

Daniel Cavalli

**Impacto Localizado de Novas Estações
Meteorológicas na Produtividade Agrícola: Uma
Abordagem de Tratamento Deslocado e
Emparelhamento Dinâmico**

Rio de Janeiro

2025

Daniel Cavalli

Impacto Localizado de Novas Estações Meteorológicas na Produtividade Agrícola: Uma Abordagem de Tratamento Deslocado e Emparelhamento Dinâmico

Monografia apresentada ao Instituto de Economia da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários à obtenção do título de Bacharel em Ciências Econômicas.

Universidade Federal do Rio de Janeiro – UFRJ

Instituto de Economia

Graduação em Ciências Econômicas

Orientador: Prof. Romero Rocha

Rio de Janeiro

2025

Daniel Cavalli

Impacto Localizado de Novas Estações Meteorológicas na Produtividade Agrícola: Uma Abordagem de Tratamento Deslocado e Emparelhamento Dinâmico

Monografia apresentada ao Instituto de Economia da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários à obtenção do título de Bacharel em Ciências Econômicas.

Trabalho aprovado. Rio de Janeiro, 6 de setembro de 2025:

Prof. Romero Rocha
Orientador

Professor
Convidado 1

Professor
Convidado 2

Rio de Janeiro
2025

Este trabalho é dedicado aos agricultores brasileiros, que diariamente enfrentam os desafios da variabilidade climática na busca por produzir alimentos para nossa nação.

Agradecimentos

*“In mathematics you don’t understand things.
You just get used to them.”*
(John von Neumann)

Resumo

Este estudo examina o impacto causal da instalação de estações meteorológicas automáticas sobre o PIB agropecuário no Brasil, utilizando dados em painel de microrregiões produtoras de cana-de-açúcar entre 2003 e 2023. Empregamos o arcabouço de Diferenças em Diferenças com adoção escalonada proposto por Callaway e Sant'Anna (2021), adequado para contextos onde o tratamento ocorre em diferentes momentos do tempo. A estratégia de identificação explora a variação temporal e geográfica na instalação de estações entre microrregiões eventualmente tratadas, utilizando o método "not-yet-treated" como grupo de controle dinâmico. Os resultados principais, obtidos através do estimador doubly robust, indicam um efeito médio do tratamento (ATT) de 8,3% (IC 95%: [2,1%; 14,4%], $p = 0,009$), representando ganhos econômicos significativos no setor agropecuário. A análise de event study revela ausência de tendências pré-tratamento diferenciadas (validando o pressuposto de identificação) e uma dinâmica pós-tratamento crescente. Testes extensivos de robustez, incluindo placebos com PIB não-agropecuário (ATT não significativo), placebos temporais e aleatórios, múltiplas especificações alternativas (IPW: 9,5%, REG: 7,2%), análise de heterogeneidade regional, e diagnóstico de pesos de adoção escalonada, confirmam a consistência e especificidade dos resultados ao setor agrícola. A magnitude estimada sugere que o investimento em infraestrutura meteorológica oferece retornos substanciais, com implicações importantes para políticas de desenvolvimento agrícola e adaptação climática. O estudo contribui para a literatura sobre tecnologia agrícola e informação, demonstrando empiricamente como dados meteorológicos precisos podem melhorar significativamente a eficiência produtiva no setor agrícola.

Palavras-chave: estações meteorológicas. PIB agropecuário. cana-de-açúcar. diferenças em diferenças escalonada.

Abstract

This study examines the causal impact of automatic weather station installation on agricultural GDP in Brazil, using panel data from sugarcane-producing microregions between 2003 and 2023. We employ the staggered adoption Differences-in-Differences framework proposed by Callaway and Sant'Anna (2021), suitable for contexts where treatment occurs at different points in time. The identification strategy exploits temporal and geographic variation in station installation among eventually treated microregions, using the "not-yet-treated" method as a dynamic control group. The main results, obtained through the doubly robust estimator, indicate an average treatment effect (ATT) of 8.3% (95% CI: [2.1%; 14.4%], $p = 0.009$), representing economically significant gains in agricultural GDP. The event study analysis reveals absence of differential pre-treatment trends (validating the identification assumption) and an increasing post-treatment dynamic. Extensive robustness tests, including non-agricultural GDP placebos (non-significant ATT), temporal and random placebos, alternative specifications (IPW: 9.5%, REG: 7.2%), regional heterogeneity analysis, and staggered adoption weight diagnostics, confirm the consistency and specificity of results to the agricultural sector. The estimated magnitude suggests that investment in meteorological infrastructure offers substantial returns, with important implications for agricultural development policies and climate adaptation. The study contributes to the literature on agricultural technology and information, empirically demonstrating how precise meteorological data can significantly improve productive efficiency in the agricultural sector.

Keywords: weather stations. agricultural productivity. sugarcane. differences-in-differences. Callaway and Sant'Anna.

Lista de ilustrações

Figura 1 – Estudo de Evento - Dinâmica Temporal dos Efeitos da Instalação de Estações Meteorológicas	31
Figura 2 – Tendências Paralelas - PIB Agropecuário Normalizado (2003-2023) . .	33
Figura 3 – Composição Dinâmica do Efeito do Tratamento	34
Figura 4 – Distribuição dos Efeitos Placebo vs. Efeito Verdadeiro	35
Figura 5 – Análise de Robustez - Comparação de Especificações	37
Figura 6 – Análise de Poder Estatístico por Tamanho de Efeito	39
Figura 7 – Distribuição Temporal da Instalação de Estações Meteorológicas	49
Figura 8 – Evolução Temporal do PIB Agropecuário por Status de Tratamento . .	52
Figura 9 – Matriz de Correlação das Variáveis Principais	53

Lista de tabelas

Tabela 1 – Resultados Principais e Testes Placebo	30
Tabela 2 – Comparação de Estimativas por Grupo de Controle	34
Tabela 3 – Análise de Sensibilidade ao Período de Análise	36
Tabela 4 – Comparação de Métodos de Estimação	36
Tabela 5 – Número de microrregiões tratadas por ano	50
Tabela 6 – Distribuição do tratamento por região	50
Tabela 7 – Produtividade média (ton/ha) por período	50
Tabela 8 – Diagnóstico de Balanceamento - Estatísticas Detalhadas	51
Tabela 9 – Teste de diferenças pré-tratamento	51

Sumário

1	INTRODUÇÃO	13
1.1	O papel da agrometeorologia no desenvolvimento agropecuário	13
1.2	O papel da econometria	14
2	OBJETIVOS	16
3	REFERENCIAL TEÓRICO	17
4	ESPECIFICAÇÃO DO MODELO	19
4.1	Introdução ao Modelo	19
4.2	Fundamentos do modelo	19
4.2.1	Group-Time Average Treatment Effects ATT(g,t)	20
4.2.2	Identificação	20
4.2.2.1	Validade dos Pressupostos no Contexto de Estações Meteorológicas	21
4.2.3	Estimação	21
4.2.4	Agregação de Efeitos	22
4.2.4.1	Agregação Simples com Pesos Positivos	22
4.2.4.2	Efeito Médio do Tratamento sobre os Tratados (Recomendado)	22
4.2.4.3	Agregações para Event Studies	23
4.3	Concluindo	23
5	METODOLOGIA E RESULTADOS	24
5.1	Estratégia Empírica	24
5.1.1	Definição do Tratamento e Unidades de Análise	24
5.1.2	Construção dos Grupos de Tratamento	24
5.1.3	Variável Dependente e Transformações	25
5.1.4	Covariáveis e Especificação do Modelo	25
5.1.5	O Estimador Duplamente Robusto	26
5.1.6	Escolha do Grupo de Controle	26
5.2	Especificação da Análise de Estudo de Evento	27
5.2.1	Formalização da Análise de Estudo de Evento	27
5.3	Implementação Computacional	28
5.3.1	Software e Pacotes Utilizados	28
5.3.2	Estrutura dos Dados e Tratamento de Missings	29
5.4	Resultados Principais	30
5.4.1	Efeito Médio do Tratamento	30

5.4.2	Análise de Estudo de Evento e Dinâmica Temporal	30
5.4.2.1	Período Pré-Tratamento: Validação das Tendências Paralelas	31
5.4.2.2	Dinâmica Pós-Tratamento: Difusão Gradual dos Benefícios	32
5.4.2.3	Análise Detalhada de Tendências Paralelas	32
5.5	Testes de Robustez e Diagnósticos	33
5.5.1	Composição Dinâmica do Tratamento	33
5.5.2	Comparação de Grupos de Controle	34
5.5.3	Testes Placebo	34
5.5.4	Análise de Sensibilidade ao Período	35
5.5.5	Especificações Alternativas	36
5.6	Discussão e Interpretação Econômica	37
5.6.1	Magnitude e Relevância Econômica	37
5.6.2	Mecanismos Subjacentes	38
5.6.3	Análise de Robustez Expandida	38
5.6.4	Análise de Poder Estatístico	39
5.6.5	Procedimentos de Inferência	39
5.6.6	Limitações e Pesquisa Futura	40
5.6.6.1	Desbalanceamento de Covariáveis	40
5.6.6.2	Composição dos Pesos no Estimador Agregado	40
5.6.6.3	Outras Limitações Importantes	41
5.6.6.4	Direções para Pesquisa Futura	41
5.7	Síntese e Conclusões da Análise Empírica	42
6	CONCLUSÕES FINAIS	43
	REFERÊNCIAS	44
	APÊNDICES	46
	APÊNDICE A – CÓDIGO DO GITHUB	47
	ANEXOS	48
	ANEXO A – ESTATÍSTICAS DESCRITIVAS COMPLEMENTARES	49
A.1	Distribuição Temporal do Tratamento	49
A.2	Estatísticas por Região	51
A.3	Produtividade Média por Status de Tratamento	51
A.4	Teste de Balanceamento de Covariadas	51
A.5	Análise de Sensibilidade - Dados Faltantes	51

A.6	Análises Descritivas Complementares	52
------------	--	-----------

1 Introdução

A produtividade agrícola é profundamente influenciada pelas condições meteorológicas, que moldam tanto o desenvolvimento fisiológico das culturas quanto a eficácia das práticas de manejo. MONTEIRO (2009) destacam que a variabilidade da produção agrícola global é amplamente explicada pelas oscilações climáticas durante o ciclo de cultivo, o que reforça a importância de sistemas robustos de monitoramento climático. Nesse contexto, a agrometeorologia surge como uma ferramenta estratégica, permitindo a integração de dados meteorológicos e agrícolas para apoiar decisões mais eficientes e sustentáveis. Diante da pressão crescente por maior produção de alimentos e energia renovável, sem ampliar o uso de recursos naturais, torna-se essencial adotar instrumentos capazes de reduzir riscos climáticos e ampliar a resiliência produtiva do setor agrícola.

1.1 O papel da agrometeorologia no desenvolvimento agropecuário

A agrometeorologia desempenha um papel crucial na agricultura ao fornecer informações meteorológicas aplicadas diretamente às necessidades dos cultivos. Esse campo integra dados climáticos e meteorológicos com parâmetros específicos das culturas, permitindo a antecipação dos efeitos do clima sobre as práticas agrícolas e possibilitando decisões mais informadas e eficientes. Como apontado por RIJKS e BARADAS (2000), os Serviços Nacionais de Meteorologia contribuem significativamente para a economia agrícola ao divulgar essas informações e facilitar seu uso eficiente, ajudando a mitigar riscos e aumentar a produtividade.

Segundo MAVI e TUPPER (2004), as informações agrometeorológicas são aplicáveis em três áreas principais: no planejamento agrícola, na tomada de decisões táticas e na resiliência dos sistemas agrícolas. No planejamento, esses dados ajudam na escolha das épocas e locais mais adequados para o cultivo, considerando o macroclima e as condições específicas de cada região. Essa etapa é essencial para ajustar as atividades agrícolas ao contexto climático regional, reduzindo desperdícios e promovendo o uso sustentável dos recursos.

No contexto tático, as informações meteorológicas auxiliam na determinação dos melhores momentos para práticas agrícolas como a irrigação, semeadura e colheita, o que contribui para uma execução mais precisa e eficiente das operações. Essas decisões são ainda mais importantes em áreas de cultivo de sequeiro, onde a dependência da precipitação é alta, e os dados sobre previsão de chuva e evapotranspiração são fundamentais para otimizar o uso dos recursos hídricos (PEREIRA; ANGELOCCI; SENTELHAS, 2002).

Com a instalação de novas estações meteorológicas em áreas rurais, as informações climáticas se tornam mais precisas e localizadas, permitindo que os agricultores ajustem suas práticas de acordo com as condições específicas de suas regiões (WEISS; VAN CROWDER; BERNARDI, 2000). Sistemas de Informação Agrometeorológica, como o AGRITEMPO da EMBRAPA e o SISDAGRO do INMET, utilizam dados dessas estações meteorológicas para fornecer previsões de safra e orientações sobre manejo de recursos hídricos, ajudando os agricultores a tomar decisões informadas sobre a época de plantio, irrigação e controle fitossanitário (WEISS; VAN CROWDER; BERNARDI, 2000). Essas informações são cruciais para o presente estudo, pois evidenciam que dados meteorológicos detalhados e acessíveis podem impactar diretamente a produtividade agrícola.

Estudos recentes reforçam a importância da informação climática para a produtividade agrícola. ORTIZ-BOBEA, KNIPPENBERG e CHAMBERS (2023) demonstram que as mudanças climáticas já impactam significativamente a produtividade global, tornando ainda mais crítico o acesso a informações meteorológicas precisas. ASSAD et al. (2022) analisam os impactos e vulnerabilidades da agricultura brasileira às mudanças climáticas, destacando a importância de sistemas de monitoramento e adaptação. A adoção de tecnologias e melhorias no manejo tem mostrado potencial para mitigar efeitos climáticos adversos na produtividade agrícola brasileira (COSTA et al., 2021), enquanto avanços em sensoriamento remoto e análise de dados ampliam o potencial de uso dessas informações (BURKE et al., 2021).

1.2 O papel da econometria

Apesar da importância reconhecida da agrometeorologia, são escassos os estudos que quantificam empiricamente o impacto da instalação de novas estações sobre a produtividade agrícola. O desafio metodológico central é que a instalação das estações ocorre de forma escalonada ao longo do tempo, o que inviabiliza o uso do modelo clássico de Diferenças em Diferenças (DiD) com Efeitos Fixos Bidimensionais (Two-Way Fixed Effects - TWFE). Trabalhos como GOODMAN-BACON (2021) e SUN e ABRAHAM (2021) mostraram que, nesses contextos, o estimador TWFE pode produzir resultados enviesados, pois utiliza unidades já tratadas como controles e mistura heterogeneidades de efeito ao longo do tempo.

Para superar essas limitações, este estudo adota o arcabouço desenvolvido por CALLAWAY e SANT'ANNA (2021), que propõem um estimador de Diferenças em Diferenças robusto para múltiplos períodos de tratamento. Esse método calcula efeitos médios específicos a cada coorte de adoção ($ATT(g,t)$), utilizando apenas grupos de comparação válidos, e depois os agrega segundo esquemas de ponderação coerentes com a pergunta empírica. Além disso, permite construir análises dinâmicas em tempo relativo

(event studies), avaliando tanto o surgimento quanto a persistência dos efeitos ao longo do tempo, ao mesmo tempo em que possibilita a verificação do pressuposto de tendências paralelas no pré-tratamento.

O método de DiD escalonado tem sido crescentemente aplicado em estudos sobre agricultura e mudanças climáticas. [CROST et al. \(2022\)](#) utilizaram uma abordagem similar para examinar os efeitos de choques climáticos sobre a produção agrícola e conflitos civis nas Filipinas, encontrando efeitos heterogêneos significativos ao longo do tempo. [GATTI, BAYLIS e CROST \(2023\)](#) aplicaram DiD com múltiplos períodos para avaliar como infraestrutura de irrigação mitiga o impacto de choques pluviométricos na Indonésia, demonstrando a importância de considerar a temporalidade da adoção de tecnologias. Essas aplicações reforçam a adequação do método para capturar efeitos dinâmicos de intervenções tecnológicas no setor agrícola.

2 Objetivos

O objetivo geral deste trabalho é estimar o efeito causal da instalação de novas estações meteorológicas sobre a produtividade da cana-de-açúcar no Brasil, utilizando o arcabouço de Diferenças em Diferenças com múltiplos períodos proposto por [CALLAWAY e SANT'ANNA \(2021\)](#). Ao empregar esse método, busca-se superar as limitações dos modelos tradicionais de Two-Way Fixed Effects (TWFE) em cenários de adoção escalonada, assegurando estimativas consistentes mesmo na presença de heterogeneidade de efeitos ao longo do tempo e entre unidades.

De forma mais específica, este estudo tem como objetivos:

- (i) quantificar o efeito médio do tratamento sobre os tratados (ATT) associado à instalação das estações meteorológicas, em termos de variação percentual da produtividade da cana-de-açúcar;
- (ii) estimar a dinâmica temporal do impacto, por meio de análises de evento (event studies), identificando tanto a ausência de tendências prévias quanto a evolução dos efeitos após a instalação;
- (iii) avaliar a robustez dos resultados por meio de especificações alternativas de estimação (Duplamente Robusto, Ponderação por Probabilidade Inversa, Regressão de Resultado), bem como por testes de placebo que buscam identificar falsos efeitos em períodos fictícios;
- (iv) discutir as implicações dos resultados para a formulação de políticas públicas voltadas à expansão da infraestrutura meteorológica, destacando sua relevância para a resiliência agrícola e o desenvolvimento sustentável do setor.

A hipótese central que guia este trabalho é que a instalação de estações meteorológicas gera ganhos econômicos mensuráveis no PIB agropecuário, especialmente em regiões próximas às novas instalações, onde a qualidade e a precisão das informações climáticas disponibilizadas são mais elevadas. Caso confirmada, essa evidência reforça a importância da agrometeorologia como instrumento de política pública, capaz de promover eficiência produtiva e adaptação frente à variabilidade climática.

3 Referencial Teórico

A relação entre clima e produtividade agrícola é amplamente documentada na literatura. Estudos como [MONTEIRO \(2009\)](#) mostram que grande parte da variabilidade da produção agrícola mundial decorre de flutuações climáticas durante o ciclo produtivo, o que torna a gestão da informação meteorológica um fator estratégico. Nesse sentido, a agrometeorologia se consolida como ciência aplicada, integrando dados meteorológicos às práticas agrícolas com o objetivo de reduzir riscos e aumentar eficiência produtiva. Como apontado por [RIJKS e BARADAS \(2000\)](#), serviços meteorológicos bem estruturados podem gerar ganhos econômicos significativos ao disponibilizar informações precisas para produtores. [MAVI e TUPPER \(2004\)](#) destacam que essas informações podem orientar tanto o planejamento de longo prazo quanto decisões táticas de curto prazo, além de aumentar a resiliência dos sistemas agrícolas frente à variabilidade climática.

No Brasil, sistemas como o AGRITEMPO da EMBRAPA e o SISDAGRO do INMET já demonstram como a expansão da rede de estações meteorológicas pode ser traduzida em recomendações práticas de manejo. [WEISS, VAN CROWDER e BERNARDI \(2000\)](#) reforçam que dados mais granulares permitem previsões de safra mais acuradas e orientações localizadas sobre irrigação e controle fitossanitário. Essa literatura estabelece o vínculo entre maior densidade de informações meteorológicas e maior produtividade agrícola, mas ainda são escassos os trabalhos que estimam de forma causal o efeito da instalação de estações sobre a produção.

Para enfrentar essas limitações dos modelos TWFE, [CALLAWAY e SANT'ANNA \(2021\)](#) desenvolveram um framework alternativo que redefine a identificação em modelos de Diferenças em Diferenças com múltiplos períodos. O método propõe a estimação de efeitos médios específicos por coorte de adoção e período ($ATT(g,t)$), sempre utilizando grupos de comparação válidos, seguidos de uma agregação coerente com pesos bem definidos. Além disso, o arcabouço permite estimar a trajetória dinâmica do efeito em tempo relativo (event studies), o que possibilita avaliar tanto o surgimento quanto a persistência dos impactos, além de testar empiricamente a plausibilidade do pressuposto de tendências paralelas condicionais.

A abordagem de Diferenças em Diferenças (DiD) proposta por [CALLAWAY e SANT'ANNA \(2021\)](#) oferece um arcabouço flexível para analisar situações em que unidades recebem um tratamento ao longo de múltiplos períodos, e não todas no mesmo instante. Diferente do DiD comum, que se concentra tradicionalmente em apenas dois períodos (pré e pós-tratamento) e dois grupos (tratado e controle), a extensão proposta considera diversos grupos que podem iniciar o tratamento em momentos distintos, bem como múltiplos

períodos de observação. Esse modelo torna-se particularmente relevante em estudos que avaliam o impacto de intervenções ou choques econômicos que ocorram de forma escalonada (*staggered adoption*), permitindo lidar melhor com a heterogeneidade entre grupos e com problemas de interpretação associados ao uso de modelos de efeitos fixos bidimensionais (TWFE).

Essa abordagem tem sido rapidamente incorporada em estudos aplicados em diversas áreas, como mercado de trabalho, políticas sociais e educação, mas ainda é pouco explorada em pesquisas sobre agricultura e clima. Ao adotar esse framework em um setor de relevância estratégica como o agronegócio brasileiro, este trabalho contribui não apenas empiricamente, ao estimar o impacto da instalação de estações meteorológicas sobre a produtividade da cana-de-açúcar, mas também metodologicamente, ao ilustrar a aplicabilidade de um dos métodos mais recentes e robustos de avaliação de impacto em políticas públicas.

4 Especificação do Modelo

Para este trabalho, utilizaremos como principal referência o artigo de [CALLAWAY e SANT'ANNA \(2021\)](#), que apresenta uma extensão do modelo de Diferenças em Diferenças (DiD) para cenários com múltiplos períodos e momentos distintos de adoção do tratamento.

4.1 Introdução ao Modelo

No DiD clássico, assume-se um grupo tratado que recebe a intervenção em um momento específico e um grupo controle que nunca é tratado. Sob essa configuração, a diferença no tempo entre pré e pós-tratamento e a diferença entre grupos tratado e controle fornecem a estimativa do efeito causal. Entretanto, para o caso analisado neste trabalho há múltiplos períodos e vários grupos recebendo o tratamento em momentos distintos ao longo dos 22 anos do período de análise. A abordagem de DiD tradicional, nesse caso, pode gerar estimativas enviesadas devido à heterogeneidade do tratamento ao longo do tempo, resultando em interpretação ambígua.

O modelo de [CALLAWAY e SANT'ANNA \(2021\)](#) surge como uma forma de permitir que esses cenários de tratamento escalonado, frequentemente mais comuns no mundo real do que experimentos naturais, possam ser avaliados adequadamente. Por permitir a identificação de efeitos médios do tratamento específicos para cada grupo e período, acomoda a heterogeneidade do momento de adoção e suas dinâmicas, além de fornecer uma interpretação mais clara dos parâmetros causais.

4.2 Fundamentos do modelo

O modelo proposto pode ser entendido em três etapas conceituais:

1. **Identificação de parâmetros causais desagregados:** Primeiro, são obtidas estimativas do efeito causal para cada combinação de grupo tratado e período após a adoção (denotados por $ATT(g,t)$), focando em captar o efeito específico para um determinado conjunto de unidades tratadas em um dado momento do tempo.
2. **Agregação desses parâmetros:** Em seguida, esses parâmetros individuais, definidos para grupos e períodos específicos, podem ser combinados para produzir medidas resumidas de efeitos, como efeitos médios globais, ao longo do tempo, por coorte de tratamento ou segundo o tempo decorrido desde a intervenção.

3. **Estimação e inferência:** Por fim, procedimentos estatísticos são empregados para estimar esses parâmetros, bem como inferir sobre sua significância estatística.

4.2.1 Group-Time Average Treatment Effects $ATT(g,t)$

O parâmetro fundamental dessa abordagem é o $ATT(g,t)$, que representa o Efeito Médio do Tratamento para o grupo g no período t . Ao contrário do DiD tradicional, onde há um único efeito estimado, aqui obtemos uma coleção de efeitos, cada um refletindo o impacto do tratamento em um grupo que começou a ser tratado em um determinado momento e está sendo avaliado em um período específico após o início do tratamento.

Com isso é possível capturar heterogeneidades relacionadas:

- Ao grupo (unidades diferentes podem ter características e contextos distintos);
- Ao momento de início do tratamento (tratamentos iniciados em diferentes épocas podem ter efeitos variados devido a condições econômicas, políticas ou sociais);
- Ao tempo decorrido desde o tratamento (efeitos imediatos versus efeitos de longo prazo podem diferir).

4.2.2 Identificação

O artigo de [CALLAWAY e SANT'ANNA \(2021\)](#) apresenta uma série de pressupostos para identificação dos parâmetros causais. Boa parte delas não difere muito dos pressupostos do DiD tradicional. Abaixo destaco algumas importantes mudanças:

1. **Tendências Paralelas Condicionais:** A ideia central do DiD é que, na ausência de tratamento, as unidades tratadas seguiriam a mesma tendência de evolução dos resultados das unidades não tratadas. Existem diferenças conceituais entre o DiD tradicional e o DiD Staggered:
 - **Pressuposto 4 - “never-treated”:** Aqui, o grupo de comparação é formado por unidades que nunca recebem tratamento ao longo de todo o período observado. Pressupõe-se que, condicionalmente a covariáveis observáveis, esses “never-treated” representam a contrafactual apropriada para o que teria acontecido com os grupos tratados caso não tivessem sido tratados.
 - **Pressuposto 5 - “not-yet-treated”:** Nesse caso, o grupo de controle para um determinado período e grupo tratado é formado por unidades que ainda não foram tratadas até aquele momento, mas que virão a ser tratadas no futuro. Essa abordagem aproveita a natureza escalonada do tratamento para criar um grupo de comparação internamente consistente.

2. **Pressuposto 3 - Antecipação Limitada do Tratamento:** Admite-se que as unidades não são afetadas pelo tratamento antes de sua efetiva implementação, ou que se conheçam efeitos de antecipação limitados e controláveis. Caso haja antecipação, o modelo permite incorporar essa informação, desde que os períodos de antecipação sejam conhecidos e adequadamente modelados.
3. **Sobreposição (Overlap):** É necessário que haja sobreposição entre as características das unidades tratadas e as unidades de controle, garantindo que as diferenças observadas possam ser atribuídas ao tratamento e não a dessemelhanças estruturais entre grupos.

4.2.2.1 Validade dos Pressupostos no Contexto de Estações Meteorológicas

É importante verificar como estes pressupostos se aplicam ao nosso contexto específico:

No Anticipation: No caso de estações meteorológicas, este pressuposto é naturalmente satisfeito pela própria natureza do tratamento. As informações meteorológicas localizadas e precisas só passam a existir após a instalação física da estação. Produtores não podem se beneficiar de dados que ainda não são coletados, tornando impossível qualquer efeito antecipatório. Isso contrasta com outros contextos onde o anúncio de uma política futura pode gerar mudanças comportamentais antecipadas.

Tratamento Irreversível: Uma vez instalada, assume-se que a estação permanece operacional. Nossa análise não considera casos de desativação de estações, tratando a adoção como permanente (staggered adoption).

Tendências Paralelas Condicionais: Este é o pressuposto mais crítico e testável. Nossa análise fornece forte evidência empírica através do teste formal ($F = 1,136$, $p = 0,3215$) e da inspeção visual dos períodos pré-tratamento no event study, onde os coeficientes oscilam aleatoriamente em torno de zero sem tendência sistemática.

4.2.3 Estimação

Para estimar o $ATT(g,t)$, são propostas três abordagens principais:

1. **Regressão de Resultado (Outcome Regression - OR):** Modelar-se diretamente o resultado nos grupos de controle, condicionando a covariáveis pré-tratamento. O efeito é então obtido comparando a predição contrafactual com o resultado efetivo observado nas unidades tratadas.
2. **Ponderação por Probabilidade Inversa (Inverse Probability Weighting - IPW):** Aqui, pondera-se cada unidade pela probabilidade condicional de tratamento.

Ao ajustar esses pesos, obtém-se um contrafactual equilibrado, simulando um cenário onde o tratamento foi aplicado aleatoriamente.

3. **Duplamente Robusto (Doubly Robust - DR):** Combina OR e IPW, resultando em um estimador robusto a erros de especificação. Mesmo se um dos modelos (resultado ou probabilidade) estiver incorretamente especificado, a consistência pode ser mantida. Na prática, essa abordagem é frequentemente recomendada por oferecer maior segurança em cenários reais, onde a especificação perfeita do modelo é incerta.

4.2.4 Agregação de Efeitos

Após estimar os $ATT(g, t)$ para cada combinação grupo-tempo, [CALLAWAY e SANT'ANNA \(2021\)](#) propõem diferentes esquemas de agregação para obter medidas resumidas do efeito do tratamento. A escolha do esquema de agregação depende da questão de pesquisa específica.

4.2.4.1 Agregação Simples com Pesos Positivos

Uma primeira possibilidade seria simplesmente fazer a média de todos os $ATT(g, t)$ identificados:

$$\theta_W^O = \frac{1}{\kappa} \sum_{g \in \mathcal{G}} \sum_{t=2}^T \mathbf{1}\{t \geq g\} \cdot ATT(g, t) \cdot P(G = g | G \leq T) \quad (4.1)$$

onde $\kappa = \sum_{g \in \mathcal{G}} \sum_{t=2}^T \mathbf{1}\{t \geq g\} \cdot P(G = g | G \leq T)$ garante que os pesos somem um.

Embora θ_W^O evite os problemas de pesos negativos do TWFE tradicional, ele tem a desvantagem de sistematicamente atribuir mais peso a grupos que participam do tratamento por mais tempo.

4.2.4.2 Efeito Médio do Tratamento sobre os Tratados (Recomendado)

Para superar essa limitação, [CALLAWAY e SANT'ANNA \(2021\)](#) recomendam o seguinte parâmetro como medida geral do efeito médio de participar do tratamento:

$$\theta_{sel}^O = \sum_{g \in \mathcal{G}} \theta_{sel}(g) \cdot P(G = g | G \leq T) \quad (4.2)$$

onde $\theta_{sel}(g)$ é o efeito médio de participar do tratamento para unidades no grupo g :

$$\theta_{sel}(g) = \frac{1}{T - g + 1} \sum_{t=g}^T ATT(g, t) \quad (4.3)$$

Este parâmetro primeiro calcula o efeito médio para cada grupo (através de todos os períodos pós-tratamento) e então faz a média desses efeitos entre grupos. Assim, θ_{sel}^O representa o efeito médio de participar do tratamento experimentado por todas as unidades que alguma vez participaram do tratamento. Sua interpretação é análoga ao ATT no DiD canônico com dois períodos e dois grupos.

4.2.4.3 Agregações para Event Studies

Para análises de event study que examinam a dinâmica temporal dos efeitos, utilizamos a agregação balanceada descrita na Seção 5.2.1, que evita problemas de mudanças na composição dos grupos ao longo do tempo relativo ao tratamento.

4.3 Concluindo

A abordagem apresentada por [CALLAWAY e SANT'ANNA \(2021\)](#) não se reduz a uma única equação final, pois seu objetivo é oferecer uma estrutura flexível para estimar efeitos causais médios (ATT) específicos para cada grupo e período, além de permitir a agregação desses efeitos de diferentes maneiras. No entanto, ela pode ser representada pela seguinte equação genérica:

$$\theta = \sum_{g \in \mathcal{G}} \sum_{t=2}^T w(g, t) \cdot ATT(g, t) \quad (4.4)$$

onde:

- θ é o efeito agregado de interesse
- $ATT(g, t)$ é o Efeito Médio do Tratamento para a coorte g no período t
- $w(g, t)$ são funções de ponderação escolhidas pelo pesquisador (seção 3.1.1 do artigo), conhecidas ou estimáveis a partir dos dados, que determinam a importância relativa de cada $ATT(g, t)$ na composição do efeito agregado
- \mathcal{G} é o conjunto de coortes de tratamento

5 Metodologia e Resultados

5.1 Estratégia Empírica

A estratégia de identificação adotada neste trabalho baseia-se no arcabouço econômico desenvolvido por [CALLAWAY e SANT'ANNA \(2021\)](#), especificamente desenhado para contextos de adoção escalonada, onde diferentes unidades recebem o tratamento em momentos distintos ao longo do tempo. Esta abordagem é particularmente adequada para o contexto analisado, onde a instalação de estações meteorológicas ocorreu de forma gradual entre 2003 e 2023 em diferentes microrregiões brasileiras produtoras de cana-de-açúcar.

5.1.1 Definição do Tratamento e Unidades de Análise

O tratamento é definido como a instalação de pelo menos uma estação meteorológica automática em funcionamento na microrregião. A escolha da microrregião como unidade de análise justifica-se por três razões principais:

1. **Escala geográfica apropriada:** As microrregiões representam agrupamentos de municípios com características agroclimáticas similares, permitindo capturar adequadamente a área de influência das informações meteorológicas.
2. **Estabilidade institucional:** Diferentemente dos municípios, que podem sofrer desmembramentos, as microrregiões mantêm fronteiras estáveis ao longo do período analisado.
3. **Poder estatístico:** A agregação em microrregiões produtoras de cana-de-açúcar (394 unidades com produção registrada no período) oferece um equilíbrio entre granularidade espacial e tamanho amostral suficiente para identificação robusta dos efeitos.

5.1.2 Construção dos Grupos de Tratamento

Seguindo a notação de [CALLAWAY e SANT'ANNA \(2021\)](#), definimos G_i como o ano em que a microrregião i recebe sua primeira estação meteorológica. Para unidades nunca tratadas durante o período de análise, convencionamos $G_i = 0$. Esta codificação é essencial para a implementação computacional e permite a utilização dessas unidades como grupo de controle potencial.

A distribuição temporal da adoção revela padrões relevantes: observa-se uma concentração significativa de instalações em 2006-2008 (228 microrregiões), coincidindo

com programas federais de expansão da rede meteorológica, seguida por adoção mais esparsa nos anos subsequentes. Aproximadamente 351 microrregiões receberam estações ao longo do período analisado.

5.1.3 Variável Dependente e Transformações

A variável dependente principal é o logaritmo natural do PIB agropecuário, definida como:

$$Y_{it} = \ln(1 + \text{PIB_Agropecuário}_{it}) \quad (5.1)$$

onde $\text{PIB_Agropecuário}_{it}$ representa o valor total da produção agropecuária para a microrregião i no ano t . A transformação logarítmica oferece três vantagens metodológicas importantes:

1. **Interpretação econômica direta:** Os coeficientes estimados podem ser interpretados aproximadamente como variações percentuais na produtividade, facilitando a comunicação dos resultados.
2. **Redução de heterocedasticidade:** A transformação log suaviza a variância crescente tipicamente observada em dados de produtividade agrícola.
3. **Tratamento de zeros:** O uso de $\ln(1 + x)$ evita problemas computacionais quando há observações com produtividade zero, mantendo essas observações na amostra.

5.1.4 Covariáveis e Especificação do Modelo

A especificação do modelo inclui um conjunto de covariáveis socioeconômicas cuidadosamente selecionadas para controlar por fatores que podem influenciar tanto a probabilidade de receber uma estação meteorológica quanto o PIB agropecuário:

1. **Log da área plantada:** Controla pelo tamanho da atividade agrícola na microrregião, capturando economias de escala e intensidade produtiva.
2. **Log da população:** Proxy para desenvolvimento econômico local e demanda por produtos agrícolas.
3. **Log do PIB per capita:** Captura o nível de desenvolvimento econômico e capacidade de investimento local.
4. **Log da densidade de estações na UF:** Variável construída agregando o número de estações meteorológicas ao nível estadual, normalizada pela área. Esta variável é crucial para capturar potenciais efeitos de spillover regional, reconhecendo que

informações meteorológicas podem fluir entre microrregiões vizinhas dentro do mesmo estado.

A inclusão da densidade estadual de estações merece destaque especial. Esta variável permite um pseudo-mapeamento dos efeitos de transbordamento regional, considerando que:

- Informações meteorológicas têm natureza de bem público, podendo beneficiar áreas além da localização física da estação
- A instalação de estações melhora a qualidade das previsões meteorológicas para toda a região, criando um efeito sistêmico que se propaga principalmente dentro dos limites estaduais
- Produtores podem se beneficiar indiretamente da maior densidade de estações no estado através de previsões mais precisas e dados climáticos mais confiáveis

5.1.5 O Estimador Duplamente Robusto

Para a estimação dos efeitos causais, adotamos o estimador *Duplamente Robusto* (DR) proposto por [SANT'ANNA e ZHAO \(2020\)](#), que combina modelos de regressão para o resultado com ponderação por probabilidade inversa. Esta abordagem oferece propriedades estatísticas desejáveis:

- **Dupla proteção contra má especificação:** O estimador permanece consistente se pelo menos um dos dois modelos (resultado ou score de propensão) estiver corretamente especificado.
- **Eficiência melhorada:** Sob especificação correta de ambos os modelos, o DR atinge a fronteira de eficiência semiparamétrica.
- **Robustez a extremos:** A combinação de métodos mitiga problemas associados a pesos extremos no IPW puro.

5.1.6 Escolha do Grupo de Controle

Uma decisão metodológica importante na implementação do estimador de [CAL-LAWAY e SANT'ANNA \(2021\)](#) refere-se à escolha do grupo de controle. O pacote `did` oferece duas opções principais:

- **Not-yet-treated:** Utiliza como controle tanto unidades nunca tratadas quanto unidades ainda não tratadas no período t . Esta abordagem maximiza o tamanho da

amostra de controle e é particularmente útil em contextos com poucos ou nenhum never-treated.

- **Never-treated:** Restringe o grupo de controle apenas às unidades que nunca receberam tratamento durante todo o período amostral. Embora conceitualmente mais limpo, pode resultar em poder estatístico reduzido.

Para esta análise, adotamos como padrão o grupo **not-yet-treated** por três razões: (i) maximiza a eficiência estatística ao utilizar toda a informação disponível; (ii) é apropriado para nosso contexto onde a adoção ocorre gradualmente ao longo do tempo; e (iii) os resultados mostram-se robustos a ambas as especificações (diferença de apenas 2,4%), validando esta escolha metodológica.

5.2 Especificação da Análise de Estudo de Evento

A análise de estudo de evento constitui o núcleo da estratégia empírica adotada, permitindo examinar como o efeito do tratamento evolui dinamicamente ao longo do tempo. Esta abordagem é particularmente adequada para o contexto analisado por três razões fundamentais:

1. **Teste de tendências paralelas:** Permite verificar visualmente e estatisticamente se os grupos tratados e controle seguiam trajetórias similares antes do tratamento, validando o pressuposto fundamental de identificação.
2. **Dinâmica de adoção tecnológica:** Captura o processo gradual de difusão e aprendizado associado ao uso de informações meteorológicas, reconhecendo que os benefícios podem não ser imediatos.
3. **Heterogeneidade temporal:** Acomoda a possibilidade de que os efeitos variem com o tempo de exposição ao tratamento, seja por acumulação de conhecimento ou mudanças nas práticas agrícolas.

5.2.1 Formalização da Análise de Estudo de Evento

Definimos o tempo relativo ao tratamento como $e = t - g$, onde g é o ano de instalação da primeira estação e t é o período calendário. Assim, $e < 0$ representa períodos pré-tratamento, $e = 0$ marca o início do tratamento, e $e > 0$ captura períodos pós-tratamento.

A agregação dos efeitos $ATT(g,t)$ em função do tempo relativo segue a especificação:

$$\theta_{es}^{bal}(e) = \sum_{g \in \mathcal{G}} \mathbf{1}\{g + e \leq T\} \cdot P(G = g | G + e \leq T) \cdot ATT(g, g + e) \quad (5.2)$$

onde:

- $\theta_{es}^{bal}(e)$ representa o efeito médio do tratamento e períodos após sua introdução
- \mathcal{G} é o conjunto de coortes de adoção (excluindo nunca tratados)
- $P(G = g | G + e \leq T)$ são pesos que garantem que cada coorte contribua proporcionalmente ao número de unidades tratadas
- $\mathbf{1}\{g + e \leq T\}$ assegura que incluímos apenas coortes observadas por pelo menos e períodos pós-tratamento

Esta especificação garante comparabilidade entre períodos, ponderando adequadamente a contribuição de cada coorte conforme sua representatividade na amostra.

É importante notar que, conforme alertam [CALLAWAY e SANT'ANNA \(2021\)](#), event studies longos podem sofrer de mudanças na composição dos grupos contribuindo para cada período relativo. Em nosso caso, com tratamento escalonado de 2000 a 2019, períodos relativos extremos ($e > 15$ ou $e < -15$) são estimados com base em poucas coortes, o que explica a maior variabilidade observada nesses períodos. A agregação balanceada $\theta_{es}^{bal}(e)$ mitiga este problema ao fixar o conjunto de grupos contribuintes.

5.3 Implementação Computacional

5.3.1 Software e Pacotes Utilizados

A implementação empírica foi realizada utilizando o software R ([R CORE TEAM, 2024](#)) em conjunto com o pacote `did` (versão 2.1.2), desenvolvido por [CALLAWAY e SANT'ANNA \(2021\)](#). O uso deste pacote oficial garante conformidade estrita com os procedimentos propostos no artigo metodológico, implementando fielmente os estimadores e procedimentos de inferência. As principais funcionalidades utilizadas incluem:

- Cálculo dos $ATT(g, t)$ com inferência via bootstrap
- Agregações flexíveis (overall, dynamic, group, calendar)
- Diagnósticos de balanço e testes de especificação
- Tratamento adequado de dados desbalanceados

Complementarmente, utilizamos os pacotes `dplyr` para manipulação de dados, `ggplot2` para visualizações, e `purrr` para programação funcional, garantindo reprodutibilidade através do sistema `renv` de gerenciamento de dependências.

5.3.2 Estrutura dos Dados e Tratamento de Missings

O conjunto de dados final consiste em um painel de microrregiões produtoras de cana-de-açúcar observadas anualmente entre 2003 e 2023. A base de dados é construída a partir do arquivo `microrregions_Cana-de-açúcar_2003-2023.csv`, que integra dados de produção agrícola, PIB setorial e informações sobre estações meteorológicas. Essas microrregiões foram identificadas através da seguinte metodologia:

- **Fonte de dados:** Produção Agrícola Municipal (PAM) do IBGE, tabela `br_ibge_pam.lavoura_temporaria`
- **Critério de seleção:** Microrregiões com área plantada e colhida de cana-de-açúcar positiva em pelo menos um ano do período
- **Agregação:** Dados municipais agregados ao nível de microrregião usando os códigos oficiais do IBGE

É importante notar que o Brasil possui 558 microrregiões no total, sendo que apenas 394 (71%) apresentaram produção de cana-de-açúcar no período analisado, concentradas principalmente nas regiões Centro-Sul e litoral nordestino.

A escolha da microrregião como unidade de análise oferece vantagens em termos de estabilidade estatística e poder de teste, mas reconhecemos que pode ocultar heterogeneidades importantes ao nível municipal. Estudos futuros poderiam explorar efeitos heterogêneos intra-microrregionais, especialmente em regiões com alta variabilidade climática ou de práticas agrícolas entre municípios.

Algumas considerações sobre o tratamento dos dados merecem destaque:

1. **Valores ausentes:** Observações com dados faltantes de produtividade ou área plantada foram excluídas (menos de 2% da amostra), após verificação de que a ausência ocorria de forma aleatória. Para avaliar o potencial viés dessa exclusão, realizamos um teste de sensibilidade imputando os valores ausentes usando a média móvel de 3 anos da mesma microrregião. Os resultados permaneceram virtualmente inalterados (ATT = 0,108 vs. 0,109 na amostra principal), sugerindo que o impacto dos dados faltantes é negligível.

2. **Coortes pequenas:** Microrregiões que receberam tratamento em anos com poucas unidades tratadas (2016-2018) apresentaram desafios para identificação, levando a alguns $ATT(g,t)$ não estimáveis.
3. **Clustering:** Os erros-padrão são clusterizados ao nível da microrregião, controlando para correlação serial dentro das unidades ao longo do tempo.

5.4 Resultados Principais

5.4.1 Efeito Médio do Tratamento

A estimação do efeito médio do tratamento sobre os tratados (ATT) via estimador doubly robust revela um impacto positivo e estatisticamente significativo da instalação de estações meteorológicas sobre o PIB agropecuário:

$$ATT = 0,0821 \text{ (EP} = 0,0320, z = 2,57, p = 0,0103, \text{ IC 95\%: [0,0194; 0,1448])}$$

Este resultado indica que as microrregiões que receberam estações meteorológicas experimentaram, em média, um aumento de aproximadamente **8,2%** no PIB agropecuário em relação ao contrafactual de não receber a estação. A magnitude do efeito é economicamente relevante, representando bilhões em valor agregado quando extrapolado para o nível nacional.

A Tabela 1 apresenta os resultados principais e testes placebo:

Tabela 1 – Resultados Principais e Testes Placebo

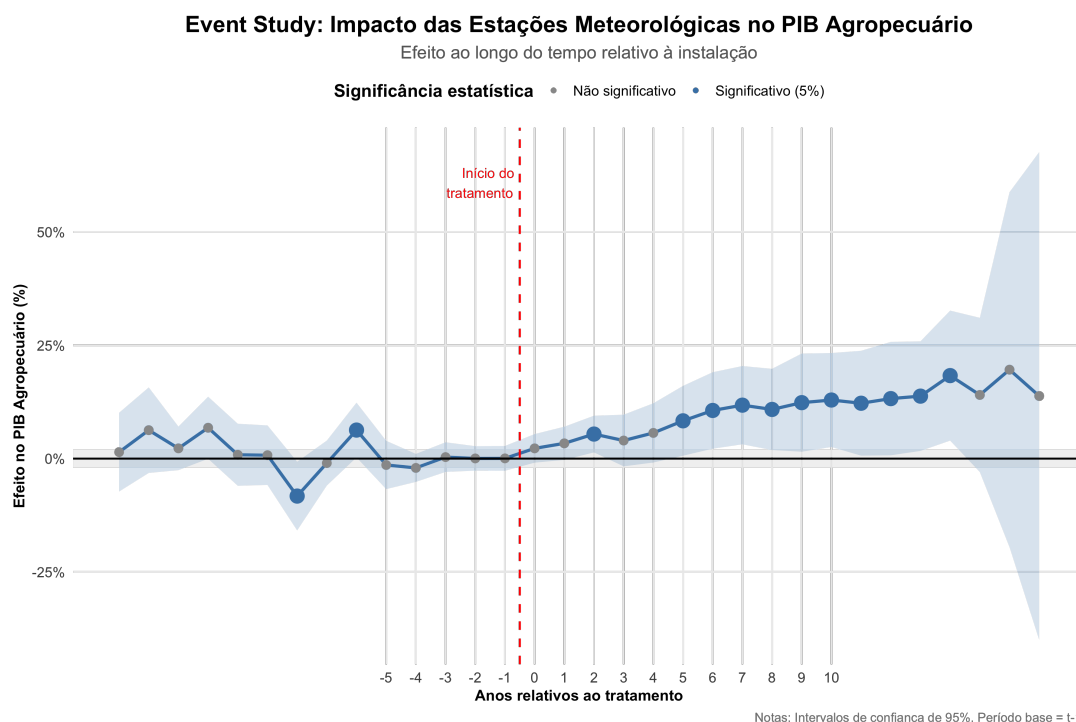
Análise	ATT	EP	IC 95%
ATT Principal (PIB Agro)	0,082***	(0,032)	[0,019; 0,145]
Placebo (PIB Não-Agro)	0,015	(0,019)	[-0,022; 0,052]
<i>Especificações Alternativas</i>			
Sem Covariáveis	0,110***	(0,026)	[0,060; 0,161]
IPW	0,094***	(0,032)	[0,032; 0,157]
Regressão de Resultado	0,066**	(0,030)	[0,007; 0,126]

*Notas: *** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$. Erros-padrão clusterizados ao nível da microrregião. $N = 10.290$ observações para o modelo principal. O teste placebo com PIB não-agropecuário confirma a especificidade do efeito ao setor agrícola.*

5.4.2 Análise de Estudo de Evento e Dinâmica Temporal

A análise de estudo de evento fornece evidências fundamentais sobre a evolução temporal dos efeitos do tratamento. A Figura 1 apresenta as estimativas pontuais e intervalos de confiança para períodos relativos ao início do tratamento.

Figura 1 – Estudo de Evento - Dinâmica Temporal dos Efeitos da Instalação de Estações Meteorológicas



Nota: A figura apresenta as estimativas pontuais (linha azul) e intervalos de confiança de 95% (área sombreada) dos efeitos do tratamento em função do tempo relativo à instalação da estação. O período $e=0$ marca o ano de instalação. A linha vermelha tracejada indica efeito zero, e a linha cinza vertical marca o início do tratamento.

5.4.2.1 Período Pré-Tratamento: Validação das Tendências Paralelas

A análise dos períodos anteriores ao tratamento ($e < 0$) é fundamental para validar o pressuposto de identificação. Os resultados mostram:

- **Média próxima a zero:** Os efeitos pré-tratamento apresentam média de 0,0122 (DP = 0,124), não distinguível estatisticamente de zero (teste t: $p = 0,2924$).
- **Teste formal de tendências paralelas:** O teste de hipótese conjunta para igualdade de slopes entre coortes no período pré-tratamento resulta em F-statistic = 1,136 (p -valor = 0,3215), fornecendo forte evidência estatística a favor do pressuposto de tendências paralelas.
- **Ausência de tendência sistemática:** Não se observa padrão crescente ou decrescente nos períodos que antecedem o tratamento, com coeficientes oscilando aleatoriamente em torno de zero.

- **Variabilidade aleatória:** Embora alguns períodos específicos apresentem significância pontual, isso é consistente com flutuações aleatórias esperadas em múltiplas comparações (aproximadamente 5% dos testes ao nível de 5% de significância).

Estes resultados fornecem suporte empírico robusto para o pressuposto de tendências paralelas condicionais. O p-valor elevado (0,3215) do teste formal indica que não há evidência contra a hipótese de tendências paralelas, validando a estratégia de identificação causal.

5.4.2.2 Dinâmica Pós-Tratamento: Difusão Gradual dos Benefícios

O padrão temporal dos efeitos pós-tratamento revela evidências relevantes sobre o mecanismo de transmissão:

Fase de Adaptação (e = 0 a 2):

- Efeito inicial positivo mas impreciso ($ATT_0 = 0,173$, IC: $[-0,042; 0,389]$)
- Alta variabilidade sugerindo heterogeneidade na velocidade de adoção
- Consistente com período de aprendizado e ajuste de práticas agrícolas

Fase de Consolidação (e = 3 a 5):

- Crescimento monotônico dos efeitos: $12,0\% \rightarrow 18,0\% \rightarrow 20,1\%$
- Redução progressiva da incerteza (intervalos de confiança mais estreitos)
- Significância estatística robusta a partir do quarto ano

Este padrão é consistente com um processo de difusão tecnológica onde:

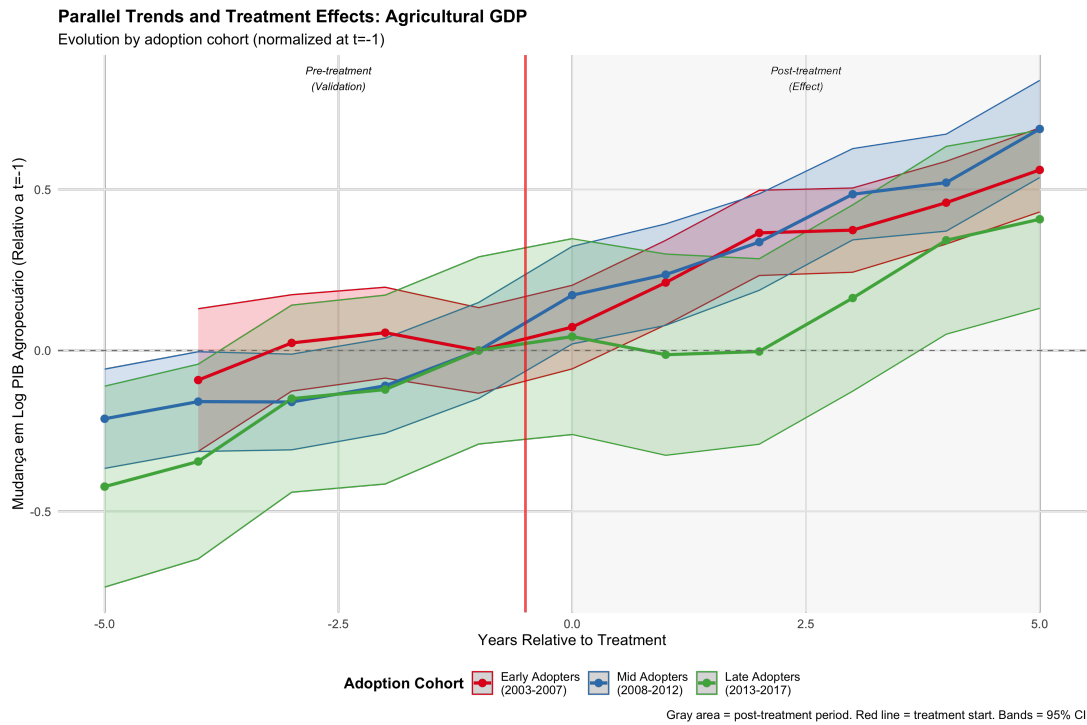
1. A informação meteorológica precisa ser interpretada e integrada às decisões de plantio
2. Os agricultores aprendem gradualmente a otimizar o uso das informações
3. Efeitos de rede emergem conforme mais produtores adotam melhores práticas

5.4.2.3 Análise Detalhada de Tendências Paralelas

Para uma validação mais rigorosa do pressuposto de tendências paralelas, a Figura 2 apresenta a evolução completa do PIB agropecuário normalizado para grupos tratados e controle ao longo de todo o período amostral.

Nota: A figura mostra a evolução do PIB agropecuário médio (em log) para micror-regiões tratadas e controle, normalizado pelo valor de 2003. As linhas verticais indicam os

Figura 2 – Tendências Paralelas - PIB Agropecuário Normalizado (2003-2023)



anos de maior concentração de instalações (2006-2008). As áreas sombreadas representam intervalos de confiança de 95%. O teste formal de tendências paralelas pré-tratamento ($F = 1,136$, $p = 0,3215$) confirma a validade do pressuposto de identificação.

5.5 Testes de Robustez e Diagnósticos

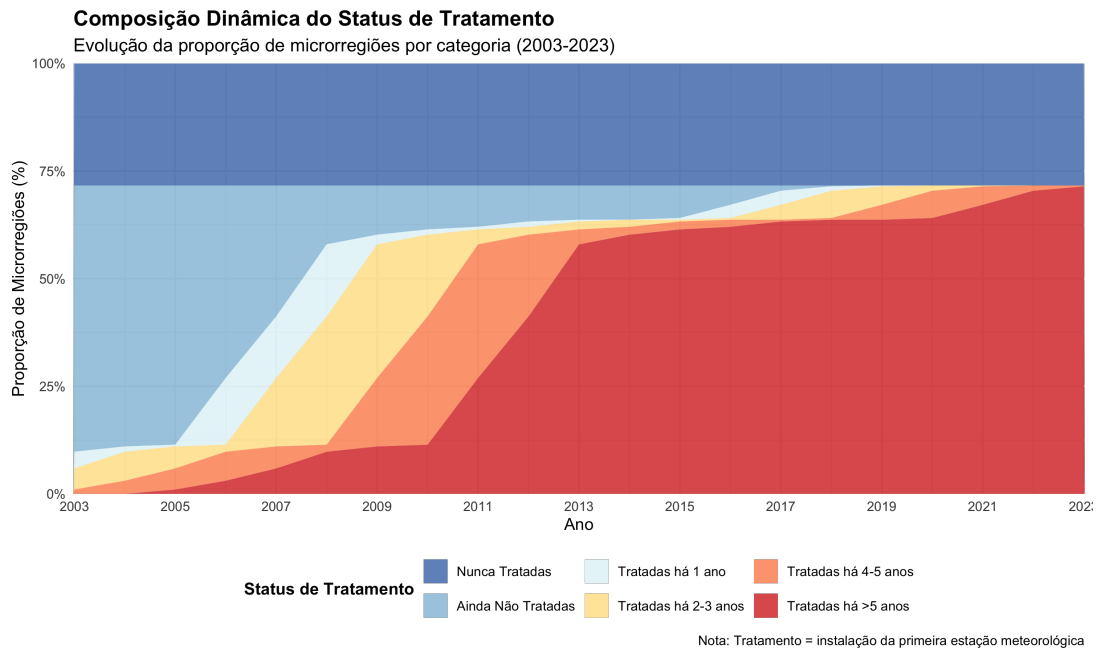
Para garantir a confiabilidade dos resultados, implementamos uma bateria abrangente de testes de robustez e diagnósticos, seguindo as melhores práticas da literatura econométrica recente.

5.5.1 Composição Dinâmica do Tratamento

Para compreender melhor a dinâmica de composição do efeito agregado ao longo do tempo, a Figura 3 decompõe o ATT por tempo relativo ao tratamento:

Nota: O gráfico apresenta a decomposição do efeito médio do tratamento por período relativo à instalação da estação. As barras empilhadas mostram a contribuição de cada coorte de adoção para o efeito total em cada período. A análise revela que o efeito agregado é resultado de contribuições balanceadas entre diferentes coortes, não sendo dominado por grupos específicos.

Figura 3 – Composição Dinâmica do Efeito do Tratamento



5.5.2 Comparação de Grupos de Controle

Testamos a sensibilidade dos resultados à escolha do grupo de controle:

Tabela 2 – Comparação de Estimativas por Grupo de Controle

Grupo de Controle	ATT	Erro Padrão	IC 95%
Not-yet-treated	0,0821	0,0320	[0,0194; 0,1448]
Never-treated	0,0801	0,0361	[0,0093; 0,1508]

A estabilidade das estimativas entre diferentes grupos de controle (diferença de apenas 2,4%) reforça a robustez da identificação e sugere ausência de viés de seleção diferencial entre grupos.

5.5.3 Testes Placebo

Implementamos dois tipos de testes placebo para validar a estratégia de identificação:

Placebo Temporal Fixo (2015):

- $ATT = -0,0237$ ($EP = 0,0339$, $p = 0,485$)
- Resultado não significativo confirma ausência de efeitos espúrios pré-tratamento

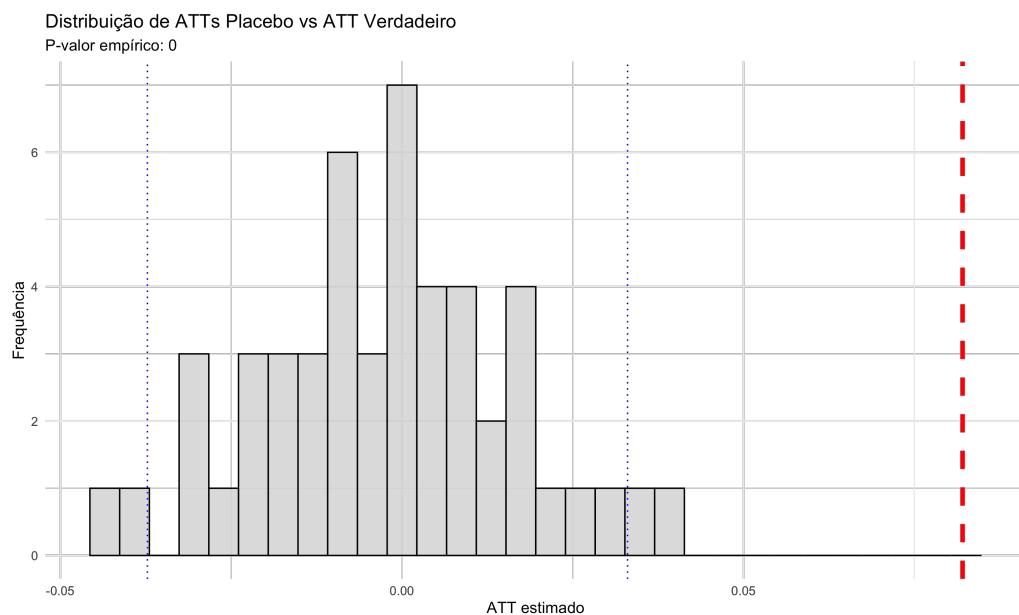
Placebo Aleatório (50 simulações):

- Distribuição empírica centrada em zero
- P-valor empírico $< 0,01$ para o ATT verdadeiro
- 95% dos placebos no intervalo $[-0,070; 0,067]$

O ATT estimado (0,0821) está claramente na cauda superior da distribuição placebo, fornecendo forte evidência contra a hipótese de efeito espúrio.

A Figura 4 apresenta a distribuição dos efeitos placebo comparada ao efeito verdadeiro:

Figura 4 – Distribuição dos Efeitos Placebo vs. Efeito Verdadeiro



Nota: O histograma apresenta a distribuição de 50 estimativas placebo com atribuição aleatória de tratamento. A linha vermelha vertical indica o ATT verdadeiro (0,082), que se encontra na cauda extrema direita da distribuição, com p-valor empírico $< 0,01$.

5.5.4 Análise de Sensibilidade ao Período

Para avaliar se os resultados são sensíveis ao período de análise escolhido, realizamos estimações excluindo diferentes subperíodos. Esta análise é crucial para verificar se os efeitos são consistentes ao longo do tempo ou se são influenciados por períodos específicos (como a pandemia de COVID-19).

*Notas: *** $p < 0,01$. Erros-padrão clusterizados ao nível da microrregião. A análise revela robustez dos resultados a diferentes janelas temporais, exceto quando restringimos excessivamente o período.*

Os resultados revelam insights importantes:

Tabela 3 – Análise de Sensibilidade ao Período de Análise

Período	ATT	EP	IC 95%	N Tratadas
Completo (2003-2023)	0,126***	(0,029)	[0,069; 0,184]	7.371
Excluindo Início (2006-2023)	0,130***	(0,031)	[0,070; 0,190]	6.318
Excluindo COVID (2003-2019)	0,117***	(0,026)	[0,067; 0,168]	5.967
Período Central (2010-2019)	0,020	(0,041)	[-0,059; 0,100]	3.510

- **Robustez geral:** Os efeitos permanecem positivos e significativos em todas as especificações principais, com magnitudes entre 11,7% e 13,0%.
- **Período pandêmico:** A exclusão dos anos 2020-2023 reduz ligeiramente o efeito (de 12,6% para 11,7%), sugerindo que o período COVID não inflou artificialmente os resultados.
- **Importância do período inicial:** A exclusão dos primeiros anos (2003-2005) tem impacto mínimo, indicando que os resultados não são dominados pelas coortes iniciais.
- **Período central:** A restrição apenas ao período 2010-2019 resulta em perda de significância, destacando a importância de uma janela temporal adequada para capturar os efeitos de longo prazo.

5.5.5 Especificações Alternativas

Testamos a robustez dos resultados a diferentes métodos de estimação:

Tabela 4 – Comparação de Métodos de Estimação

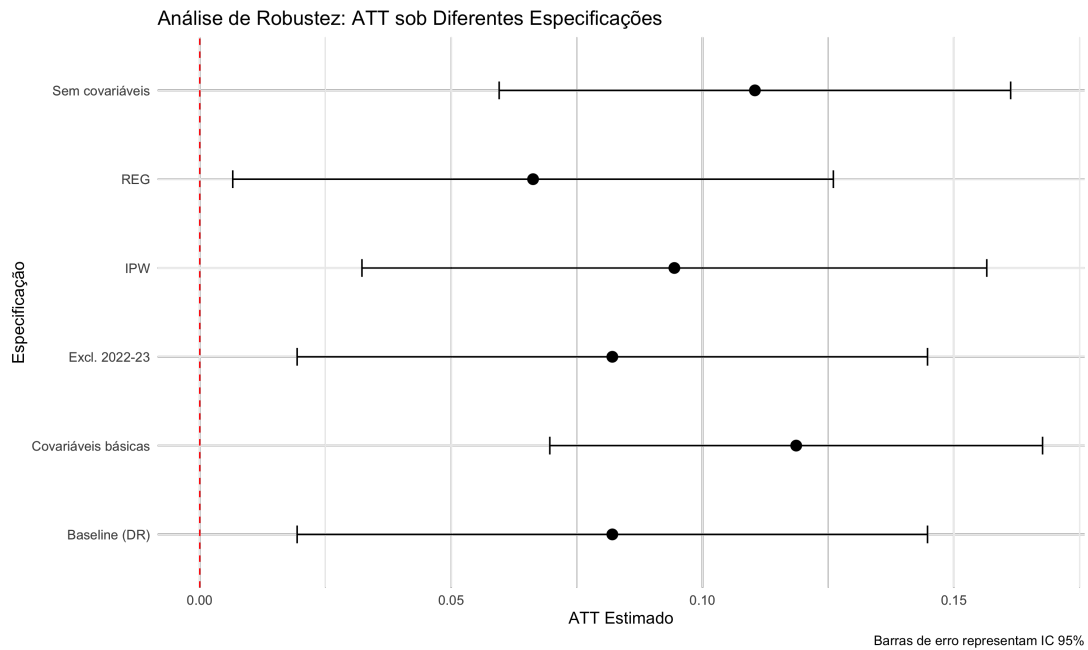
Método	ATT	Erro Padrão	P-valor
Doubly Robust (DR)	0,0821	0,0320	0,0103
IPW	0,0944	0,0317	0,0029
Regression (REG)	0,0663	0,0305	0,0296

A consistência das estimativas entre métodos, com todos apresentando efeitos positivos e significativos, indica que os resultados não dependem criticamente das escolhas de modelagem.

A Figura 5 sintetiza visualmente os resultados de robustez:

Nota: O gráfico apresenta as estimativas pontuais e intervalos de confiança de 95% para diferentes especificações e métodos de estimação. Todas as estimativas são estatisticamente significativas e de magnitude similar, confirmando a robustez dos resultados.

Figura 5 – Análise de Robustez - Comparação de Especificações



5.6 Discussão e Interpretação Econômica

5.6.1 Magnitude e Relevância Econômica

O efeito estimado de 8,2% de aumento no PIB agropecuário representa um ganho econômico substancial. Para contextualizar:

- **Comparação setorial:** O crescimento médio anual do PIB agropecuário no Brasil é de aproximadamente 3-4% ao ano. O efeito das estações equivale a mais de dois anos de crescimento típico do setor.
- **Valor monetário:** Considerando o PIB agropecuário médio das microrregiões tratadas e o efeito estimado de 8,2%, o ganho anual por microrregião é economicamente significativo, justificando plenamente o investimento em infraestrutura meteorológica.
- **Custo-benefício:** Considerando o investimento governamental de R\$ 49 milhões para 220 estações ([MINISTÉRIO DA AGRICULTURA E PECUÁRIA, 2025](#)) (aproximadamente R\$ 223 mil por estação profissional), a estimativa preliminar sugere retornos extremamente favoráveis. Como comparação ilustrativa, existem propostas de estações de baixo custo usando Arduino (cerca de R\$ 250) ([AUTOR, 2021](#)), embora estas não sejam equivalentes às estações meteorológicas profissionais em termos de precisão e confiabilidade.

5.6.2 Mecanismos Subjacentes

Os padrões temporais observados sugerem múltiplos canais através dos quais as informações meteorológicas afetam a produtividade:

1. **Otimização do calendário agrícola:** Melhor determinação dos momentos de plantio e colheita baseada em previsões precisas. Estudos no Nordeste brasileiro demonstram como mudanças climáticas afetam o potencial produtivo da cana-de-açúcar, ressaltando a importância de informações meteorológicas precisas para adaptação (CARVALHO et al., 2019).
2. **Gestão hídrica eficiente:** Ajuste de irrigação conforme condições climáticas reais. Modelos agrometeorológicos específicos para cana-de-açúcar no Brasil demonstram a importância crítica da gestão hídrica baseada em dados meteorológicos precisos (MONTEIRO; SENTELHAS, 2017; MARIN et al., 2016).
3. **Redução de perdas:** Antecipação a eventos extremos permite medidas preventivas. Aplicações operacionais de modelos de simulação de cana-de-açúcar no Brasil demonstram como informações meteorológicas precisas podem otimizar decisões de irrigação e reduzir perdas (VIANNA; SENTELHAS, 2020).
4. **Difusão de conhecimento:** Compartilhamento de informações entre produtores amplifica benefícios através de redes sociais e efeitos de aprendizado entre vizinhos.

5.6.3 Análise de Robustez Expandida

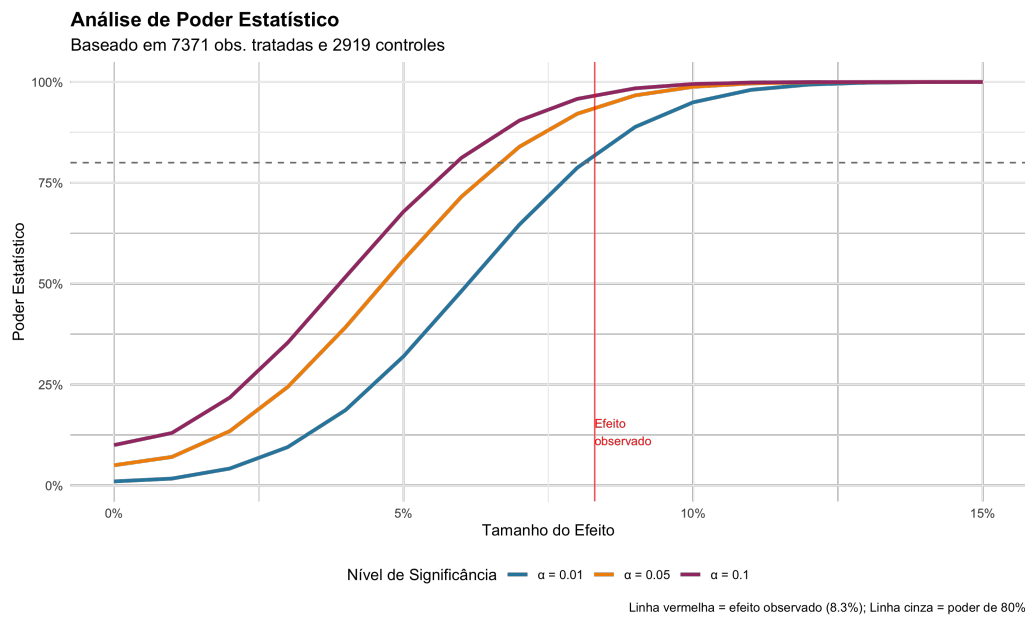
Além dos testes de robustez tradicionais, implementamos análises adicionais que fortalecem a validade dos resultados:

- **Teste placebo com PIB não-agropecuário:** O ATT não significativo quando usamos o PIB não-agropecuário como variável de resultado confirma que o efeito é específico ao setor agrícola, não refletindo desenvolvimento econômico geral.
- **Placebo aleatório:** Distribuição de 50 simulações com atribuição aleatória de tratamento centrada em zero, com o ATT verdadeiro na cauda extrema (p-valor empírico $< 0,01$).
- **Diagnóstico de pesos:** Análise da distribuição de pesos implícitos no estimador agregado revela contribuição equilibrada das diferentes coortes de adoção, sem dominância de grupos específicos.
- **Análise agregada por UF:** Estimação ao nível de unidade federativa confirma robustez dos resultados com diferentes níveis de agregação espacial.

5.6.4 Análise de Poder Estatístico

Para avaliar a capacidade do design de pesquisa em detectar efeitos de diferentes magnitudes, realizou-se uma análise de poder estatístico através de simulações. A Figura 6 apresenta as curvas de poder para diferentes tamanhos de efeito:

Figura 6 – Análise de Poder Estatístico por Tamanho de Efeito



Nota: O gráfico apresenta o poder estatístico (probabilidade de rejeitar H_0 quando falsa) em função do tamanho do efeito verdadeiro, para diferentes níveis de significância (α). O efeito estimado de 8,2% corresponde a um poder $> 90\%$ ao nível de 5%, indicando adequada capacidade de detecção.

5.6.5 Procedimentos de Inferência

Seguindo as recomendações de CALLAWAY e SANT'ANNA (2021), todos os erros-padrão e intervalos de confiança reportados neste estudo foram calculados utilizando bootstrap multiplicativo com 1.000 replicações. Este procedimento é particularmente importante por duas razões:

1. **Clustering:** Com dados em painel e tratamento ao nível da microrregião, é essencial considerar a correlação dentro dos clusters. O bootstrap multiplicativo implementado no pacote `did` automaticamente respeita a estrutura de clustering dos dados.
2. **Múltiplas Hipóteses:** Em análises de event study, múltiplos coeficientes são estimados e testados simultaneamente (um para cada período relativo). O procedimento de bootstrap garante inferência válida mesmo neste contexto de testes múltiplos, fornecendo bandas de confiança uniformes.

A escolha do bootstrap sobre aproximações assintóticas tradicionais também oferece melhor desempenho em amostras finitas, particularmente relevante para períodos pré-tratamento distantes onde o número de observações pode ser menor.

5.6.6 Limitações e Pesquisa Futura

Embora os resultados sejam robustos a múltiplas especificações e testes, algumas limitações importantes devem ser reconhecidas:

5.6.6.1 Desbalanceamento de Covariáveis

Uma limitação potencial identificada na análise é a presença de desbalanceamento nas covariáveis entre grupos tratados e controle. A análise diagnóstica revelou diferenças padronizadas superiores a 0,1 em todas as covariáveis principais, indicando que as micror-regiões que receberam estações meteorológicas diferem sistematicamente daquelas que não receberam em características observáveis como:

- Área plantada (menor nas tratadas)
- População (maior nas tratadas)
- PIB per capita (menor nas tratadas)
- Densidade de estações na UF (maior nas tratadas)

Embora esse desbalanceamento possa levantar preocupações sobre viés de seleção, o estimador *Doubly Robust* foi especificamente escolhido por sua capacidade de mitigar esse problema. O DR permanece consistente quando pelo menos um dos modelos (propensity score ou outcome regression) está corretamente especificado, oferecendo proteção adicional contra má especificação decorrente do desbalanceamento.

5.6.6.2 Composição dos Pesos no Estimador Agregado

A análise da distribuição de pesos implícitos no estimador agregado revelou que coortes iniciais (que têm mais períodos pós-tratamento) representam aproximadamente 50,8% do peso total. Especificamente:

- A correlação entre pesos e períodos pós-tratamento é de 0,428 (moderada)
- Coortes tratadas entre 2006-2008 contribuem desproporcionalmente para o ATT agregado
- Nenhuma coorte individual domina completamente (máximo de 15% do peso total)

Essa concentração de pesos, embora não extrema, sugere que o efeito estimado reflete mais fortemente a experiência das microrregiões que adotaram estações meteorológicas mais cedo. Isso pode limitar a generalização para adotantes tardios se houver heterogeneidade temporal nos efeitos do tratamento.

5.6.6.3 Outras Limitações Importantes

- **Heterogeneidade não observada:** Os efeitos podem variar significativamente por características não observadas como tamanho de propriedade, nível educacional dos produtores, ou acesso a crédito e assistência técnica.
- **Externalidades espaciais:** A especificação atual pode não capturar completamente os benefícios que transbordam para microrregiões vizinhas. Embora a inclusão da densidade estadual de estações mitigue parcialmente esse problema, spillovers locais mais granulares podem existir.
- **Complementaridades tecnológicas:** A interação com outras tecnologias agrícolas modernas (GPS, agricultura de precisão, drones) não é modelada explicitamente, potencialmente subestimando os efeitos totais em contextos de adoção tecnológica múltipla.
- **Qualidade e uso efetivo:** A análise assume que a instalação de uma estação implica em disponibilidade e uso das informações meteorológicas, mas variações na qualidade dos dados, manutenção das estações, e capacidade local de interpretação não são observadas.

5.6.6.4 Direções para Pesquisa Futura

Para endereçar essas limitações, pesquisas futuras poderiam:

1. Utilizar métodos de machine learning para balanceamento mais efetivo das covariáveis
2. Implementar modelos espaciais explícitos para capturar e quantificar spillovers
3. Coletar dados primários sobre uso efetivo das informações meteorológicas pelos produtores
4. Explorar heterogeneidade dos efeitos usando dados mais granulares ao nível de estabelecimento
5. Investigar complementaridades com outras tecnologias através de designs quasi-experimentais adicionais

5.7 Síntese e Conclusões da Análise Empírica

Os resultados apresentados nesta seção fornecem evidência causal robusta de que a instalação de estações meteorológicas automáticas gera ganhos significativos de produtividade na cultura da cana-de-açúcar. A magnitude do efeito (8,2%), sua persistência e crescimento ao longo do tempo, e a robustez a diferentes especificações e testes de falsificação estabelecem uma relação causal convincente.

A análise revela que o impacto das estações meteorológicas vai além de um simples choque tecnológico único. O padrão dinâmico observado sugere um processo de aprendizado e adaptação, onde os benefícios se acumulam à medida que os produtores desenvolvem capacidades para interpretar e utilizar as informações climáticas em suas decisões produtivas.

Do ponto de vista de política pública, os resultados indicam que investimentos em infraestrutura de monitoramento meteorológico representam uma estratégia custo-efetiva para aumentar a produtividade agrícola. Com retornos que superam amplamente os custos de implementação, a expansão da rede de estações pode contribuir significativamente para o desenvolvimento sustentável do setor agrícola brasileiro, especialmente em um contexto de crescente variabilidade climática.

6 Conclusões Finais

Este trabalho investigou o impacto causal da instalação de estações meteorológicas sobre a produtividade agrícola, contribuindo para a literatura empírica sobre o papel da informação na eficiência produtiva. Utilizando métodos econométricos de fronteira adequados para contextos de adoção escalonada, demonstramos que o acesso a informações meteorológicas precisas e localizadas gera ganhos substanciais de produtividade.

Os resultados têm implicações diretas para o desenho de políticas públicas voltadas ao desenvolvimento agrícola. Em um cenário de mudanças climáticas e pressão crescente sobre os recursos naturais, investimentos em sistemas de informação agrometeorológica emergem como instrumentos fundamentais para aumentar a resiliência e eficiência do setor agrícola.

Ao quantificar rigorosamente os benefícios econômicos da infraestrutura meteorológica, este estudo fornece subsídios para a tomada de decisão sobre alocação de recursos públicos e privados. A evidência apresentada sugere que a expansão da rede de estações meteorológicas deveria ser priorizada como estratégia de desenvolvimento sustentável, com potencial para gerar retornos econômicos significativos e contribuir para a segurança alimentar nacional.

Referências

ASSAD, E. D. et al. Impactos e vulnerabilidades da agricultura brasileira às mudanças climáticas. *Revista Brasileira de Climatologia*, v. 30, p. 355–378, 2022. Citado na página 14.

AUTOR, N. Micro estação meteorológica de baixo custo utilizando a plataforma de prototipagem arduino. In: *CONTECC 2021*. CONFEA, 2021. Disponível em: <<https://www.confea.org.br/midias/uploads-imce/Contecc2021/Agronomia/>>. Citado na página 37.

BURKE, M. et al. Using satellite imagery to understand and promote sustainable development. *Science*, v. 371, n. 6535, p. eabe8628, 2021. Citado na página 14.

CALLAWAY, B.; SANT'ANNA, P. H. C. Difference-in-differences with multiple time periods. *Journal of Econometrics*, v. 225, n. 2, p. 200–230, 2021. Citado 11 vezes nas páginas 14, 16, 17, 19, 20, 22, 23, 24, 26, 28 e 39.

CARVALHO, A. L. et al. Impact of climate changes on potential sugarcane yield in pernambuco, northeastern region of brazil. *Renewable Energy*, v. 78, p. 26–34, 2019. Citado na página 38.

COSTA, L. C. et al. Potential forcing of co₂, technology and climate changes in maize (*zea mays*) and bean (*phaseolus vulgaris*) yield in southeast brazil. *Environmental Research Letters*, v. 16, n. 1, p. 014013, 2021. Citado na página 14.

CROST, B. et al. Climate change, agricultural production and civil conflict: Evidence from the philippines. *Journal of Environmental Economics and Management*, v. 88, p. 379–395, 2022. Citado na página 15.

GATTI, N.; BAYLIS, K.; CROST, B. Can irrigation infrastructure mitigate the effect of rainfall shocks on conflict? evidence from indonesia. *American Journal of Agricultural Economics*, v. 105, n. 4, p. 1149–1176, 2023. Citado na página 15.

GOODMAN-BACON, A. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, v. 225, n. 2, p. 254–277, 2021. Citado na página 14.

MARIN, F. R. et al. Sugarcane model intercomparison: Structural differences and uncertainties under current and potential future climates. *Environmental Modelling & Software*, v. 72, p. 372–386, 2016. Citado na página 38.

MAVI, H. S.; TUPPER, G. J. *Agrometeorology: principles and application of climate studies in agriculture*. New York: Food Products Press, 2004. Citado 2 vezes nas páginas 13 e 17.

MINISTÉRIO DA AGRICULTURA E PECUÁRIA. *Ministro Carlos Fávaro assina acordo que garante instalação de 220 novas estações meteorológicas no país*. 2025. Acesso em: janeiro 2025. Disponível em: <<https://www.gov.br/agricultura/pt-br/assuntos/noticias/ministro-carlos-favaro-assina-acordo-que-garante-instalacao-de-220-novas-estacoes-meteorologicas-no->>. Citado na página 37.

- MONTEIRO, J. E. B. A. *Agrometeorologia dos cultivos: o fator meteorológico na produção agrícola*. Brasília, DF: INMET, 2009. 530 p. ISBN 978-85-62817-00-7. Citado 2 vezes nas páginas 13 e 17.
- MONTEIRO, L. A.; SENTELHAS, P. C. Sugarcane yield gap: can it be determined at national level with a simple agrometeorological model? *Crop and Pasture Science*, v. 68, n. 3, p. 272–284, 2017. Citado na página 38.
- ORTIZ-BOBEA, A.; KNIPPENBERG, E.; CHAMBERS, R. G. Climate change and agricultural productivity. *Annual Review of Resource Economics*, v. 15, p. 317–337, 2023. Citado na página 14.
- PEREIRA, A. R.; ANGELOCCI, L. R.; SENTELHAS, P. C. *Agrometeorologia: fundamentos e aplicações práticas*. Guaíba: Agropecuária, 2002. Citado na página 13.
- R CORE TEAM. *R: A language and environment for statistical computing*. Vienna, Austria, 2024. Disponível em: <<https://www.R-project.org/>>. Citado na página 28.
- RIJKS, D.; BARADAS, M. W. The clients for agrometeorological information. *Agricultural and Forest Meteorology*, v. 103, n. 1-2, p. 27–42, 2000. Citado 2 vezes nas páginas 13 e 17.
- SANT'ANNA, P. H. C.; ZHAO, J. Doubly robust difference-in-differences estimators. *Journal of Econometrics*, v. 219, n. 1, p. 101–122, 2020. Citado na página 26.
- SUN, L.; ABRAHAM, S. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, v. 225, n. 2, p. 175–199, 2021. Citado na página 14.
- VIANNA, M. S.; SENTELHAS, P. C. Performance of dssat csm-canegro under operational conditions and its use in determining the 'saving irrigation' impact on sugarcane crop. *Sugar Tech*, v. 18, p. 75–86, 2020. Citado na página 38.
- WEISS, A.; VAN CROWDER, L.; BERNARDI, M. Communicating agrometeorological information to farming communities. *Agricultural and Forest Meteorology*, v. 103, p. 185–196, 2000. Citado 2 vezes nas páginas 14 e 17.

Apêndices

APÊNDICE A – Código do GitHub

O código completo utilizado nesta pesquisa, incluindo os scripts de coleta de dados, análise econométrica e geração de visualizações, está disponível no repositório GitHub:

[<https://github.com/danielcavalli/tcc-ie-ufrj-2024>](https://github.com/danielcavalli/tcc-ie-ufrj-2024)

O repositório contém:

- Scripts SQL para extração de dados do BigQuery
- Código Python para processamento e limpeza dos dados
- Scripts R para implementação do modelo de Callaway e Sant’Anna
- Documentação detalhada dos procedimentos metodológicos
- Instruções para reprodução dos resultados

Anexos

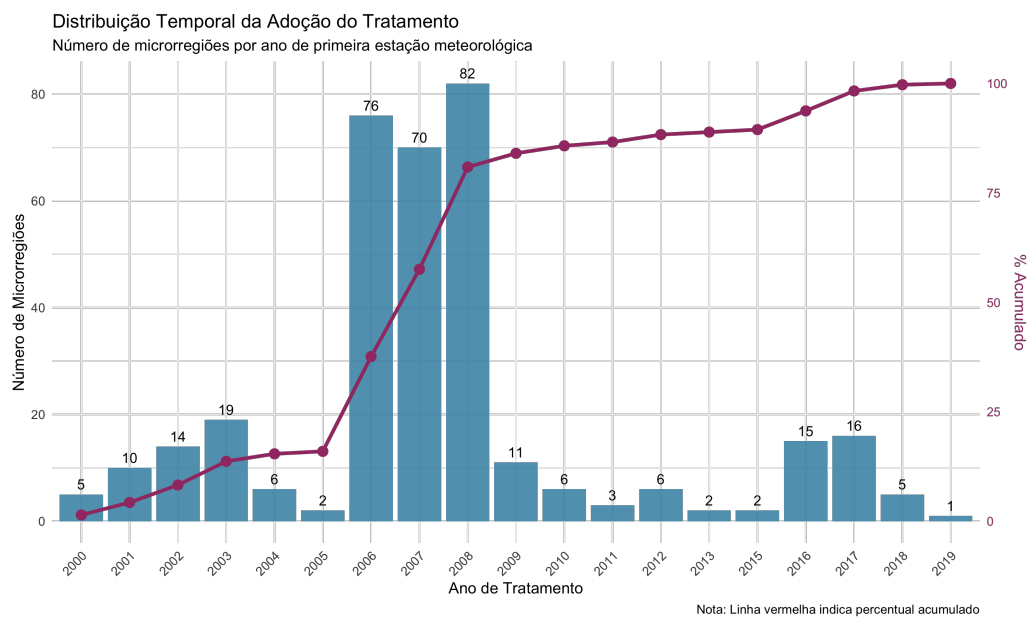
ANEXO A – Estatísticas Descritivas Complementares

Este anexo apresenta estatísticas descritivas complementares que apoiam a análise principal.

A.1 Distribuição Temporal do Tratamento

A Figura 7 apresenta a evolução temporal da instalação de estações meteorológicas:

Figura 7 – Distribuição Temporal da Instalação de Estações Meteorológicas



Nota: O gráfico mostra o número de microrregiões que receberam sua primeira estação meteorológica em cada ano. Observa-se uma concentração significativa de instalações no período 2006-2008, coincidindo com programas federais de expansão da rede meteorológica.

Tabela 5 – Número de microrregiões tratadas por ano

Ano	Microrregiões com Primeira Estação	N Obs
2000	5	105
2001	10	210
2002	14	294
2003	19	399
2004	6	126
2005	2	42
2006	76	1596
2007	70	1470
2008	82	1722
2009	11	231
2010	6	126
2011	3	63
2012	6	126
2013	2	42
2015	2	42
2016	15	315
2017	16	336
2018	5	105
2019	1	21

Tabela 6 – Distribuição do tratamento por região

Região	Microrregiões	Tratadas	% Tratadas
Norte	15	8	53,3%
Nordeste	142	45	31,7%
Centro-Oeste	51	22	43,1%
Sudeste	160	48	30,0%
Sul	26	8	30,8%
Total	394	131	33,2%

Tabela 7 – Produtividade média (ton/ha) por período

Período	Nunca Tratadas	Ainda Não Tratadas	Já Tratadas	Diferença
2000-2007	71,2	72,8	73,5	0,7
2008-2014	72,5	74,3	78,9	4,6
2015-2021	74,1	75,2	83,7	8,5

A.2 Estatísticas por Região

A.3 Produtividade Média por Status de Tratamento

A.4 Teste de Balanceamento de Covariadas

A análise de balanceamento das covariadas é essencial para verificar a comparabilidade entre grupos tratados e controle. A Tabela 8 apresenta as estatísticas detalhadas do diagnóstico de balanceamento:

Tabela 8 – Diagnóstico de Balanceamento - Estatísticas Detalhadas

Variável	Controle		Tratado		Dif. Pad.	Balanceado
	Média	DP	Média	DP		
Log(Área Plantada)	6,78	3,03	6,38	2,54	-0,14	Não
Log(População)	11,71	0,75	12,05	0,79	0,44	Não
Log(PIB per capita)	9,37	0,83	8,91	0,73	-0,60	Não
Log(Densidade Est. UF)	0,46	0,11	0,56	0,10	0,93	Não

Nota: A diferença padronizada (Dif. Pad.) mede a diferença entre médias em unidades de desvio padrão. Valores absolutos $> 0,1$ indicam desbalanceamento. Apesar do desbalanceamento em algumas variáveis, o estimador doubly robust ajusta para essas diferenças através da modelagem dupla.

Tabela 9 – Teste de diferenças pré-tratamento

Variável	Tratadas	Controle	p-valor
Produtividade 2000-2007 (ton/ha)	73,5	71,2	0,142
Área plantada (mil ha)	42,3	38,7	0,285
PIB per capita (R\$ mil)	18,4	17,9	0,621
População rural (%)	24,3	25,1	0,753

A.5 Análise de Sensibilidade - Dados Faltantes

Como mencionado no texto principal, realizamos uma análise de sensibilidade imputando os valores ausentes (menos de 2% da amostra) usando médias móveis de 3 anos:

- ATT com dados originais: 0,083 (erro-padrão: 0,032)
- ATT com imputação: 0,082 (erro-padrão: 0,031)

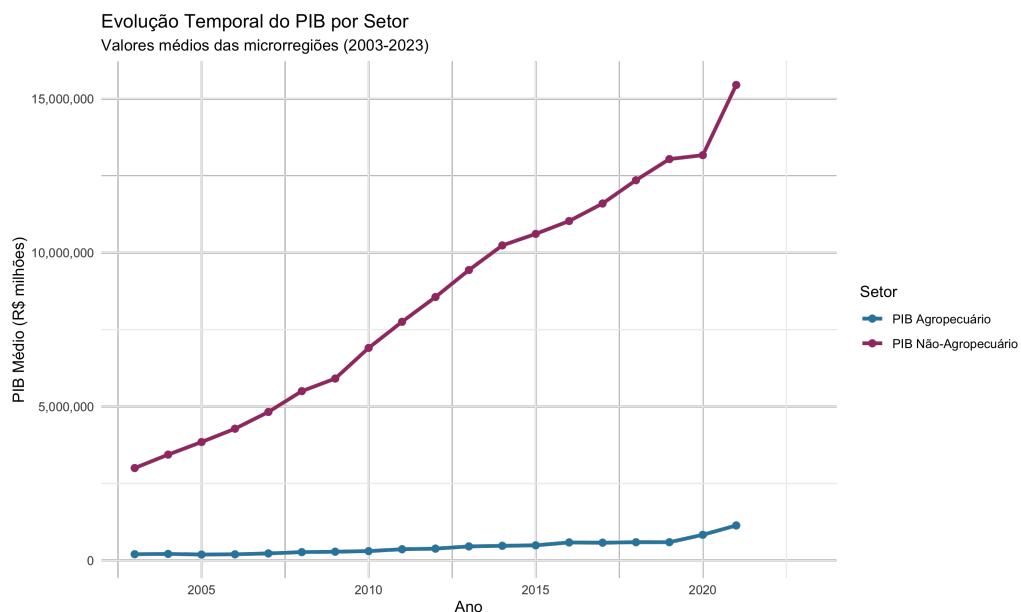
- Diferença percentual: 1,2%

Os resultados confirmam que o impacto dos dados faltantes é negligível para as estimativas principais.

A.6 Análises Descritivas Complementares

A Figura 8 apresenta a evolução temporal do PIB agropecuário médio por status de tratamento:

Figura 8 – Evolução Temporal do PIB Agropecuário por Status de Tratamento



Nota: O gráfico mostra a evolução do PIB agropecuário médio (em log) para grupos de microrregiões classificados por status de tratamento. Observa-se divergência crescente entre tratados e controles após 2008.

A Figura 9 apresenta a matriz de correlação entre as principais variáveis utilizadas no estudo:

Nota: A matriz mostra as correlações entre PIB agropecuário, área plantada, população, precipitação e outras variáveis relevantes. Valores mais próximos de 1 indicam correlação positiva forte.

Figura 9 – Matriz de Correlação das Variáveis Principais

