são importantes para entender como as variáveis ajustam-se a desvios do equilíbrio de longo prazo. Quando uma variável  $y_i$  está acima de seu valor de equilíbrio de longo prazo (ou seja, a relação  $\sum_{i=1}^n \alpha_i y_i$  está acima de  $\alpha_0$ ), o coeficiente de correção de erro  $\lambda_i$  será negativo, forçando  $y_i$  a diminuir para retornar ao seu valor de equilíbrio. Caso contrário, se a variável estiver abaixo de seu valor de longo prazo, o coeficiente de correção de erro será positivo, levando a variável a aumentar em direção ao equilíbrio conforme Judge (1991)[7].

Em um sistema com várias variáveis, as interações entre elas influenciam a dinâmica de curto e longo prazo, e os coeficientes de correção de erro fornecem uma medida da velocidade e da direção desse ajuste. A presença de múltiplas variáveis cointegradas implica que os coeficientes de correção de erro  $\lambda_1, \lambda_2, \dots, \lambda_n$  podem variar de acordo com a magnitude dos desvios de cada variável em relação ao seu equilíbrio de longo prazo.

Em um modelo com  $n$  variáveis, as estimativas dos coeficientes  $\beta_{ij}$  e  $\lambda_i$  são realizadas simultaneamente, levando em consideração as relações de longo prazo entre as variáveis. Além disso, é necessário realizar testes de cointegração, como o teste de Johansen, para determinar o número de relações de cointegração presentes no sistema, ou seja, quantos vetores de cointegração existem. Esses testes ajudam a estabelecer as dependências de longo prazo entre as variáveis e a identificar o número de relações de cointegração que devem ser incorporadas ao modelo.

A equação a seguir descreve o modelo vetorial de correção de erros para a Formação Bruta de Capital Fixo (FBCF) em relação às variáveis explicativas selecionadas:

$$\Delta FBCF_t = \alpha_0 + \sum_{i=1}^n \beta_i \Delta x_{i,t-1} + \lambda \left( FBCF_{t-1} - \gamma_0 - \sum_{i=1}^n \gamma_i x_{i,t-1} \right) + \varepsilon_t \quad (10)$$

Onde:

- $\Delta FBCF_t$ : Variação da Formação Bruta de Capital Fixo no período  $t$ , representando mudanças no nível de investimentos produtivos.
- $x_{i,t-1}$ : Variáveis explicativas com defasagens:
  - PIB $_{t-1}$ : Produto Interno Bruto no período  $t - 1$ , usado como proxy do nível de atividade econômica.
  - Selic $_{t-1}$ : Taxa básica de juros no período  $t - 1$ , que influencia o custo de financiamento.
  - ICEI $_{t-1}$ : Índice de Confiança do Empresário Industrial no período  $t - 1$ , indicando o otimismo industrial.
  - DP $_{t-1}$ : Depósitos de Poupança no período  $t - 1$ , representando a oferta de recursos financeiros.
- $\alpha_0$ : Intercepto do modelo, que captura o valor médio da variação de FBCF $_t$  na ausência de mudanças nas variáveis explicativas.
- $\beta_i$ : Coeficientes das diferenças defasadas das variáveis explicativas, que medem os impactos de curto prazo.
- $\lambda$ : Coeficiente de ajuste de longo prazo, indicando a velocidade com que a FBCF $_t$  retorna ao equilíbrio de longo prazo após um choque.
- $\gamma_0$  e  $\gamma_i$ : Parâmetros do equilíbrio de longo prazo para o nível das variáveis explicativas.
- $\varepsilon_t$ : Termo de erro, que representa os fatores não explicados pelo modelo.

A equação captura tanto os efeitos de curto prazo ( $\beta_i$ ) quanto o mecanismo de correção de longo prazo ( $\lambda$ ) entre a Formação Bruta de Capital Fixo e as variáveis explicativas selecionadas.

A escolha do VECM justifica-se pela natureza das séries temporais utilizadas neste estudo, as quais, conforme indicado pelos testes de raiz unitária (ADF), são integradas de ordem um, I(1), e apresentaram evidências de cointegração com base no teste de Johansen.

O VECM é a modelagem adequada para esse tipo de estrutura de dados, pois permite capturar tanto a dinâmica de curto prazo quanto as relações de equilíbrio de longo prazo entre as variáveis macroeconômicas analisadas.

Diferentemente do modelo VAR em primeiras diferenças, que ignora as relações de cointegração, o VECM incorpora termos de correção de erro que refletem desvios do equilíbrio de longo prazo, promovendo um ajuste mais preciso e teoricamente consistente.

Além disso, o modelo é amplamente utilizado na literatura empírica brasileira sobre investimento (ver Gonzales et al., 2014 [4]; Da Rocha Gonçalves et al., 2019 [3]; Ribeiro, 2021 [11]), o que reforça sua adequação ao objetivo desta pesquisa: investigar os determinantes macroeconômicos da Formação Bruta de Capital Fixo (FBCF) no Brasil entre 1999 e 2023.

## 4 Descrição das Variáveis e Fontes de Dados

A seleção das variáveis fundamenta-se nos estudos abordados na seção de Revisão da Literatura. Para a análise dos determinantes da Formação Bruta de Capital Fixo (FBCF) no Brasil, o modelo econômétrico emprega as seguintes variáveis: Formação Bruta de Capital Fixo (FBCF), que representa o investimento em ativos fixos como máquinas, equipamentos e construções; Produto Interno Bruto (PIB), um indicador macroeconômico que mede a soma de todos os bens e serviços finais produzidos em um país, refletindo a atividade econômica geral; Taxa Selic (Sistema Especial de Liquidação e Custódia), a taxa básica de juros da economia brasileira, crucial para a política monetária e o custo de captação de recursos para investimentos; Depósito de Poupança (DP), que reflete a disponibilidade de recursos para financiamento da economia e pode impactar a liquidez do mercado; Utilização da Capacidade Instalada (UCI), um indicador que mede o percentual da capacidade produtiva utilizada pelas empresas, onde uma alta UCI pode sinalizar a necessidade de novos investimentos; e o Índice de Confiança do Empresário Industrial (ICEI), que mede as expectativas dos empresários da indústria em relação ao futuro da economia e de suas próprias empresas, influenciando decisões de investimento.

As fontes de dados para FBCF e Depósito de Poupança (DP) foram o Sistema IBGE de Recuperação Automática (SIDRA) do Instituto Brasileiro de Geografia e Estatística (IBGE). PIB, a Taxa Selic e a Utilização da Capacidade Instalada (UCI) foram obtidos do Sistema Gerenciador de Séries Temporais (SGS) do Banco Central do Brasil. Por fim, o Índice de Confiança do Empresário Industrial (ICEI) foi fornecido pela Confederação Nacional da Indústria (CNI). O período analisado compreende o primeiro trimestre de 1999 (1999.T1) até o último trimestre de 2023 (2023.T4), totalizando 100 observações. Embora o PIB e os Depósitos de Poupança sejam originalmente disponibilizados em periodicidade trimestral, as demais séries foram ajustadas para a periodicidade trimestral para garantir a compatibilidade e a consistência do conjunto de dados.

Na tabela 1, são apresentadas as estatísticas descritivas para as variáveis analisadas, oferecendo um panorama inicial de observação das características e comportamentos ao longo do período. Com base nos resultados, destacam-se as seguintes observações: O PIB e FBCF apresentam médias positivas, o que é esperado e reflete a tendência de crescimento econômico e de investimento no Brasil ao longo do período analisado. No entanto, a FBCF exibe um desvio padrão significativamente maior (0,090) em comparação com o PIB (0,033). Essa maior variabilidade na FBCF é um indicativo importante de que os investimentos produtivos são mais voláteis e sensíveis a choques econômicos e de confiança, o que justifica a

investigação de seus determinantes.

A Taxa Selic destaca-se pela elevada curtose (6,283) e assimetria positiva (1,006). Uma curtose superior a 3 indica a presença de valores extremos (outliers), o que é consistente com períodos de política monetária restritiva com taxas de juros muito elevadas (como em 2002-2003 e 2015-2016). A assimetria positiva sugere que a distribuição da Selic tem uma "cauda" mais longa para valores altos, refletindo a prevalência de períodos de juros mais elevados em relação à média.

A Utilização da Capacidade Instalada (UCI), com uma curtose elevada (16,534), indica uma concentração de valores próximos à média, mas com a presença de eventos raros de utilização muito alta ou muito baixa, sugerindo flutuações mais acentuadas em momentos específicos da atividade industrial. A média próxima de zero indica que a variável pode ter sido tratada como taxa de variação ou desvio de um valor de referência.

O Índice de Confiança do Empresário Industrial (ICEI) apresenta o menor desvio padrão (0,153) entre as variáveis mais voláteis (FBCF, Selic, UCI), sugerindo uma menor volatilidade nas expectativas dos empresários em comparação com as flutuações de outras variáveis econômicas. Sua curtose elevada (5,502) ainda aponta para a presença de alguns períodos de otimismo ou pessimismo extremos.

O Depósito de Poupança (DP) tem uma média positiva (0,090). Sua curtose (1,948) inferior a 3 e assimetria levemente negativa (-0,077) sugerem uma distribuição mais próxima da normalidade em comparação com as outras variáveis, mas com uma pequena concentração de valores ligeiramente acima da média. Os valores extremos (máximo e mínimo) para todas as variáveis, especialmente para a Selic (1,618 e -1,099), FBCF (0,255 e -0,215) e UCI (0,291 e -0,263), reforçam a existência de períodos de grande instabilidade e mudanças significativas na dinâmica econômica brasileira, fator que pode influenciar as decisões por parte dos investidores.

Dessa forma, a análise das estatísticas descritivas indica diferentes níveis de volatilidade e estabilidade entre as variáveis econômicas analisadas. A presença de assimetria e curtose elevadas em algumas séries sugere que a distribuição dos dados pode não ser normal, o que representa um aspecto relevante para a escolha e aplicação dos testes econométricos, tema que será aprofundado nas seções seguintes.

Tabela 1: Resultado das Estatísticas das Variáveis

Estatística	PIB	FBCF	Capacidade	Taxa Selic	ICEI	DP
Média	0,022	0,016	-0,001	-0,029	0,001	0,090
Desvio Padrão	0,033	0,090	0,053	0,422	0,153	0,073
Variância	0,001	0,008	0,003	0,178	0,023	0,005
Mediana	0,023	0,030	0,001	-0,034	-0,005	0,091
Máximo	0,117	0,255	0,291	1,618	0,515	0,225
Mínimo	-0,107	-0,215	-0,263	-1,099	-0,592	-0,048
Curtose	4,987	3,029	16,534	6,283	5,502	1,948
Assimetria	-0,610	-0,120	0,399	1,006	0,147	-0,077

Fonte: elaborado pelos autores.

A seguir, apresentamos os gráficos das séries temporais com a finalidade de observar o formato das variáveis e fornecer uma ideia preliminar sobre seu comportamento, incluindo a identificação de possíveis tendências, sazonalidades e outras características que possam indicar ou sugerir a presença de estacionariedade. No entanto, cabe ressaltar que a confirmação formal da estacionariedade das séries será feita posteriormente, por meio do teste de Dickey-Fuller Aumentado (ADF), o qual permitirá avaliar estatisticamente a presença de raízes unitárias.

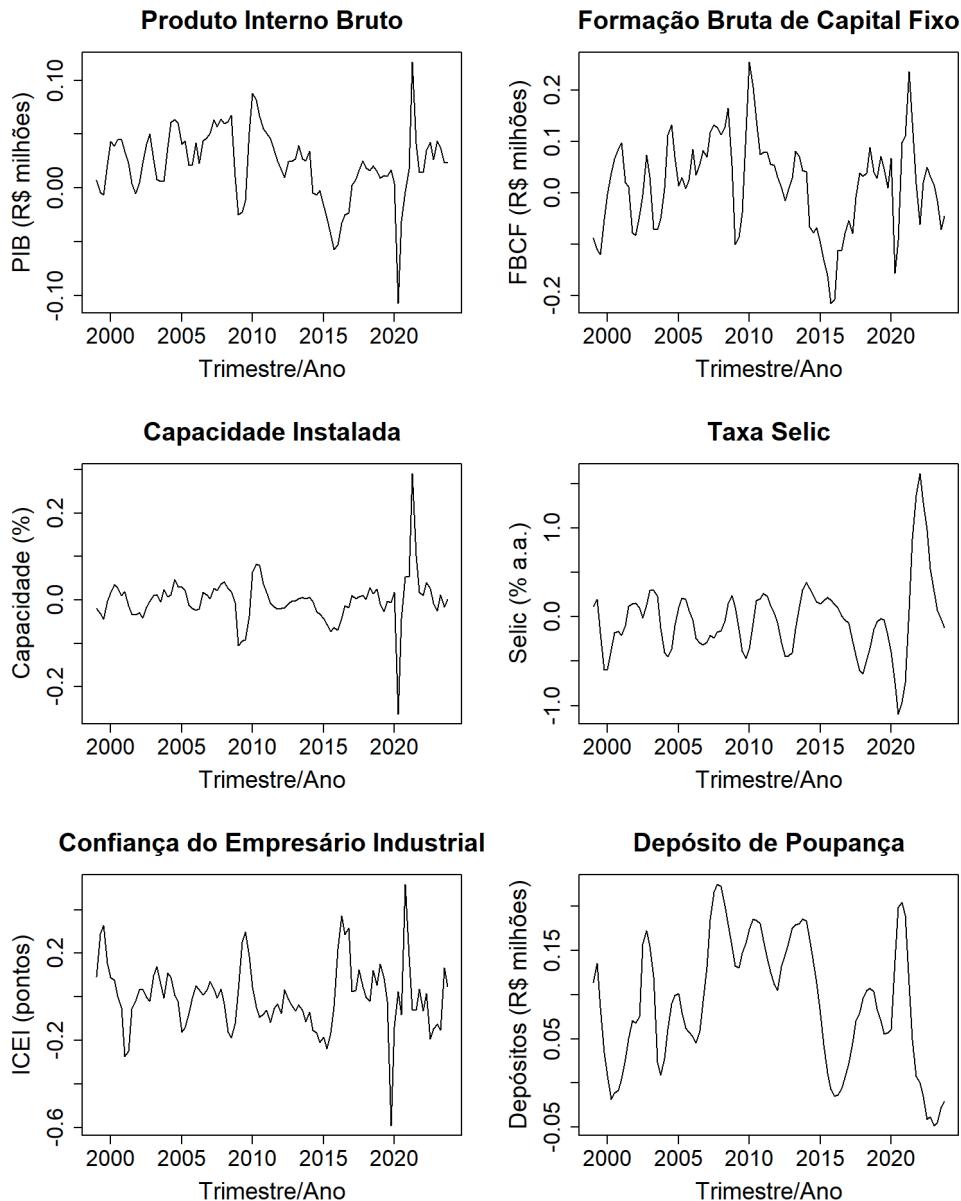
Na figura 1 a plotagem dos gráficos revela padrões distintos para cada série, o que é fundamental para a análise de estacionariedade. Para o Produto Interno Bruto (PIB) e a Formação Bruta de Capital Fixo (FBCF) observa-se tendências de crescimentos ao longo do período analisado, com presença de períodos de flutuações marcadas por desacelerações significativas em períodos de crise econômica, como assinalado nos anos de 2009, 2015-2016 e a acentuada queda em 2020, influenciado pela ocorrência da pandemia do covid-19. Essa tendência de alta e as flutuações acentuadas sugerem fortemente a presença de não estacionariedade nessas séries.

No tocante a representação da Taxa Selic demonstra uma elevada volatilidade, com picos notáveis e patamares elevados entre 2002 e 2005, e novamente em 2015-2016, refletindo períodos de combate à inflação e política monetária mais restritiva. Após 2016, observa-se uma tendência de queda, antes de uma nova elevação no final do período. Essa dinâmica de mudança de regime e picos extremos também sugere um comportamento de não estacionariedade na série temporal.

O gráfico da Capacidade Instalada (UCI) mostra flutuações em torno de uma média, mas com quedas durante os anos de 2009, 2015-2016 e especialmente 2020, indicando desuso da capacidade produtiva. Embora não apresente uma tendência clara de crescimento ou declínio de longo prazo como PIB ou FBCF, a amplitude das oscilações e as possíveis quebras estruturais demandam atenção para a estacionariedade.

O Índice de Confiança do Empresário Industrial (ICEI) e os Depósitos de Poupança (DP), à primeira vista, parecem variar em torno de uma média mais constante, sem uma tendência de longo prazo tão pronunciada quanto as outras variáveis. O ICEI mostra oscilações que refletem os ciclos de otimismo e pessimismo empresarial, com quedas bruscas em momentos de crise. O DP, por sua vez, apresenta um comportamento relativamente mais estável, mas com períodos de crescimento e queda que acompanham a dinâmica econômica geral. Apesar da aparente estacionariedade visual para estas duas últimas, a confirmação estatística com o teste ADF será essencial para todas as séries.

Figura 1: Gráficos das Séries Temporais



A análise visual dos gráficos reforça a suspeita de não estacionariedade para a maioria das séries, especialmente PIB, FBCF, Selic e UCI, enquanto ICEI e DP podem ser estacionárias.

## 5 Análise das Relações Lineares entre as Variáveis do Modelo

O objetivo desta seção é explorar as relações lineares entre as variáveis selecionadas para a análise econômica e avaliar a presença de multicolinearidade. A multicolinearidade refere-se à extensão em que uma variável pode ser explicada pelas outras variáveis na análise. À medida que a multicolinearidade aumenta, fica mais complicada a interpretação da variável estatística, uma vez que se torna mais difícil verificar o efeito de qualquer variável, devido a suas inter-relações,[?]. Portanto, esta análise é crucial para compreender as interdependências entre as variáveis e assegurar a correta especificação e interpretação do Modelo Vetorial de Correção de Erros (VECM).

Na figura 2 apresenta a matriz de correlação entre a Formação Bruta de Capital Fixo (FBCF), a variável dependente, e as principais variáveis explicativas: Produto Interno Bruto (PIB), Utilização da Capacidade Instalada (UCI), Taxa Selic, Índice de Confiança do Empresário Industrial (ICEI) e Depósitos de Poupança (DP). Os coeficientes variam de -1 a 1, representando a força e a direção das associações entre as variáveis,[10].

Sendo assim, a análise da Figura 2 revela importantes relações lineares entre as variáveis. O PIB apresenta uma forte correlação positiva com a Formação Bruta de Capital Fixo (FBCF) (0,86) e com a Utilização da Capacidade Instalada (Capacidade) (0,77), evidenciando que investimentos produtivos e a utilização do estoque de capital já existente são determinantes importantes para o crescimento econômico. Este resultado é altamente consistente com a teoria keynesiana, que postula o investimento como um motor fundamental do crescimento econômico, e com a literatura empírica sobre ciclos econômicos, onde períodos de expansão do PIB são tipicamente acompanhados por um aumento nos investimentos produtivos.

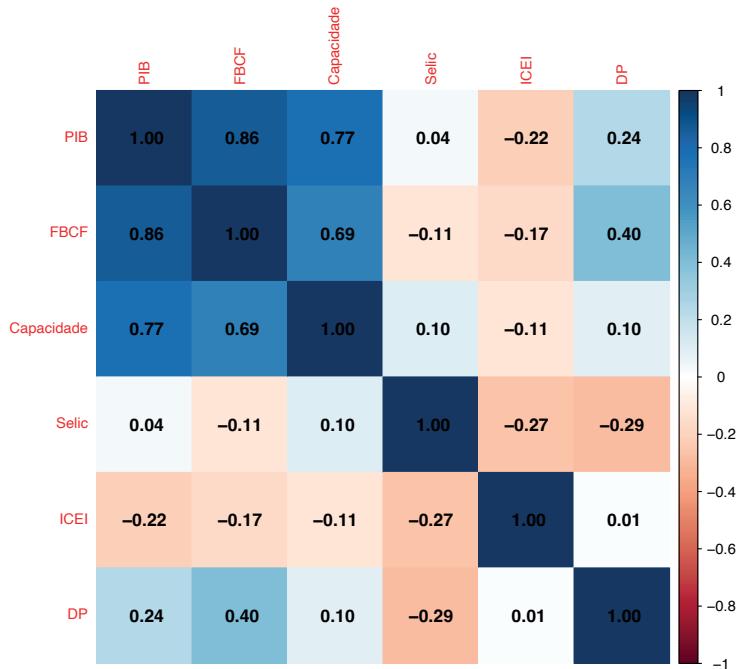
A FBCF contribui para o PIB, e um PIB crescente sinaliza expectativas positivas que estimulam novos investimentos. A FBCF também demonstra uma correlação positiva significativa (0,69) com a Utilização da Capacidade Instalada (UCI). Isso sugere que, à medida que a capacidade produtiva existente se aproxima de seu limite de uso, as empresas tendem a aumentar seus investimentos em ativos fixos para expandir a produção e atender à demanda crescente. Essa relação é esperada, pois a capacidade instalada serve como um sinal para decisões de investimento. Há ainda uma correlação positiva moderada (0,40) entre FBCF e Depósito de Poupança (DP). Embora não seja tão forte quanto a relação com o PIB, indica que uma maior disponibilidade de recursos poupançados na economia pode, em tese, facilitar o financiamento de novos investimentos, ao tornar o capital mais acessível ou de menor custo.

O Índice de Confiança do Empresário Industrial (ICEI) apresenta uma correlação negativa de -0,17 com a FBCF. Apesar de surpreendente à primeira vista, uma interpretação para essa baixa e inversa correlação pode residir na dinâmica da confiança: em um ambiente de alta incerteza, mesmo com confiança flutuante, a materialização do investimento pode depender de fatores mais tangíveis como taxa de juros ou demanda. Alternativamente, essa baixa correlação pode sugerir que a confiança do empresário, isoladamente, pode não

ser o principal driver de investimento no curto prazo no contexto brasileiro analisado, ou que outros fatores macroeconômicos exercem influência mais direta.

Além das relações com a FBCF, é importante notar as correlações entre as próprias variáveis explicativas. O PIB, por exemplo, possui forte correlação positiva com a UCI (0,77), o que é esperado, pois o crescimento econômico impulsiona a utilização da capacidade produtiva. A Taxa Selic apresenta correlações negativas com ICEI (-0,27) e DP (-0,29), indicando que políticas monetárias mais restritivas (Selic alta) podem impactar negativamente a confiança empresarial e a disponibilidade de poupança, o que é consistente com a teoria econômica. A existência dessas correlações entre as variáveis explicativas sinaliza a necessidade de avaliar a multicolinearidade.

Figura 2: Matriz Correlação das variáveis independentes e o FBCF



Além disso, foi realizada uma análise da multicolinearidade entre as variáveis independentes para verificar possíveis redundâncias que poderiam comprometer a precisão das estimativas no modelo econômético. O Índice de Variância Inflacionada (VIF) é calculado a partir da matriz  $R = (X^T X)^{-1}$ , onde  $X$  representa as variáveis independentes centradas e padronizadas. Essa matriz reflete as correlações entre as variáveis, e os elementos da diagonal correspondem ao VIF, indicando o aumento da variância dos coeficientes de regressão devido à multicolinearidade. O VIF de uma variável  $X_i$  é definido como  $VIF_i = \frac{1}{1-R_i^2}$ , onde  $R_i^2$  é o coeficiente de determinação da regressão de  $X_i$  sobre as demais variáveis. Valores elevados de VIF sugerem uma dependência linear forte entre variáveis independentes, podendo comprometer a estabilidade do modelo e as inferências estatísticas, de acordo com Montgomery (2021) [10].

O Índice de Variância Inflacionada (VIF) foi utilizado para identificar essa questão. Os resultados indicaram que todas as variáveis apresentaram valores de VIF abaixo de 5, o que está dentro do limite considerado aceitável pela literatura, indicando ausência de multicolinearidade severa. Os valores estão assinalados na Tabela 2: Ainda assim, é pertinente uma análise individual dos valores: o VIF mais elevado foi observado para o PIB (2,79). Embora bem abaixo do limite crítico, esse valor indica uma colinearidade moderada com outras variáveis explicativas do modelo, principalmente a Utilização da Capacidade Instalada (UCI) (com a qual o PIB tem uma forte correlação de 0,77). Isso significa que uma parte da variância do PIB é explicada pelas outras variáveis independentes no modelo. A Utilização da Capacidade Instalada (UCI) apresenta um VIF de 2,57, também refletindo uma colinearidade moderada, principalmente com o PIB, dada a forte correlação anteriormente discutida.

Tabela 2: Índice de Variância Inflacionada (VIF) das Variáveis Independentes

Variável	Descrição	VIF
PIB	Produto Interno Bruto	2,79
Capacidade	Utilização da Capacidade Instalada	2,57
Selic	Taxa Básica de Juros	1,20
ICEI	Índice de Confiança do Empresário Industrial	1,15
DP	Depósitos de Poupança	1,19

## 6 Discussão

A presente seção buscar demonstrar e discutir os resultados realizado sob auxílio dos critérios expostos na metodologia, sendo o modelo escolhido para esse trabalho foi o VECM (modelo vetorial de correção do erro). Vale ressaltar que o processo de modelagem foi inteiramente realizado com a utilização do software chamado *R Studio*. O ambiente do software escolhido foi devido a familiaridade e de ampla utilização para verificar análises envolvendo séries temporais. A seguir observa-se os principais resultados atingidos.

### 6.1 Resultados

Dessa forma sob o uso do software e coleta de dados. Observa-se que primeiramente para a construção do modelo VECM, o primeiro passo é realizar testes de raiz unitária, fundamentais para avaliar a estacionariedade das séries em análise. O teste selecionado para esse propósito foi o Dickey-Fuller Aumentado (ADF), conforme os critérios apresentados anteriormente no capítulo de revisão metodológica.

Nesse sentido, no teste de Dickey-Fuller Aumentado (ADF), a hipótese nula indica a presença de raiz unitária na série, ou seja, que ela não é estacionária. Por outro lado, a hipótese alternativa sugere a ausência de raiz unitária, indicando que a série é estacionária. Para isso, a seguir será apresentado quais variáveis são ou não estacionárias na 3, sendo posteriormente aplicado a diferenciação, na 4, dado a finalidade de torná-la estacionária, sendo um requisito essencial na execução do modelo econométrico da análise.

Tabela 3: Resultados do Teste de Dickey-Fuller Aumentado (ADF)

Série	p-valor	Hipótese Nula	Conclusão
PIB	0,267	Não rejeita	Não estacionária
FBCF	0,346	Não rejeita	Não estacionária
Capacidade	<0,01	Rejeita	Estacionária
Taxa Selic	0,046	Rejeita	Estacionária
ICEI	<0,01	Rejeita	Estacionária
DP	0,570	Não rejeita	Não estacionária

Fonte: elaborado pelos autores.

O Teste ADF foi aplicado às séries temporais, conforme mostrado na tabela 3. Em seguida, foi realizado a diferenciação das séries e aplicamos o mesmo teste. Os resultados após a diferenciação são apresentados na tabela 4. Como pode ser observado, as séries temporais, em sua forma original, não são todas que rejeitaram a hipótese nula de não estacionariedade. Dessa forma, conclui-se que as todas as séries se tornam estacionárias ao serem transformadas pela primeira diferença. A partir desse ponto, passamos a utilizar as séries diferenciadas para a estimação dos modelos e para a avaliação das respectivas estatísticas.

Tabela 4: Resultados do Teste de Dickey-Fuller Aumentado (ADF)  
Após a Diferenciação

Série Diferenciada	p-valor	Hipótese Nula	Conclusão
$\Delta$ PIB	<0,01	Rejeita	Estacionária
$\Delta$ FBCF	<0,01	Rejeita	Estacionária
$\Delta$ Capacidade	<0,01	Rejeita	Estacionária
$\Delta$ Taxa Selic	<0,01	Rejeita	Estacionária
$\Delta$ ICEI	<0,01	Rejeita	Estacionária
$\Delta$ DP	<0,01	Rejeita	Estacionária

Fonte: elaborado pelos autores.

Após garantir a estacionariedade por meio da diferenciação e verificação das séries temporais, é possível proceder com os testes de cointegração. Entre esses, destaca-se o Teste de Johansen, que tem como objetivo identificar possíveis relações de longo prazo que poderiam ser perdidas devido à transformação das séries.

Os resultados apresentados nas tabelas incluem os seguintes critérios de seleção de defasagens: Critério de Informação de Akaike (AIC), Critério de Informação de Hannan-Quinn (HQ), Critério de Informação de Schwarz (SC). Para este caso, a escolha do número de defasagens foi baseada no Critério de Informação de Schwarz (SC) como é demonstrado na tabela 5, com 2 defasagens.

Tabela 5: Critérios para Seleção de Defasagens

Critério	AIC	HQ	SC	FPE
Lags	10	3	2	10

Fonte: elaborado pelos autores.

Nota: Os valores de lags foram definidos com base no Critério de Informação de Schwarz (SC), conforme o teste realizado.

Baseado no critério SC como explicitado anteriormente, selecionam-se 2 lags para o modelo. A partir desta adequação dos lags, pode-se prosseguir para o teste de Johansen, com a finalidade de investigar a existência ou não de cointegração entre as séries temporais, além de identificar a quantidade de vetores de cointegração estão presentes em cada modelo.

Os resultados do teste de Johansen observados na tabela 6, indicam a presença de cointegração entre as variáveis do modelo, dada a estatística de teste para  $r = 0$  (231,39), que excede amplamente os valores críticos para os níveis de significância de 1% (104,20), 5% (90,39) e 10% (85,18). Esses resultados sugerem que existe uma relação de longo prazo entre as variáveis em questão. Portanto, considerando que o teste de Johansen indicou múltiplas relações de cointegração significativas, o VECM é a escolha mais adequada para a análise.

Tabela 6: Resultados do Teste de Johansen - Coeficientes e Estatísticas de Teste

Estágio	Estatística do Teste	Valores Críticos (10%)	Valores Críticos (5%)	Valores Críticos (1%)
$r \leq 5$	7,83	6,50	8,18	11,65
$r \leq 4$	21,10	15,66	17,95	23,52
$r \leq 3$	48,68	28,71	31,52	37,22
$r \leq 2$	86,95	45,23	48,28	55,43
$r \leq 1$	144,75	66,49	70,60	78,87
$r = 0$	231,39	85,18	90,39	104,20

Fonte: elaborado pelos autores.

Nota: A tabela mostra os resultados do teste de Johansen para diferentes estágios de cointegração. Os valores críticos indicam os limites para o teste de cointegração com níveis de significância de 10%, 5% e 1%.

Completado as etapas anteriores, o modelo a ser analisado é definido, considerando a estacionariedade das séries temporais, o número ideal de lags para garantir robustez e a compreensão da dinâmica dos vetores de cointegração. Porém antes de encerrar esta seção, serão analisados dois testes importantes para avaliar a qualidade do modelo: o teste de Shapiro-Wilk, para verificar a normalidade dos resíduos, e o teste de Breusch-Pagan, para analisar se os erros possuem homocedasticidade.

Na Tabela 7, os resultados indicam que os resíduos do modelo são normais, conforme o teste de Shapiro-Wilk ( $p$ -valor = 0,172). Além disso, o teste de Breusch-Pagan ( $p$ -valor = 0,059) sugere que os erros são homocedásticos. Conclui-se que não há rejeição das hipóteses nulas em ambos os casos, ao nível de significância de 5%, ou seja, o  $p$ -valor encontrado é maior do que 0,05.

Tabela 7: Resultados dos Testes de Normalidade e Homocedasticidade dos Resíduos

Teste	p-valor	Hipótese Nula	Conclusão
Shapiro-Wilk	0,172	Não rejeita	Normal
Breusch-Pagan	0,059	Não rejeita	Homocedástico

Fonte: elaborado pelos autores.

Nota: A conclusão baseia-se no nível de significância de 5%.

Na próxima subseção, serão apresentados os resultados das equações do modelo VECM e as funções de impulso-resposta.

## 6.2 Equações Estimadas do Modelo VECM

Esta seção apresenta as equações de curto prazo estimadas do Modelo de Correção de Erros Vetorial (VECM) para as variáveis em análise. Cada subseção detalha a dinâmica da variação de uma variável, revelando como os desvios de longo prazo, representados pelos Termos de Correção de Erro (ECTs), e as variações defasadas das demais variáveis do sistema influenciam o comportamento presente. Os coeficientes obtidos são apresentados com seus respectivos erros padrão entre parênteses, e a significância estatística é indicada por: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ , ·  $p < 0.1$ .

### 6.3.1. Equação para $\Delta\text{PIB}$

$$\begin{aligned}
 \Delta\text{PIB}_t = & 0.0323(0.0916)\text{ECT1}_{t-1} - 0.0225(0.0503)\text{ECT2}_{t-1} \\
 & + 0.0187(0.0929)\text{ECT3}_{t-1} - 0.0037(0.0063)\text{ECT4}_{t-1} \\
 & + 0.0660^{***}(0.0193)\text{ECT5}_{t-1} + 0.0001(0.0038)\text{Intercepto} \\
 & + 0.0041(0.2070)\Delta\text{PIB}_{t-1} + 0.0110(0.0684)\Delta\text{FBCF}_{t-1} \\
 & - 0.0453(0.0772)\Delta\text{Capacidade}_{t-1} - 0.0257(0.0140)\Delta\text{Selic}_{t-1} \\
 & - 0.0540^{*}(0.0210)\Delta\text{ICEI}_{t-1} + 0.3002^{**}(0.0894)\Delta\text{DP}_{t-1} \\
 & - 0.2183(0.1974)\Delta\text{PIB}_{t-2} - 0.0372(0.0552)\Delta\text{FBCF}_{t-2} \\
 & + 0.1485^{*}(0.0695)\Delta\text{Capacidade}_{t-2} + 0.0097(0.0170)\Delta\text{Selic}_{t-2} \\
 & + 0.0782^{***}(0.0178)\Delta\text{ICEI}_{t-2} - 0.1873^{*}(0.0898)\Delta\text{DP}_{t-2} + \varepsilon_{\text{PIB},t}
 \end{aligned}$$

A variação do PIB ( $\Delta\text{PIB}_t$ ) é significativamente influenciada pelo termo de correção de erro  $\text{ECT5}_{t-1}$  (0.0660\*\*\*). Esse resultado indica que o PIB se ajusta ao seu equilíbrio de longo prazo, particularmente em resposta a desvios na relação entre o Índice de Confiança do Empresário Industrial (ICEI) e os Depósitos de Poupança (DP). O coeficiente positivo sugere que se houver um desequilíbrio onde o ICEI está relativamente alto em relação ao DP, o PIB tende a crescer para corrigir essa relação.

No curto prazo, notamos que o PIB é sensível a choques defasados. Uma redução na Taxa Selic (-0.0257) ou na confiança do empresário (ICEI) (-0.0540\*) no período anterior tendem a reduzir o PIB. Por outro lado, um aumento nos Depósitos de Poupança

(DP) (0.3002\*\*) no período anterior impulsiona o crescimento. Com dois trimestres de defasagem, um aumento na Capacidade Instalada (0.1485\*) e na confiança do empresário (ICEI) (0.0782\*\*\*) estão associados a um PIB mais alto.

### 6.3.2. Equação para $\Delta FBCF$

$$\begin{aligned}
 \Delta FBCF_t = & 0.5359^*(0.2399)ECT1_{t-1} - 0.3219^*(0.1318)ECT2_{t-1} \\
 & - 0.0323(0.2432)ECT3_{t-1} - 0.0304^*(0.0165)ECT4_{t-1} \\
 & + 0.1209^*(0.0505)ECT5_{t-1} - 0.0123(0.0100)\text{Intercepto} \\
 & + 1.0062^*(0.5422)\Delta PIB_{t-1} + 0.0219(0.1791)\Delta FBCF_{t-1} \\
 & - 0.2406(0.2023)\Delta \text{Capacidade}_{t-1} - 0.0574(0.0367)\Delta \text{Selic}_{t-1} \\
 & - 0.1448^*(0.0551)\Delta ICEI_{t-1} + 0.7032^{**}(0.2342)\Delta DP_{t-1} \\
 & - 0.2229(0.5171)\Delta PIB_{t-2} - 0.2786^*(0.1447)\Delta FBCF_{t-2} \\
 & + 0.2928(0.1822)\Delta \text{Capacidade}_{t-2} + 0.0360(0.0445)\Delta \text{Selic}_{t-2} \\
 & + 0.1013^*(0.0466)\Delta ICEI_{t-2} - 0.4497^*(0.2352)\Delta DP_{t-2} + \varepsilon_{FBCF,t}
 \end{aligned}$$

A variação da Formação Bruta de Capital Fixo (FBCF) ( $\Delta FBCF_t$ ) mostra significativa correção por meio dos termos de erro. O coeficiente negativo de  $ECT2_{t-1}$  ( $-0.3219^*$ ) é esperado para um mecanismo de correção de erro, indicando que a FBCF se ajusta para baixo quando está acima de seu equilíbrio de longo prazo. Além disso,  $ECT1_{t-1}$  ( $0.5359^*$ ),  $ECT4_{t-1}$  ( $-0.0304^*$ ) e  $ECT5_{t-1}$  ( $0.1209^*$ ) também exercem influência significativa ou marginal, sinalizando múltiplos caminhos para o retorno ao equilíbrio de longo prazo.

No curto prazo, a FBCF é marginalmente impulsionada pelo crescimento do PIB no período anterior ( $1.0062^*$ ). A confiança do empresário (ICEI) tem um efeito misto: uma queda no ICEI ( $-0.1448^*$ ) no lag 1 reduz a FBCF, enquanto um aumento no ICEI ( $0.1013^*$ ) no lag 2 a impulsiona. Um aumento nos Depósitos de Poupança (DP) ( $0.7032^{**}$ ) no período anterior eleva a FBCF, mas, no segundo lag, um aumento nos DP ( $-0.4497^*$ ) está marginalmente associado a uma queda, o que pode refletir dinâmicas de alocação de capital ao longo do tempo.

### 6.3.3. Equação para $\Delta \text{Capacidade}$

$$\begin{aligned}
 \Delta \text{Capacidade}_t = & 0.3060(0.1876)ECT1_{t-1} + 0.0199(0.1031)ECT2_{t-1} \\
 & - 0.4975^*(0.1902)ECT3_{t-1} - 0.0109(0.0129)ECT4_{t-1} \\
 & + 0.0624(0.0395)ECT5_{t-1} - 0.0026(0.0078)\text{Intercepto} \\
 & + 0.6238(0.4240)\Delta PIB_{t-1} - 0.0736(0.1401)\Delta FBCF_{t-1} \\
 & - 0.1859(0.1582)\Delta \text{Capacidade}_{t-1} + 0.0114(0.0287)\Delta \text{Selic}_{t-1} \\
 & - 0.1004^*(0.0431)\Delta ICEI_{t-1} + 0.2012(0.1832)\Delta DP_{t-1} \\
 & - 0.3676(0.4044)\Delta PIB_{t-2} + 0.0434(0.1132)\Delta FBCF_{t-2} \\
 & + 0.2287(0.1424)\Delta \text{Capacidade}_{t-2} + 0.0061(0.0348)\Delta \text{Selic}_{t-2} \\
 & + 0.1836^{***}(0.0365)\Delta ICEI_{t-2} + 0.0640(0.1839)\Delta DP_{t-2} + \varepsilon_{\text{Capacidade},t}
 \end{aligned}$$

A variação na Utilização da Capacidade Instalada ( $\Delta\text{Capacidade}_t$ ) é significativamente corrigida pelo termo  $\text{ECT3}_{t-1}$  ( $-0.4975^*$ ). O sinal negativo e a significância deste coeficiente indicam que a Utilização da Capacidade se ajusta para retornar ao seu equilíbrio de longo prazo: se a capacidade estiver acima do equilíbrio, tende a diminuir. Os outros termos de correção de erro não mostram significância.

No curto prazo, a Utilização da Capacidade é negativamente impactada por uma queda no Índice de Confiança do Empresário Industrial (ICEI) ( $-0.1004^*$ ) no período anterior. Contudo, um aumento no ICEI ( $0.1836^{***}$ ) dois trimestres antes gera um impacto positivo e altamente significativo, sugerindo um efeito defasado da confiança do empresário sobre a produção.

#### 6.3.4. Equação para $\Delta\text{Selic}$

$$\begin{aligned}
 \Delta\text{Selic}_t = & 0.7785(0.6363)\text{ECT1}_{t-1} - 0.6842(0.3496)\text{ECT2}_{t-1} \\
 & + 1.3894^*(0.6452)\text{ECT3}_{t-1} - 0.1960^{***}(0.0438)\text{ECT4}_{t-1} \\
 & - 0.2650\cdot(0.1339)\text{ECT5}_{t-1} - 0.0395(0.0265)\text{Intercepto} \\
 & + 2.2644(1.4383)\Delta\text{PIB}_{t-1} + 0.0812(0.4751)\Delta\text{FBCF}_{t-1} \\
 & - 0.5393(0.5367)\Delta\text{Capacidade}_{t-1} + 0.7055^{***}(0.0975)\Delta\text{Selic}_{t-1} \\
 & + 0.2239(0.1462)\Delta\text{ICEI}_{t-1} + 0.4830(0.6214)\Delta\text{DP}_{t-1} \\
 & - 0.4595(1.3717)\Delta\text{PIB}_{t-2} + 1.0415^{**}(0.3839)\Delta\text{FBCF}_{t-2} \\
 & - 0.3959(0.4832)\Delta\text{Capacidade}_{t-2} + 0.0011(0.1182)\Delta\text{Selic}_{t-2} \\
 & + 0.3051^*(0.1237)\Delta\text{ICEI}_{t-2} - 0.7927(0.6238)\Delta\text{DP}_{t-2} + \varepsilon_{\text{Selic},t}
 \end{aligned}$$

A variação da Taxa Selic ( $\Delta\text{Selic}_t$ ) demonstra uma forte e complexa reação a diversos desequilíbrios de longo prazo. Sendo significativamente corrigida pelos termos  $\text{ECT2}_{t-1}$  ( $-0.6842\cdot$ ),  $\text{ECT3}_{t-1}$  ( $1.3894^*$ ),  $\text{ECT4}_{t-1}$  ( $-0.1960^{***}$ ), e  $\text{ECT5}_{t-1}$  ( $-0.2650\cdot$ ). O coeficiente negativo e altamente significativo de  $\text{ECT4}_{t-1}$  é particularmente importante, pois indica que a Selic se ajusta para retornar ao seu próprio equilíbrio de longo prazo, um resultado crucial para entender a dinâmica da política monetária.

Em relação à dinâmica de curto prazo, a Taxa Selic exibe uma alta persistência, com sua variação no período anterior ( $0.7055^{***}$ ) sendo um preditor altamente significativo da variação atual. A ocorrência de aumento na FBCF dois trimestres antes ( $1.0415^{**}$ ) e na confiança do empresário (ICEI) dois trimestres antes ( $0.3051^*$ ) estão associados a um aumento na Selic, sugerindo uma possível reação do Banco Central a sinais de aquecimento econômico.

### 6.3.5. Equação para $\Delta ICEI$

$$\begin{aligned}
\Delta ICEI_t = & 0.2168(0.7030)ECT1_{t-1} - 0.5080(0.3862)ECT2_{t-1} \\
& + 0.1089(0.7128)ECT3_{t-1} - 0.0623(0.0483)ECT4_{t-1} \\
& - 0.5814^{***}(0.1479)ECT5_{t-1} + 0.0057(0.0293)Intercepto \\
& - 0.3121(1.5890)\Delta PIB_{t-1} - 0.2152(0.5249)\Delta FBCF_{t-1} \\
& + 0.8386(0.5930)\Delta Capacidade_{t-1} + 0.0562(0.1077)\Delta Selic_{t-1} \\
& + 0.1317(0.1616)\Delta ICEI_{t-1} + 0.9314(0.6865)\Delta DP_{t-1} \\
& + 0.2711(1.5154)\Delta PIB_{t-2} - 0.3422(0.4241)\Delta FBCF_{t-2} \\
& - 0.0937(0.5338)\Delta Capacidade_{t-2} + 0.0182(0.1306)\Delta Selic_{t-2} \\
& + 0.0773(0.1367)\Delta ICEI_{t-2} - 0.5941(0.6891)\Delta DP_{t-2} + \varepsilon_{ICEI,t}
\end{aligned}$$

A variação do Índice de Confiança do Empresário Industrial (ICEI) ( $\Delta ICEI_t$ ) é predominantemente influenciada por sua própria correção em relação ao equilíbrio de longo prazo, manifestada pelo termo  $ECT5_{t-1}$  ( $-0.5814^{***}$ ). Este coeficiente negativo e altamente significativo indica que o ICEI tem uma forte e rápida tendência de se ajustar para retornar ao seu nível de equilíbrio quando há desvios.

### 6.3.6. Equação para $\Delta DP$

$$\begin{aligned}
\Delta DP_t = & -0.1589(0.1076)ECT1_{t-1} + 0.1263^*(0.0591)ECT2_{t-1} \\
& - 0.1895^*(0.1091)ECT3_{t-1} - 0.0098(0.0074)ECT4_{t-1} \\
& + 0.0143(0.0226)ECT5_{t-1} + 0.0104^*(0.0045)Intercepto \\
& + 0.2328(0.2431)\Delta PIB_{t-1} - 0.0442(0.0803)\Delta FBCF_{t-1} \\
& - 0.0942(0.0907)\Delta Capacidade_{t-1} + 0.0018(0.0165)\Delta Selic_{t-1} \\
& - 0.0056(0.0247)\Delta ICEI_{t-1} + 0.5801^{***}(0.1050)\Delta DP_{t-1} \\
& + 0.4024(0.2319)\Delta PIB_{t-2} - 0.1101(0.0649)\Delta FBCF_{t-2} \\
& - 0.1214(0.0817)\Delta Capacidade_{t-2} - 0.0037(0.0200)\Delta Selic_{t-2} \\
& - 0.0620^{**}(0.0209)\Delta ICEI_{t-2} - 0.0598(0.1054)\Delta DP_{t-2} + \varepsilon_{DP,t}
\end{aligned}$$

A variação dos Depósitos de Poupança (DP) ( $\Delta DP_t$ ) é influenciada principalmente pelos termos de correção de erro  $ECT2_{t-1}$  ( $0.1263^*$ ) e, marginalmente, por  $ECT3_{t-1}$  ( $-0.1895^*$ ). Um Intercepto positivo e significativo ( $0.0104^*$ ) sugere um crescimento autônomo médio dos depósitos, mesmo na ausência de variações em outras variáveis, indicando uma tendência inerente de expansão da poupança.

Tabela 8: Coeficientes Estimados do Modelo VECM

Equação	Termo	Coeficiente (Erro Padrão)
PIB	ECT1	0.0323 (0.0916)
	ECT2	-0.0225 (0.0503)
	ECT3	0.0187 (0.0929)
	ECT4	-0.0037 (0.0063)
	ECT5	0.0660**** (0.0193)
	Intercepto	0.0001 (0.0038)
FBCF	ECT1	0.5359** (0.2399)
	ECT2	-0.3219** (0.1318)
	ECT3	-0.0323 (0.2432)
	ECT4	-0.0304* (0.0165)
	ECT5	0.1209** (0.0505)
Capacidade	ECT1	0.3060 (0.1876)
	ECT2	0.0199 (0.1031)
	ECT3	-0.4975** (0.1902)
	ECT4	-0.0109 (0.0129)
	ECT5	0.0624 (0.0395)
Selic	ECT1	0.7785 (0.6363)
	ECT2	-0.6842* (0.3496)
	ECT3	1.3894** (0.6452)
	ECT4	-0.1960**** (0.0438)
	ECT5	-0.2650* (0.1339)
ICEI	ECT1	0.2168 (0.7030)
	ECT2	-0.5080 (0.3862)
	ECT3	0.1089 (0.7128)
	ECT4	-0.0623 (0.0483)
	ECT5	-0.5814**** (0.1479)
DP	ECT1	-0.1589 (0.1076)
	ECT2	0.1263** (0.0591)
	ECT3	-0.1895* (0.1091)
	ECT4	-0.0098 (0.0074)
	ECT5	0.0143 (0.0226)

Nota: Valores entre parênteses representam os erros padrão.

Significância: \* $p < 0,10$ , \*\* $p < 0,05$ , \*\*\* $p < 0,01$ , \*\*\*\* $p < 0,001$ .

Fonte: Elaboração própria.

No curto prazo, os Depósitos de Poupança mostram forte persistência, com sua própria variação no período anterior (0.5801\*\*\*) sendo o principal fator explicativo. Com dois trimestres de defasagem, um aumento no PIB (0.4024) está marginalmente associado a um aumento nos depósitos, enquanto um aumento na FBCF (-0.1101) se relaciona mar-

ginalmente a uma queda. Uma redução significativa na confiança do empresário (ICEI) ( $-0.0620^{**}$ ) no segundo lag também se associa a uma diminuição nos Depósitos de Poupança, destacando a sensibilidade da poupança à percepção do ambiente econômico.

A apresentação das equações do Modelo de Correção de Erros Vetorial (VECM) trouxe entendimentos importantes sobre as dinâmicas de curto e longo prazo que afetam a Formação Bruta de Capital Fixo (FBCF) e as variáveis relacionadas do modelo, entre 1999 e 2023. Os resultados destacam a interconectividade das variáveis econômicas e fornecem uma compreensão aprofundada dos propulsores do investimento produtivo.

Do ponto de vista econômico, a FBCF não opera de forma isolada. A FBCF respondeativamente a desequilíbrios de longo prazo e a choques de curto prazo em outras variáveis macroeconômicas. A capacidade da FBCF de se ajustar a relações de cointegração com os Depósitos de Poupança (ECT1, ECT2) é uma descoberta significativa. Isso sugere que a disponibilidade de fundos para investimento e a estrutura de financiamento (representada pelos depósitos) são essenciais para o reequilíbrio do capital fixo no longo prazo. A sensibilidade da FBCF à confiança do empresário (ICEI) no curto prazo, tanto no primeiro quanto no segundo período de defasagem, ressalta a importância das expectativas e do ambiente de negócios nas decisões de investimento. O efeito variado dos Depósitos de Poupança na FBCF em diferentes defasagens indica a complexidade da intermediação financeira e os caminhos pelos quais a poupança se transforma em investimento produtivo.

Os resultados das outras equações reforçam essa perspectiva. A forte correção automática do ICEI e da Selic em direção aos seus respectivos equilíbrios de longo prazo demonstra que a confiança e a política monetária possuem dinâmicas internas robustas que influenciam o ambiente para o investimento. A ligação do PIB com a confiança (ICEI) e a poupança (DP) confirma que o crescimento econômico e as escolhas de investimento estão intimamente ligados às expectativas e à capacidade de financiamento da economia. A relevância dos termos de correção de erro para diversas variáveis aponta para a existência de relacionamentos de equilíbrio de longo prazo. Isso sugere que as políticas econômicas devem considerar não apenas os efeitos imediatos, mas como as variáveis retornam às posições de equilíbrio.

A análise envolve a reação a desvios do equilíbrio de longo prazo e a resposta a choques de curto prazo em variáveis como confiança, poupança e política monetária. Para promover um aumento duradouro do investimento, as políticas devem focar não só em estimular a FBCF no curto prazo, mas em assegurar a estabilidade e o reequilíbrio das relações macroeconômicas de longo prazo que a sustentam. A seguir, para complementar a análise das relações e dinâmicas encontradas, a seção de impulso-resposta detalhará como choques específicos em uma variável afetam as demais ao longo do tempo.

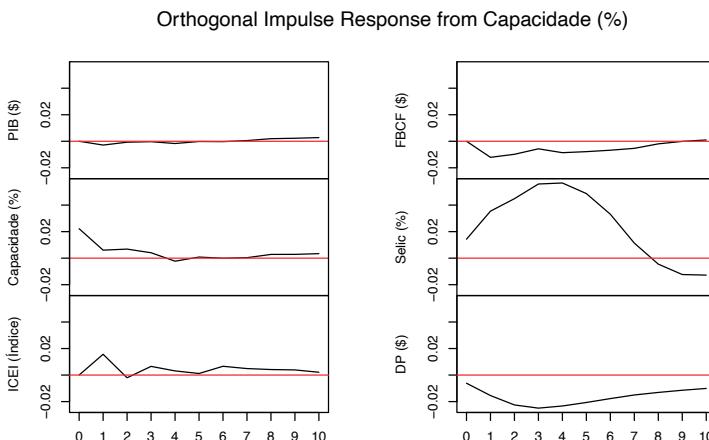
## 6.3 Análise das Funções Impulso-Resposta

Uma forma de compreender e visualizar os resultados econômicos geridos pelo modelo é através da execução das funções de impulso nas variáveis utilizadas no modelo. Nesse contexto, a utilização da função de impulso-resposta visa avaliar como um choque em uma variável no período  $t$  pode afetar o comportamento de outra ao longo do tempo.

A seguir pode-se encontrar os gráficos de impulso-resposta referente as variáveis analisadas para até 10 trimestres a frente de quando ocorreu o impulso:

Na Figura 3, é analisada a resposta das variáveis a um choque positivo e ortogonal na Utilização da Capacidade Instalada. A análise abrange um horizonte temporal de dez trimestres. Um choque positivo na Capacidade Utilizada gera um efeito positivo de curto prazo no Índice de Confiança do Empresário Industrial (ICEI) e na própria Capacidade Utilizada, com essa resposta gradualmente atenuando-se em períodos mais longos. A Taxa Selic exibe uma resposta inicial de maior potência, culminando em um efeito negativo no longo prazo, embora a trajetória inicial seja de um pico positivo antes da reversão. Por fim, tanto a Formação Bruta de Capital Fixo (FBCF) quanto os Depósitos de Poupança (DP) apresentam um efeito negativo inicial que decai ao longo do tempo. É notável que a resposta dos Depósitos de Poupança manifesta um grau de impacto superior em comparação à resposta da FBCF.

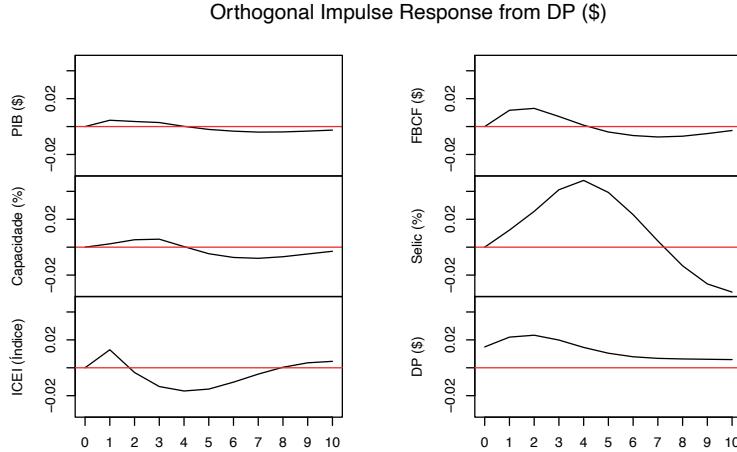
Figura 3: Impulso-Resposta Capacidade



Na Figura 4, é ilustrada a resposta das variáveis a um choque positivo e ortogonal nos Depósitos de Poupança (DP), ao longo de um horizonte temporal de dez trimestres. Similarmente, a Capacidade Utilizada demonstra um efeito de baixa magnitude, com uma pequena variação que se dissipava rapidamente. O Índice de Confiança do Empresário Industrial (ICEI) apresenta uma resposta inicialmente positiva, que se reverte para negativa antes de desaparecer em períodos mais longos. A Formação Bruta de Capital Fixo (FBCF) reage com um efeito positivo inicial mais pronunciado. Contudo, essa resposta positiva é

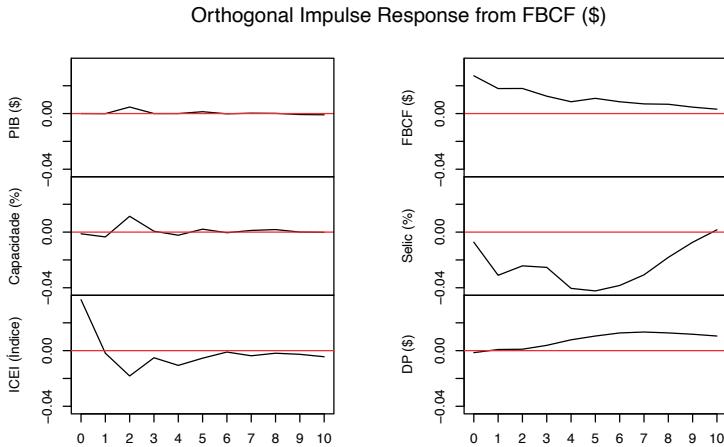
revertida, culminando em um efeito negativo que se atenua com o passar do tempo. Para a própria variável Depósitos de Poupança, o efeito inicial positivo dissipa-se gradualmente ao longo do horizonte. A Taxa Selic, por sua vez, exibe um efeito positivo no curto e médio prazo, o qual se transforma em um impacto negativo e persistente no longo prazo.

Figura 4: Impulso-Resposta DP



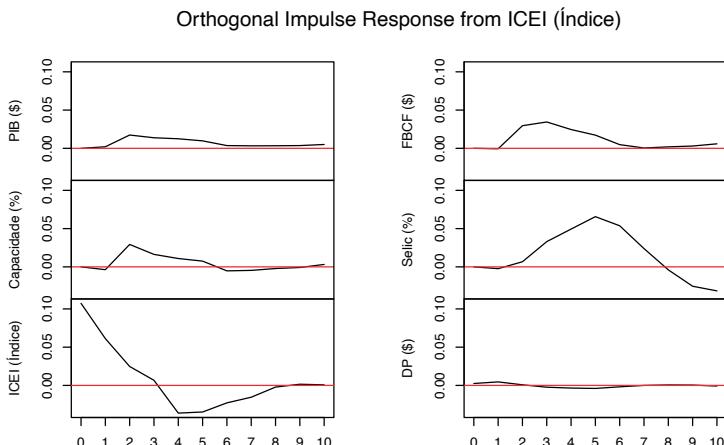
Na Figura 5, é analisada a resposta das variáveis a um choque positivo e ortogonal na Formação Bruta de Capital Fixo (FBCF), que representa a variável dependente do modelo. Esta análise é conduzida ao longo de um horizonte temporal de dez trimestres. A Capacidade Utilizada, por sua vez, exibe um pequeno efeito positivo inicial, que rapidamente se atenua e oscila em torno de zero nos trimestres subsequentes. O Índice de Confiança do Empresário Industrial (ICEI) manifesta uma resposta inicial negativa, revertendo-se para um impacto positivo de curta duração antes de se dissipar. A Taxa Selic apresenta um efeito notavelmente negativo no curto prazo, que gradualmente decai ao longo do tempo. Em contraste, a própria FBCF, alvo do choque, demonstra um impacto inicial amplamente positivo que decresce de forma persistente. Os Depósitos de Poupança (DP) recebem um ganho positivo e persistente em resposta ao choque na FBCF, estabilizando-se em um patamar positivo.

Figura 5: Impulso-Resposta FBCF



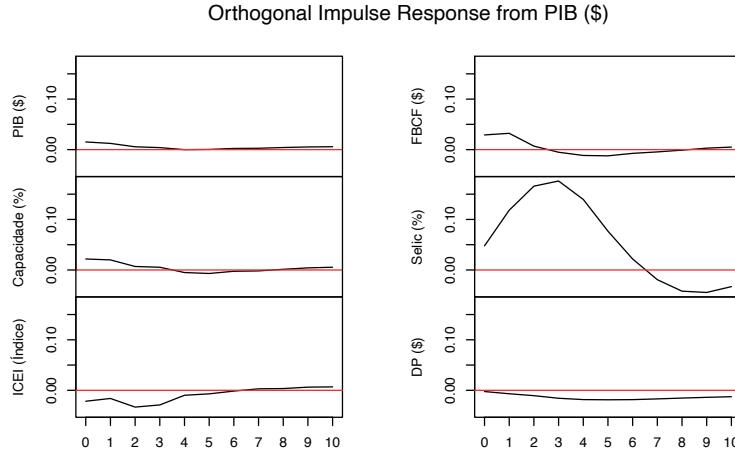
A Figura 6 ilustra a resposta das variáveis a um choque positivo e ortogonal no Índice de Confiança do Empresário Industrial (ICEI), considerando um horizonte temporal de dez trimestres. A própria variável ICEI reage inicialmente com um efeito amplamente positivo que decai rapidamente, transformando-se em um impacto negativo transitório antes de se dissipar no longo prazo. A Capacidade Utilizada e a Formação Bruta de Capital Fixo (FBCF) apresentam ganhos iniciais positivos em resposta ao choque no ICEI, com esses efeitos se atenuando e desaparecendo em períodos mais longos. Por fim, a Taxa Selic exibe um efeito positivamente elevado no curto e médio prazo, o qual se inverte para um impacto negativo que persiste nos períodos finais do horizonte de análise.

Figura 6: Impulso-Resposta ICEI



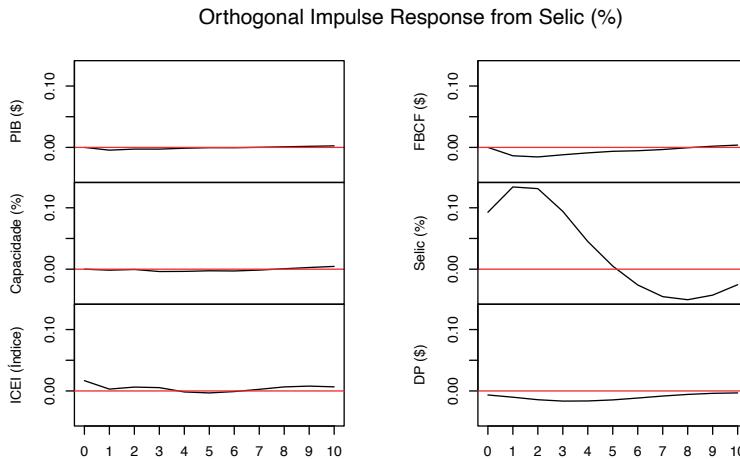
A Figura 7 apresenta a função impulso-resposta a um choque positivo e ortogonal no Produto Interno Bruto (PIB), abrangendo um horizonte temporal de dez trimestres. Observam-se efeitos positivos de pequena dimensão e de curta duração na Formação Bruta de Capital Fixo (FBCF) e na Capacidade Utilizada, os quais se dissipam rapidamente. O Índice de Confiança do Empresário Industrial (ICEI) exibe um pequeno efeito inicial negativo que igualmente não persiste. Em contraste com as demais, a Taxa Selic demonstra um efeito positivamente elevado no curto e médio prazo, atingindo um pico antes de reverter para um impacto negativo que se estende aos períodos finais da análise.

Figura 7: Impulso-Resposta PIB



A Figura 8 conclui a análise das funções impulso-resposta, apresentando os efeitos de um choque positivo e ortogonal na Taxa Selic, ao longo de um horizonte temporal de dez trimestres. Observam-se efeitos negativos, de pequena dimensão e de curta duração, tanto para os Depósitos de Poupança (DP) quanto para a Formação Bruta de Capital Fixo (FBCF); ambos os impactos se dissipam rapidamente nos trimestres subsequentes. O Índice de Confiança do Empresário Industrial (ICEI) também exibe um pequeno efeito negativo inicial que igualmente não persiste. Em contraste, a própria Taxa Selic, alvo do choque, demonstra um impacto inicial amplamente positivo que, após atingir um pico, reverte-se para um efeito negativo e persistente no longo prazo.

Figura 8: Impulso-Resposta SELIC



As análises das Funções Impulso-Resposta demonstram a complexidade dos determinantes da Formação Bruta de Capital Fixo (FBCF). Observa-se que o investimento reage tanto a desvios de longo prazo quanto a choques de curto prazo na confiança e em outras variáveis macroeconômicas. A capacidade de resposta da política monetária a esses choques constitui, ademais, um fator relevante. Para sustentar o investimento, as políticas econômicas devem considerar não apenas estímulos de curto prazo, mas também a estabilidade e o reequilíbrio das relações macroeconômicas de longo prazo que suportam a formação de capital.

## 7 Conclusão

Em síntese, os resultados estimados indicam que variáveis como PIB, ICEI, Depósitos de Poupança e a própria FBCF apresentaram influência positiva sobre a Formação Bruta de Capital Fixo no curto prazo. A Capacidade Instalada mostrou-se negativamente relacionada, enquanto a Selic não apresentou significância estatística. O termo de correção de erros evidenciou relações de longo prazo consistentes com cointegração.

Os achados estão alinhados com parte da literatura nacional e internacional, especialmente em relação à influência da poupança interna e da confiança empresarial. Contudo, os resultados sobre a Selic contrastam com alguns estudos anteriores.

Entre as limitações do estudo, destaca-se a ausência de variáveis fiscais e de choques institucionais, bem como o uso de agregados nacionais que podem omitir efeitos heterogêneos.

Como agenda futura, sugere-se explorar modelos estruturais, ampliar o conjunto de variáveis, e realizar análises regionais ou setoriais para identificar nuances do comportamento do investimento no Brasil.

# Referências

- [1] BUENO, R. D. L. D. S. *Econometria de séries temporais*. Editora, 2018.
- [2] CLARK, J. M. Business acceleration and the law of demand: A technical factor in economic cycles. In *Classics in Institutional Economics, Part II, Volume 6*. Routledge, 2024, pp. 24–43.
- [3] DA ROCHA GONÇALVES, R., AND DE OLIVEIRA PASSOS, M. Determinantes macroeconômicos da formação bruta de capital fixo no brasil: 1996-2014.
- [4] GONZALES, E. O., SBARDELLATI, E. C. A., AND SANTOS, A. S. D. Uma investigação empírica sobre os determinantes do investimento no brasil (1995-2013). *Anais do VII Encontro Internacional da Associação Keynesiana Brasileira, São Paulo* (2014).
- [5] GUJARATI, D. N., AND PORTER, D. C. Econometria básica. *Porto Alegre: AMGH* 1, 1 (2011), 1–856.
- [6] JORGENSON, D. W. Capital theory and investment behavior. *The American economic review* 53, 2 (1963), 247–259.
- [7] JUDGE, G. G., ET AL. Theory of econometrics: An introductory exposition.
- [8] LÉLIS, M. T. C., BREDOW, S. M. S., AND CUNHA, A. M. Determinantes macroeconômicos dos investimentos no brasil: um estudo para o período 1996-2012. *Revista de Economia Contemporânea* 19, 2 (2015), 203–234.
- [9] MEURER, M. E. K. Determinantes da formação bruta de capital fixo no brasil: uma análise para o período entre 1996e 2016.
- [10] MONTGOMERY, D. C., PECK, E. A., AND VINING, G. G. *Introduction to linear regression analysis*. John Wiley & Sons, 2021.
- [11] RIBEIRO, V. D. F. Determinantes da formação bruta de capital fixo no brasil: um estudo econométrico de 1996-2020. *Trabalho de Conclusão de Curso (Bacharelado em Ciências Econômicas)-Instituto de Economia, Universidade Federal do Rio de Janeiro, Rio de Janeiro* (2021).

## SOBRE OS AUTORES

**FERNANDO ANTONIO LUCENA AIUBE:** É professor associado da Faculdade de Economia da UERJ. Possui graduação em Engenharia Elétrica pela Universidade Federal de Goiás (1980), mestrado (1995) e doutorado (2005) ambos em Engenharia de Produção, PUC-Rio. Tem experiência e interesse nas seguintes áreas/temas: finanças, derivativos, econometria, opções reais, IA aplicada a economia, avaliação econômica de projetos de investimento.

**EDISON AMERICO HUARSAYA TITO:** É professor adjunto do Instituto de Matemática e Estatística da Universidade do Estado do Rio de Janeiro (UERJ). Graduado em Engenharia Eletrônica, é mestre e doutor em Métodos de Apoio à Decisão pela PUC-Rio, doutor em Processamento de Sinais pela UFRJ, mestre em Finanças e Economia Empresarial pela FGV e pós-doutor em Administração pela PUC-Rio. Atua nas áreas de finanças, economia empresarial, avaliação de projetos, gestão de riscos, inteligência artificial, ciência de dados e blockchain.

# ECONOMETRIA EM FOCO

ENTRE JUROS, CONFIANÇA E ENERGIA NO BRASIL



[www.atenaeditora.com.br](http://www.atenaeditora.com.br)



[contato@atenaeditora.com.br](mailto:contato@atenaeditora.com.br)



[@atenaeditora](https://www.instagram.com/atenaeditora)



[www.facebook.com/atenaeditora.com.br](https://www.facebook.com/atenaeditora.com.br)

# ECONOMETRIA EM FOCO

**ENTRE JUROS, CONFIANÇA E ENERGIA NO BRASIL**



- 🌐 [www.atenaeditora.com.br](http://www.atenaeditora.com.br)
- ✉️ [contato@atenaeditora.com.br](mailto:contato@atenaeditora.com.br)
- 📷 [@atenaeditora](https://www.instagram.com/atenaeditora)
- ⬇️ [www.facebook.com/atenaeditora.com.br](https://www.facebook.com/atenaeditora.com.br)