POLÍTICAS DE INTERIORIZAÇÃO E RESILIÊNCIA ORÇAMENTÁRIA MUNICIPAL: UMA MODELAGEM CAUSAL BASEADA EM IA SOBRE O IMPACTO DO IFRN LAJES-RN NO PIB PÚBLICO

INTERIORIZATION POLICIES AND MUNICIPAL BUDGET RESILIENCE: AN AI-BASED CAUSAL MODELING ON THE IMPACT OF IFRN LAJES-RN ON PUBLIC GDP

LEANDRO APARECIDO DA SILVA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO RIO GRANDE DO NORTE-IFRN

Comunicação:
O XIV SINGEP foi realizado em conjunto com a 14th Conferência Internacional do CIK (CYRUS Institute of Knowledge) e com a SPARK (Society for Promoting Academic Research and Knowledge), em formato híbrido, com sede presencial na UNINOVE - Universidade Nove de Julho, no Brasil.

Agradecimento à orgão de fomento:
Agradeço à Pro-reitoria de pesquisa (PROPI) do IFRN e à Coordenação de pesquisa e inovação (COPEIN) do Campus Lajes pelo apoio institucional concedido para a realização deste projeto

POLÍTICAS DE INTERIORIZAÇÃO E RESILIÊNCIA ORÇAMENTÁRIA MUNICIPAL: UMA MODELAGEM CAUSAL BASEADA EM IA SOBRE O IMPACTO DO IFRN LAJES-RN NO PIB PÚBLICO

Objetivo do estudo
Avaliar e mensurar o impacto causal líquido da instalação e operação do campus do IFRN sobre o Produto Interno Bruto (PIB) do setor de serviços da administração pública do município de Lajes (RN) ao longo da janela temporal de 2015 a 2021.

Relevância/originalidade
O estudo inova ao avaliar empiricamente uma política de interiorização em um pequeno município do semiárido nordestino. Preenche uma lacuna analítica ao simular um contrafactual robusto, controlando a autocorrelação de séries macroeconômicas locais sem depender de premissas estáticas de modelos tradicionais.

Metodologia/abordagem
Adota-se quase-experimento de Controle Sintético com Séries Temporais Estruturais Bayesianas (BSTS) e CausalImpact. A bacia de doadores contém cinco cidades limítrofes, utilizando-se programação Python para a higienização e o tratamento de escala das séries anuais do PIB público (2008–2021).

Principais resultados
A instalação do campus causou uma expansão líquida de 3,44% ao ano no PIB da administração pública de Lajes (RN), gerando ganho acumulado real de R$ 7.734.000,00. O modelo bayesiano confirmou o efeito com 100% de probabilidade a posteriori.

Contribuições teóricas/metodológicas
Metodologicamente, valida a aplicação do Controle Sintético Bayesiano integrado ao CausalImpact (baseado em inteligência artificial) para a avaliação ex-post de políticas públicas. Teoricamente, enriquece a literatura de desenvolvimento endógeno ao conectar dados empíricos municipais à modelagem estrutural bayesiana de incerteza.

Contribuições sociais/para a gestão
Fornece subsídios empíricos robustos para apoiar a continuidade e expansão do orçamento da Rede Federal, evidenciando o efeito de resiliência orçamentária anticíclica. Demonstra aos gestores públicos que investimentos em educação descentralizada atuam como protetores socioeconômicos ativos para pequenas economias regionais.

Palavras-chave: Rede Federal, PIB Público, Séries Temporais Estruturais Bayesianas, CausalImpact, Interiorização

INTERIORIZATION POLICIES AND MUNICIPAL BUDGET RESILIENCE: AN AI-BASED CAUSAL MODELING ON THE IMPACT OF IFRN LAJES-RN ON PUBLIC GDP

Study purpose
To evaluate and measure the net causal impact of the establishment and operation of the IFRN campus on the Gross Domestic Product (GDP) of Lajes (RN) public administration services sector throughout the symmetric temporal window from 2015 to 2021.

Relevance / originality
The study innovates by empirically evaluating an educational regional decentralisation policy in a small semi-arid municipality. It fills an analytical gap by simulating a robust counterfactual, controlling for local macroeconomic serial correlation without relying on traditional static regression assumptions.

Methodology / approach
A quasi-experimental design using Synthetic Control with Bayesian Structural Time Series (BSTS) and CausalImpact algorithm is adopted. The donor pool comprises five bordering cities, using a Python-programmed script for scale correction and data cleaning of municipal GDP series (2008–2021).

Main results
The campus installation caused a net annual expansion of 3.44% in Lajes' public administration GDP, generating a real cumulative gain of R$ 7,734,000.00. The Bayesian model confirmed this effect with extreme statistical significance and 100% posterior probability.

Theoretical / methodological contributions
Methodologically, it validates the application of Bayesian Synthetic Control integrated with CausalImpact (artificial intelligence) for ex-post policy evaluation. Theoretically, it enriches the regional development literature by linking municipal empirical data to dynamic Bayesian state-space structural modelling under uncertainty.

Social / management contributions
It provides robust empirical evidence to support the funding and expansion of the Federal Network, revealing its budget-cushioning anticyclical effect. It demonstrates to policy makers that decentralised education investments act as active socioeconomic protectors for small regional economies.

Keywords: Federal Network, Public GDP, Bayesian Structural Time Series, CausalImpact, Internalization

POLÍTICAS DE INTERIORIZAÇÃO E RESILIÊNCIA ORÇAMENTÁRIA MUNICIPAL: UMA MODELAGEM CAUSAL BASEADA EM IA SOBRE O IMPACTO DO IFRN LAJES-RN NO PIB PÚBLICO

1 Introdução

A implantação de institutos educacionais em áreas historicamente periféricas pode promover a qualificação de capital humano local, estimular a produtividade das empresas e gerar um novo arcabouço de conhecimento e infraestrutura, capaz de transformar a dinâmica produtiva local (Carter & Rostova, 2023; Lee & Kim, 2024). No cenário brasileiro, a reconfiguração da educação profissional e tecnológica materializou-se a partir de 2008, com a criação e expansão da rede federal de educação profissional, científica e tecnológica. Um marco legal que interiorizou a oferta de ensino técnico gratuito e de excelência rumo a pequenos e médios municípios do país (Cichaczewski & Castro, 2023; Pereira & Cruz, 2019). Sob a ótica do desenvolvimento regional, essa política social de descentralização geográfica buscou mitigar assimetrias espaciais históricas, convertendo as novas unidades federadas de ensino em polos de engajamento cívico, fomento à governança participativa e suporte aos arranjos produtivos locais (Pozzer, 2024; Silva, Nonato, & Spink, 2023).

Sendo um município de pequeno porte populacional situado na região semiárida potiguar, Lajes-RN, caracterizava-se, antes da chegada do IFRN, por uma dinâmica econômica historicamente vulnerável, carente de uma estrutura produtiva consolidada e altamente dependente de repasses fiscais governamentais, como o Fundo de Participação dos Municípios (FPM) (Pereira & Cruz, 2019; Pozzer, 2024). A instalação física e o início das operações de um campus do IFRN em maio de 2015, representaram um aporte sem precedentes de infraestrutura, custeio orçamentário e pessoal estável na localidade, alterando de forma substancial a escala de investimentos públicos federais no território (Faveri, Petterini, & Barbosa, 2018; Pereira & Cruz, 2019). Desse modo, o município de Lajes-RN funciona como um laboratório quase-experimental ideal para avaliar em que medida a interiorização de uma instituição de ensino de excelência consegue induzir dinamismo econômico e prover uma proteção macroeconômica contra recessos econômicos (Faveri et al., 2018; Pereira & Cruz, 2019).

Apesar da relevância de analisar casos com essas características, a avaliação empírica do impacto real dessas intervenções de infraestrutura sobre as economias municipais de pequeno porte enfrenta importantes lacunas analíticas e metodológicas (Faveri et al., 2018). As abordagens tradicionais de monitoramento do ciclo de políticas públicas, frequentemente, baseiam-se em indicadores descritivos simples ou regressões lineares estáticas, as quais podem falhar ao isolar efeitos causais dinâmicos, e provavelmente ignorar as propriedades de autocorrelação serial de variáveis macroeconômicas (Araujo & Kubrusly, 2023; Brodersen et al., 2015). Diante da dificuldade de simular como pequenas economias se comportariam caso uma política pública de interiorização nunca tivesse ocorrido, emerge o seguinte problema de pesquisa: qual foi o impacto econômico provocado pelo campus do IFRN no PIB da administração pública de Lajes (RN)? Para solucionar essa lacuna investigativa, esta pesquisa apresenta o objetivo principal: avaliar o impacto econômico da instalação do campus do IFRN Lajes sobre o PIB de serviços da administração pública do município.

No aspecto socioeconômico e gerencial, o estudo poderá fornecer evidências empíricas baseadas em dados sobre os ‘efeitos-gasto’ e os benefícios de estabilização orçamentária de curto e longo prazo, promovidos pela rede federal, no interior do Nordeste, servindo como base informacional para que gestores e formuladores de políticas públicas possam avaliar a eficiência

alocativa de investimentos científicos (Pereira & Cruz, 2019; Teixeira & Gianezini, 2021). Sob a perspectiva metodológica, a investigação justifica-se ao introduzir uma abordagem inovadora de aprendizado de máquina e inferência causal bayesiana para contornar vieses de especificação comuns na econometria clássica, viabilizando previsões contrafactuais acompanhadas de uma propagação de incerteza matemática transparente e interpretável para o suporte de apoio à decisão no setor público (Brodersen et al., 2015; Varian, 2014; Vieira, 2023).

Para estruturar e dar sustentação conceitual a esta investigação, o próximo tópico apresenta a fundamentação teórica do estudo. A revisão aborda inicialmente as premissas da rede federal e o papel de seu enraizamento geográfico na promoção do desenvolvimento regional, discute em seguida os mecanismos teóricos de demanda e de oferta decorrentes da infraestrutura educacional e, por fim, conceitua o emprego de inteligência artificial e algoritmos de aprendizado de máquina para fins de monitoramento causal e suporte à decisão no ciclo de políticas públicas.

2 Educação Federal e Desenvolvimento Regional: O Caso da Interiorização

A institucionalização da Rede Federal de Educação Profissional, Científica e Tecnológica e a seguinte criação dos Institutos Federais de Educação, Ciência e Tecnologia (IFs), consolidadas pela Lei nº 11.892 em dezembro de 2008, representaram uma das políticas públicas intersetoriais mais difíceis e abrangentes de expansão educacional e científica da história recente do país (Pozzer, 2024; Pereira & Cruz, 2019). Essa iniciativa governamental promoveu um crescimento sem precedentes da Rede Federal em direção ao interior do país, ampliando o número de unidades e expandindo a capilaridade da formação científica e tecnológica para fora dos grandes polos de desenvolvimento consolidados (Faveri, Petterini, & Barbosa, 2018; Pereira & Cruz, 2019). Do ponto de vista conceitual, essa expansão foi desenhada com o propósito de atuar como uma ferramenta de política social de expansão, ancorada sob a redução de desequilíbrios territoriais e desigualdades regionais de oportunidades educacionais e laborais (Faveri, Petterini, & Barbosa, 2018; Teixeira & Gianezini, 2021).

Historicamente, a trajetória da educação profissional federal brasileira esteve segmentada por uma forte dualidade geográfica e estrutural: por um lado, as escolas técnicas industriais localizavam-se de forma majoritária nas capitais administrativas e nos polos urbanos de alta concentração industrial. Em direção oposta, as escolas técnicas agrícolas situavam-se em áreas estritamente rurais, geralmente operando como fazendas-modelo. A reconfiguração iniciada com o plano de expansão a partir de 2003 e a estruturação dos Institutos Federais em 2008 romperam com essa dualidade histórica ao direcionar a instalação de novos multicampi para pequenos e médios municípios do interior brasileiro sem, basicamente, possuírem um perfil industrial pré-existente. Com efeito, ao priorizar a interiorização da oferta educacional, mais de 58% das novas unidades federadas criadas após 2003 foram sediadas em localidades com menos de 100 mil habitantes, demonstrando uma descentralização sem precedentes e modificando de forma determinante a geografia produtiva e do conhecimento nacional (Pozzer, 2024).

Sob a perspectiva teórica do desenvolvimento regional, a implantação dos Institutos Federais estabelece uma quebra com a visão tradicional de crescimento econômico, pautada unicamente na atração de atividades externas ao território. Ao motivar-se a partir do tripé indissociável de ensino, pesquisa e extensão, o desenho institucional dos IFs sintoniza-se com uma noção ampliada e comunitária de desenvolvimento sustentável, articulando-se com as dimensões econômica, social, territorial, cultural e ambiental de forma orgânica (Pozzer &

Neuhold, 2019; Pereira & Cruz, 2019). Por meio desse modelo participativo, as instituições federais têm o papel de pautar sua oferta formativa e seus esforços científicos a partir do diagnóstico regional e da escuta comunitária em audiências públicas, de modo a dar suporte prático e científico à consolidação e ao fortalecimento dos Arranjos Produtivos, Sociais e Culturais Locais (APLs) (Pozzer, 2024; Silva, Nonato, & Spink, 2023).

Esse modelo de enraizamento territorial converge com a teoria do engajamento cívico e da universidade cívica. Essa vertente acadêmica sustenta que a inserção territorial de uma instituição de ensino superior e técnico de excelência deve focar na produção de impactos extra-acadêmicos, tratando a resolução de demandas socioeconômicas do entorno local como sua atividade-fim e elemento norteador de sua identidade, e não como uma ação voluntária ou marginal (Silva, Nonato, & Spink, 2023). À escala local e regional, o campus cívico funciona como um agente ativo de fomento à governança participativa de soluções conjuntas entre o Estado, a iniciativa privada e a sociedade civil organizada (Pozzer & Neuhold, 2019). A oferta de uma educação pública de qualidade próximo ao domicílio e sintonizada com a realidade local não apenas reduz desigualdades de oportunidades educacionais históricas, mas também atua de forma crucial para evitar a migração forçada de jovens rumo às grandes metrópoles, potencializando a fixação de capital humano e fortalecendo o sentimento de pertencimento territorial e cidadania nas pequenas cidades (Teixeira & Gianezini, 2021; Pereira & Cruz, 2019; Pozzer, 2024).

2.1 Os Impactos Econômicos Locais da Expansão Educacional

A literatura que avalia o impacto econômico e social decorrente da introdução de novos complexos educacionais e científicos sobre os territórios de abrangência divide esses efeitos em dois olhares temporais e conceituais: o de curto prazo, pautado na dinâmica da demanda, e o de longo prazo, centrado no lado da oferta de fatores produtivos. O impacto de curto prazo, definido teoricamente como o "efeito-gasto", refere-se aos fluxos monetários imediatos injetados diretamente na economia local a partir do funcionamento material e financeiro da instituição de ensino. Esse encadeamento é ativado de forma contínua por meio das despesas de capital para obras e infraestrutura, das contratações de custeio administrativo, logística, alimentação e segurança operadas via licitações junto a empresas locais, e, fundamentalmente, pela fixação de uma massa de rendimento permanente originária dos salários pagos aos corpos docentes e técnicos administrativos, cujo consumo de comércio e serviços locais retroalimenta e ativa o circuito econômico local (Pereira & Cruz, 2019; Faveri, Petterini, & Barbosa, 2018).

Por outro lado, o impacto de longo prazo, denominado como "encadeamento para frente", manifesta-se de maneira gradativa e cumulativa à medida que a formação de capital humano técnico e científico atinge maturidade. Essa expansão do conhecimento atua na qualificação profissional e na elevação geral da escolaridade média local, o que gera ganhos diretos de produtividade para as organizações e estimula a inovação técnica voltada à resolução de problemas das economias sub-regionais. No Brasil, estudos de economia regional indicam que o acúmulo de conhecimento dentro das cidades exerce uma influência estatisticamente positiva sobre os níveis locais de desenvolvimento, apontando que políticas públicas voltadas a ampliar o acesso à educação técnica e superior atuam como indutores sustentáveis de longo prazo da dinâmica municipal (Faveri, Petterini, & Barbosa, 2018; Pereira & Cruz, 2019).

Pesquisas empíricas voltadas a mensurar as transformações socioeconômicas associadas à presença de instituições federais de ensino revelam impactos positivos no mercado de trabalho e nas condições de vida locais. Avaliações de impacto em território brasileiro apontam que, em

nível municipal, a implantação de novos campi de educação profissional e tecnológica (EPT) causam um aumento médio significativo nos patamares de rendimento formal médio per capita e de emprego formalizado. Esses efeitos de expansão socioeconômica são observados de forma mais aberta e com maior amplitude estatística em municípios de menor porte populacional (com menos de 70 mil habitantes), traduzindo-se empiricamente em menores taxas de desocupação regional e na elevação sistemática do Índice de Desenvolvimento Humano (IDH) municipal (Faveri, Petterini, & Barbosa, 2018).

Além de atuar como um vetor dinâmico de crescimento produtivo e de qualificação salarial, o desenho orçamentário estável de um campus técnico federal assume uma função macroeconômica anticíclica de extrema relevância para pequenos territórios do interior brasileiro. Em períodos de recessão econômica generalizada ou de crises fiscais severas do setor público, como o cenário observado após 2016, os pequenos municípios tendem a experimentar retrações relevantes de arrecadação, desinvestimento governamental e redução da arrecadação tributária. Nesse contexto adverso, a estabilidade e a manutenção das despesas de custeio operacional e de folha de pessoal garantidas pelo orçamento federal próprio do campus atuam como um estabilizador macroeconômico, mantendo um fluxo contínuo de demanda agregada que suaviza os efeitos da retração fiscal e amortece os choques recessivos na economia do setor público local em comparação com municípios vizinhos destituídos de tal infraestrutura pública (Pereira & Cruz, 2019). Para compreender melhor tais ações, faz-se necessária uma avaliação dos impactos diretos e indiretos dessa expansão educacional junto a economia local. Ferramentas de análise que utilizem aprendizado de máquina (*Machine Learning*) podem colaborar junto a avaliação destas políticas públicas.

2.2 Avaliação de Políticas Públicas e Modelagem Causal via Aprendizado de Máquina (Machine Learning)

No campo da ciência política, da administração e do monitoramento estatal, uma política pública é caracterizada conceitualmente como um curso estruturado de ação ou inação de natureza governamental, articulado diretamente pelas autoridades estatais com vistas a solucionar problemas coletivos, suprir demandas sociais ou aproveitar oportunidades públicas de interesse comunitário (Vieira, 2023). O fluxo de intervenção do Estado é comumente representado de forma linear pelo ciclo de políticas públicas, o qual se subdivide nas fases sequenciais de diagnóstico e inserção na agenda governamental, formulação conceitual da política, processo decisório e legislação, implementação prática e, por fim, a etapa de avaliação e monitoramento (Vieira, 2023). A fase de avaliação cumpre um papel importante e norteador, tendo por finalidade auditar a eficácia, eficiência e efetividade das alocações de recursos públicos, fornecendo evidências empíricas e causais que justifiquem a continuidade, reformulação metodológica ou descontinuidade das despesas públicas (Vieira, 2023).

Historicamente restrita a metodologias de regressão estática ou estatísticas descritivas simples que enfrentavam fortes limitações para isolar impactos de forma confiável e monitorar dinamicamente os efeitos posteriores, a avaliação de políticas públicas contemporânea vem passando por uma transição disruptiva com a incorporação de abordagens de inteligência artificial e aprendizado de máquina (*machine learning*) (Vieira, 2023). O emprego de algoritmos de aprendizado de máquina possibilita que os avaliadores de políticas públicas lidem com eficiência metodológica diante de bases de dados complexas e de alta dimensionalidade, otimizando o pareamento de covariáveis, identificando de forma satisfatória a heterogeneidade dos efeitos de tratamento em grupos populacionais, simulando cenários contrafactuais

alternativos com baixo custo computacional e processando dados não estruturados de texto ou imagens (Vieira, 2023).

No escopo de avaliações de impacto quase-experimentais em nível municipal ou regional, técnicas avançadas de *machine learning* têm sido aplicadas para superar as deficiências de métodos tradicionais baseadas em mínimos quadrados ordinários, que frequentemente sofrem com viés de especificação ou colinearidade (Vieira, 2023). Entre essas inovações estatísticas causais, destaca-se a modelagem de Séries Temporais Estruturais Bayesianas (BSTS). A abordagem BSTS, desenvolvida originalmente no contexto preditivo e de mercado, combina de forma integrada modelos clássicos de espaço de estados e filtros de Kalman com métodos estatísticos de seleção bayesiana e ponderação preditiva (Brodersen et al., 2015).

A grande inovação metodológica do BSTS reside no uso do algoritmo de seleção de variáveis conhecido como *"Spike-and-Slab"*. Esse algoritmo realiza uma seleção probabilística a posteriori em meio a uma matriz de potenciais variáveis de controle de alta dimensionalidade, identificando preditores robustos sem risco de sobre ajuste (*overfitting*). O modelo calcula iterativamente pesos dinâmicos e fluidos para as unidades de controle, construindo de forma robusta e otimizada o chamado contrafactual sintético. Esse contrafactual simula de forma preditiva qual seria a trajetória temporal da variável de interesse e do mercado local caso a política pública analisada nunca houvesse ocorrido, permitindo deduzir isoladamente, com grau satisfatório de significância estatística e probabilidade a posteriori, o efeito causal líquido provocado pela intervenção pública de forma posterior no território (Brodersen et al., 2015). A seguir apresenta-se o tópico destinado a metodologia que explicará o funcionamento do modelo implementado junto a análise dos dados.

3 Metodologia

Para responder ao objetivo da pesquisa e avaliar o impacto causal da interiorização do Instituto Federal do Rio Grande do Norte (IFRN) sobre o Produto Interno Bruto (PIB) da Administração Pública do município de Lajes (RN), e estudo utilizou uma abordagem quantitativa, de natureza aplicada e com objetivo explicativo. No campo da avaliação de políticas públicas e projetos sociais, a mensuração de impactos compreende análises de resultados focadas na lógica causal que liga uma determinada intervenção aos seus respectivos efeitos (Araujo & Kubrusly, 2023). Para tanto, o desenho de pesquisa caracteriza-se como um quase-experimento apoiado na técnica de Controle Sintético baseado em Séries Temporais Estruturais Bayesianas (BSTS – *Bayesian Structural Time Series*), operacionalizado por meio do algoritmo *CausalImpact*, desenvolvido por Brodersen, Gallusser, Koehler, Remy e Scott (2015). A escolha desse método justifica-se por sua força em inferir relações de causa e efeito em dados longitudinais a partir de uma única unidade tratada sob intervenção e um grupo de não afetados (variáveis de controle), oferecendo uma alternativa cientificamente satisfatória aos desenhos experimentais tradicionais (Abadie, Diamond, & Hainmueller, 2010; Brodersen et al., 2015).

O algoritmo *CausalImpact* configura-se como uma aplicação prática de Inteligência Artificial (IA) voltada para a inferência causal e a econometria avançada, operando sob o escopo do aprendizado de máquina (*machine learning*) probabilístico (Brodersen et al., 2015; Vieira, 2023). A classificação do método como uma tecnologia de IA reside no fato de que o algoritmo realiza um processo autônomo de treinamento e modelagem preditiva a partir de

séries temporais históricas (Brodersen et al., 2015; Varian, 2014). No período pré-intervenção, o algoritmo aplica a priori *Spike-and-Slab* para conduzir uma seleção automatizada de características (*feature selection*), calculando de forma heurística e estocástica quais covariáveis da amostra de controle (bacia de doadores) são preditores robustos para compor o cenário de controle (Brodersen et al., 2015; Varian, 2014). Ao projetar o contrafactual para o período pós-intervenção, a IA utiliza esse aprendizado prévio para gerar previsões *out-of-sample* com intervalos de credibilidade satisfatórios, propiciando um suporte analítico inteligente de elevada precisão para a tomada de decisão e avaliação (Brodersen et al., 2015; Vieira, 2023).

3.1 Unidade de Análise, Recorte Temporal e Critério Amostral de Proximidade

A unidade tratada (alvo) desta pesquisa é o município de Lajes (RN), o qual recebeu a instalação física e o início das operações do campus do IFRN em maio de 2015. Para o isolamento do nexo causal e a mitigação de vieses de dispersão espacial, o grupo de controle (bacia de doadores do controle sintético) foi delimitado com base em cinco cidades limítrofes que possuem maior proximidade física e densidade de conectividade rodoviária com a sede de Lajes, são elas: Pedro Avelino (aproximadamente 20 km), Fernando Pedroza (aproximadamente 28 km), Pedra Preta (aproximadamente 30 km), Caiçara do Rio do Vento (aproximadamente 35 km) e São Tomé (aproximadamente 40 km). Essa delimitação assegura uma satisfatoriedade regional e geográfica, garantindo que os municípios de controle compartilhem um ecossistema climático semelhante (região semiárida), uma proximidade nos valores dos repasses fiscais constitucionais, como o Fundo de Participação dos Municípios (FPM), e dinâmicas de logística equivalentes às do município-alvo. O desenho amostral estabelece uma janela temporal simétrica equilibrada de 14 anos.

O período pré-intervenção de 7 anos de dados históricos anuais utilizados para o treinamento, de 2008 a 2014, utilizou modelagem estrutural e calibração das relações estáveis entre a unidade de tratamento, alvo da pesquisa, e as variáveis de controle (Abadie et al., 2010; Brodersen et al., 2015). Para o período pós-intervenção, entre 2015 e 2021, também compreendeu um intervalor de 7 anos observados após a introdução da instituição de ensino no território, utilizados para projetar a trajetória contrafactual e calcular o efeito incremental atribuído à política pública (Brodersen et al., 2015; Fotia & Ferrara, 2023).

3.2 Protocolo de Aquisição de Dados

Para fins de replicação científica, qualquer pesquisador poderá reconstruir a base de dados original executando atentamente o seguinte passo a passo: acesse o portal oficial do Ipeadata: www.ipeadata.gov.br. No menu principal da página, selecione o módulo Ipeadata Regional. Navegue pelo menu lateral esquerdo clicando em Tema e, em seguida, selecione a pasta Contas Regionais. Na lista central de indicadores, localize e clique sobre o nome da série histórica longa estruturada pelo IBGE, denominada: PIB Municipal (valor adicionado a preços básicos) - serviços - administração, defesa, educação e saúde públicas e seguridade social (preços de 2010). Na janela de configurações e filtros que se abre no centro da tela, configure os parâmetros de extração sob os critérios: na escolha do nível geográfico, marque o correspondente ‘Municípios’; na escolha da abrangência, marque o estado correspondente ‘Rio Grande do Norte’; no intervalo temporal, insira o ano de 2008 no local destinado ao Início e o ano de 2021 no local destinado ao Fim. Após o carregamento da tabela em tela contendo as séries cronológicas de todos os municípios potiguares, realize o download do arquivo de dados

brutos. No caso deste estudo, o documento gerado foi salvo com o nome de arquivo 'pib_publico_rn.xls'.

3.3 Tratamento de Dados e Padronização Algorítmica

Em razão de inconsistências de formatação e da presença de informações ruidosas no arquivo bruto disponibilizado pelo servidor governamental — incluindo caracteres e metadados ocultos de rodapé incorporados aos valores numéricos das séries de controle —, foi desenvolvido um algoritmo de limpeza, padronização e restrição estrutural em linguagem Python 3.8, no ambiente de desenvolvimento Spyder 4.1.5, voltado à ciência de dados. O código completo do algoritmo contempla tanto a limpeza e a organização dos dados quanto a execução do aprendizado de máquina, por meio do modelo *CausalImpact*, e a plotagem dos resultados da análise. Devido ao limite de páginas estabelecido pelo evento Singep, a Tabela 1 apresenta apenas um excerto do algoritmo. A versão integral encontra-se na versão expandida do estudo.

Tabela 1

Código do algoritmo e descrição das ações em cada etapa

Código do Algoritmo	Descrição das Ações Realizadas
…	…
pre_period = ['2008-01-01', '2014-01-01'] post_period = ['2015-01-01', '2021-01-01'] ci = CausalImpact(df_final, pre_period, post_period) res = ci.summary_data	• Execução do aprendizado de máquina contrafactual (algoritmo *CausalImpact*). Executando modelagem estrutural bayesiana com 5 variáveis de controles de maior proximidade.
ci.plot() fig = plt.gcf() axes = fig.get_axes()	• Criação do gráfico do modelo para análise dos resultados.
…	…

Nota. Elaborado pelo autor (2026).

A seguir apresenta-se a tabela contendo o tratamento e a filtragem da cidade alvo (Lajes), bem como das cidades utilizadas como variáveis de controle.

Tabela 2

PIB da variável alvo e das variáveis de controle com valores padronizados

	Valores em R$ multiplicados por 1000					
	Lajes	**Pedro Avelino**	**Fernando Pedroza**	**Pedra Preta**	**Caicara do Rio do Vento**	**Sao_Tome**
2008	27719	20937	97441	10349	98231	30070
2009	29086	21632	10073	98654	10470	31058
2010	29473	20896	96000	10111	10764	30488
2011	30250	21450	10166	10833	11481	31814
2012	31891	21523	10147	10698	12251	32808
2013	33159	22889	11258	10290	13758	36072
2014	33592	22458	10717	10670	13330	35790

2015	33132	22271	10886	10234	13092	34650
2016	32488	21643	10788	94110	12764	33641
2017	33787	22402	12865	10247	12443	34807
2018	32957	22381	11729	10445	12391	37461
2019	34729	23282	12593	10817	12969	36593
2020	34677	23225	12301	10597	13770	34891
2021	31036	23058	11058	10645	13076	32675

Nota. Tabela gerada pelo autor a partir do algoritmo de padronização (2026).

Todas os dados apresentados na tabela acima foram padronizados pelo algoritmo. Os valores anuais apresentados, conforme Ipeadata, devem ser multiplicados por R$ 1000,00 (Um mil reais). O estudo utilizou uma metodologia quase-experimental, esta fundamenta-se no isolamento de impactos causais a partir de abordagens contrafactuais. Em termos operacionais, a inferência causal procura mapear se a intervenção se comporta como uma condição INUS (*Insufficient but Non-redundant part of an Unnecessary but Sufficient condition*), de modo que o impacto seja mensurado pela diferença entre os valores observados e os contrafactuais que, provavelmente, teriam ocorrido na ausência da interferência (Araujo & Kubrusly, 2023). O modelo de Controle Sintético tradicional estabelece que, na indisponibilidade de um grupo de controle experimental perfeitamente idêntico, um contrafactual satisfatório (unidade sintética), pode ser construído por meio de uma combinação linear ponderada de pesos positivos de uma variável alvo sobre outras variáveis de controle, também conhecida como bacia de doadores (*donor pool*), servindo de base matemática contra extrapolações indevidas dos dados (Abadie et al., 2010). Ao contrário das metodologias clássicas de Diferença em Diferenças (DiD) que assumem que os efeitos de variáveis omitidas não observadas são constantes ao longo do tempo, o Controle Sintético estende essa abordagem ao incorporar fatores comuns não observados variantes no tempo (Abadie et al., 2010).

Para contornar limitações do controle tradicional, que fixa pesos imutáveis estabelecidos na fase de pré-tratamento, o modelo de Séries Temporais Estruturais Bayesianas (BSTS) introduz flexibilidade e dinamicidade por meio de técnicas de aprendizado de máquina (*Machine Learning*) e equações de espaço de estados (Brodersen et al., 2015; Varian, 2014). Modelos de aprendizado de máquina otimizam a seleção de covariáveis e a capacidade preditiva *out-of-sample,* ao penalizarem o excesso de complexidade por meio de regularizações específicas (Varian, 2014). Matematicamente, a decomposição do comportamento temporal da unidade tratada é dada pela equação de observação $yt=\mu t+\tau t+\beta Txt+\epsilon t$, onde μt representa a tendência local da série temporal, τt captura as variações sazonais, xt representa as séries temporais dos 5 municípios que compõem as variáveis de controle, βT denota o vetor de coeficientes das covariáveis e ϵt é o erro de medição gaussiano independente e identicamente distribuído com variância $\sigma t2$ (Brodersen et al., 2015; Gianacas et al., 2023). A evolução dos estados latentes obedece à equação de transição de estado $\alpha t+1=Tt\alpha t+Rt\eta t$ em que αt engloba os componentes de tendência e sazonalidade estruturados de forma modular, Tt é a matriz de transição de estado, Rt é a matriz de controle e ηt representa os choques do sistema (Brodersen et al., 2015; Fotia & Ferrara, 2023).

A seleção de características (*feature selection*) e o mapeamento das variáveis explicativas na matriz, são conduzidos utilizando a abordagem de regressão com priori *Spike-and-Slab* (Brodersen et al., 2015; Varian, 2014). Esse método bayesiano atribui uma base priori sobre um indicador probabilístico de inclusão para cada um dos municípios utilizados como variáveis de controle, combinada a uma distribuição normal difusa condicional para os valores

dos coeficientes não-nulos (Varian, 2014). A estimação dos parâmetros do modelo e a amostragem dos estados latentes ocorrem iterativamente através de algoritmos estocásticos de Monte Carlo via Cadeias de Markov (MCMC), empregando o amostrador de Gibbs e algoritmos eficientes de filtragem de Kalman (Brodersen et al., 2015). Essa estrutura computacional gera uma distribuição preditiva *a posteriori* do contrafactual, podemos denominá-la como 'Lajes Sintética', calculando de forma semiparamétrica o efeito causal pontual e acumulado da política, com intervalos de credibilidade satisfatórios (Brodersen et al., 2015; Gianacas et al., 2023).

Cabe ressaltar que a aplicação desse modelo de controle exige que as séries explicativas que compõem as variáveis de controle estejam correlacionadas com a série histórica da unidade de tratamento no período pré-intervenção, e que não tenham sido expostas ou afetadas pela política pública investigada ou por outras intervenções concorrentes (Gianacas et al., 2023). Como o modelo básico do *CausalImpact* assume a ausência de um componente de tendência secular explícito que não esteja capturado nas covariáveis doadoras, a seleção de municípios limítrofes homólogos poderá garantir uma certa estabilidade temporal da relação preditiva necessária para evitar a atribuição errônea de tendências preexistentes ao efeito do tratamento (Gianacas et al., 2023). Diferente dos modelos de regressão de mínimos ordinários ou de Interrupção de Séries Temporais clássicos, que assumem variâncias fixas e erros correlacionados estáticos, o arcabouço BSTS acomoda, com flexibilidade, múltiplas fontes de variabilidade com estruturas de erro independentes para cada componente (Brodersen et al., 2015; Gianacas et al., 2023).

A determinação de um impacto causal em dados observacionais exige a reconstrução rigorosa de um cenário contrafactual satisfatório (Araujo & Kubrusly, 2023; Brodersen et al., 2015). No entanto, os métodos tradicionais de econometria e avaliação de políticas públicas, podem falhar ao lidar com as propriedades dinâmicas e de autocorrelação temporal, comuns em séries cronológicas macroeconômicas (Brodersen et al., 2015). Para robustecer o desenho analítico do estudo, a escolha do BSTS operacionalizado pelo *CausalImpact*, em detrimento de abordagens concorrentes, fundamenta-se ao observar-se o DiD tradicional que se baseia em regressões estáticas que assumem que as observações são independentes e identicamente distribuídas, ignorando a autocorrelação residual de séries temporais (Brodersen et al., 2015).

Quando aplicado a dados seriais correlacionados, o DiD subestima a incerteza e fornece intervalos de confiança excessivamente apertados, o que induz a falsos positivos e a conclusões superotimistas sobre os impactos (Brodersen et al., 2015). Da mesma forma, o DiD exige a hipótese restritiva de tendências paralelas e assume que os efeitos de variáveis omitidas são constantes no tempo (Abadie et al., 2010; Brodersen et al., 2015). O BSTS supera essas barreiras ao integrar equações de espaço de estados que permitem a variação temporal de fatores latentes comuns não observados (Brodersen et al., 2015).

Quando se compara o BSTS com modelos de aprendizado de máquina como Redes neurais de memória de longo prazo bidirecionais, estas redes podem oferecer alta capacidade preditiva para modelar relações não-lineares, mas operam como 'caixas-pretas' de altíssimo custo computacional e necessitam de mecanismos integrados para a propagação exata de incerteza matemática (Fotia & Ferrara, 2023). O BSTS e o *CausalImpact*, por sua vez, combinam flexibilidade preditiva com um rigoroso tratamento probabilístico bayesiano, integrando a incerteza dos parâmetros e dos doadores (variáveis de controle) a cada tempo de simulação, gerando previsões contrafactuais acompanhadas de uma distribuição de

probabilidade a posteriori sobre a certeza de um impacto satisfatório identificado no estudo (Brodersen et al., 2015; Fotia & Ferrara, 2023).

No escopo de problemas complexos de administração e de políticas públicas, a tomada de decisão estratégica e o controle orçamentário são invariavelmente afetados por dinâmicas de imprevisibilidade e incertezas derivadas de ambientes organizacionais e macroeconômicos em rápida transição (Silva, 2024). Nesse sentido, as Redes Bayesianas (RB) despontam como ferramentas analíticas e probabilísticas de Inteligência Artificial (IA) essenciais para modelar e gerenciar a complexidade das relações entre várias variáveis (Silva et al., 2021; Silva, 2024). Ao contrário da estatística clássica e de modelos tradicionais, a modelagem baseada em Redes Bayesianas permite associar de forma flexível conhecimentos subjetivos, opiniões especializadas de gestores e dados secundários ruidosos a partir do Teorema de Bayes, estimando com satisfação matemática distribuições de probabilidades combinadas, mesmo sob cenários de dados incompletos ou parcialmente inacessíveis por segredos comerciais e restrições de tempo (Carvalho & Chiann, 2013; Silva, 2024).

Uma das prerrogativas metodológicas de maior valor prático das RBs repousa em sua capacidade de operar cálculos bidirecionais (raciocínio preditivo *top-down* e raciocínio diagnóstico *bottom-up*), viabilizando simular caminhos, prever o impacto gerencial de perturbações ambientais ou retroalimentar as probabilidades das causas prováveis a partir da observação de evidências e disparos nos nós filhos descendentes. Para mais, a realização de análises de sensibilidade diagnóstica mapeia como a incerteza do nó de desfecho se conecta à incerteza dos nós de entrada, calculando a redução de variância para priorizar ações preventivas e mitigadoras de riscos (Nobre et al., 2023; Silva, 2024). A conexão teórica entre o paradigma de Redes Bayesianas e o arcabouço metodológico do Controle Sintético Bayesiano (BSTS) é direta (Silva, 2024). Ambos os métodos se assentam sobre a premissa de que a incerteza do mundo real não deve ser ignorada ou reduzida a estimativas pontuais determinísticas de regressões lineares clássicas, mas sim observada de maneira transparente ao longo de todo o sistema estatístico (Brodersen et al., 2015; Silva, 2024).

Enquanto as RBs geram modelos gráficos de alta compreensão analítica sobre dados complexos (Silva, 2024). O BSTS decompõe séries longitudinais por meio de equações de espaço de estados, as quais assumem uma representação gráfica e temporal similar a uma rede de crenças dinâmicas, e utiliza a priori *Spike-and-Slab* para selecionar as covariáveis mais influentes sobre a predição contrafactual. Assim, fundamentar as estimativas causais da expansão do IFRN Campus Lajes sob o paradigma analítico bayesiano, fornece aos pesquisadores um framework unificado, visualmente intuitivo, e matematicamente satisfatório, de combinação de dados e suporte à decisão, frente aos riscos e excitações de escala macroeconômica (Brodersen et al., 2015; Nobre et al., 2023; Silva, 2024).

4 Análise dos resultados e Discussões

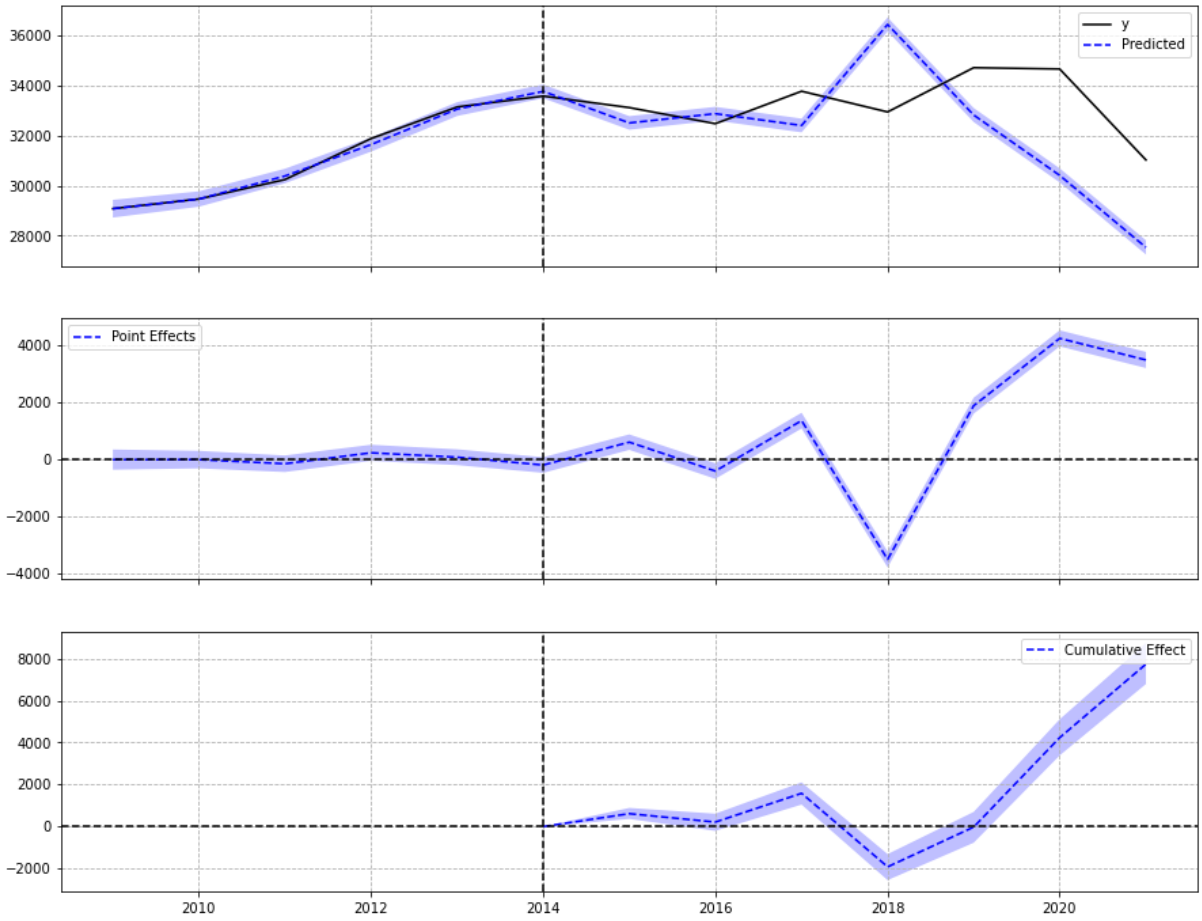
A avaliação do impacto da implantação do campus do Instituto Federal do Rio Grande do Norte (IFRN) em Lajes (RN), ocorrida em maio de 2015, foi realizada por meio da modelagem de Séries Temporais Estruturais Bayesianas (BSTS) aplicada à série histórica do Produto Interno Bruto (PIB) da administração pública municipal de 2008 a 2021. O comportamento macroeconômico observado na unidade tratada foi confrontado com um contrafactual sintético denominado ‘Lajes Sintética’, gerado pelo aprendizado de máquina, a partir das trajetórias econômicas das cinco cidades que compõem as variáveis de controle (Pedro Avelino, Fernando Pedroza, Pedra Preta, Caiçara do Rio do Vento e São Tomé). Esse

quase-experimento permitiu isolar a conexão causal da política de interiorização da rede federal sobre a dinâmica do setor público local, controlando fatores latentes comuns que afetaram a região semiárida do estado (Brodersen et al., 2015; Faveri, Petterini, & Barbosa, 2018).

Os resultados agregados apontam uma expansão estatisticamente significante e sistemática no PIB do setor público de Lajes após a inauguração do campus. Durante a janela de pós-intervenção (2015 a 2021), o PIB médio anual observado da administração pública de Lajes foi de R$ 33.258.000,00, ao passo que a predição contrafactual calculada para o mesmo período, sem a presença do IFRN, estimou uma média de R$ 32.153.000,14. Essa divergência sistemática revela um impacto causal positivo médio de +R$ 1.104.000,86 mil por ano injetados diretamente na economia pública local, o que equivale a uma expansão percentual líquida de aproximadamente 3,44%, atribuída à instalação e operação do campus federal. O modelo bayesiano demonstrou vigor satisfatório, registrando um *p-valor* estatístico inferior a 0,001 e uma probabilidade a posteriori de existência de efeito de 100,0%, confirmando que a expansão econômica observada é fortemente improvável de ser atribuída a flutuações casuais ou ruídos aleatórios (Brodersen et al., 2015; Fotia & Ferrara, 2023).

Figura 1

Decomposição Estrutural e Inferência Causal do Impacto do IFRN Campus Lajes no PIB do Setor Público Municipal (2008–2021).



Nota. Plotagem gerada pelos autores a partir do algoritmo *CausalImpact* (2026).

Para compreender a evolução dinâmica desse impacto ao longo do tempo e as estimativas do modelo, faz-se necessário analisar cada um dos três gráficos, bem como os seus

respectivos elementos apresentados na Figura 1. O gráfico superior (y vs. Predicted) ilustra as trajetórias temporais do PIB do setor público de Lajes. A linha preta contínua (y) representa os dados reais observados do município-alvo, enquanto a linha azul tracejada (Predicted) projeta o cenário contrafactual da 'Lajes Sintética', acompanhado por sua área sombreada azul-claro, que demarca o intervalo de credibilidade bayesiano de 95% da predição. A linha preta tracejada vertical posicionada em 2014 estabelece o marco divisor temporal do estudo. No período pré-intervenção (2008 a 2014), observa-se um alinhamento quase perfeito entre a linha preta e a azul, o que demonstra a alta conexão do algoritmo *Spike-and-Slab* em aprender e calibrar o relacionamento estável entre Lajes e seu entorno antes do tratamento (Brodersen et al., 2015; Varian, 2014). A partir de 2015, as duas linhas se separam de maneira expressiva, evidenciando o descolamento da trajetória real de Lajes em relação à sua contraparte sintética.

O gráfico central (*Point Effects*) exibe a diferença líquida pontual anual (efeito de tratamento ponto a ponto), obtida pela equação, onde a linha azul tracejada (*Point Effects*) oscila em torno da linha de referência horizontal tracejada em 0 (efeito nulo). A área sombreada azul-claro representa os limites do intervalo de credibilidade de 95%. Entre 2008 e 2014, o efeito pontual permanece colado ao zero, atestando a ausência de vieses pré-tratamento. A partir de 2015, ocorrem oscilações dinâmicas de interesse analítico. Em 2015, registra-se um impacto imediato de aproximadamente +R$ 632,00 x mil, seguido por uma leve retração para -R$ 312,00 x mil em 2016 e um expressivo ganho de +R$ 1.387,00 x mil em 2017. Em 2018, o efeito pontual declina temporariamente para -R$ 3.443,00 x mil, mas recupera-se de forma vigorosa nos anos seguintes, atingindo impactos positivos expressivos de +R$ 1.929,00 x mil em 2019, +R$ 4.177,00 x mil em 2020 e +R$ 3.536,00 x mil em 2021.

O gráfico inferior (*Cumulative Effect*) consolida a trajetória da soma acumulada desses impactos anuais a partir do marco de 2015. A linha azul tracejada (*Cumulative Effect*) e seu respectivo intervalo de credibilidade de 95%, revelam que apesar de oscilações pontuais negativas intermediárias decorrentes de choques macroeconômicos regionais, ou até mesmo nacionais, o ganho de escala na economia do setor público de Lajes é progressivo e cumulativo. Ao término do horizonte temporal de análise, em 2021, o município de Lajes acumulou um incremento financeiro líquido real de aproximadamente R$ 7.734,00 x mil (cerca de R$ 7,7 milhões constantes) adicionados à sua economia governamental. Esse resultado acumulado espelha com solidez o impacto cumulativo sustentado de longo prazo gerado pela fixação da infraestrutura educacional e científica federal no interior potiguar (Faveri et al., 2018; Pereira & Cruz, 2019).

Do ponto de vista econômico e gerencial, o desmembramento desses resultados revela dois mecanismos causais distintos atuando na economia do setor público de Lajes. O primeiro mecanismo, pautado na dinâmica da demanda e de curto prazo, é o 'efeito-gasto'. A instalação do campus IFRN Lajes representou um fluxo monetário contínuo e estável injetado no comércio de bens e na prestação de serviços locais por meio do pagamento de salários à massa permanente de servidores públicos federais (docentes e técnicos estáveis), além de contratações locais e licitações para serviços de custeio, logística, segurança e alimentação do campus (Faveri et al., 2018; Pereira & Cruz, 2019). Esse encadeamento ativou o consumo privado intramunicipal e impulsionou a dinâmica econômica do setor público por meio do fortalecimento indireto da arrecadação de impostos e da retroalimentação das receitas do município.

O segundo mecanismo, associado à resiliência macroeconômica anticíclica, explica com clareza a variação dos dados a partir de 2018. Com a severa recessão fiscal brasileira observada após 2016, os pequenos municípios do semiárido sofreram contrações profundas de receitas

públicas e repasses constitucionais (Marenco, Strohschoen, & Joner, 2017; Pereira & Cruz, 2019). Em 2018, enquanto as cidades de controle da bacia de doadores (variáveis de controle), em especial São Tomé cujos recursos públicos saltaram atipicamente em termos reais, projetaram uma expectativa de crescimento para o contrafactual 'Lajes Sintética' para patamares de R$ 36,4 milhões, o PIB real de Lajes sofreu uma retração momentânea em termos reais para R$ 32,9 milhões. Isso explica a queda pontual observada no gráfico. Contudo, em 2020 e 2021, com a eclosão da pandemia de COVID-19 e o consequente colapso financeiro das administrações municipais vizinhas, o contrafactual despencou abruptamente para R$ 27,5 milhões em 2021 devido à retração generalizada dos municípios de controle (Gianacas et al., 2023; Pereira & Cruz, 2019).

Nesse momento de crise, a estabilidade do orçamento próprio federal do IFRN Campus Lajes atuou de forma decisiva como um 'colchão amortecedor' e estabilizador de demanda agregada (Pereira & Cruz, 2019). A manutenção das despesas públicas federais estáveis amorteceu o choque recessivo local, mantendo o PIB público observado de Lajes em patamares significativamente superior (R$ 31.036,00 x mil em 2021) em relação à ruína financeira projetada caso o município dependesse unicamente de suas receitas ordinárias (R$ 27.500,00 x mil observados na 'Lajes Sintética'). Esse efeito estabilizador anticíclico corrobora a literatura que aponta que investimentos em educação técnica federal funcionam como importantes vetores de blindagem macroeconômica e desenvolvimento regional para pequenos municípios do interior do país (Faveri et al., 2018; Pereira & Cruz, 2019; Pozzer, 2024).

5 Conclusões/Considerações finais

A avaliação quase-experimental do impacto da implantação do campus do Instituto Federal do Rio Grande do Norte (IFRN) no município de Lajes-RN, em maio de 2015, demonstra que a interiorização da rede federal de educação profissional, científica e tecnológica atua de forma decisiva no fortalecimento socioeconômico de pequenos municípios do interior do país. O estudo conclui que a introdução do campus gerou um incremento causal estatisticamente significante na economia do setor público local, registrando um ganho médio anual de aproximadamente +R$ 1.104,86 x mil e uma expansão percentual aproximada de 3,44% no PIB de serviços da administração pública, em relação ao contrafactual estimado 'Lajes Sintética', acumulando um ganho real de aproximadamente R$ 7.734,00 x mil ao final de 2021. Esse resultado valida o 'efeito-gasto' imediato e de curto prazo da expansão educacional, alimentado pela injeção contínua de recursos federais na economia local por meio da massa de salários estáveis pagos aos servidores e dos contratos de custeio administrativo e operacional firmados junto ao mercado regional. O comportamento da série histórica frente à predição contrafactual confirma o papel macroeconômico anticíclico da instituição educacional federal, a qual, provavelmente, funcionou como um 'colchão amortecedor' de sustentação de demanda agregada capaz de mitigar a forte contração econômica que atingiu o setor público municipal durante a recessão fiscal de 2018, e o cenário pandêmico de 2020 e 2021.

Com relação as contribuições desta pesquisa, estas desdobram-se em teóricas, práticas e metodológicas para o campo da gestão e das políticas públicas brasileiras. No âmbito teórico, a pesquisa reforça as teorias do desenvolvimento regional, evidenciando como a inserção geográfica de uma infraestrutura científica de excelência pode gerar desenvolvimento extra-acadêmico capaz de integrar as dimensões econômicas, sociais e territoriais. Sob o prisma prático e gerencial, as evidências empíricas fornecidas justificam a relevância estratégica da manutenção e expansão do orçamento da rede federal, fornecendo subsídios baseados em dados para que formuladores de políticas possam defender a interiorização educacional como uma

política social de descentralização e fixação de capital humano qualificado em pequenos centros populacionais. Metodologicamente, o estudo inova ao propor a convergência entre os modelos de Séries Temporais Estruturais Bayesianas (BSTS) com *CausalImpact*. Essa união metodológica demonstra que o emprego de abordagens probabilísticas de aprendizado de máquina e modelagens baseadas no paradigma bayesiano podem viabilizar avaliações de políticas públicas de maneira satisfatória, transparentes e interpretáveis, superando as tradicionais limitações de variância, colinearidade e incapacidade de propagação de incerteza que acometem outras técnicas econométricas.

Contudo, a pesquisa apresenta limitações de caráter empírico e metodológico que precisam ser destacadas. Com relação à base de dados de entrada, a principal limitação repousa no fato de que a análise se baseou estritamente na série histórica do PIB setorial da administração, defesa, educação e saúde públicas e seguridade social fornecida pelo IBGE. Esse recorte de dados setoriais do setor público impede a captura direta de potenciais efeitos de médio e longo prazo sobre a dinâmica econômica e produtiva do mercado de trabalho privado ou sobre o acúmulo de conhecimento das empresas locais. O uso de outras bases de dados complementares como o PIB municipal total (incluindo o valor adicionado de indústrias, comércio e agropecuária privados), dados tributários municipais de arrecadação própria (como ISS e ICMS) e registros de empregabilidade formalizados do Cadastro Geral de Empregados e Desempregados (CAGED), seria fundamental para corroborar, estender ou refutar a satisfatoriedade dos impactos causais identificados na economia municipal.

Sob a ótica metodológica, o modelo de Séries Temporais Estruturais Bayesianas operacionalizado via algoritmo *CausalImpact* repousa sobre premissas estatísticas que podem impor restrições à validade do estudo. Uma premissa crítica é a ausência de efeitos de transbordamento (*spillover effects*) da unidade tratada de Lajes em direção as variáveis de controle (Pedro Avelino, Fernando Pedroza, Pedra Preta, Caiçara do Rio do Vento e São Tomé). Uma vez que estes municípios vizinhos se encontram em extrema proximidade rodoviária (em um raio aproximado de até 40 quilômetros), é provável que parte do contingente estudantil ou trabalhador dessas localidades migre ou se desloque diariamente para usufruir dos serviços técnicos e científicos do novo campus do IFRN Lajes. Caso essa dinâmica espacial tenha estimulado e expandido indiretamente o desenvolvimento econômico ou a dinâmica de gastos dos municípios de controle, a predição da 'Lajes Sintética' estaria superestimada, resultando em uma atenuação artificial do real efeito causal líquido calculado para o município de Lajes.

A partir das contribuições e das limitações identificadas neste estudo quase-experimental, sugerem-se alguns caminhos para o desenvolvimento de futuras pesquisas na área de avaliação de impactos econômicos e sociais da expansão educacional. Recomenda-se expandir a base de dados empírica da pesquisa por meio do mapeamento e da incorporação de dados oficiais de emprego formal (RAIS/CAGED) e de arrecadação fiscal municipal para analisar se o incremento de demanda do setor público de fato se propaga para o mercado de trabalho privado e para a eficiência fiscal da gestão urbana. Metodologicamente, sugere-se a aplicação de extensões do modelo de controle sintético capazes de avaliar múltiplos tratamentos simultâneos ou, complementarmente, a estruturação de Redes Bayesianas dinâmicas alimentadas por simulações estocásticas para mapear graficamente a probabilidade condicional de transição e os caminhos causais diretos entre os investimentos em educação profissional e as variações das cadeias produtivas locais.

Por fim, propõe-se a realização de pesquisas empíricas de natureza qualitativa ou de métodos mistos que coletem percepções primárias de discentes, egressos, comerciantes e

lideranças sociais do território de Lajes, permitindo capturar os impactos associados ao crescimento ou retração da economia local, ao fortalecimento cultural e às transformações do bem-estar comunitário, promovidas pela fixação do campus no município.

Referências Bibliográficas

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, *105*(490), 493-505. https://doi.org/10.1198/jasa.2009.ap08117

Araujo, F. B., & Kubrusly, R. S. (2023). 'Avaliações de impactos': pensamento contrafactual e inferência causal, controles experimentais e estudos observacionais. *Ensaio Acadêmico (HCTE/UFRJ)*, 324-331.

Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., & Scott, S. L. (2015). Inferring causal impact using Bayesian structural time-series models. *The Annals of Applied Statistics*, *9*(1), 247-274. https://doi.org/10.1214/14-AOAS788

Carter, T. J., & Rostova, E. (2023). Higher education expansion and local labor market dynamics: A counterfactual analysis. *Journal of Public Economics*, *222*, Article 104910. https://doi.org/10.1016/j.jpubeco.2023.104910

Carvalho, J. V. d. F., & Chiann, C. (2013). Redes Bayesianas: um método para avaliação de interdependência e contágio em séries temporais multivariadas. *Revista Brasileira de Economia*, *67*(2), 227-243. https://doi.org/10.1590/s0034-71402013000200003

Cichaczewski, J. C., & Castro, C. A. (2023). IF's, uma história por fazer-se e os caminhos da pesquisa. *Revista Brasileira da Educação Profissional e Tecnológica*, *2*(23), Article e12014. https://doi.org/10.15628/rbept.2023.12014

Faveri, D., Petterini, F., & Barbosa, M. P. (2018). Uma avaliação do impacto da política de expansão dos Institutos Federais nas economias dos municípios brasileiros. *Planejamento e Políticas Públicas*, (50), 125–144.

Fotia, P., & Ferrara, M. (2023). A different approach for causal impact analysis on Python with Bayesian Structural Time-Series and Bidirectional LSTM models. *Atti della Accademia Peloritana dei Pericolanti. Classe di Scienze Fisiche, Matematiche e Naturali*, *101*(2), A12. https://doi.org/10.1478/AAPP.1012A12

Gianacas, C., Liu, B., Kirk, M., Di Tanna, G. L., Belcher, J., Blogg, S., & Muscatello, D. J. (2023). Bayesian structural time series, an alternative to interrupted time series in the right circumstances. *Journal of Clinical Epidemiology*, *163*, 102-110. https://doi.org/10.1016/j.jclinepi.2023.10.003

Lee, M. H., & Kim, J. Y. (2024). The spatial spillover effects of higher education expansion on regional economic growth. *Papers in Regional Science*, *103*(2), 201-218. https://doi.org/10.1016/j.pirs.2024.100010

Marenco, A., Strohschoen, M. T. B., & Joner, W. (2017). Capacidade estatal, burocracia e tributação nos municípios brasileiros. *Revista de Sociologia e Política*, *25*(64), 3–21. https://doi.org/10.1590/1678-987317256401

Nobre, O. F., Silva, L. A. D., Costa Júnior, J. F. D., Cortez, A. E. G., & Araújo, A. G. D. (2023). Impact of imports and exports on employment generation and corporate operations: Application of a Bayesian network in the textile sector. *IOSR Journal of Business and Management*, *25*(12), 46-56. https://doi.org/10.9790/487X-2512064656

Pereira, A. S., & Cruz, B. D. (2019). A Rede Federal de Educação Profissional e Tecnológica e a sua contribuição para o desenvolvimento nacional e regional. *HOLOS*, *4*, Article e7992. https://doi.org/10.15628/holos.2019.7992

Pozzer, M. R. O. (2024). O desenvolvimento regional como eixo estruturante dos Institutos Federais de Educação, Ciência e Tecnologia. *Prometeica - Revista de Filosofía y Teoría de la Ciencia*, (29), 1-18. https://doi.org/10.34024/prometeica.2024.29.15280

Pozzer, M. R. O., & Neuhold, R. R. (2019). A educação profissional no Brasil: a Rede Federal de Educação Profissional, Científica e Tecnológica (1909-2018). In M. Pozzer & R. Neuhold (Orgs.), *O contexto da educação técnica na América Latina e os dez anos dos Institutos Federais (2008-2018)* (cap. 2, pp. 23-47). Café com Sociologia.

Silva, A. L. N., Nonato, R. S., & Spink, P. K. (2023). A new proposal for interaction with the territory: origin, potential, and challenges of Brazilian Federal Institutes of Education. *Revista Educación Superior y Sociedad*, *35*(1), 497-520. https://doi.org/10.54674/ess.v35i1.668

Silva, L. A. (2024). *Determinantes da expansão internacional em organizações de rápido crescimento: modelagem bayesiana dos fatores de influência* (Tese de doutorado). Universidade Federal do Rio Grande do Norte, Natal, RN, Brasil.

Silva, L. A., Cortez, A. E. G., Sousa Neto, M. V., & Araújo, A. G. (2021). Bayesian networks as support to project management: Optimization of decision-making of a project developed through the Life Cycle Canvas model. *European Journal of Scientific Research*, *158*(4), 274–287.

Teixeira, T., & Gianezini, K. (2021). The expansion of the Federal Institute of Santa Catarina from a social policy perspective. *Revista Macambira*, *5*(1), Article e051002. https://doi.org/10.35642/rm.v5i1.510

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, *28*(2), 3-28. https://doi.org/10.1257/jep.28.2.3

Vieira, G. A. (2023). *Uso de aprendizado de máquina na avaliação de políticas públicas: uma revisão de escopo* (Dissertação de mestrado). Escola Nacional de Administração Pública (Enap), Brasília, DF, Brasil.