

UNIVERSIDADE FEDERAL DE ALAGOAS
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
CIÊNCIAS ECONÔMICAS

LUCAS MATHEUS FALCÃO DA SILVA

TAXA DE CESÁREA EM HOSPITAIS PÚBLICOS E PRIVADOS NO BRASIL: UMA
ANÁLISE DA DEMANDA INDUZIDA PELA OFERTA, 2014–2022

MACEIÓ
2026

LUCAS MATHEUS FALCÃO DA SILVA

Taxa de cesárea em hospitais públicos e privados no Brasil: uma análise da demanda induzida
pela oferta, 2014–2022

Trabalho de Conclusão de Curso apresentado para
o Curso de Ciências Econômicas da Universidade
Federal de Alagoas, como requisito parcial para
a obtenção do grau de Bacharel em Ciências
Econômicas.

Orientador: Prof. Dr. Keuler Hissa Teixeira

Catálogo na fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico

Bibliotecária: Helena Cristina Pimentel do Vale CRB-4/661

- S586t Silva, Lucas Matheus Falcão da.
Taxa de cesárea em hospitais públicos e privados no Brasil : uma análise da demanda induzida pela oferta, 2014–2022 / Lucas Matheus Falcão da Silva. – 2026.
127 f. : il.
- Orientador: Keuler Hissa Teixeira.
Monografia (Trabalho de Conclusão de Curso em Ciências Econômicas) –
Universidade Federal de Alagoas, Faculdade de Economia. Administração e Contabilidade.
Maceió, 2026.
- Bibliografia: f. 98-103.
Apêndices: f. 104-127.
1. Cesárea. 2. Demanda induzida pela oferta. 3. Classificação de Robson. 4. Diferenças em diferenças. I. Título.

CDU: 330.163:614

LUCAS MATHEUS FALCÃO DA SILVA

Taxa de cesárea em hospitais públicos e privados no Brasil: uma análise da demanda induzida pela oferta, 2014–2022

Trabalho de Conclusão de Curso apresentado para o Curso de Ciências Econômicas da Universidade Federal de Alagoas, como requisito parcial para a obtenção do grau de Bacharel em Ciências Econômicas.

Data de Aprovação:

Banca Examinadora



Documento assinado digitalmente
KEULER HISSA TEIXEIRA
Data: 01/07/2026 15:18:53-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Keuler Hissa Teixeira
Universidade Federal de Alagoas
Orientador



Documento assinado digitalmente
ADHEMAR RANCIARO NETO
Data: 01/07/2026 15:50:23-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Adhemar Ranciaro Neto
Universidade Federal de Alagoas
Examinador



Documento assinado digitalmente
ANA CAROLINA DA CRUZ LIMA
Data: 01/07/2026 16:19:11-0300
Verifique em <https://validar.iti.gov.br>

Prof.^a Dra. Ana Carolina da Cruz Lima
Universidade Federal do Rio de Janeiro
Examinadora

Dedico este trabalho ao meu avô Belo, que
ainda em vida contava a todos que eu seria um
Cientista.

AGRADECIMENTOS

Agradeço, antes de tudo, às mulheres da minha vida, que me forjaram, me criaram e me deram todo o apoio: minha mãe, Mirta Falcão, e minha avó, Mariza Falcão. Com uma força imensa para lutar, foram elas o alicerce de toda esta longa trajetória. Ao lado da minha avó Mariza, meu avô, Abelardo Falcão, cuidou de mim nos anos em que minha mãe precisava trabalhar. Minha mãe, batalhadora, sustentou sozinha uma casa e um filho, como mãe solo. Dela vem o meu maior exemplo.

Agradeço à minha companheira, Beatriz Rodrigues, fundamental do início ao fim desta jornada. Esteve ao meu lado nos momentos mais difíceis, trouxe luz quando eu mais precisava e me ajudou a tomar uma das decisões mais importantes da minha vida acadêmica: a reopção para o curso de Ciências Econômicas.

Ao meu primo, Miguel Falcão, que me fez companhia em tantas noites de estudo.

Ao meu orientador, professor, tutor e amigo, Prof. Dr. Keuler Hissa Teixeira, por quem tenho o maior carinho, respeito e admiração. Keuler foi fundamental em toda a minha trajetória no curso de Economia; levo comigo os seus ensinamentos e espero repassá-los adiante.

Aos meus amigos Vinícius Ventura e João Victor, que me deram muita força na reta final deste estudo, me tranquilizando a cada dia. Aos meus amigos Gustavo Parizzato e Mateus Amaral, que deram todo o apoio.

Ao Programa de Educação Tutorial (PET), do qual fiz parte durante quase toda a graduação. A bolsa do PET foi o meu sustento nesse período; sem esse apoio, teria sido muito difícil continuar no curso. Mas o PET foi além da bolsa: foi um divisor de águas na minha formação. Com a tutoria do Prof. Dr. Keuler Hissa, aprendi muito além da sala de aula, e tudo o que vivi ali foi essencial para o meu amadurecimento. Agradeço, em especial, aos petianos Beatriz Duarte, Gabriel Raymond, João Pedro, Wellington Domingos e Laila Januário, e a todos os demais colegas.

À Faculdade de Economia, Administração e Contabilidade da UFAL e, além do meu orientador, aos professores que marcaram a minha formação: Prof. Dr. Cid Olival, Profa. Dra. Camila Herminda, Prof. Dr. Anderson Aristides, Profa. Dra. Verônica Antunes, Prof. Dr. Roberto Simiqueli e Prof. Dr. Márcio Porangaba.

Aos colegas de turma, em especial Mateus Andrade e Nathaly Oliveira, que dividiram comigo grande parte da caminhada da graduação.

À Universidade Federal de Alagoas, minha segunda casa durante toda esta trajetória. Entrei em 2019, no curso de Química, e mudei para as Ciências Econômicas pelo programa de reopção, onde encontrei o meu caminho.

À Governança Corporativa e aos colegas de trabalho que foram fundamentais nesta jornada de estudos: João Victor, Fabrício Almeida, Fabrícia Casado, Poliana Santana e Vinícius Ventura.

Por fim, agradeço de modo especial aos professores que gentilmente aceitaram compor a banca examinadora deste trabalho, Prof. Dr. Adhemar Ranciaro Neto e Profa. Dra. Ana Carolina da Cruz Lima, pela disponibilidade e pelas contribuições que certamente o enriquecerão.

RESUMO

A taxa brasileira de cesárea está entre as mais altas do mundo e se concentra, de forma persistente, nos hospitais privados lucrativos, mesmo entre gestantes de baixo risco clínico. Este estudo investiga em que medida esse diferencial entre hospitais é compatível com a hipótese de demanda induzida pela oferta (*supplier-induced demand*, SID) e qual o efeito causal de uma intervenção institucional sobre esse diferencial. Para isso, combina o Sistema de Informações sobre Nascidos Vivos (SINASC) e o Cadastro Nacional de Estabelecimentos de Saúde (CNES) no período de 2014 a 2022, com amostra principal nos Grupos 1 e 3 da Classificação de Robson, a população em que a cesárea é menos indicada clinicamente. A estratégia empírica combina modelos de *Machine Learning*, um conjunto de estimadores observacionais com pressupostos distintos e um conjunto moderno de estimadores de diferenças-em-diferenças aplicados à adesão de hospitais ao Projeto Parto Adequado da Agência Nacional de Saúde Suplementar. Os resultados convergem: a estrutura do hospital pesa mais do que o perfil da gestante na previsão da via de parto; o diferencial dos hospitais privados lucrativos permanece após controles e é compatível com a demanda induzida pela oferta; e a adesão ao Parto Adequado reduz a cesárea nos hospitais aderentes, com efeito concentrado no setor privado lucrativo. As conclusões indicam que o excesso de cesárea na população de baixo risco tem origem mais institucional do que individual, e que uma política dirigida ao arranjo institucional produz efeito.

Palavras-chave: Cesárea; Demanda induzida pela oferta; *Machine Learning*; Classificação de Robson; Diferenças em diferenças.

ABSTRACT

Brazil's cesarean section rate is among the highest in the world and is persistently concentrated in for-profit private hospitals, even among women at low clinical risk. This study investigates the extent to which this gap between hospitals is compatible with the supplier-induced demand (SID) hypothesis and the causal effect of an institutional intervention on that gap. To this end, it combines the National Live Birth Information System (SINASC) and the National Registry of Health Establishments (CNES) for the 2014–2022 period, with a main sample of Robson Groups 1 and 3, the population with the least clinical indication for cesarean. The empirical strategy combines *Machine Learning* models, a set of observational estimators with distinct assumptions, and a modern set of difference-in-differences estimators applied to hospital adherence to the Adequate Childbirth Project (*Projeto Parto Adequado*) of the National Supplementary Health Agency. The results converge: hospital structure weighs more than the maternal profile in predicting the mode of delivery; the differential of for-profit private hospitals persists after controls and is compatible with supplier-induced demand; and adherence to the Adequate Childbirth Project reduces cesarean rates in adopting hospitals, with the effect concentrated in the for-profit private sector. The conclusions indicate that the excess of cesareans in the low-risk population is more institutional than individual in origin, and that a policy targeting the institutional arrangement produces an effect.

Keywords: Cesarean section; Supplier-induced demand; *Machine Learning*; Robson Classification; Difference in differences.

LISTA DE FIGURAS

Figura 1 – Taxa de cesárea por unidade da federação. Brasil, partos hospitalares, 2010 e 2022.	39
Figura 2 – Taxa de cesárea por natureza jurídica do hospital. Brasil, 2012 e 2022. . . .	42
Figura 3 – Taxa de cesárea por grupo de Robson, com Grupos 1 e 3 (amostra principal) em destaque. Brasil, 2022.	45
Figura 4 – Distribuição da idade materna por via de parto. Robson 1+3, Brasil, 2014–2022.	46
Figura 5 – Estratégia empírica: as quatro perguntas, as famílias de instrumentos que respondem a cada uma e a leitura conjunta triangulada.	53
Figura 6 – Ganho de previsibilidade da via de parto (AUROC) por bloco cumulativo de variáveis. Amostra principal (Robson 1+3).	70
Figura 7 – Razão de chance de cesárea por natureza jurídica do hospital, em três especificações. Amostra principal (Robson 1+3), 2014–2022.	73
Figura 8 – Balanço das 20 covariáveis mais desbalanceadas (SMD) entre privado e público, antes e depois do IPW. Robson 1+3, 2014–2022.	75
Figura 9 – Distribuição do escore de propensão por setor, amostra principal Robson 1+3.	76
Figura 10 – Triangulação dos estimadores observacionais do diferencial de cesárea entre privado lucrativo e público (p.p.), com IC 95%. Robson 1+3.	78
Figura 11 – Efeito da mudança de natureza jurídica sobre a taxa de cesárea, por subgrupo de transição (<i>event study</i>), com IC 95%.	84
Figura 12 – Efeito do Projeto Parto Adequado sobre a taxa de cesárea por estimador de diferença-em-diferenças, com IC 95%. Robson 1+3.	89
Figura 13 – Efeito do Projeto Parto Adequado sobre a taxa de cesárea por região, com IC 95%. Robson 1+3.	91
Figura 14 – Efeito do PPA por tempo relativo à adesão (<i>event study</i>): estimadores de Sun e Abraham (2021) e Callaway e Sant’Anna (2021). Robson 1+3, 2014–2022.	92
Figura A1 – Taxa de cesárea por perfil de atendimento do hospital.	109
Figura A2 – Distribuição da apresentação fetal e da idade gestacional.	111
Figura A3 – Distribuição do peso ao nascer por via de parto.	112
Figura B1 – Distribuição do escore de propensão por setor, amostra Robson 1–4 (baixo risco amplo).	114
Figura B2 – Distribuição do escore de propensão por setor, amostra Robson 1–10 (inclui o Grupo 5).	115

LISTA DE TABELAS

Tabela 1 – Distribuição de hospitais e de nascidos vivos por natureza jurídica, Brasil, 2012 e 2022.	41
Tabela 2 – Taxa de cesárea por grupo de Robson, Brasil, 2022. Os Grupos 1 e 3 (em destaque) compõem a amostra principal deste trabalho.	43
Tabela 3 – Balance pré-tratamento entre tratados e controles, bruto e após reponderação por escore de propensão, Robson 1+3, 2014–2022.	48
Tabela 4 – Fluxo amostral do SINASC–CNES integrado (2014–2022) às amostras analíticas.	50
Tabela 5 – Razão de chances (odds ratio) de cesárea por natureza jurídica, três especificações de logit, Robson 1+3, Brasil, 2014–2022.	71
Tabela 6 – ATE e ATT do parto em hospital privado vs. público sobre a probabilidade de cesárea, estimados por inverse probability weighting (IPW), Brasil, 2014–2022.	74
Tabela 7 – Tratamento, covariáveis, estimando e estimativa de cada estimador da Pergunta 3.	77
Tabela 8 – Robustez causal: diferencial privado lucrativo vs. público entre estimadores, Robson 1+3, Brasil, 2014–2022.	79
Tabela 9 – Heterogeneidade do diferencial privado lucrativo: logit Especificação 3 estratificado, Robson 1+3.	80
Tabela 10 – Painel hospital $\times$ ano da taxa de cesárea por natureza jurídica, WLS, Robson 1+3, Brasil, 2014–2022.	82
Tabela 11 – Efeito da informação obstétrica da gestante sobre a probabilidade de cesárea, exercício de médicas como pacientes (Johnson e Rehavi (2016)), setor privado, escolaridade Superior, Robson 1+3, Brasil, 2014–2022.	86
Tabela 12 – Efeito do Projeto Parto Adequado sobre a taxa de cesárea, diferenças em diferenças com efeitos fixos de hospital e ano (TWFE), Robson 1+3, Brasil, 2014–2022.	88
Tabela 13 – ATT do Projeto Parto Adequado sobre a taxa de cesárea por tempo relativo à adesão, estimador de Sun e Abraham (2021), Robson 1+3, Brasil, 2014–2022.	88
Tabela 14 – ATT do Projeto Parto Adequado sobre a taxa de cesárea por tempo relativo à adesão, estimador de Callaway e Sant’Anna (2021), Robson 1+3, Brasil, 2014–2022.	92
Tabela A1 – Nascimentos hospitalares e tipo de parto, Brasil, 2010–2022.	105
Tabela A2 – Nascimentos hospitalares e taxa de cesárea por ano, Brasil, 2010–2022.	105
Tabela A3 – Taxa de cesárea por região, Brasil: 2010, 2022, variação 2010–2022 e agregado do período.	106
Tabela A4 – Taxa de cesárea por unidade da federação, Brasil: 2010, 2022, variação 2010–2022 e agregado do período.	107
Tabela A5 – Estabelecimentos com nascimentos registrados, nascidos vivos e densidade obstétrica por região, Brasil, 2010 e 2022.	108

Tabela A6 – Distribuição de hospitais e de nascidos vivos por perfil de atendimento ^a , Brasil, 2010 e 2022.	108
Tabela A7 – Indicadores de infraestrutura instalada por perfil de atendimento, hospitais com nascimentos registrados, Brasil, 2010 e 2022.	108
Tabela A8 – Indicadores de infraestrutura instalada por perfil de atendimento ^a e natureza jurídica ^b , hospitais com nascimentos registrados, Brasil, 2022.	109
Tabela A9 – Locomoção da gestante e tipo de parto, Brasil: 2010, 2022, variação 2010–2022 e agregado do período.	110
Tabela A10–Distribuição de partos e taxa de cesárea por grupo de Robson, Brasil, 2022.	110
Tabela A11–Decomposição da variação da taxa de cesárea entre 2012 e 2022 nos grupos de Robson 1–10, Brasil.	111
Tabela A12–Taxa de cesárea segundo número de cesáreas anteriores declaradas, Brasil, 2022.	111
Tabela A13–Distribuição de APGAR ao 1º e ao 5º minuto por tipo de parto, Brasil, 2022.	112
Tabela A14–Decomposição preditiva por subconjuntos de variáveis (ablation), XGBoost, holdout temporal 2021–2022, Robson 1+3.	113
Tabela C1 – Bounds de viés de variáveis omitidas no LPM (estabilidade de coeficientes), Robson 1+3, Brasil, 2014–2022.	116
Tabela C2 – E-value para confundimento não observado, privado lucrativo vs. público (logit Especificação 3).	116
Tabela C3 – Estabilidade temporal do diferencial privado lucrativo: logit Especificação 3 por ano e por janela, Robson 1+3, 2014–2022.	117
Tabela C4 – Teste de Wald conjunto sobre tendências paralelas, cohort 2017, Robson 1+3, Brasil, 2014–2022.	117
Tabela C5 – Bateria de robustez do efeito do Projeto Parto Adequado sobre a taxa de cesárea: estimadores modernos de DiD, placebos, elegibilidade do controle e heterogeneidade, Robson 1+3, Brasil, 2014–2022.	118
Tabela C6 – Efeito do Projeto Parto Adequado sobre a taxa de cesárea por nível de informação obstétrica da gestante (L0–L4), setor privado, Robson 1+3, Brasil, 2014–2022.	119
Tabela D1 – Perfil dos 151 hospitais que mudaram de natureza jurídica entre 2014 e 2022, por subgrupo de transição, frente aos nunca tratados, Robson 1+3.	120
Tabela D2 – Event study TWFE dos 151 hospitais que mudaram de natureza jurídica entre 2014 e 2022, taxa de cesárea (p.p.), Brasil.	121
Tabela E1 – Bateria de robustez do exercício de médicas como pacientes, setor privado, escolaridade Superior, Robson 1+3, Brasil, 2014–2022.	122
Tabela F1 – Tempo de execução por etapa na amostra principal Robson 1+3 (8,9 milhões de partos).	125

LISTA DE QUADROS

Quadro 1 – Síntese da literatura empírica sobre cesárea e demanda induzida pela oferta.	34
Quadro 2 – Agrupamento dos códigos de natureza jurídica (IBGE/Receita Federal) em categorias analíticas.	40
Quadro 3 – Roteiro metodológico: as quatro perguntas, os métodos e a finalidade de cada passo.	51

LISTA DE ABREVIATURAS E SIGLAS

ACOG	<i>American College of Obstetricians and Gynecologists</i>
AIPW	<i>Augmented Inverse Probability Weighting</i>
AME	Efeito Marginal Médio (<i>Average Marginal Effect</i>)
ANS	Agência Nacional de Saúde Suplementar
ARP	Procedimento de inferência de Andrews-Roth-Pakes (<i>Honest DiD</i>)
ATE	Efeito Médio do Tratamento (<i>Average Treatment Effect</i>)
ATT	Efeito Médio do Tratamento sobre os Tratados (<i>Average Treatment Effect on the Treated</i>)
AUPRC	Área Sob a Curva Precisão-Revocação (<i>Area Under the Precision-Recall Curve</i>)
AUROC	Área Sob a Curva ROC (<i>Area Under the Receiver Operating Characteristic Curve</i>)
CBO	Classificação Brasileira de Ocupações
CDC	<i>Centers for Disease Control and Prevention</i>
CI2	Ciclo Intensivo 2 (terceira fase do PPA, 2020–2022)
CNES	Cadastro Nacional de Estabelecimentos de Saúde
CNPJ	Cadastro Nacional da Pessoa Jurídica
CPU	Unidade Central de Processamento (<i>Central Processing Unit</i>)
DATASUS	Departamento de Informática do SUS
DiD	Diferenças em Diferenças (<i>Difference in Differences</i>)
EB	<i>Entropy Balancing</i>
EF	Efeitos Fixos
FLCI	Intervalos de Confiança de Comprimento Fixo (<i>Fixed Length Confidence Intervals</i>)
GLMM	<i>Generalized Linear Mixed Model</i>
GPU	Unidade de Processamento Gráfico (<i>Graphics Processing Unit</i>)
IBGE	Instituto Brasileiro de Geografia e Estatística
IC	Intervalo de Confiança
ICC	Coeficiente de Correlação Intraclasse (<i>Intraclass Correlation Coefficient</i>)
IESS	Instituto de Estudos de Saúde Suplementar
IHI	<i>Institute for Healthcare Improvement</i>
IPW	<i>Inverse Probability Weighting</i>
IW	<i>Interaction-Weighted</i> (estimador de Sun e Abraham)

LPM	Modelo de Probabilidade Linear (<i>Linear Probability Model</i>)
MDE	Mínimo Efeito Detectável (<i>Minimum Detectable Effect</i>)
NJ	Natureza Jurídica
OMS	Organização Mundial da Saúde
OR	Razão de Chances (<i>Odds Ratio</i>)
OVB	Viés por Variáveis Omitidas (<i>Omitted Variable Bias</i>)
PPA	Projeto Parto Adequado
PS	Escore de Propensão (<i>Propensity Score</i>)
PSM	<i>Propensity Score Matching</i>
RM	<i>Relative Magnitudes</i> (Honest DiD de Rambachan e Roth)
RR	Risco Relativo
SE	Erro-Padrão (<i>Standard Error</i>)
SFL	Sem Fins Lucrativos
SHAP	<i>SHapley Additive exPlanations</i>
SID	Demanda Induzida pela Oferta (<i>Supplier-Induced Demand</i>)
SINASC	Sistema de Informações sobre Nascidos Vivos
SMD	Diferença Padronizada de Médias (<i>Standardized Mean Difference</i>)
SUS	Sistema Único de Saúde
TWFE	Efeitos Fixos de Duas Vias (<i>Two-Way Fixed Effects</i>)
UF	Unidade da Federação

SUMÁRIO

1	INTRODUÇÃO	14
2	REVISÃO DE LITERATURA	17
2.1	Cesárea: história, importância clínica e a faixa adequada	17
2.2	O custo do excesso e sua escala global	20
2.3	O caso brasileiro	21
2.3.1	O lado da oferta: por que a Economia da Saúde tem algo a dizer	23
2.3.2	A política pública e seus limites: o caso do Parto Adequado	25
2.4	Revisão de literatura empírica	26
2.4.1	Estudos internacionais	26
2.4.2	Estudos nacionais	29
2.4.3	Estudos regionais e estaduais	32
3	METODOLOGIA E BASE DE DADOS	38
3.1	Base de dados	38
3.2	A heterogeneidade espacial	39
3.3	A estrutura institucional da rede hospitalar	40
3.4	A Classificação de Robson e a amostra principal Robson 1+3	42
3.5	Perfil sociodemográfico das gestantes	46
3.6	Balanceamento entre tratados e controles (pré- e pós-reponderação) . .	47
3.7	Síntese da análise exploratória	48
3.8	Base integrada e amostras analíticas	49
3.9	Estratégia empírica: quatro perguntas, quatro famílias de instrumentos	51
3.10	Modelos estimados e especificações	55
3.10.1	Estimadores de escore de propensão	55
3.10.2	Estruturas de painel e efeito aleatório por hospital	59
3.10.3	Sensibilidade a inobservados e decomposição preditiva	61
3.10.4	Médicas como pacientes: teste de assimetria informacional	63
3.11	Identificação por choque regulatório: desenho DiD escalonado	64
4	ESTIMAÇÃO E RESULTADOS	67
4.1	Visão geral dos achados	67
4.2	A previsibilidade da decisão	68
4.3	O peso relativo de oferta e demanda	69

4.4	O diferencial público × privado e a hipótese de demanda induzida pela oferta	70
4.4.1	O diferencial após controles	71
4.4.2	Heterogeneidade do diferencial	80
4.4.3	Compatibilidade com SID e a ameaça do painel <i>within-hospital</i>	81
4.4.4	Médicas como pacientes: um teste do canal informacional	84
4.5	O efeito causal quase-experimental do Projeto Parto Adequado	86
4.6	Síntese: a leitura conjunta das quatro perguntas	92
4.7	Limitações e fronteiras da inferência	93
5	CONCLUSÃO	96
	REFERÊNCIAS	98

	APÊNDICES	104
--	----------------------------	------------

	APÊNDICE A – ESTATÍSTICA DESCRITIVA COMPLEMENTAR	105
--	---	------------

A.1	Série histórica e totais nacionais	105
A.2	Distribuição espacial	106
A.3	Estrutura institucional da rede	107
A.4	Locomoção da gestante	109
A.5	Classificação de Robson	110
A.6	Perfil sociodemográfico e desfechos neonatais	112
A.7	Decomposição preditiva por blocos	112

	APÊNDICE B – DIAGNÓSTICOS DE SUPORTE COMUM	114
--	---	------------

	APÊNDICE C – ROBUSTEZ COMPLEMENTAR DA ESTIMAÇÃO CAUSAL	116
--	---	------------

C.1	Sensibilidade a inobservados	116
C.2	Estabilidade temporal do diferencial	116
C.3	Robustez do DiD do Projeto Parto Adequado	117
C.4	Comparabilidade: restrição vs. ajuste	119

	APÊNDICE D – PERFIL DOS HOSPITAIS QUE MUDARAM DE NATUREZA JURÍDICA	120
--	---	------------

	APÊNDICE E – MÉDICAS COMO PACIENTES: TESTES DE ROBUSTEZ	122
--	--	------------

	APÊNDICE F – REPRODUTIBILIDADE COMPUTACIONAL	124
--	---	------------

F.1	Ambiente de software	124
F.1.1	Python	124

F.1.2	R	124
F.1.3	Hardware	124
F.2	Reprodutibilidade numérica	125
F.3	Tempos de execução típicos	125
F.4	Limitações computacionais reconhecidas	125
F.4.1	Entropy balancing na amostra Robson 1–10	125
F.4.2	GLMM e PSM na amostra completa	126
F.4.3	Gestão de memória na amostra completa	126

1 INTRODUÇÃO

A taxa brasileira de cesárea está entre as mais altas do mundo. Em 2022, 58,7% dos partos hospitalares no país foram cirúrgicos, mais do que o triplo da faixa que a Organização Mundial da Saúde reconhece como referência populacional (Betrán *et al.*, 2016; Boerma *et al.*, 2018). O excesso persiste há pelo menos três décadas e se reparte de forma desigual entre os subsistemas de saúde: em 2022, a taxa alcançou 80,7% no setor privado lucrativo e ficou em 47,1% na rede pública, um diferencial bruto de 33,6 pontos percentuais que resistiu às sucessivas campanhas federais e regulatórias pela redução da cirurgia (Brasil. Ministério da Saúde, 2022; Agência Nacional de Saúde Suplementar (ANS), 2017).

Quando a comparação se volta para hospitais públicos e privados, o problema sai do terreno estritamente clínico e entra no da economia da saúde. A cesárea é remunerada de forma distinta do parto vaginal, em geral em valor maior por unidade de tempo médico no segmento pago por procedimento¹. A decisão é tomada por um profissional que detém todo o conhecimento clínico relevante, em forte assimetria de informação com a paciente (Arrow, 1963).

Parte do custo do excesso é socializada, por meio de internações neonatais, de complicações em gestações futuras e de isenções tributárias concedidas a planos de saúde. E o sistema é dual: o Sistema Único de Saúde (SUS) opera ao lado de uma saúde suplementar regulada pela Agência Nacional de Saúde Suplementar (ANS), e é na fronteira entre os dois subsistemas que o excesso se concentra. Reunidos, esses elementos colocam a cesárea no problema clássico da demanda induzida pela oferta², na tradição de Arrow (1963), Evans (1974) e Gruber e Owings (1996).

A preferência das gestantes também não explica o excesso. Hopkins (2000), Béhague, Victora e Barros (2002) e Domingues *et al.* (2014) mostraram que a maioria das mulheres brasileiras prefere o parto vaginal e que a distância entre a vontade declarada e o desfecho efetivo é regular, concentrada em hospitais privados e crescente ao longo do pré-natal, independentemente de complicações clínicas. Resta, então, examinar o lado da oferta: o estabelecimento, sua estrutura de propriedade, o circuito de financiamento a que responde e a margem de decisão clínica que cada arranjo institucional preserva.

A partir desse quadro, o trabalho pergunta em que medida o diferencial sistemático de

¹ No Brasil, o reembolso de operadoras privadas pelo parto cesáreo costuma ficar acima do reembolso pelo parto vaginal, e o tempo médico exigido pela cirurgia é menor e mais previsível que o tempo de assistência a um trabalho de parto. Esse diferencial de remuneração e de tempo cria incentivo à substituição do parto vaginal pela cesárea no segmento remunerado por procedimento (Gruber; Owings, 1996).

² Demanda induzida pela oferta (*supplier-induced demand*, SID): situação em que o prestador, valendo-se da assimetria de informação que o separa do paciente, influencia a quantidade de serviço consumida para além do que a condição clínica justificaria (Evans, 1974; Dranove; Wehner, 1994).

cesárea entre os hospitais brasileiros é compatível com a hipótese de demanda induzida pela oferta e qual o efeito causal de uma intervenção institucional sobre esse diferencial. A pergunta tem duas partes que se completam. A primeira é interpretativa. Ela examina se o diferencial entre as categorias institucionais (público, privado sem fins lucrativos e privado lucrativo) permanece depois de controlar pelo perfil clínico e sociodemográfico das gestantes, numa população de baixo risco elegível ao parto vaginal. Se permanecer, o diferencial é compatível com a demanda induzida pela oferta. Ainda assim, a compatibilidade observada não basta para provar o canal causal.

A segunda é quase-experimental: examina se a adesão gradual de hospitais ao Projeto Parto Adequado (PPA) da ANS, entre 2015 e 2022, produziu redução causalmente identificável da cesárea nos hospitais que aderiram, e em quais deles essa redução se concentrou. A primeira parte estabelece que há um diferencial sistemático a reduzir; a segunda testa se uma política dirigida a esse diferencial entrega resultado.

Responder às duas partes exige quatro passos encadeados, em que cada um prepara o seguinte. O primeiro pergunta quanto da decisão sobre a via de parto é previsível a partir das informações administrativas disponíveis. É um exercício preditivo, que fixa o limite máximo de previsibilidade e mostra quanto da decisão escapa a qualquer variável observada. O segundo, quanto dessa previsibilidade vem do perfil da gestante e quanto vem da estrutura do hospital, para identificar qual deles importa mais para o desfecho. O terceiro, se o diferencial entre as categorias institucionais sobrevive depois de igualar o perfil das gestantes, no subgrupo clinicamente mais comparável da Classificação de Robson³; é o ponto em que a evidência passa a ser compatível com a demanda induzida pela oferta. O quarto, se a adesão ao Projeto Parto Adequado de fato reduziu a cesárea, e em que tipo de hospital.

Os dois primeiros passos são descritivos; o terceiro qualifica a leitura como compatível com o mecanismo; o quarto testa o efeito de uma política sobre esse mecanismo.

A resposta empírica combina o Sistema de Informações sobre Nascidos Vivos (SINASC) ao Cadastro Nacional de Estabelecimentos de Saúde (CNES), entre 2014 e 2022, cerca de 25,3 milhões de partos hospitalares em aproximadamente cinco mil hospitais. A amostra principal restringe-se aos Grupos 1 e 3 da Classificação de Robson, 8,9 milhões de partos. Os dois primeiros passos empregam modelos de *Machine Learning*; o terceiro, um conjunto de estimadores observacionais com pressupostos distintos; o quarto, um conjunto de estimadores

³ Grupos 1 e 3 da Classificação de Robson (Robson, 2001; Organização Mundial da Saúde (OMS), 2017): gestantes de feto único, cefálico, a termo e com trabalho de parto iniciado de forma espontânea.

modernos de diferenças-em-diferenças, apropriados à entrada gradual dos hospitais no programa. O detalhamento de cada modelo, com as fórmulas e os pressupostos de identificação, está na seção de metodologia.

Quatro escolhas complementares tornam o estudo relevante: a janela de 2014 a 2022, que cobre as três fases do programa; o recorte clínico de Robson 1+3, que isola a população de maior discricionabilidade; a tipologia tripartite das categorias institucionais, em lugar do binário público *versus* privado; e a leitura por coortes de adesão, que permite examinar as tendências paralelas.

Além desta introdução, o trabalho está organizado da seguinte forma. A seção 2 revisa a literatura: reconstrói a trajetória clínica e histórica da cesárea, a faixa adequada reconhecida pela OMS e o caso brasileiro como ponta extrema do fenômeno, e percorre a literatura empírica internacional, nacional e regional sobre a cesárea e a demanda induzida pela oferta. A seção 3 apresenta as bases SINASC e CNES, a estrutura institucional da rede hospitalar e o perfil das gestantes, e formaliza a estratégia empírica e os modelos estimados. A seção 4 estima os modelos e apresenta os resultados. A seção 5 sintetiza as conclusões, discute as implicações e registra a agenda de pesquisa.

2 REVISÃO DE LITERATURA

Esta seção revisa o que se sabe sobre a cesárea em dois planos. O primeiro reconstrói a trajetória clínica e histórica do procedimento, das indicações em que a cirurgia salva vidas ao excesso observado hoje. O segundo percorre a literatura empírica que mediu esse excesso e investigou suas causas, com ênfase na hipótese de demanda induzida pela oferta.

2.1 Cesárea: história, importância clínica e a faixa adequada

Poucos procedimentos cirúrgicos carregam um peso simbólico tão denso quanto a cesárea. Conta-se há séculos que o termo derivaria do nascimento de Júlio César, mas a etimologia provavelmente é falsa: a mãe do estadista, Aurélia Cotta, sobreviveu por décadas ao parto do filho, em uma época em que a operação era quase sempre fatal para a mulher (Todman, 2007). A história real da cesárea, antes do século XIX, é menos lendária e mais sombria. Nos registros compreendidos entre a Antiguidade e o período medieval, o procedimento aparece tipicamente como gesto póstumo, feito em mulheres já mortas ou agonizantes, em uma tentativa última de salvar o feto, com frequência por motivos religiosos como o batismo do recém-nascido (Todman, 2007). A cesárea, por longo tempo, não foi uma intervenção a serviço da gestante: foi uma intervenção feita *apesar* dela.

A virada, isto é, a transformação da cesárea de procedimento quase sempre fatal em cirurgia rotineira e segura, é uma história do século XIX e da primeira metade do século XX, e ocorreu por acúmulo de inovações técnicas que se reforçaram. Em 1876, o italiano Eduardo Porro descreveu uma técnica que combinava a cesárea com a retirada do útero, contendo a hemorragia e a infecção que matavam quase todas as parturientes operadas⁴; a mortalidade materna, altíssima no final do século XVIII, recuou bastante após as inovações cirúrgicas e a difusão da antissepsia ao longo do século XIX (Antoine; Young, 2020). Em 1882, o alemão Max Sänger introduziu a sutura uterina em duas camadas, dispensando a retirada do órgão e produzindo uma queda adicional, “espetacular”, da mortalidade, razão pela qual é citado na literatura como o pai da cesárea moderna (Hem; Bordahl, 2003; Antoine; Young, 2020). Em 1926, o britânico John Munro Kerr popularizou a incisão transversa baixa do segmento uterino inferior⁵, técnica que,

⁴ A técnica original ficou conhecida como “operação de Porro” (cesariana com histerectomia subtotal subsequente). Era mutilante, porque implicava retirada do útero, mas reduziu drasticamente a mortalidade materna por hemorragia e sepse pós-cirúrgica.

⁵ A incisão de Kerr é horizontal, feita na parte fina do útero próxima ao colo, e cicatriza melhor que a clássica vertical no corpo uterino. Reduz risco de ruptura em gestações futuras e permite, em parte dos casos, parto vaginal posterior.

com pequenos ajustes, ainda é o padrão dominante no século XXI (Antoine; Young, 2020).

Essas inovações cirúrgicas, contudo, não bastariam sozinhas. Foi a sua convergência com um conjunto mais amplo de transformações da medicina e da saúde pública (anestesia obstétrica moderna, antissepsia, sulfamidas em 1937, penicilina disponível a partir de 1940, bancos de sangue para transfusão e acompanhamento pré-natal sistemático) que produziu o resultado epidemiológico que hoje pode parecer banal: a quase erradicação da morte materna nos países ricos. Loudon (2000) descreve esse processo como a saída de uma fase longa e estável de mortalidade materna alta que, no mundo ocidental, durou de meados do século XIX até a década de 1930; a partir de 1937, as taxas começaram a cair de forma acentuada e simultânea em quase todos os países desenvolvidos, e em vinte anos as diferenças entre nações praticamente se dissolveram.

A magnitude do salto é rara na história da saúde pública. Nos Estados Unidos, segundo o levantamento histórico do *Centers for Disease Control and Prevention* (CDC), a mortalidade materna caiu cerca de 99% ao longo do século XX, passando de cerca de 850 mortes por 100 mil nascidos vivos no início do período para 7,7 por 100 mil em 1997 (CDC, 1999). A cesárea segura é uma parte desse mosaico, não a única, mas uma das que tornaram tecnicamente possível interromper a tempo gestações que antes seriam quase sempre fatais.

Estabelecida a virada técnica, abre-se a pergunta operacional: quais gestações são essas em que a cirurgia é efetivamente necessária? A pergunta importa porque este estudo não é contra a cesárea. Existem condições clínicas em que o parto vaginal é tecnicamente impraticável ou perigoso demais, e em que a cesárea é a *única* via segura. Existem condições clínicas, como a placenta prévia (com prevalência de 0,5% a 1% dos partos), em que a cesárea é a indicação correta⁶ e o parto vaginal é incompatível. A apresentação pélvica do feto a termo é outro caso reconhecido⁷. A esses dois exemplos somam-se outras situações que a literatura clínica reconhece como indicações firmes: ruptura uterina, descolamento prematuro de placenta, sofrimento fetal agudo, certas formas graves de pré-eclâmpsia e desproporção céfalo-pélvica absoluta⁸. Em todos

⁶ Placenta prévia: a placenta se implanta sobre o orifício interno do colo do útero, fechando a saída pela qual o feto desceria. Está associada a sangramento antes do parto em 40% a 60% dos casos e hemorragia pós-parto em 20% a 35%; é indicação absoluta de cesárea programada (Cassardo *et al.*, 2026).

⁷ Apresentação pélvica: posição em que as nádegas ou os pés do feto se apresentam ao canal de parto, e não a cabeça. O *Term Breech Trial*, ensaio clínico randomizado multicêntrico no *Lancet*, com 2.088 mulheres em 121 centros de 26 países, mostrou redução de mortalidade perinatal e morbidade neonatal grave de 5,0% para 1,6% com cesárea planejada vs. tentativa de parto vaginal (RR=0,33; IC 95%: 0,19–0,56) (Hannah *et al.*, 2000).

⁸ Em linguagem operacional: a pelve materna é estreita demais para a cabeça do feto. Os demais termos descrevem, respectivamente: ruptura da parede do útero durante o trabalho de parto; separação prematura da placenta da parede uterina; sinais de oxigenação fetal comprometida; e elevação grave de pressão arterial associada à gestação. Em todos, recusar a cirurgia equivale a aceitar risco evitável de morte materna ou fetal.

esses contextos, recusar a cirurgia equivale, na prática, a aceitar um risco evitável de morte materna ou fetal.

Reconhecida a centralidade clínica da cesárea nesses casos, a história subsequente é a de um descompasso. Em abril de 1985, a Organização Mundial da Saúde (OMS) reuniu em Fortaleza, no Brasil, um painel internacional de especialistas em saúde reprodutiva, em conferência conjunta de seus escritórios europeu e das Américas. Sob o título *Appropriate Technology for Birth*, o grupo discutiu o que já era visto como aumento injustificado das taxas de cesárea em escala global. O documento síntese da conferência, publicado no *Lancet*, fixou a frase que se tornaria a referência canônica do debate pelas três décadas seguintes: “*there is no justification for any region to have a caesarean section rate higher than 10–15%*” (OMS, 1985). A recomendação foi formulada com base nos dados disponíveis na época, sobretudo de países do norte europeu que apresentavam bons desfechos materno-perinatais nessa faixa, e tinha intenção explicitamente populacional: era um parâmetro para regiões e sistemas de saúde, não para hospitais individualmente.

Trinta anos depois, com base empírica bem mais ampla, a OMS revisou oficialmente sua posição. No documento *WHO Statement on Caesarean Section Rates*, de 2015, e em revisão sistemática publicada no mesmo ano, a organização reconheceu duas coisas. Primeiro, o corte único de 15% foi aplicado de modo equivocado com frequência: projetado para populações, foi usado como meta hospitalar. Segundo, sob análises mais sólidas, acima de cerca de 10% deixa de existir associação estatística entre a frequência de cesáreas e a redução de mortalidade materna ou neonatal (Betrán *et al.*, 2016). A nova formulação substituiu a meta numérica rígida por um princípio flexível: **o que importa é que toda mulher que precise de uma cesárea a receba**, nem mais, nem menos. O problema deslocou-se. Não há mais, no mundo, subuso generalizado da cirurgia. Há, em parcela crescente do mundo, excesso.

A revisão da posição da OMS não desautorizou Fortaleza: refinou o que aquela recomendação queria dizer, com base em evidências empíricas acumuladas ao longo das duas décadas anteriores. Althabe *et al.* (2006), em estudo ecológico com dados de países de baixa, média e alta renda, foi um dos primeiros a documentar que a relação entre taxa de cesárea e mortalidade materno-neonatal não é monotônica: cresce até um certo ponto e depois perde significância estatística ou estabiliza. Ye *et al.* (2016), em estudo ecológico longitudinal com dados de 194 países do *World Health Statistics*, identificaram que o aumento da taxa reduz a mortalidade materna e neonatal apenas até cerca de 10 partos cirúrgicos por 100 nascidos vivos. Molina *et al.* (2015), em análise transversal com dados de 2005–2012 dos 194 estados-membros da

OMS, situaram o ponto de inflexão em valor mais alto, próximo de 19 cesáreas por 100 nascidos vivos. Vogel *et al.* (2015), em análise secundária de duas pesquisas multipaís coordenadas pela OMS sobre 21 países, aplicou a Classificação de Robson em escala internacional⁹ e mostrou que o crescimento global se concentra nos Grupos 1, 2 e 5, justamente os de maior espaço para julgamento clínico. Essas estimativas usam metodologia ecológica e estão sujeitas à conhecida distância entre o agregado populacional e a indicação individual¹⁰. Em conjunto, contudo, esses estudos situam o ponto de inflexão da relação entre cerca de 10 e 19 cesáreas por 100 nascimentos: acima dele, em corte populacional, não há associação com melhor desfecho materno-perinatal. A faixa é coerente com a referência populacional de cerca de 10% a 15% reconhecida pela OMS após a revisão de 2015 (Betrán *et al.*, 2016). O que está fora dessa faixa, em qualquer das duas pontas, deixa de ser apenas estatística e vira problema clínico.

2.2 O custo do excesso e sua escala global

Identificada a faixa adequada, cabe perguntar o que ocorre quando o sistema opera muito acima dela. A face clínica desse problema, do lado do excesso, foi sintetizada pela revisão sistemática de Sandall *et al.* (2018) na série do *Lancet* sobre uso apropriado da cesárea. Os autores documentam um padrão de risco que cresce em gestações subsequentes conforme aumenta o número de cesáreas anteriores: ruptura uterina, placenta prévia, acretismo placentário¹¹, gravidez ectópica e prematuridade iatrogênica¹² aumentam de forma progressiva, a cada nova cesárea. Para a criança, o estudo identifica associação da via cesárea com maior incidência de asma até os 12 anos e de obesidade até os 5 anos. O mecanismo biológico mais estudado é a microbiota: Reyman *et al.* (2019) mostram que o recém-nascido por cesárea perde o contato com as bactérias vaginais maternas que colonizam o intestino e ajudam a calibrar o sistema imune nos primeiros meses de vida, uma das vias propostas para o excesso de doenças atópicas e metabólicas na infância, como asma e obesidade (Rusconi *et al.*, 2017; Goldani *et al.*, 2011). Do lado materno, a

⁹ Classificação de Robson: sistema de dez grupos mutuamente exclusivos que classifica gestantes a partir de paridade, idade gestacional, tipo de início do trabalho de parto, número de fetos, apresentação fetal e cesárea anterior (Robson, 2001). Adotada pela OMS como padrão para comparação internacional de taxas de cesárea (Organização Mundial da Saúde (OMS), 2017).

¹⁰ Falácia ecológica: associação observada no nível agregado (entre países, entre regiões) não se transporta automaticamente para o nível individual (Robinson, 1950). A taxa populacional ótima de cesáreas não diz qual mulher específica deve ou não ser operada.

¹¹ Acretismo placentário: aderência anormal da placenta à parede do útero, podendo invadir o miométrio (acreta), atingir camadas profundas (increteta) ou perfurar o útero (percreta). Está associado a hemorragia grave no parto e, em parcela dos casos, exige histerectomia (retirada do útero) de emergência.

¹² Prematuridade iatrogênica: nascimento prematuro decorrente de decisão médica de antecipar o parto, e não de trabalho de parto espontâneo. Em cesáreas eletivas, agendamentos antes de 39 semanas elevam o risco de complicações respiratórias e de internação em UTI neonatal.

revisão sistemática de Mascarello, Horta e Silveira (2017) documenta maior morbidade materna grave associada à cesárea sem indicação clínica, e Getahun *et al.* (2006) estimam que a cesárea prévia eleva o risco de placenta prévia e de descolamento prematuro da placenta na gestação seguinte. Esses achados não retiram da cesárea seu papel de cirurgia que salva vidas em indicações estabelecidas. Mostram, no entanto, que a cirurgia carrega um custo biológico próprio. É contra esse fundo que se deve ler o caso brasileiro, descrito a seguir como ponta extrema do excesso global.

A magnitude do excesso é o que justifica chamar o fenômeno de epidemia. A série em três artigos publicada pelo *Lancet* em 2018 sobre a otimização do uso da cesárea oferece o retrato mais completo disponível. O estudo de Boerma *et al.* (2018) cobriu 169 países, responsáveis por 98,4% dos nascimentos mundiais. Em 2015, 21,1% dos partos no mundo foram cesáreas (29,7 milhões de procedimentos), número que praticamente dobrou em relação aos 12,1% de 2000. Os autores estimam que cerca de dois terços do crescimento são atribuíveis à expansão do parto institucional (mais mulheres tendo seus filhos em hospitais), enquanto cerca de um terço decorre do aumento da taxa de cesárea *dentro* dos hospitais, ou seja, de mudanças na própria prática obstétrica. As desigualdades observadas são igualmente fortes: a taxa varia de 0,6% no Sudão do Sul a 58,1% na República Dominicana; em países de baixa e média renda, as cesáreas são cerca de cinco vezes mais frequentes no quintil mais rico do que no mais pobre; e, no mundo, são 1,6 vez mais frequentes em estabelecimentos privados do que em públicos (Boerma *et al.*, 2018). Brasil e China são citados pelos autores como casos em que o uso é alto justamente entre gestações de baixo risco e mulheres de maior escolaridade, ou seja, no subgrupo em que a indicação clínica seria, em tese, *menos* sustentável.

2.3 O caso brasileiro

O Brasil ocupa, nesse retrato, uma posição extrema. A pesquisa Nacer no Brasil, coordenada por Maria do Carmo Leal na Fundação Oswaldo Cruz e considerada o maior estudo já realizado sobre parto e nascimento no país, com 23.940 mulheres acompanhadas em 266 hospitais entre 2011 e 2012, registrou taxa geral de cesárea de 52%, com 46% no setor público e 88% no setor privado (Leal *et al.*, 2014). Estudos posteriores reforçaram e atualizaram esses números, sem alterar muito o diagnóstico: o Brasil figura, junto à Nicarágua, no topo do *ranking* mundial de cesáreas (Filho; Rissin, 2018). A literatura nacional aponta, ainda, um paradoxo perinatal incômodo: apesar da melhoria de quase todos os indicadores de saúde materna nas últimas décadas (ampliação do acesso ao pré-natal, quase universalização do parto hospitalar,

qualificação do quadro profissional), a mortalidade materna brasileira não acompanhou a queda esperada, ficando em níveis incompatíveis com o grau de medicalização da assistência (Diniz, 2009). O excesso de intervenções, longe de ser sintoma colateral inofensivo, parece compor o próprio mecanismo desse paradoxo. O Brasil opera, assim, em uma configuração contraintuitiva: não se fazem poucas cesáreas em quem precisa, e sim muitas em quem não precisaria.

A descrição agregada (Brasil em torno de 56% de cesáreas, com setor privado próximo de 90%) é apenas o ponto de entrada do problema. Quando se examina a estrutura desse excesso, surgem regularidades que descartam, uma a uma, as explicações fáceis e levam a discussão ao terreno desta pesquisa. Os dados oficiais do Ministério da Saúde, vindos do Sistema de Informações sobre Nascidos Vivos (SINASC), confirmam a estabilidade do fenômeno: a taxa nacional atingiu 55,5% em 2021 (Brasil. Ministério da Saúde, 2022), sustentando a tendência estável de cerca de 56% observada entre 2014 e 2020 (Pereira *et al.*, 2024), e mantém, há mais de uma década, forte estratificação por região, fonte de financiamento e perfil socioeconômico. Em São Paulo, Estado mais populoso da Federação, 156 mil partos normais e 153 mil cesarianas foram realizados pelo SUS em 2023 (taxa de 49,4% nesse subsistema) (Brasil. Ministério da Saúde, 2022), ao passo que, na rede suplementar do mesmo Estado, beneficiárias registraram taxa de cesárea de 80% no período 2015–2021, com 757.307 internações analisadas e gasto total de R\$ 7,7 bilhões em internação por parto (Silva; Pazin-Filho, 2024). Dados setoriais do Instituto de Estudos de Saúde Suplementar mostram que, embora exista tendência de queda lenta (de 84,5% em 2013 para 82,7% em 2020 nas operadoras de planos de saúde), o nível permanece muito alto e estável (Instituto de Estudos de Saúde Suplementar (IESS), 2021). O padrão, portanto, não é um pico estatístico restrito a um ano ou a uma região: é sustentado, escalonado, financeiramente relevante e desigual entre subsistemas.

A explicação mais imediata para essa configuração (a de que as próprias gestantes brasileiras “preferem” a cesárea) é a primeira a desmoronar quando submetida ao teste empírico. Hopkins (2000), em estudo etnográfico-epidemiológico, combinou *survey* pós-parto, observação participante em enfermarias obstétricas e entrevistas em profundidade, em hospitais públicos e privados de Natal e Porto Alegre. O estudo mostrou que a maioria das mulheres brasileiras, sobretudo as primíparas¹³, *não* deseja ativamente ter um filho por cesárea. Os mecanismos pelos quais essas mulheres terminam sendo operadas envolvem, com frequência, a autoridade médica reorganizando a decisão durante o trabalho de parto. Béhague, Victora e Barros (2002),

¹³ Primípara: mulher que está dando à luz pela primeira vez. É um dos critérios que define o Grupo 1 da Classificação de Robson e, neste estudo, o subgrupo de maior interesse, por ter espaço pequeno para indicação clínica imediata de cesárea.

em coorte de base populacional em Pelotas com 5.304 mulheres, chegaram a uma conclusão complementar e relevante: parte do desejo declarado por cesárea entre mulheres de baixa renda decorre não de preferência estética ou de medo da dor, mas do medo de receber cuidado de qualidade insuficiente no parto vaginal, uma “preferência” que é, na verdade, reação ao que a gestante antecipa do funcionamento do sistema. A pesquisa Nascir no Brasil documentou esse padrão na escala nacional: Domingues *et al.* (2014) registraram preferência inicial por cesárea de apenas 27,6%, e taxa observada no setor privado de 87,5%, muito acima do que as próprias gestantes desejavam. Silva *et al.* (2017) e Rocha e Ferreira (2020), em estudo transversal com puérperas e em revisão integrativa, sustentam o mesmo achado central para a década seguinte: a preferência declarada por parto vaginal predomina, e a divergência entre vontade e desfecho é regular, não fruto de escolha individual. A demanda das gestantes, portanto, não fecha a equação.

2.3.1 O lado da oferta: por que a Economia da Saúde tem algo a dizer

Se a preferência declarada das mulheres não basta para explicar o quadro, é preciso olhar para o outro lado da relação clínica. A economia da saúde tem, há mais de meio século, um vocabulário preciso para o que ocorre quando o agente que decide o que o paciente vai consumir é também quem se beneficia desse consumo. O ponto de partida é Arrow (1963): o mercado de cuidados em saúde não se comporta como um mercado comum porque a doença é incerta e o paciente não consegue avaliar, por conta própria, se o tratamento indicado é o melhor disponível. O paciente depende, para esse juízo, do mesmo profissional que oferece o serviço. Dessa dependência nasce uma relação de agência, em que o médico decide, em nome do paciente, o que será consumido. Quando a agência é perfeita, o profissional escolhe exatamente o que o paciente escolheria se tivesse a mesma informação; quando não é, abre-se espaço para que a decisão incorpore também os interesses de quem decide. Arrow (1963) observou que boa parte das instituições do setor (a ética profissional, a confiança, a regulação) existe justamente para conter os efeitos dessa assimetria, sinal de que o mercado, sozinho, não a resolve.

Evans (1974) deu nome e forma a essa possibilidade: a **demand induzida pela oferta** (*supplier-induced demand*, SID). O prestador, amparado no diferencial de informação, exerce influência direta, e não pela via do preço, sobre a quantidade e a composição dos serviços que o paciente consome. Evans (1974) formalizou duas especificações alternativas do comportamento médico e mostrou que ambas geram respostas anômalas, em que a oferta passa a determinar a demanda em vez do contrário: na evidência canadense e norte-americana que reuniu, cerca de 80% da variação na quantidade de serviços respondia à oferta de profissionais, não à necessidade

dos pacientes. A consequência de política é direta: instrumentos pensados para conter preços ou utilização excessiva não funcionam quando é o próprio prestador quem organiza a demanda. A indução não pressupõe má-fé; basta que a remuneração ou as condições de trabalho do profissional dependam do tipo e da quantidade de procedimentos realizados.

A cesárea é o exemplo de manual desse mecanismo, por uma razão concreta. A cesárea e o parto vaginal são, em larga medida, substitutos clínicos para a mesma gestante, mas têm remuneração e custo de tempo muito diferentes para o médico: a cirurgia é agendável, mais rápida e mais previsível do que a espera por um trabalho de parto que pode durar horas em qualquer horário. Onde a remuneração premia o procedimento e o tempo do profissional é escasso, o incentivo aponta para a substituição do parto vaginal pela cesárea. Gruber e Owings (1996) ofereceram o teste empírico mais influente dessa hipótese: usaram a queda da fertilidade nos Estados Unidos entre 1970 e 1982 como um choque exógeno de renda sobre os obstetras (menos partos significam menor receita potencial) e mostraram que quedas estaduais de fertilidade foram seguidas por aumentos estaduais na taxa de cesárea, na direção prevista pela substituição em favor de uma alternativa mais bem remunerada. O resultado é consistente diante de explicações concorrentes e simétrico entre períodos de queda e de alta da fertilidade, o que reforça a leitura de que parte da decisão clínica responde a incentivos financeiros do prestador.

A mesma literatura advertiu, desde cedo, contra identificações frágeis. Dranove e Wehner (1994) aplicaram a técnica então usual de variável instrumental ao próprio número de nascimentos, que o obstetra não controla, e encontraram indução onde a indução não poderia existir: o teste acusava demanda induzida até naquilo que é fixado pela demografia. A advertência elevou o padrão de identificação exigido nessa literatura. É uma das razões pelas quais este trabalho adota, mais adiante, seleção em observáveis combinada a um desenho quase-experimental, em vez de instrumentos inadequados. Levado ao Brasil, o argumento tem dois canais que a estratégia empírica procurará distinguir. O primeiro é financeiro: no setor privado, remunerado por procedimento, a cesárea costuma pagar mais por unidade de tempo médico do que o parto vaginal. O segundo é informacional: como na formulação de Arrow (1963) e Evans (1974), a gestante depende do obstetra para avaliar a real necessidade da cirurgia, e essa assimetria é tanto maior quanto menos a paciente conhece a obstetrícia. Trabalhos etnográficos brasileiros qualificam essa leitura, situando a decisão na dinâmica concreta da relação clínica e advertindo contra reduzi-la tanto à preferência da mulher quanto à conveniência do obstetra (McCallum, 2005). A leitura econômica do excesso brasileiro de cesárea não é, portanto, uma imputação moral a médicos individuais; é uma afirmação sobre incentivos agregados que operam de forma

sistemática onde o arranjo institucional os sustenta.

Esses canais admitem uma formalização mínima, na tradição de Evans (1974) sobre a agência médica. Seja $p \in [0, 1]$ a probabilidade de uma gestação clinicamente elegível ao parto vaginal terminar em cesárea, escolhida pelo prestador (médico ou hospital) para maximizar

$$U(p) = \underbrace{(R_c - R_v) p}_{\text{receita líquida da via}} - \underbrace{C(p)}_{\text{custo da cirurgia injustificada}}, \quad (1)$$

em que $R_c - R_v \geq 0$ é o diferencial de remuneração da cesárea sobre o parto vaginal (e captura também o ganho de tempo e de previsibilidade da cirurgia agendada), e $C(p)$ é o custo crescente e convexo de operar além da indicação clínica p^* (risco assistencial, escrutínio do pagador e da paciente, custo reputacional), com $C'(p^*) = 0$ e $C'' > 0$. A condição de primeira ordem $R_c - R_v = C'(\hat{p})$ entrega a taxa escolhida $\hat{p} \geq p^*$, crescente no diferencial de remuneração e decrescente na inclinação do custo de injustificação. Três previsões testáveis seguem: (i) prestadores que enfrentam $R_c - R_v$ maior ou C' menor (remuneração por procedimento, menor escrutínio) escolhem $\hat{p}$ maior, o diferencial do privado lucrativo captado adiante na análise observacional do diferencial institucional; (ii) uma intervenção que eleva C' (monitoramento, retorno de indicadores, metas), como o PPA, desloca $\hat{p}$ para baixo, o que o desenho quase-experimental do Parto Adequado testa; (iii) pacientes mais informadas elevam o custo marginal de operar sem indicação, reduzindo $\hat{p}$ para esse grupo, o que o exercício das médicas verifica. O modelo é mínimo: não microfundamenta a barganha clínica, apenas mapeia os sinais que a evidência deve exibir se o canal de demanda induzida pela oferta estiver operando.

2.3.2 A política pública e seus limites: o caso do Parto Adequado

A política pública brasileira reconheceu o problema e tentou agir sobre o excesso. Em 2015, a Agência Nacional de Saúde Suplementar (ANS) lançou o **Projeto Parto Adequado**, em parceria com o Hospital Israelita Albert Einstein e o *Institute for Healthcare Improvement* (IHI), organizado em quatro componentes, um deles focado em informação e participação das gestantes nas decisões sobre o parto (Agência Nacional de Saúde Suplementar (ANS), 2017). Sob a metodologia da Ciência da Melhoria do IHI, baseada em sessões de aprendizado, visitas técnicas e ciclos iterativos de melhoria, o programa foi estruturado em três fases (Agência Nacional de Saúde Suplementar (ANS), 2023). A fase 1 (abril de 2015 a outubro de 2016) foi o projeto-piloto, com 35 hospitais. A fase 2, de expansão, desenvolveu-se em duas etapas: a primeira (maio de 2017 a maio de 2019) reuniu 108 hospitais privados e públicos, e a segunda, o Ciclo Intensivo da Fase 2 (outubro de 2020 a abril de 2022), reuniu 27 hospitais sob apoio metodológico mais

próximo. A fase 3 foi lançada em outubro de 2019 como Movimento Parto Adequado, voltada à disseminação em larga escala, sem coorte fechada de hospitais aderentes (Mangolim, 2024).

2.4 Revisão de literatura empírica

Esta seção reúne os estudos empíricos que sustentam a investigação sobre a taxa de cesárea e a demanda induzida pela oferta no cuidado obstétrico. Os trabalhos estão organizados em três blocos: estudos internacionais, que firmam o arcabouço de identificação da SID e os estimadores quase-experimentais empregados adiante; estudos nacionais, que mensuram o fenômeno em escala brasileira e avaliam as políticas de contenção da cesárea; e estudos regionais e estaduais, que aproximam a evidência das magnitudes de custo e das particularidades locais. Um quadro-resumo ao final consolida desenho, dados e principais achados de cada estudo.

2.4.1 Estudos internacionais

Dranove e Wehner (1994) mostram que o instrumento clássico para identificar demanda induzida pela oferta produz indução falsa quando aplicado ao mercado de partos nos Estados Unidos, mercado em que a indução deveria ser próxima de zero porque o volume é fixado pela demografia. O instrumento testado é o estimador de mínimos quadrados em dois estágios, e os autores atribuem o resultado a duas falhas do desenho convencional: a regressão de primeiro estágio não é identificada (falha de exclusão) e o trânsito de pacientes entre áreas de mercado vizinhas é mal tratado, contaminando a estimativa por *spillover*¹⁴. O estudo é um alerta metodológico relevante para qualquer aplicação de variável instrumental em SID e justifica, na seção 4, a opção deste estudo por identificação por seleção em observáveis.

Gruber e Owings (1996) oferecem o teste empírico mais influente da hipótese de SID aplicada à cesárea: quando obstetras são atingidos por choque negativo de renda, a frequência de cesárea aumenta de modo estatisticamente significativo, padrão consistente com a substituição de parto vaginal por uma alternativa mais bem remunerada. A identificação explora a queda da fertilidade nos Estados Unidos no início dos anos 1980 como variação exógena da demanda enfrentada pelos obstetras-ginecologistas e compara estados em diferentes graus de exposição a esse choque. A análise usa microdados nacionalmente representativos de partos do período e regressões de efeitos fixos por estado que isolam a correlação intraestadual entre a evolução da fertilidade e a utilização de cesárea, com controles temporais e geográficos. O resultado central é

¹⁴ O termo *spillover* (transbordamento) designa o trânsito de pacientes através das fronteiras entre áreas de mercado vizinhas para receber atendimento: como parte da demanda de uma área é atendida em outra, a oferta de uma região afeta os desfechos da região contígua e contamina a comparação entre mercados.

a correlação forte e significativa entre quedas estaduais de fertilidade e elevações estaduais na frequência de cesárea: o prestador, diante da queda de demanda, explora a relação de agência ofertando cuidado excessivo.

Epstein e Nicholson (2009) mostram que a variação na taxa de cesárea entre médicos do mesmo mercado é cerca de duas vezes maior do que a variação entre mercados, sinal de que o estilo individual do prestador pesa mais do que diferenças de contexto regional. Programas de residência explicam não mais que 4% da variação ajustada por risco, mesmo entre médicos recém-formados, e o estilo de tratamento de cada profissional muda pouco ao longo da carreira.¹⁵ O estudo é referência empírica para a leitura de que o diferencial entre estabelecimentos não vem só de seleção de pacientes, mas também de prática clínica idiossincrática moldada institucionalmente.

Johnson e Rehavi (2016) oferecem o desenho que este estudo adapta ao SINASC–CNES brasileiro: comparar o tratamento de pacientes médicas com o de pacientes não médicas, pareadas em características clínicas e socioeconômicas, para isolar o canal informacional do canal financeiro da SID. O resultado central é que médicas têm cerca de 10% menos probabilidade de receber cesárea do que pacientes comparáveis e apresentam melhores desfechos clínicos; apenas um quarto desse efeito vem de seleção diferencial entre serviços. Incentivos financeiros afetam o tratamento de pacientes não médicas, mas não afetam o tratamento de médicas quando estas são pacientes, leitura compatível com a hipótese de que o diferencial de informação entre médico e paciente é elemento central da SID no caso obstétrico.¹⁶ A operacionalização desse desenho está em §4.4.4.

Han *et al.* (2017) mostram que, na Coreia do Sul, a propriedade, o porte, a especialidade e o perfil de demanda do hospital se associam de modo significativo à decisão sobre via de parto, padrão que os autores leem como compatível com decisões orientadas pela manutenção da margem financeira diante de competição.¹⁷ O estudo oferece contraste asiático ao quadro brasileiro e reforça a leitura de que o componente institucional do excesso de cesáreas opera além das características individuais da paciente.

Goodman-Bacon (2021) formaliza o diagnóstico de viés do estimador clássico TWFE quando o tratamento é escalonado e heterogêneo entre *cohorts*. O resultado, teórico-metodológico,

¹⁵ O recorte cobre o universo de partos realizados nos estados da Flórida e Nova York por quinze anos. A estratégia decompõe a variação ajustada por risco em três componentes: dentro do mercado, entre mercados e atribuível à formação de residência ou ao aprendizado entre pares.

¹⁶ O estudo usa microdados de partos nos Estados Unidos em período em que a base permite distinguir ocupação e cobertura das pacientes, e explora a presença de hospitais HMO como fonte de variação no incentivo financeiro do prestador.

¹⁷ Regressões logísticas sobre dados administrativos do *National Health Insurance*, ajustadas por idade, paridade, comorbidades e indicadores clínicos.

decompõe a estimativa do TWFE com efeitos fixos de unidade e tempo em soma ponderada de comparações implícitas dois a dois entre *cohorts* tratados em momentos distintos. Comparações em que o *cohort* tratado mais tarde é confrontado com *cohort* tratado mais cedo (que já está sob efeito) entram com peso positivo na estimativa e atenuam o coeficiente quando há heterogeneidade dinâmica. A decomposição é insumo direto do conjunto de testes de robustez aplicado à adesão ao PPA na seção 4.

Chaisemartin e D’Haultfœuille (2020) propõem o estimador DID_M como alternativa ao TWFE em painéis com tratamento escalonado heterogêneo. A construção é uma média ponderada de diferenças em diferenças 2×2 sobre unidades que mudam *status* entre dois períodos consecutivos (*switchers*), o que evita por desenho as comparações implícitas entre *cohorts* já tratados. O estimador é não viesado para o efeito médio sobre os *switchers* e tende a entregar magnitudes menores que TWFE quando há atraso na resposta, porque capta sobretudo o efeito imediato.

Sun e Abraham (2021) propõem o estimador *interaction-weighted* (IW), que estima coeficientes saturados por par (*cohort*, *event-time*) e agrega ponderando pelo tamanho relativo dos *cohorts*; hospitais nunca-tratados entram como grupo de comparação puro. O estimador é consistente para o efeito médio sob tendências paralelas, recupera a trajetória dinâmica completa do efeito e corrige o viés escalonado documentado por Goodman-Bacon.

Callaway e Sant’Anna (2021) propõem um arcabouço para diferenças em diferenças com múltiplos períodos que admite três estimadores: regressão de desfecho simples, IPW e duplamente robusto. Em qualquer das três variantes, estima-se $ATT(g, t)$ como diferenças em diferenças 2×2 separadas para cada par *cohort* g e tempo t , com nunca-tratados como contrafactual, agregando por *event-time*. A vantagem é não impor a forma funcional do TWFE saturado e tornar transparente quais comparações entram em cada agregado; sob tendências paralelas, o estimador converge com o IW de Sun-Abraham.

Rambachan e Roth (2023) propõem a abordagem *Honest DiD* como ferramenta de sensibilidade a violações parciais da hipótese de tendências paralelas. Define-se o conjunto $\Delta^{RM}(M)$ de violações pós-tratamento que não excedem M vezes a maior violação observada no pré-tratamento e calcula-se o IC de 95% robusto do ATT sob essa restrição; o *breakdown* M é o menor multiplicador para o qual o intervalo passa a incluir zero. A abordagem é a base do quadro de robustez da seção 4.

2.4.2 Estudos nacionais

Knau, Triaca e Santos (2023) apresentam a evidência empírica mais direta do efeito quase-experimental do Programa Parto Adequado em escala nacional: a fase um (2015–2016) reduziu a taxa de cesárea em 4,1 pontos percentuais nas cesáreas totais e 7,6 nas intraparto em hospitais públicos, e em 8,1 nas eletivas em hospitais privados, todas significativas a 1%. Os efeitos sobre cesáreas eletivas em públicos e sobre parto e intraparto em privados violam tendências paralelas e são tratados pelos autores com cautela.¹⁸

Mangolim (2024) oferece a avaliação quase-experimental mais próxima da *fase 2* do Programa (2017–2019), lacuna que os estudos anteriores, concentrados na fase 1, não cobriam. A autora seleciona o grupo de controle por *propensity score matching* e estima um modelo de diferenças em diferenças com efeitos fixos de hospital e tempo sobre a taxa de cesárea agregada por hospital, encontrando redução de cerca de 2,5% no modelo canônico e de 3,0% nos hospitais privados. Este estudo aproxima-se diretamente desse desenho e avança em dois pontos. Primeiro, substitui o DiD canônico, sabidamente enviesado sob adesão escalonada (Goodman-Bacon, 2021; Sun; Abraham, 2021), por um conjunto de estimadores robustos a heterogeneidade (Sun-Abraham, Callaway-Sant’Anna, De Chaisemartin-D’Haultfoeuille, com decomposição de Goodman-Bacon e *Honest DiD*). Segundo, restringe a inferência a uma população clinicamente comparável (Robson 1+3), em vez da taxa hospitalar agregada, o que isola melhor a margem discricionária da composição de pacientes.

Melo e Menezes-Filho (2023) estimam que a proibição nacional brasileira de cesáreas eletivas antes da 39ª semana gestacional reduziu a taxa de cesárea em 1,6 ponto percentual e elevou marginalmente o tempo gestacional, o peso ao nascer e o Apgar¹⁹ do primeiro minuto. A identificação usa, como contrafactual, nascimentos em apresentação pélvica ou transversa, grupo em que a cesárea é a via predominante independentemente da regulação, o que isola o efeito da política sobre a margem discricionária do prestador.²⁰ O artigo é referência metodológica direta para a identificação causal de regulação obstétrica em escala nacional.

Domingues *et al.* (2014) mostram que a proporção final de cesáreas no setor privado

¹⁸ O recorte é nacional, com dados do SINASC integrados ao CNES via Plataforma de Ciência de Dados, no período 2006–2019. A estratégia é TWFE com efeitos fixos de hospital e ano, controles maternos individuais (idade, escolaridade, paridade, peso ao nascer) e erros-padrão clusterizados por hospital; o grupo de controle reproduz os critérios da ANS para a fase um.

¹⁹ O índice de Apgar avalia a vitalidade do recém-nascido no 1º e no 5º minuto de vida, somando de 0 a 10 a pontuação de cinco sinais clínicos — frequência cardíaca, respiração, tônus muscular, irritabilidade reflexa e cor da pele; valores mais altos indicam melhor condição ao nascer.

²⁰ O estimador é diferenças em diferenças sobre microdados nacionais do SINASC. Sobre apresentação pélvica e transversa, ver §3.4.

brasileiro chegou a 87,5%, muito acima da preferência inicial de 27,6% das gestantes, e que essa divergência cresce ao longo do pré-natal independentemente do diagnóstico de complicações, evidência compatível com indução da demanda pelo prestador. Em ambos os setores, a proporção efetiva superou de forma substancial o desejo inicial, com magnitude muito maior no privado.²¹ O estudo é uma das raras evidências brasileiras que rastreiam a decisão da via de parto entre o desejo inicial da mulher e a via de fato realizada.

Hopkins, Amaral e Mourão (2014) mostram que a taxa de cesárea no setor privado brasileiro cresceu de forma muito mais acentuada do que no setor público entre 1998 e 2008 e que a fonte de financiamento é variável central para entender o excesso. O estudo separa dois fenômenos que o dado agregado nacional confunde. No setor público, o crescimento da cesárea pode refletir expansão de cobertura e incorporação de populações de maior risco clínico. No setor privado, o crescimento ocorre sobre níveis já elevados e em população com menor risco obstétrico médio.²²

Nakamura-Pereira *et al.* (2016), em suplemento da pesquisa Nacer no Brasil, mostram que a taxa de cesárea no setor privado brasileiro chegou a 87,9% contra 42,9% no setor público (51,9% no agregado nacional), com todos os dez grupos da Classificação de Robson registrando taxas superiores no setor privado.²³ O Grupo 5 (multíparas com cesárea anterior, feto único cefálico) foi o maior contribuinte para a taxa privada, enquanto os Grupos 1 e 2 concentraram a contribuição no setor público. Os autores concluem que a alta prevalência de cesárea nos grupos de baixo risco obstétrico (Grupos 1, 2 e 4) no setor privado é o componente mais crítico do diferencial entre setores, leitura que este estudo retoma ao recortar Robson 1+3 como amostra principal.

Freitas e Fernandes (2016) são o precedente estadual mais direto do desenho observacional desta pesquisa: sobre 61 maternidades de Santa Catarina em 2012, estimam por regressão de Cox robusta, em modelo hierarquizado, uma razão de prevalência de cesárea de 1,44 (IC95% [1,39; 1,50]) no setor privado frente ao público, com taxa bruta de 89% no privado contra 45,1% no SUS. Sem a Classificação de Robson, ainda não consolidada nos dados de 2012, os autores aproximam a comparabilidade clínica estratificando por grupos de risco e tratam explicitamente a cesárea anterior e a idade acima de 35 anos como *indicações relativas*, não estritas. O recorte

²¹ Os dados vêm da pesquisa Nacer no Brasil, coorte hospitalar nacional com puérperas entrevistadas em 2011 e 2012; a metodologia combina análise bivariada e teste qui-quadrado, estratificada por fonte de pagamento e paridade.

²² O desenho combina análise descritiva e modelos de regressão sobre dados nacionais da Pesquisa Nacional por Amostra de Domicílios para identificar a contribuição diferencial de cada tipo de estabelecimento ao crescimento agregado da taxa.

²³ O estudo aplica a Classificação de Robson (definição em §3.4) ao universo da coorte nacional de 2011 e 2012.

por Robson 1+3 adotado aqui é a versão padronizada dessa mesma estratégia de comparar, entre os setores, mulheres clinicamente semelhantes.

Paixão *et al.* (2021b) oferecem uma ampla aplicação em escala nacional da Classificação de Robson sobre o SINASC: cerca de 24 milhões de registros entre 2011 e 2017 confirmam que a taxa nacional é fortemente influenciada por gestações com cesárea anterior (Grupo 5) e por nulíparas com cesárea eletiva (Grupo 2), com o diferencial social esperado por escolaridade da mãe e variação regional. O estudo é a referência metodológica direta deste trabalho, que adota a mesma operacionalização de Robson sobre o SINASC e estende a análise pela integração formal com o CNES.

Paixão *et al.* (2021a) examinam a associação entre cesárea e mortalidade infantil em escala nacional brasileira: sem ajuste, a associação é praticamente nula nos grupos de baixo risco, mas o balanceamento por escore de propensão revela uma associação positiva antes mascarada por confundimento, com heterogeneidade por subgrupo. O recorte cobre 17,8 milhões de nascimentos brasileiros vinculados aos registros de óbito do SIM, configurando coorte longitudinal retrospectiva.²⁴ O estudo consolida o escore de propensão como instrumento padrão de inferência observacional sobre microdados SINASC vinculados, abordagem retomada na seção 4.

Rocha *et al.* (2023) retomam o desenho de escore de propensão sobre 17 milhões de partos brasileiros entre 2012 e 2019 para mostrar probabilidade substancialmente elevada de nascimento *early-term* (37 a 38 semanas) em gestações que terminam em cesárea, presente em todos os principais grupos de Robson, ainda que com magnitude variável entre os grupos e diferencial associado ao tipo de financiamento.²⁵ O resultado liga a margem de cesárea sem indicação à antecipação do parto, externalidade clínica relevante para a leitura econômica do excesso.

Pereira *et al.* (2024) mostram que, em todos os dez grupos da Classificação de Robson, a proporção de cesáreas permanece mais alta no setor privado do que no público durante todo o período de 2014 a 2020. No setor privado, a tendência é de queda; no público, há tendência de alta. Grupos com perfil clínico favorável ao parto vaginal (Robson 1, 4 e 5) registram proporções elevadas de cesárea no privado, achado que os autores leem como indicativo de cesárea sem

²⁴ O *record linkage* é probabilístico. A estratégia constrói escore de propensão a partir de características clínicas, sociodemográficas e geográficas observáveis e estima razões de hazard ajustadas por inverso da probabilidade de tratamento (IPW) para diferentes janelas etárias.

²⁵ *Early-term* designa partos entre 37 semanas e 38 semanas + 6 dias, classificação clínica utilizada pelo American College of Obstetricians and Gynecologists. A estratégia aplica a Classificação de Robson para garantir comparabilidade clínica e estima escore de propensão com base em características pré-tratamento.

indicação clínica adequada.²⁶

Borem *et al.* (2020) conduziram a avaliação interna da Fase 1 do PPA da ANS e documentam aumento de cerca de 14 pontos percentuais na frequência de partos vaginais sobre a linha de base agregada do conjunto de 28 hospitais privados que aderiram entre 2015 e 2017. A heterogeneidade do efeito se associa a indicadores de capacidade gerencial prévia (acreditação, equipe multidisciplinar dedicada, monitoramento contínuo).²⁷ É referência convergente para o achado quase-experimental da seção 4 e levanta a hipótese rival de capacidade gerencial diferencial, qualificada de modo explícito entre as limitações.

Negrini *et al.* (2021) descrevem redução sustentada da taxa de cesárea sobre o agregado intraparto de uma maternidade privada lucrativa de São Paulo ao longo de cinco anos a partir de 2015, com magnitudes próximas às de Borem e colegas. A intervenção combinou manejo ativo do trabalho de parto, capacitação, presença de doulas e *feedback* mensal, implementada por ondas de protocolos.²⁸

Kondo *et al.* (2025) conduziram intervenção de melhoria de qualidade focada nos Grupos 1 e 3 da Classificação de Robson em uma maternidade privada lucrativa brasileira de grande porte e documentam redução estatisticamente significativa nesses grupos, da ordem de cinco a dez pontos percentuais, sem deterioração de desfechos maternos ou neonatais.²⁹ O estudo é a justificativa empírica recente para o recorte explícito Robson 1+3 como amostra principal de análise causal, alinhando a definição clínica adotada aqui à literatura de intervenção sobre o mesmo recorte.

2.4.3 Estudos regionais e estaduais

Raspantini *et al.* (2016) mostram desvio à esquerda da idade gestacional dos nascimentos no Município de São Paulo entre 2013 e 2014, com mediana de 39 semanas no SUS contra 38 semanas na rede não SUS e maior concentração de pré-termos tardios (34 a 36 semanas) na segunda. O estudo articula registros do SINASC com o CNES em escopo municipal e oferece precedente metodológico direto para a integração das duas bases adotada nesta pesquisa.

²⁶ A análise aplica regressão de Prais-Winsten para tendência em série temporal proporcional sobre microdados do SINASC, com estratificação por grupo de Robson e tipo de serviço.

²⁷ O desenho é *quality improvement* pré-pós com linha de base e medição mensal de indicadores, em parceria com o Institute for Healthcare Improvement.

²⁸ O desenho é *quality improvement* interno com análise de séries temporais interrompidas.

²⁹ O desenho clássico de melhoria de qualidade combina linha de base pré-intervenção, implementação faseada e monitoramento contínuo segmentado por grupo de Robson, complementado por séries temporais interrompidas e *control charts*; o período é de aproximadamente três anos.

Entringer, Pinto e Gomes (2019) estimam por microcusteio em três maternidades públicas brasileiras que o custo médio do parto vaginal pelo SUS é de R\$ 808,16 e o da cesárea eletiva é de R\$ 1.113,70, com diferença que sobe para 32% quando se inclui o período em alojamento conjunto; recursos humanos são o item de custo dominante em ambos os procedimentos. O estudo fixa a ordem de grandeza da diferença direta de custos entre as duas vias na ótica do financiador público.

Silva e Pazin-Filho (2024) documentam que a taxa agregada de cesárea no setor privado paulista entre 2015 e 2021 foi de 80%, 71% superior à observada no setor público para a mesma população, com custo mediano da cesárea 15% superior ao do parto vaginal e gasto total de R\$ 7,7 bilhões sobre 757.307 internações. A análise é econômica, sobre beneficiárias da saúde suplementar paulista, e oferece a magnitude monetária do excesso na ótica do financiador privado.

Campos, Rattner e Diniz (2024) avaliam a efetividade do PPA na redução de cesáreas em maternidades privadas do Município de São Paulo entre 2014 e 2019 (277.747 nascimentos do SINASC) e mostram redução em ambos os grupos, mais pronunciada nos hospitais participantes; a tendência de queda é constante a partir do lançamento em 2014 e torna-se estatisticamente significativa em 2018. A taxa nos hospitais participantes recuou de 83,8% para 72,3%, ainda muito acima da taxa esperada pela ferramenta C-model para essa população (45,2%).³⁰ O estudo dialoga diretamente com a seção 4, embora limitado ao recorte municipal paulista.

O Quadro 1 consolida os estudos discutidos nesta seção, com a abrangência, o tipo de estimação, a base de dados, o período e os principais resultados de cada um.

³⁰ O desenho é retrospectivo, com comparação entre hospitais participantes e não participantes, controlado por variáveis demográficas, maternas e de classificação de Robson; nenhuma redução foi observada antes de 2014. C-model: ferramenta desenvolvida com participação de pesquisadores da OMS para estimar a taxa de cesárea esperada em uma população, ajustada por características clínicas e demográficas (Souza *et al.*, 2016).

Quadro 1 – Síntese da literatura empírica sobre cesárea e demanda induzida pela oferta.

Autores (ano)	Abrangência	Tipo de estimação	Base de dados	Período	Principais resultados
Estudos internacionais					
Dranove e Wehner (1994)	Internacional	Mínimos quadrados em dois estágios (crítica de identificação)	Partos, EUA	n.d.	O instrumento clássico acusa indução falsa até no número de partos; eleva a exigência de identificação em SID.
Gruber e Owings (1996)	Internacional	Efeitos fixos com choque exógeno de fertilidade	Microdados de partos, EUA	1970–1982	Quedas estaduais de fertilidade seguem-se de aumentos da cesárea, compatíveis com SID.
Epstein e Nicholson (2009)	Internacional	Decomposição da variância ajustada por risco	Partos, Flórida e Nova York	~15 anos	Variação entre médicos cerca de duas vezes a variação entre mercados; o estilo do prestador domina.
Johnson e Rehavi (2016)	Internacional	Comparação médicas vs. não médicas (canal informacional)	Partos, EUA	n.d.	Médicas têm cerca de 10% menos cesárea; o diferencial de informação é central à SID.
Han <i>et al.</i> (2017)	Internacional	Regressão logística	Seguro nacional, Coreia do Sul	n.d.	Propriedade, porte e competição do hospital associam-se à decisão de cesárea.
Estudos nacionais					
Domingues <i>et al.</i> (2014)	Nacional	Bivariada; teste qui-quadrado	Nascer no Brasil	2011–2012	Preferência inicial por cesárea de 27,6% vs. 87,5% realizada no setor privado.
Hopkins, Amaral e Mourão (2014)	Nacional	Regressão logística	PNAD	1998–2008	A cesárea cresce muito mais no privado; a fonte de pagamento é determinante.
Nakamura-Pereira <i>et al.</i> (2016)	Nacional	Classificação de Robson (descritivo)	Nascer no Brasil	2011–2012	Privado 87,9% vs. público 42,9%; todos os dez grupos maiores no privado.
Paixão <i>et al.</i> (2021b)	Nacional	Classificação de Robson	SINASC (~24 milhões)	2011–2017	Taxa puxada pelos Grupos 5 e 2; diferencial social e regional.

continua na próxima página

(continuação do Quadro 1)

Autores (ano)	Abrangência	Tipo de estimação	Base de dados	Período	Principais resultados
Paixão <i>et al.</i> (2021a)	Nacional	Escore de propensão; IPW (<i>hazard</i>)	SINASC vinculado ao SIM (17,8 milhões)	n.d.	A associação cesárea–mortalidade infantil emerge após o balanceamento.
Rocha <i>et al.</i> (2023)	Nacional	Escore de propensão	SINASC (17 milhões)	2012–2019	A cesárea eleva o nascimento <i>early-term</i> em todos os grupos de Robson.
Pereira <i>et al.</i> (2024)	Nacional	Série temporal (Prais-Winsten)	SINASC/DataSUS	2014–2020	Cesárea maior no privado em todos os Robson; queda no privado, alta no público.
Melo e Menezes-Filho (2023)	Nacional	Diferenças em diferenças (contrafactual pélvico/transverso)	SINASC	n.d.	A restrição a cesáreas eletivas (sem indicação clínica) antes de 39 semanas reduziu a taxa em 1,6 p.p.
Knau, Triaca e Santos (2023)	Nacional	Diferenças em diferenças (TWFE)	SINASC–CNES (PCDaS)	2006–2019	PPA fase 1: –4,1 p.p (público, total) e –8,1 p.p (eletivas, privado).
Mangolim (2024)	Nacional	Pareamento por escore de propensão + DiD com efeitos fixos	SINASC	2017–2019	PPA fase 2: redução de ~2,5% no geral e ~3,0% no privado.
Borem <i>et al.</i> (2020)	Nacional (28 hospitais)	Melhoria de qualidade, pré-pós	Indicadores do PPA	2015–2017	+14 p.p de partos vaginais; efeito associado à capacidade gerencial.
Estudos regionais e locais					
Freitas e Fernandes (2016)	Local (SC)	Regressão de Cox hierárquica	61 maternidades, SC	2012	Razão de prevalência de cesárea de 1,44 no privado vs. público.
Raspantini <i>et al.</i> (2016)	Local (SP)	Descritivo (idade gestacional)	SINASC–CNES	2013–2014	Desvio à esquerda da idade gestacional; mediana 39 (SUS) vs. 38 (não SUS).
Entringer, Pinto e Gomes (2019)	Local (3 maternidades)	Microcusteio	Dados primários	n.d.	A cesárea eletiva custa cerca de 32% mais que o parto vaginal pelo SUS.

continua na próxima página

(continuação do Quadro 1)

Autores (ano)	Abrangência	Tipo de estimação	Base de dados	Período	Principais resultados
Silva e Pazin-Filho (2024)	Local (SP)	Análise de custos	Saúde suplementar, SP	2015–2021	Privado 80% (71% acima do público); gasto de R\$ 7,7 bilhões.
Campos, Rattner e Diniz (2024)	Local (SP)	DiD retrospectivo (participantes vs. não)	SINASC	2014–2019	PPA: privado de 83,8% para 72,3%, ainda acima do C-model (45,2%).
Negrini <i>et al.</i> (2021)	Local (1 maternidade, SP)	Séries temporais interrompidas	Dados hospitalares	~2015–2020	Redução sustentada da cesárea intraparto ao longo de cinco anos.
Kondo <i>et al.</i> (2025)	Local (1 maternidade)	Melhoria de qualidade (Robson 1+3)	Dados hospitalares	~3 anos	Redução de 5 a 10 p.p em Robson 1+3 sem piora de desfechos.

Fonte: Elaboração própria a partir dos estudos revisados.

O quadro evidencia a convergência dos três blocos de literatura em torno do diferencial público–privado e da leitura por demanda induzida pela oferta, pano de fundo empírico para as escolhas metodológicas das seções seguintes.

3 METODOLOGIA E BASE DE DADOS

Por que a taxa brasileira de cesárea se sustenta tão acima do que recomenda a literatura clínica, e como esse excesso se distribui pela rede hospitalar do país? Para responder, é preciso mostrar antes como a rede está organizada e quem nasce em cada parte dela. Esta seção faz isso: apresenta as duas fontes administrativas que sustentam toda a análise empírica, descreve a estrutura institucional da oferta hospitalar brasileira (com a separação tripartite entre público, privado sem fins lucrativos e privado lucrativo), localiza o recorte clínico que define a amostra principal e fecha com a variação sociodemográfica do perfil das gestantes.

3.1 Base de dados

O estudo utiliza duas bases administrativas brasileiras: o Sistema de Informações sobre Nascidos Vivos (SINASC), que registra cada nascimento vivo ocorrido no país, e o Cadastro Nacional de Estabelecimentos de Saúde (CNES), que descreve cada hospital. As duas são unidas pelo identificador do estabelecimento, no nível do parto, no recorte 2014–2022.³¹ Cada linha do conjunto resultante representa um parto: do SINASC vêm as informações da mãe e do recém-nascido; do CNES, os atributos do hospital onde o parto ocorreu. A lista completa de variáveis, a integração formal e os filtros estão detalhados na seção Fontes e construção da base integrada (§3.8).³² A unidade observacional é o parto; o hospital é a unidade de *cluster* para a inferência estatística das seções seguintes.

A integração formal entre SINASC e CNES no nível do parto é o passo que permite ao trabalho deslocar o foco do indivíduo para o estabelecimento. O SINASC, sozinho, descreve o que aconteceu com a gestante; o CNES, por sua vez, descreve as características do hospital. Juntos, os dois registros permitem perguntar por que duas gestantes com perfis clínicos comparáveis recebem indicações de parto diferentes em hospitais com regimes de propriedade diferentes. Essa pergunta estrutura todo o restante do trabalho.

A taxa nacional de cesárea sobre o agregado intraparto fica em torno de 56% desde 2014 (Pereira *et al.*, 2024), e atinge 55,5% em 2021 (Brasil. Ministério da Saúde, 2022). O número

³¹ O recorte temporal é justificado pela cobertura das duas variáveis-chave da pesquisa: o grupo da Classificação de Robson só atinge cobertura populacional consistente em 2014, e a escolaridade da mãe na codificação *2010-agregada* só a partir de 2013. Cortar em 2014 preserva nove safras anuais, suficientes para efeito fixo de ano. A janela e os filtros estão formalizados na seção de construção da base integrada (§3.8).

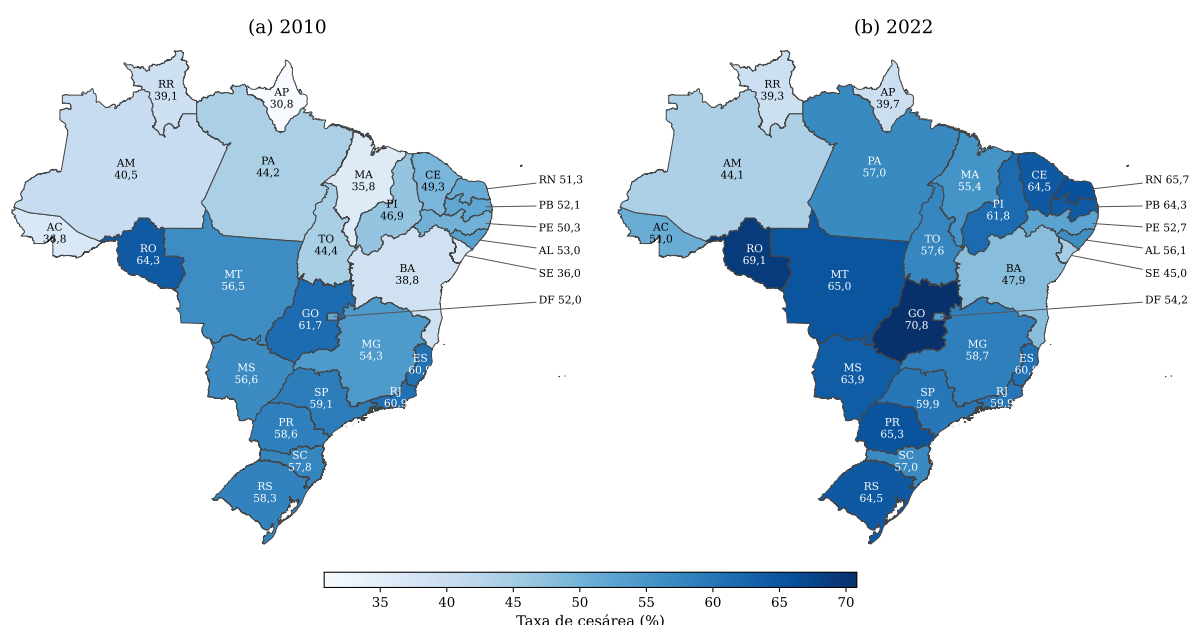
³² Descartam-se nascimentos sem correspondência no CNES (perda inferior a 0,5%), partos com tipo inválido, a categoria residual *Outros* de natureza jurídica (1,78%) e o perfil *Sem registro* (2,52%); a base resultante reúne cerca de 25,3 milhões de partos em aproximadamente cinco mil hospitais.

agregado, porém, esconde mais do que mostra: resulta da soma de hospitais que operam em níveis muito diferentes da taxa, e a estrutura interna dessa soma é o objeto desta seção.

3.2 A heterogeneidade espacial

A taxa de cesárea é muito desigual entre as regiões brasileiras. Em 2022, Sudeste (61%), Sul (60%) e Centro-Oeste (62%) operam acima da média nacional; Norte (50%) e Nordeste (52%), abaixo (Brasil. Ministério da Saúde, 2022; Pereira *et al.*, 2024). A dispersão entre estados, dentro de cada região, é da mesma ordem da dispersão entre regiões: São Paulo opera em 64%, Pernambuco em 50%. A Figura 1 apresenta a distribuição estadual em dois cortes temporais.

Figura 1 – Taxa de cesárea por unidade da federação. Brasil, partos hospitalares, 2010 e 2022.



Fonte: Elaboração própria a partir do SINASC, 2010 e 2022.

O fato que importa ao argumento, visível na Figura 1, é que a hierarquia regional é estável no tempo: estados que operavam em taxas elevadas em 2010 seguem operando em taxas elevadas em 2022, e vice-versa. Não há reordenamento; há convergência generalizada para taxas elevadas. O diferencial entre estados sobrevive ao controle pelo perfil materno e pelo grupo de Robson, padrão coerente com Boerma *et al.* (2018), e justifica o efeito fixo de UF nas regressões da seção 4: o que se quer medir é o diferencial entre tipos de hospital, depurado da heterogeneidade regional. A série histórica nacional e a distribuição completa por região e por UF, em dois cortes temporais, estão no Apêndice A (§A.1 e §A.2).

3.3 A estrutura institucional da rede hospitalar

O CNES classifica cada hospital pela sua natureza jurídica, o regime de propriedade da entidade que o controla. Os 24 códigos previstos no cadastro são agregados aqui em três grupos clinicamente relevantes, daqui em diante referidos como **tipologia tripartite: Público** (administração pública direta e indireta, incluindo hospitais universitários e empresas públicas de saúde), **Privado sem fins lucrativos** (associações, fundações privadas, organizações religiosas, Santas Casas de Misericórdia) e **Privado lucrativo** (sociedades empresariais com fim de lucro).³³

O Quadro 2 consolida o agrupamento dos códigos de natureza jurídica nessas três categorias.

Quadro 2 – Agrupamento dos códigos de natureza jurídica (IBGE/Receita Federal) em categorias analíticas.

Categoria	Códigos incluídos	Exemplos
Público	1xxx; 2011; 2038	União, estados, municípios, autarquias e fundações públicas, hospitais universitários federais e estaduais, secretarias de saúde, empresas públicas (NJ 2011) e sociedades de economia mista (NJ 2038) de controle estatal.
Privado sem fins lucrativos	3xxx	Associações privadas, fundações privadas, organizações religiosas, Santas Casas de Misericórdia e demais hospitais beneficentes.
Privado lucrativo	2xxx (exceto 2011 e 2038)	Sociedades empresárias limitadas, sociedades anônimas, empresas individuais de responsabilidade limitada e cooperativas com fim lucrativo.

Fonte: Elaboração própria a partir da Tabela de Natureza Jurídica 2018 (IBGE/Receita Federal) e do dicionário do CNES.

A separação importa por uma razão econômica direta: as três categorias operam sob estruturas de incentivo distintas. Hospitais públicos atendem por orçamento, sem margem comercial sobre o procedimento. Hospitais privados lucrativos atendem por reembolso (de operadora de plano de saúde ou particular), com margem comercial direta. Hospitais privados sem fins lucrativos ocupam uma posição intermediária: têm a estrutura de governança privada, mas não distribuem lucro, e na prática brasileira combinam atendimento SUS com atendimento à saúde suplementar em proporções variáveis. Em 2022, a taxa mediana de cesárea por hospital

³³ Operacionalmente: Público corresponde aos códigos de natureza jurídica 1xxx do IBGE/Receita Federal mais os códigos 2011 (Empresa Pública) e 2038 (Sociedade de Economia Mista), que apesar do prefixo 2 são juridicamente parte da administração pública indireta; Privado sem fins lucrativos corresponde aos códigos 3xxx; Privado lucrativo corresponde aos demais 2xxx. Códigos 4xxx (pessoas físicas) e 5xxx (organizações internacionais) compõem uma categoria *Outros* residual, com menos de 1% dos estabelecimentos com nascimentos registrados, excluída da modelagem. A categorização segue a Portaria SAS/MS nº 1.319/2014 (Brasil. Ministério da Saúde. Secretaria de Atenção à Saúde, 2014).

nas três categorias é de 47% no Público, 63% no Privado sem fins lucrativos e 81% no Privado lucrativo. O setor sem fins lucrativos opera, portanto, em nível intermediário entre Público e Lucrativo. A dicotomia tradicional público versus privado, ao colapsar o sem fins lucrativos junto com o lucrativo, mistura realidades institucionais distintas e atenua o diferencial que se quer medir. A tipologia tripartite preserva a informação que importa.

A literatura brasileira separa os hospitais por critérios diversos, todos extraídos do próprio CNES: Freitas e Fernandes (2016) classificam as maternidades de Santa Catarina pela razão entre leitos SUS e privados (público, privado e misto), e Mangolim (2024) adota a dicotomia público *versus* privado pelo tipo de prestador. A opção pela natureza jurídica, aqui, é a que isola de forma mais limpa o segmento lucrativo, enquanto a dimensão de financiamento (vínculo com o SUS ou com a saúde suplementar) entra separadamente, como perfil de atendimento.

A Tabela 1 consolida a distribuição da rede e dos nascimentos pelos três grupos em dois cortes temporais, e três pontos organizam esta leitura.³⁴ Primeiro, o Brasil tem uma rede obstétrica majoritariamente pública por contagem de hospitais e por contagem de nascimentos, com participação pública crescente entre 2012 e 2022 (de 53% para 60% dos hospitais; de 40% para 47% dos nascimentos). Segundo, o Privado sem fins lucrativos é o maior bloco *privado*, à frente do lucrativo tanto em hospitais quanto em nascimentos. Terceiro, o Privado lucrativo perde participação relativa ao longo da década, mas permanece o segmento que opera nas taxas mais altas de cesárea. O Brasil tem, portanto, uma rede obstétrica pública dominante em volume e um segmento lucrativo minoritário em volume; é justamente esse segmento, contudo, que sustenta a posição internacional do país no mapa global da cesárea. O detalhamento da rede por região, perfil de atendimento e infraestrutura instalada está no Apêndice A (§A.3).

Tabela 1 – Distribuição de hospitais e de nascidos vivos por natureza jurídica, Brasil, 2012 e 2022.

Categoria	Hospitais (%)		Nascidos vivos (%)	
	2012	2022	2012	2022
Público	53,0%	59,5%	40,3%	47,0%
Privado sem fins lucrativos	25,9%	23,4%	36,5%	33,7%
Privado lucrativo	20,9%	16,8%	21,7%	18,0%
Outros	0,2%	0,3%	1,5%	1,4%
Total (N)	5.694	5.344	2.877.780	2.534.512

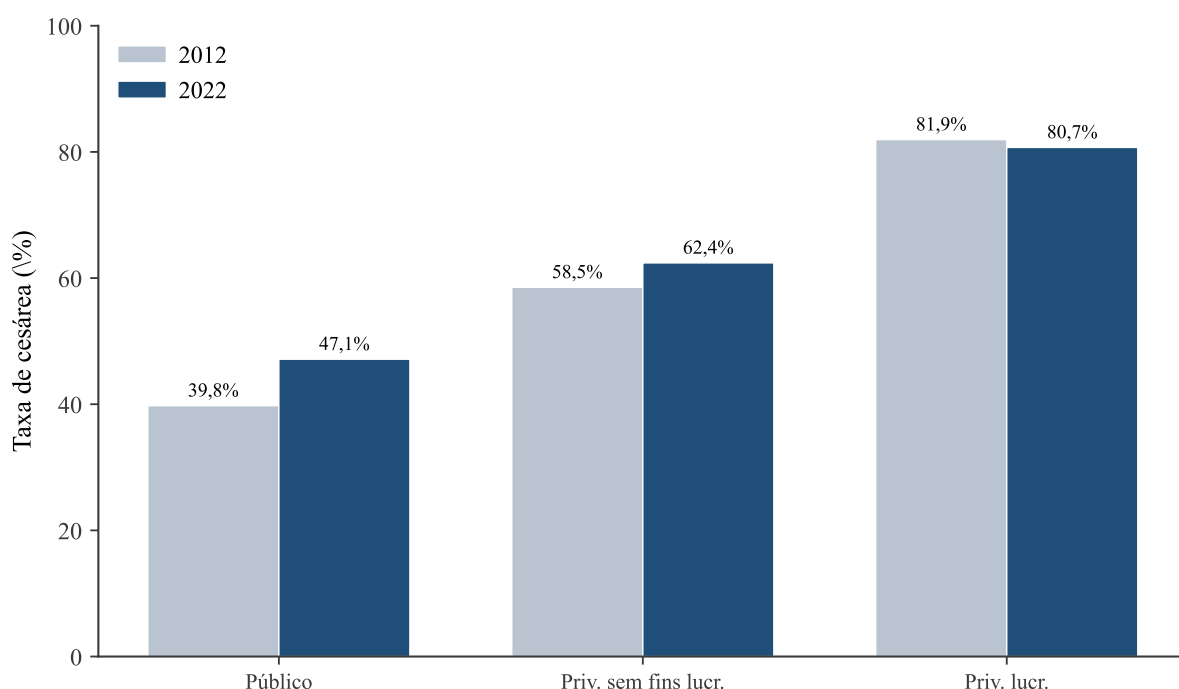
Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

³⁴ A linha-base de comparação da tabela é 2012, e não 2010, porque o CNES histórico de 2010 e 2011 não traz a variável de natureza jurídica preenchida com cobertura suficiente para a categorização.

A heterogeneidade interna da rede fica mais nítida na comparação entre categorias de natureza jurídica.

A Figura 2 compara a taxa de cesárea por categoria de natureza jurídica entre 2012 e 2022: o privado lucrativo permanece sistematicamente acima, próximo de 80%, enquanto o público sobe alguns pontos mas se mantém bem abaixo, de modo que a distância entre os setores se preserva ampla ao longo do período.

Figura 2 – Taxa de cesárea por natureza jurídica do hospital. Brasil, 2012 e 2022.



Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

A persistência desse fosso ao longo de uma década, apesar das sucessivas campanhas federais, indica que o diferencial entre os setores é estrutural, e não conjuntural; resta isolar a margem clinicamente discricionária em que o diferencial se decide, tarefa da próxima seção.

3.4 A Classificação de Robson e a amostra principal Robson 1+3

Comparar diretamente a taxa de cesárea de hospitais privados lucrativos com a de hospitais públicos é enganoso se as gestantes que chegam a cada um forem clinicamente diferentes. Para isolar o componente discricionário, é preciso comparar gestantes em situações clinicamente comparáveis.

A Classificação de Robson (Robson, 2001; Organização Mundial da Saúde (OMS), 2017) é a ferramenta padrão para esse fim. A classificação estratifica a população obstétrica em

dez grupos clinicamente comparáveis a partir de cinco características registradas na admissão: paridade (se já teve filho antes), histórico de cesárea, apresentação do feto³⁵, idade gestacional e se o trabalho de parto foi espontâneo ou induzido.

A Tabela 2 mostra por que o recorte importa: a taxa de cesárea varia enormemente entre os grupos. Nos Grupos 1 e 3 (em destaque), que concentram a maior discricionariedade clínica, a taxa é comparativamente baixa; nos grupos com perfil clínico mais complexo, a taxa se aproxima ou ultrapassa os 90%.

Tabela 2 – Taxa de cesárea por grupo de Robson, Brasil, 2022. Os Grupos 1 e 3 (em destaque) compõem a amostra principal deste trabalho.

Grupo de Robson	Taxa de cesárea (%)	Participação nos partos (%)
1 • nulípara, cefálica, termo, espontâneo	44,6	17,1
2 • nulípara, cefálica, termo, induzida/eletiva	73,9	14,0
3 • múltipara s/ ces. anterior, cefálica, termo, espontâneo	20,0	19,2
4 • múltipara s/ ces. anterior, cefálica, termo, induzida/eletiva	52,2	9,3
5 • cesárea anterior, cefálica, termo	86,1	25,0
6 • nulípara, pélvica	92,2	1,3
7 • múltipara, pélvica	89,5	2,0
8 • gestação múltipla	86,8	2,3
9 • apresentação transversa/oblíqua	97,1	0,2
10 • pré-termo, cefálica	56,0	9,7
Brasil (grupos 1–10)	58,8	100,0

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Os Grupos 1 e 3 reúnem gestantes *sem cesárea anterior*, com feto único cefálico, gestação a termo e trabalho de parto espontâneo; a única diferença entre os dois grupos é ter parido antes ou não. O Grupo 1 são nulíparas (nenhum parto anterior) e o Grupo 3 são múltiparas (já pariram antes, sempre por via vaginal). É a população de máxima discricionariedade clínica: são gestantes de baixo risco, clinicamente elegíveis ao parto vaginal, e qualquer cesárea registrada acontece durante o trabalho de parto, em situação em que o parto vaginal estava em curso. Quando se observa, nessa população, taxa elevada de cesárea, é razoável atribuir uma parcela do excesso a fatores externos à indicação clínica estrita.

A amostra principal restringe-se aos Grupos 1 e 3 por três razões que se reforçam. Primeira: comparabilidade clínica máxima. Como as gestantes dos Grupos 1 e 3 são clinicamente muito parecidas entre si, a diferença de cesárea que sobra entre os tipos de hospital reflete a conduta de cada estabelecimento, e não o perfil das pacientes que cada um atende; o diferencial

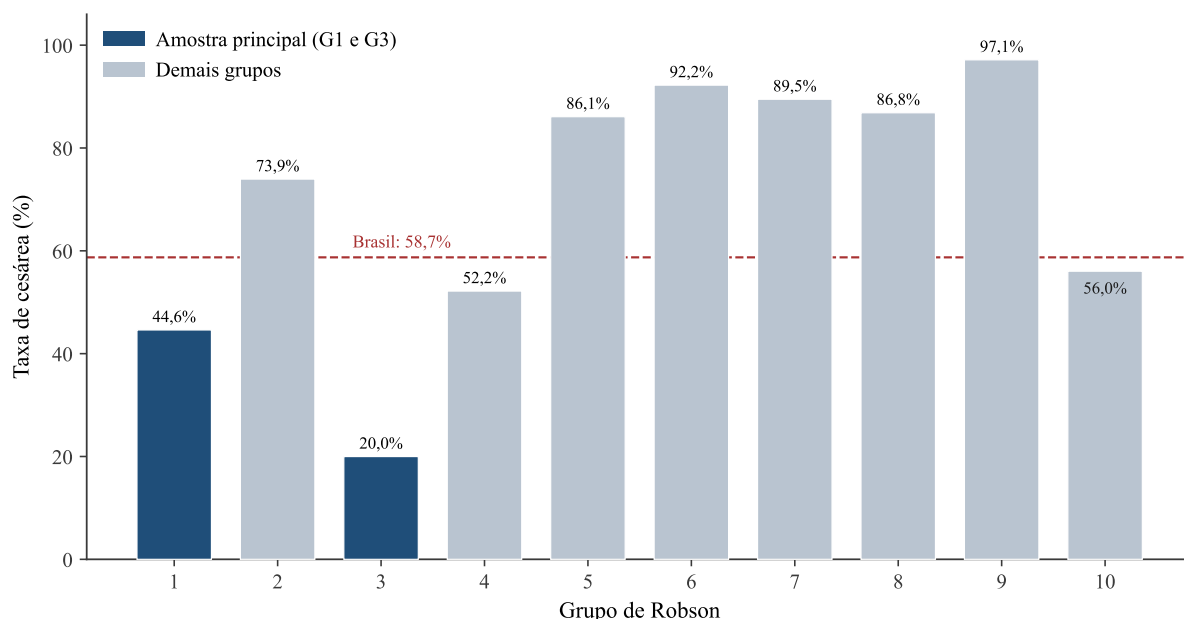
³⁵ Apresentação fetal: parte do feto que se apresenta primeiro ao canal de parto. Cefálica (cabeça) é a apresentação mais comum; nas apresentações pélvica (nádegas ou pés) e transversa, os dados mostram proporção de cesárea muito mais alta.

residual é, portanto, interpretável como discricionariedade clínica, e não como reflexo do *mix* de pacientes. Segunda: contornabilidade da endogeneidade do Grupo 5. Gestantes só entram no Grupo 5 porque já tiveram cesárea no passado; incluí-lo na amostra causal mistura efeito de tratamento atual com efeito da decisão pretérita. Terceira: alinhamento com a literatura clínica recente. Kondo *et al.* (2025), em intervenção de melhoria de qualidade em maternidade privada brasileira, adotam exatamente esse recorte como população-alvo da redução. Nakamura-Pereira *et al.* (2016), na primeira aplicação nacional da Classificação de Robson sobre o SINASC, identificam os Grupos 1 e 2 como os de contribuição mais crítica para o excesso brasileiro. Paixão *et al.* (2021b), ao replicar o exercício em escala maior, confirmam o padrão.

A amostra principal contém **8.947.809 partos** em **5.280 hospitais** no período **2014–2022**, com taxa interna de cesárea de **31,87%**. Esse percentual está mais próximo da faixa de referência discutida pela OMS (Betrán *et al.*, 2016) do que a taxa agregada brasileira; ainda assim, é alto para uma população de baixo risco, clinicamente elegível ao parto vaginal. Duas amostras adicionais sustentam exercícios de sensibilidade: **Robson 1–4** (*baixo risco amplo*, 14,7 milhões de partos), que inclui também os Grupos 2 e 4 (induzidos ou cesárea pré-trabalho de parto, na definição corrente da literatura epidemiológica brasileira (Leal *et al.*, 2014)); e **Robson 1–10** (todos os grupos, 23,8 milhões de partos), usada para verificar se o achado central sobrevive à inclusão dos grupos endógenos e dos grupos de alto risco. A distribuição dos dez grupos de Robson e a decomposição do excesso de cesárea por grupo estão no Apêndice A (§A.5, Tabelas A10 e A11).

A Figura 3 traduz em imagem o contraste da Tabela 2: os Grupos 1 e 3 (em destaque), a amostra principal, operam nas taxas mais baixas, enquanto os grupos de maior complexidade clínica se aproximam ou ultrapassam 90%.

Figura 3 – Taxa de cesárea por grupo de Robson, com Grupos 1 e 3 (amostra principal) em destaque. Brasil, 2022.



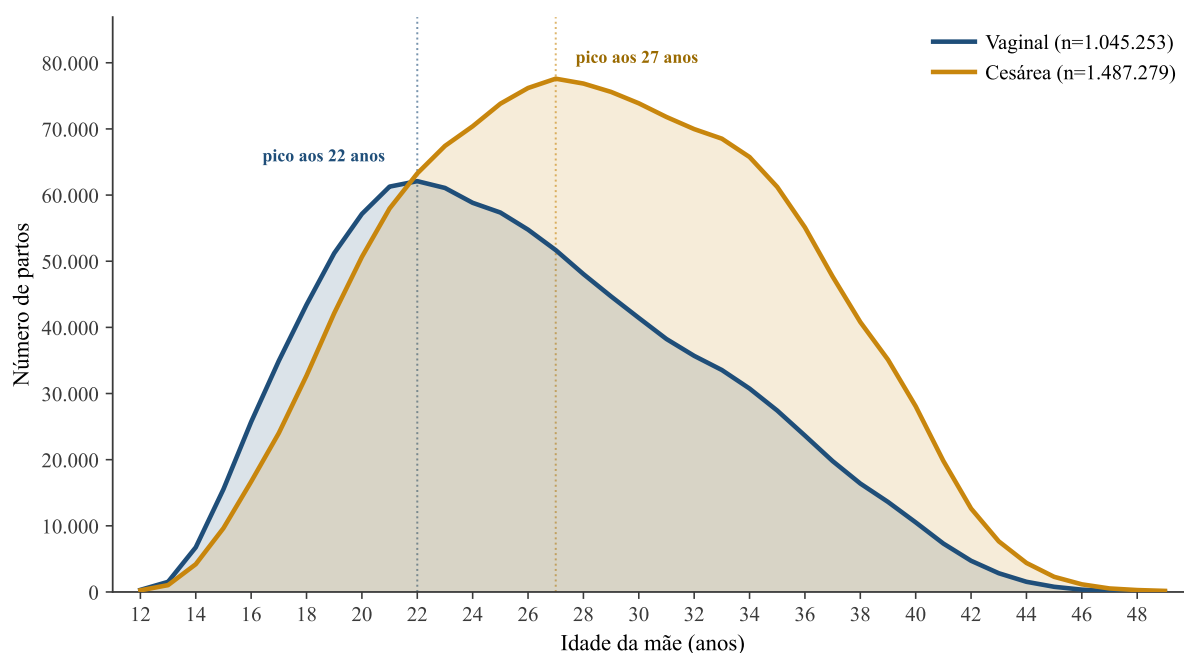
Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

A comparabilidade clínica desta amostra opera em duas camadas que se completam. A primeira é a *restrição* aqui descrita. A Classificação de Robson iguala os eixos *obstétricos*: apresentação cefálica, feto único, gestação a termo, trabalho de parto espontâneo e ausência de cesárea anterior. São justamente as dimensões em que a via de parto seria, de outro modo, quase predeterminada e em que o ajuste estatístico opera mal por falta de variação no desfecho. A segunda é o *ajuste*: dentro de Robson 1+3, os estimadores da seção 4 (logit, IPW, AIPW, *entropy balancing*, PSM e GLMM) equiparam as gestantes também nas características *sociodemográficas e maternas* que a Classificação de Robson não cobre, em particular idade, escolaridade, raça e cor, paridade e adequação do pré-natal. A análise de balanceamento (§3.6) mostra, pela diferença padronizada de médias, que essas covariáveis estão fortemente desbalanceadas entre os setores antes do ajuste (idade 0,42, escolaridade 0,58, raça/cor 0,38) e praticamente iguais depois dele (todas abaixo de 0,04). Idade, escolaridade e raça, portanto, não ficam de fora: entram como variáveis de equiparação na segunda camada, não como critério de recorte da amostra. A restrição cuida do que é clinicamente quase determinístico; o ajuste cuida do confundimento sociodemográfico genuíno. Uma sensibilidade que abandona o recorte de Robson e equipara os setores apenas por ajuste, sobre o grupo clínico mais amplo (cefálico, feto único e a termo, incluindo cesárea anterior), confirma o diferencial do privado lucrativo e mostra que esse diferencial não depende dessa escolha (Apêndice C.4).

3.5 Perfil sociodemográfico das gestantes

A taxa de cesárea, mesmo dentro do recorte clínico Robson 1+3, varia sistematicamente com as características socioeconômicas da gestante. A diferença é nítida na escolaridade: em 2022, a taxa entre gestantes com ensino fundamental (1 a 7 anos de estudo) é de 29%; entre as de ensino médio (8 a 11 anos), 32%; entre as de ensino superior (12 ou mais anos), 42%. O crescimento é aproximadamente linear e atravessa todas as regiões. Por idade, a taxa cresce de 24% entre gestantes de 15 a 19 anos para 46% entre as de 35 a 39 anos; a Figura 4 mostra esse deslocamento, com a distribuição etária das cesáreas puxada para idades mais altas que a dos partos vaginais. Por raça e cor, gestantes brancas operam em taxa mais alta (40%) do que pretas (28%) e pardas (29%).

Figura 4 – Distribuição da idade materna por via de parto. Robson 1+3, Brasil, 2014–2022.



Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

A leitura econômica dessa variação só se completa com a estrutura tripartite da oferta, que é o que organiza o padrão observado. Em hospitais privados lucrativos, a taxa de cesárea em Robson 1+3 supera 75% em todas as faixas de escolaridade, idade e raça e cor; em hospitais públicos, oscila entre 25% e 40% conforme o estrato. A maior parte da variação sociodemográfica observada no agregado nacional é, portanto, mediada pela escolha do estabelecimento: gestantes com mais escolaridade e mais idade têm acesso preferencial à saúde suplementar privada (por renda, por cobertura ocupacional, por localização), e o setor privado opera em taxa de cesárea

sistematicamente mais alta.

Parte do que parece ser efeito da educação ou da idade da mãe é, portanto, efeito do tipo de hospital ao qual essas mães chegam. Separar empiricamente os dois eixos (perfil materno e regime de propriedade do hospital) é um dos objetivos centrais do exercício causal da seção 4; a coexistência das duas dimensões mostrada aqui torna esse exercício empiricamente bem-posto. Os desfechos neonatais (Apgar) por setor institucional, que complementam esta leitura, estão no Apêndice A (§A.6).

3.6 Balanceamento entre tratados e controles (pré- e pós-reponderação)

Antes de avançar para a estimação causal da seção 4, é útil mostrar, em forma tabular, o quanto o perfil materno difere entre gestantes atendidas em hospitais públicos e em hospitais privados lucrativos e o quanto essa diferença se atenua depois da reponderação pelo escore de propensão. A medida convencional para isso é a diferença padronizada de médias (SMD³⁶), que escala a diferença bruta entre os dois grupos pelo desvio-padrão combinado e dispensa unidade. A literatura adota como limiar de balanceamento adequado o valor $SMD < 0,10$ (Austin, 2009). Olhar a SMD antes e depois da reponderação responde a uma pergunta de identificação direta: a comparação ajustada entre tipos de hospital é válida sob seleção em observáveis? A Tabela 3 apresenta o resultado para sete covariáveis-chave.

A leitura econômica é direta. Antes da reponderação, hospitais privados lucrativos atendem um perfil materno sistematicamente distinto do perfil que chega ao público: gestantes mais velhas, mais escolarizadas, mais frequentemente brancas, com pré-natal mais completo e maior proporção de múltiparas. As SMDs brutas em escolaridade Superior (0,58), idade da gestante (0,42), raça/cor Branca (0,38) e Kotelchuck³⁷ adequado (0,31) estão muito acima do limiar de 0,10 e indicam que uma comparação simples de taxa de cesárea entre os dois grupos confundiria efeito do tipo de hospital com efeito da composição das gestantes. Depois da reponderação por IPW e do pareamento por PSM, todas as SMDs caem para valores entre 0,01 e 0,04, abaixo do limiar convencional. Tecnicamente, isso significa que as duas amostras reponderadas são comparáveis nas covariáveis observáveis listadas, e que o diferencial residual de taxa de cesárea estimado nas seções seguintes pode ser lido como associação ajustada sob

³⁶ Do inglês *standardized mean difference*.

³⁷ Índice de Kotelchuck: medida de adequação da assistência pré-natal que combina o mês de início do acompanhamento e o número de consultas realizadas, ajustados à idade gestacional, classificando o pré-natal em níveis de adequação (de inadequado a adequado/mais que adequado) (Kotelchuck, 1994); captura intensidade do acompanhamento, não conduta médica.

seleção em observáveis, no sentido descrito na seção sobre o IPW (§3.10.1).

Tabela 3 – Balance pré-tratamento entre tratados e controles, bruto e após reponderação por escore de propensão, Robson 1+3, 2014–2022.

Covariável	Média Pub.	Média Lucr.	SMD bruto	SMD pós-IPW	SMD pós-PSM
Idade da gestante (anos)	26,1	30,4	0,42	0,03	0,02
Escolaridade Superior (%)	12,5	51,7	0,58	0,04	0,03
Múltipara (Robson 3, %)	41,8	49,1	0,15	0,02	0,02
Apresentação cefálica (%)	99,1	99,5	0,08	0,01	0,01
Pré-natal Kotelchuck adequado (%)	58,2	79,7	0,31	0,03	0,03
Idade gestacional ≥ 39 sem (%)	62,4	68,3	0,12	0,02	0,02
Raça/cor Branca (%)	34,6	65,2	0,38	0,03	0,03

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

A Figura 8 da seção 4 amplia esse exame para as vinte covariáveis com maior desbalanço inicial; aqui o objetivo é apenas registrar, antes de qualquer estimação, que o ajuste por escore de propensão cumpre o papel para o qual foi mobilizado.

3.7 Síntese da análise exploratória

Três escolhas metodológicas das seções seguintes saem diretamente da análise empírica descrita aqui. Primeira: a **tipologia tripartite** público / privado sem fins lucrativos / privado lucrativo (§3.3) é o tratamento das análises causais, em substituição à dicotomia tradicional. Segunda: o **recorte clínico Robson 1+3** (§3.4) é a amostra principal, isolando a população de máxima discricionariedade clínica. Terceira: o **efeito fixo de UF** é mantido nas especificações para absorver a heterogeneidade regional residual que sobrevive ao perfil materno e à classificação clínica. A coexistência de variação simultânea na oferta hospitalar e no perfil materno torna empiricamente bem-posta a separação entre os dois eixos na seção 4.

A partir daqui, o trabalho formaliza seu método. Cada etapa é aqui *definida*: que pergunta econômica responde, qual hipótese precisa valer para o resultado ser válido e como interpretar o número que cada modelo gera. A seção 4 executa o que está aqui definido, reporta as escolhas operacionais e interpreta os achados.

Cada estimador é introduzido por uma pergunta em texto corrido, em seguida formalizado, em seguida lido. Quando estimamos mais de um estimador para a mesma pergunta, a justificativa para a duplicidade está no texto: cada estimador relaxa uma hipótese exigida pelo anterior, de modo que os resultados se sustentam mutuamente.

3.8 Base integrada e amostras analíticas

As duas fontes administrativas brasileiras mobilizadas são o SINASC e o CNES, apresentadas em escopo descritivo na análise exploratória (§3.1). Do SINASC vêm: a via de parto (vaginal ou cesárea), as características clínicas da gestação (idade gestacional, apresentação fetal³⁸, paridade, tipo de gravidez), o histórico obstétrico, as variáveis sociodemográficas da mãe (idade, raça e cor, escolaridade, ocupação), o pré-natal (número de consultas, mês de início, índice de Kotelchuck³⁹), o desfecho neonatal imediato (peso ao nascer, Apgar) e o identificador do estabelecimento. Desde 2010 o SINASC registra o grupo da Classificação de Robson em variável dedicada, com cobertura populacional consistente a partir de 2014. Do CNES vêm: a natureza jurídica da entidade mantenedora, o tipo de gestão, o perfil de atendimento, o vínculo com o SUS, os indicadores de aceitação de internação por fonte de pagamento e a capacidade técnica instalada (centro obstétrico, unidade neonatal, totais de salas e leitos). A categorização de natureza jurídica segue a Portaria SAS/MS nº 1.319/2014 (Brasil. Ministério da Saúde. Secretaria de Atenção à Saúde, 2014).

A integração entre as bases é feita pelo identificador do estabelecimento, no nível do parto: cada linha contém simultaneamente o nascimento, o perfil da mãe, o histórico obstétrico, a classificação clínica e os atributos do hospital. Nascimentos sem correspondência no CNES foram excluídos, com perda inferior a 0,5% do total bruto.

A janela temporal é 2014–2022. Os anos anteriores foram excluídos por dois motivos. Primeiro, a variável de Robson tem cobertura insuficiente em 2010–2011 e está zerada em 2013, o que inviabiliza o uso da Classificação de Robson como critério amostral. Segundo, a variável de escolaridade da mãe só está preenchida com cobertura populacional a partir de 2013. Cortar em 2014 mantém uma janela suficiente para o uso de efeito fixo de ano nas especificações.

Dois filtros de integração são aplicados a montante: via de parto válida (vaginal ou cesárea, exclusão dos 0,08% com código ignorado) e estabelecimento identificável no CNES (necessário para qualquer atributo institucional e para o agrupamento na inferência, com perda inferior a 0,5%). A base integrada resultante contém **25.278.233 partos hospitalares**, distribuídos em aproximadamente cinco mil hospitais. As amostras analíticas aplicam, na sequência, dois filtros de natureza institucional: a exclusão da categoria residual *Outros* de natureza jurídica (1,78%) e do perfil *Sem registro* (2,52%), que reúne unidades básicas e prontos atendimentos com o parto

³⁸ Apresentação fetal: posição do feto em relação ao canal de parto no momento do nascimento — cefálica (cabeça), pélvica (nádegas) ou córmica/transversa; apresentações não cefálicas elevam a indicação de cesárea.

³⁹ Índice de adequação da assistência pré-natal, que combina o mês de início e o número de consultas em uma classificação de adequação do acompanhamento.

como evento raro. A Tabela 4 reconstrói o fluxo do registro integrado às três amostras.

Tabela 4 – Fluxo amostral do SINASC–CNES integrado (2014–2022) às amostras analíticas.

Etapas	Partos	Hospitais	Cesárea
Base integrada (via de parto válida; <i>match</i> CNES)	25.278.233	≈5.000	n.a.
<i>Amostras analíticas</i> (excluem a categoria <i>Outros</i> de natureza jurídica e o perfil <i>Sem registro</i>):			
Robson 1–10 (todos os grupos; sensibilidade extrema)	23.838.936	5.537	56,55%
Robson 1–4 (baixo risco amplo; sensibilidade)	14.743.401	5.351	43,44%
Robson 1+3 (principal)	8.947.809	5.280	31,87%

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

As três amostras — a principal e duas de sensibilidade — são subconjuntos da base integrada com os mesmos filtros comuns, diferindo apenas no escopo clínico de Robson.

A **amostra principal** restringe-se aos Grupos 1 e 3 da Classificação de Robson (Robson, 2001; Organização Mundial da Saúde (OMS), 2017). À comparabilidade clínica que motiva o recorte (§3.4) soma-se aqui uma segunda razão: a remoção da endogeneidade do Grupo 5, em que a cesárea atual é mecanicamente correlacionada com cesárea anterior, o que contamina qualquer leitura sobre indução. A amostra principal contém **8.947.809 partos** em **5.280 hospitais**, com taxa interna de cesárea de **31,87%**, valor próximo ao limiar populacional revisado pela OMS (Betrán *et al.*, 2016). Por construção, gestantes nessa amostra estão em trabalho de parto espontâneo: qualquer cesárea registrada ocorreu durante o trabalho de parto (intraparto), uma versão clinicamente refinada da chamada cesárea evitável. A restrição por Robson equipara os eixos obstétricos; as covariáveis sociodemográficas e maternas que a classificação não cobre (idade, escolaridade, raça e cor, paridade e adequação do pré-natal) são equiparadas por ajuste nos estimadores a seguir, com o balanço aferido pela diferença padronizada de médias (§3.6).

As duas **amostras de sensibilidade** ampliam o recorte. A primeira inclui os Grupos 1 a 4 (G2 e G4 acrescentam casos induzidos ou cesárea pré-trabalho de parto): **14.743.401 partos**, taxa de **43,44%**; é a definição corrente de *baixo risco amplo* usada pela literatura epidemiológica brasileira (Leal *et al.*, 2014). A segunda mantém todos os Grupos 1 a 10, incluindo o G5 endógeno e os grupos de alto risco: **23.838.936 partos**, taxa de **56,55%**. A interpretação principal repousa sobre Robson 1+3; as duas sensibilidades servem para julgar até que ponto o efeito sobrevive à incorporação dos canais excluídos.

Uma nota sobre escopo. As taxas que abrem este trabalho (58,7% no agregado nacional de 2022, 80,7% no setor privado lucrativo) descrevem o *problema* de saúde pública em toda a sua extensão; a amostra principal, restrita a Robson 1+3, opera sobre uma taxa interna bem menor (31,87%). A distância entre as duas não é incoerência, e sim divisão de tarefas: a estatística

agregada motiva a pergunta e dimensiona a relevância do fenômeno, enquanto o recorte clínico isola a margem em que a decisão de operar é de fato discricionária, removendo da comparação as gestações em que a cesárea é mecanicamente esperada (cesárea anterior, apresentações não cefálicas, prematuridade). O diferencial estimado nas seções seguintes é, portanto, conservador: mede a indução onde é mais difícil de justificar clinicamente, não onde a manchete é mais alta.

3.9 Estratégia empírica: quatro perguntas, quatro famílias de instrumentos

A estratégia empírica organiza-se em quatro perguntas, cada uma respondida por uma família distinta de instrumentos. O método acompanha a natureza de cada pergunta. As perguntas descritivas são respondidas por modelos preditivos, e as causais, por estimadores de inferência causal.

O Quadro 3 resume essa correspondência: para cada pergunta, o método principal, os testes de robustez e a finalidade do passo. O quadro deixa explícito, por exemplo, que os *bounds* de Oster e o E-value entram como sensibilidade a inobservados depois de estimado o diferencial institucional, e não como estimadores independentes.

Quadro 3 – Roteiro metodológico: as quatro perguntas, os métodos e a finalidade de cada passo.

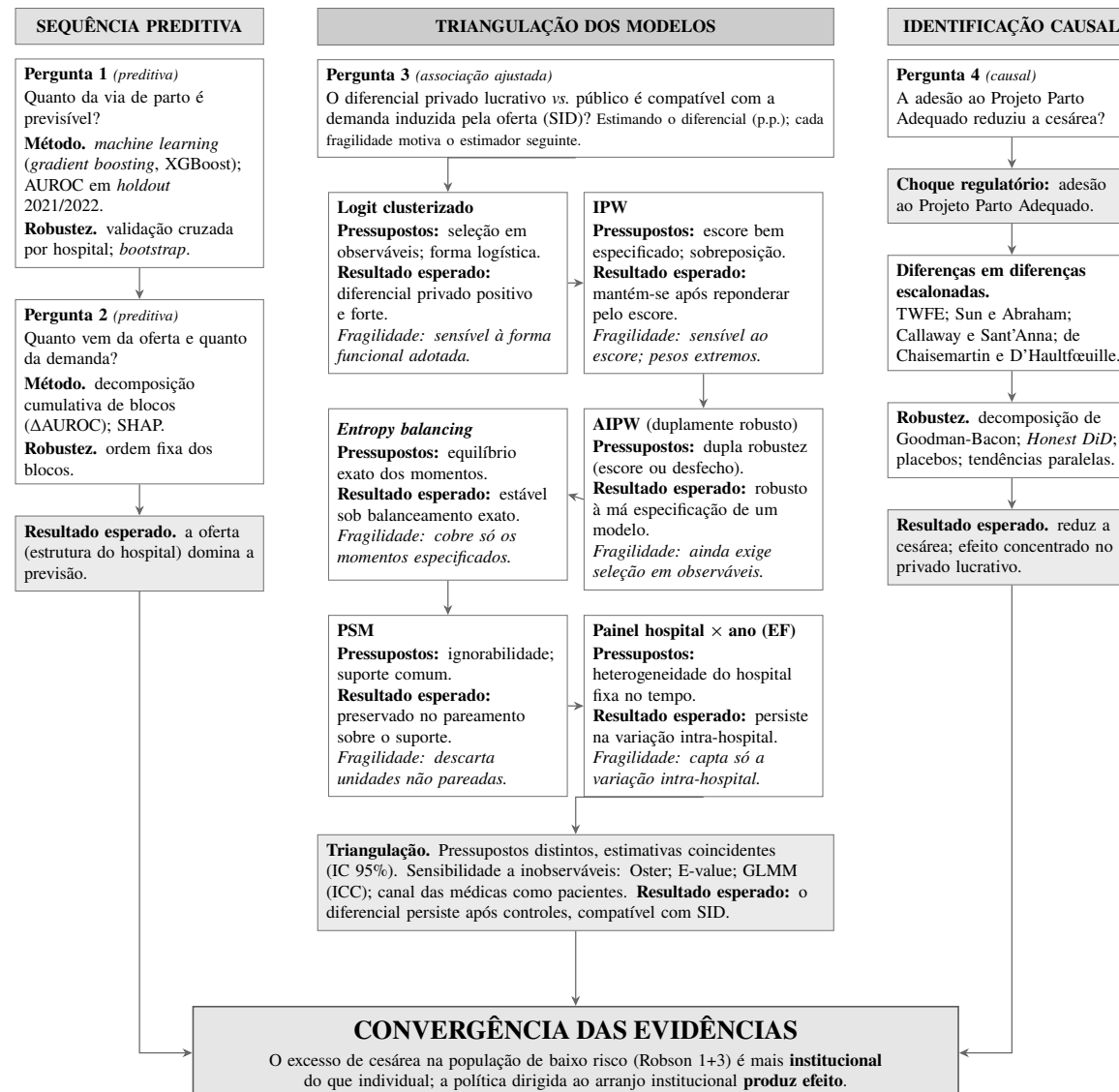
Pergunta	Método principal	Robustez e sensibilidade	Finalidade (tipo de inferência)
Quanto da via de parto é previsível?	<i>Machine Learning</i> (<i>gradient boosting</i>), avaliado em <i>holdout</i> temporal por AUROC.	Validação cruzada por hospital; intervalos por <i>bootstrap</i> .	Medir o limite máximo de previsibilidade da decisão a partir do registro administrativo. <i>Preditiva, não causal</i> .
Quanto vem da oferta e quanto da demanda?	Decomposição cumulativa de blocos de variáveis (Δ AUROC) e importância por SHAP.	Ordem fixa dos blocos, sob o mesmo recorte clínico.	Separar a contribuição preditiva da estrutura do hospital e do perfil da gestante. <i>Preditiva</i> .
O diferencial institucional é compatível com a demanda induzida pela oferta?	Logit clusterizado, IPW, AIPW, <i>entropy balancing</i> , PSM e painel hospital-ano com efeitos fixos.	<i>Bounds</i> de Oster e E-value (sensibilidade a inobservados, aplicados após os estimadores); GLMM (ICC); heterogeneidade.	Estimar o diferencial privado lucrativo vs. público sob seleção em observáveis e medir sua fragilidade a confundimento. <i>Associação ajustada, compatível com SID</i> .
(refinamento) Qual o canal da assimetria?	Comparação entre gestantes médicas e não médicas de mesmo perfil.	Pareamento; recorte por escolaridade e setor.	Distinguir o canal informacional do financeiro da demanda induzida. <i>Associação ajustada</i> .
A adesão ao Parto Adequado reduziu a cesárea?	Diferenças em diferenças escalonadas (TWFE; Sun-Abraham; Callaway-Sant'Anna; De Chaisemartin-D'Haultfoeuille).	Decomposição de Goodman-Bacon; <i>Honest DiD</i> ; placebos; teste de tendências paralelas no <i>cohort</i> 2017.	Identificar o efeito causal da adesão ao programa sobre a taxa de cesárea. <i>Causal, no escopo do PPA</i> .

Fonte: Elaboração própria.

Fixada essa correspondência, as três predições do modelo do prestador formalizado na seção 2.3.1 motivam o que cada família de estimadores testa.

A Figura 5 sintetiza a estratégia empírica: cada uma das quatro perguntas é respondida pela família de instrumentos adequada à sua natureza — preditiva, observacional ou quase-experimental —, e as leituras convergem para uma evidência triangulada.

Figura 5 – Estratégia empírica: as quatro perguntas, as famílias de instrumentos que respondem a cada uma e a leitura conjunta triangulada.



Previsibilidade da decisão e decomposição entre oferta e demanda: *gradient boosting* de árvores em duas implementações concorrentes (XGBoost e LightGBM), na tradição de *Machine Learning* aplicado à economia consolidada por Mullainathan e Spiess (2017). A avaliação usa *holdout* temporal estrito⁴⁰, com leitura por AUROC saturado para a previsibilidade e por Δ AUROC marginal por bloco para a decomposição, complementada por importância individual via SHAP (Lundberg; Lee, 2017).

Diferencial entre as categorias institucionais (associação observacional consistente com SID): conjunto de seis estimadores baseados em escore de propensão sobre tratamento tripartite (logit binário com erros-padrão clusterizados, IPW, AIPW (Robins; Rotnitzky; Zhao, 1994), *entropy balancing* (Hainmueller, 2012), PSM 1:1 (Ho *et al.*, 2011) e painel hospital–ano com efeitos fixos), complementado por *bounds* de Oster (2019) e E-value de VanderWeele e Ding (2017) para sensibilidade a inobservados, logit hierárquico com intercepto aleatório por hospital (Bates *et al.*, 2015) e exercício do canal informacional adaptado de Johnson e Rehavi (2016).

Efeito causal quase-experimental do Parto Adequado (dentro do escopo do PPA): conjunto moderno de estimadores de DiD escalonado sobre adesão ao Projeto Parto Adequado da ANS, formalizada na seção de identificação por choque regulatório (§3.11): efeitos fixos em duas vias (*two-way fixed effects*, TWFE)⁴¹ em três especificações; ponderado por interação (*interaction-weighted*, IW) de Sun e Abraham (2021); estimador de regressão de desfecho de Callaway e Sant’Anna (2021); DID_M de Chaisemartin e D’Haultfœuille (2020); decomposição de Goodman-Bacon (2021); *Honest DiD* via *Relative Magnitudes* de Rambachan e Roth (2023); dois placebos; e nove cenários de elegibilidade.

Por que estimar mais de um estimador por pergunta. No diferencial institucional, os seis estimadores observacionais respondem ao mesmo alvo conceitual (a diferença esperada na probabilidade de cesárea entre privado e público, condicional ao perfil materno), mas cada um carrega um pressuposto identificacional diferente. O logit pede que a forma funcional esteja correta; o IPW pede que o modelo do tratamento esteja correto; o AIPW pede que pelo menos um dos dois (tratamento ou desfecho) esteja correto; o *entropy balancing* dispensa a forma funcional do escore e impõe balanço exato dos primeiros momentos; o PSM 1:1 dispensa qualquer modelo paramétrico no estágio final; o painel hospital–ano absorve heterogeneidade não observada estável no tempo. A convergência dos seis estimadores entrega, portanto, uma triangulação: o achado é robusto em conjunto se sobrevive sob hipóteses de identificação diferentes, já que

⁴⁰ Treino 2014–2019, validação 2020 (uso exclusivo para *early stopping*), teste 2021–2022 para a métrica final.

⁴¹ Especificação de painel que inclui simultaneamente efeitos fixos de unidade (hospital) e de tempo (ano), absorvendo a heterogeneidade constante em cada uma das duas dimensões.

cada modelo exige uma suposição própria. A divergência entre os estimadores, quando ocorre, informa por qual canal o pressuposto foi violado.

No desenho do Parto Adequado, a lógica é a mesma. Agora a escolha entre estimadores é ditada pelo viés conhecido do TWFE clássico em desenhos escalonados heterogêneos. O IW de Sun-Abraham, a regressão de desfecho (*outcome regression*, OR) de Callaway-Sant’Anna e o DID_M de De Chaisemartin-D’Haultfoeuille corrigem esse viés por vias matemáticas distintas. A Decomposição de Goodman-Bacon quantifica diretamente a fração da estimativa TWFE que vem das comparações *later vs. already-treated*, que são justamente as problemáticas no caso escalonado.

Como as famílias se relacionam. A interação entre a previsão e a decomposição, o diferencial institucional e o desenho do Parto Adequado não é redundante. O *boosting* responde à previsibilidade e à decomposição, mas não fornece leitura causal: alta importância SHAP de uma variável de oferta significa apenas que a variável move a previsão, não que cause o desfecho. A leitura causal observacional vem do conjunto que estima o diferencial institucional, sob seleção em observáveis. A leitura causal quase-experimental vem da série DiD/PPA do Parto Adequado, sob paralelismo. O diferencial institucional não controla por inobserváveis relevantes (preferência da gestante, gravidade não codificada, cultura institucional); o desenho do Parto Adequado sofre *selection-on-gains*⁴² reconhecida no subconjunto voluntariamente aderente. As limitações são complementares: cada etapa contribui o que a outra não pode entregar.

3.10 Modelos estimados e especificações

Esta seção formaliza cada estimador, expondo, para cada um, a pergunta econômica que responde, a fórmula que o operacionaliza, o pressuposto identificacional que o sustenta e a leitura do coeficiente que produz. Os estimadores não são independentes: cada um foi escolhido para aliviar uma limitação do anterior, e a exposição segue essa ordem.

3.10.1 Estimadores de escore de propensão

O primeiro grupo reúne os cinco estimadores baseados no escore de propensão (Rosenbaum; Rubin, 1983), que identificam o diferencial institucional sob seleção em observáveis: o logit, o IPW, o AIPW, o *entropy balancing* e o PSM.

⁴² Autosseleção pelos ganhos esperados: como a adesão ao PPA é voluntária, tendem a aderir os hospitais que antecipam maior benefício/efeito, o que enviesará o ATT estimado para os aderentes e limita a validade externa.

O ponto de partida é a regressão logística binária, que estima diretamente a razão de chance de cesárea associada à natureza jurídica do hospital, depois de controlar pelo perfil clínico e sociodemográfico observável. O modelo postula

$$\Pr(Y_i = 1 \mid \mathbf{x}_i) = \frac{1}{1 + \exp(-\mathbf{x}_i^\top \boldsymbol{\beta})}, \quad (2)$$

em que $Y_i \in \{0, 1\}$ indica se o i -ésimo nascimento foi cesárea e $\mathbf{x}_i$ é o vetor de covariáveis, que inclui o tratamento (duas *dummies* de natureza jurídica privada, com o público como referência), os controles e os efeitos fixos; o coeficiente β_{priv} traduz-se na razão de chance $\exp(\beta_{\text{priv}})$, a forma usual da literatura epidemiológica de comparar grupos.

São estimadas três especificações monotonicamente cumulativas: a Especificação 1 (clínico mínimo) inclui o tratamento, faixa etária materna em seis classes, grupo de Robson em duas classes, gestação múltipla e efeitos fixos de ano e de UF; a Especificação 2 acrescenta cinco blocos maternos (raça e cor, escolaridade em cinco classes, classificação de Kotelchuck do pré-natal, apresentação fetal e prematuridade); e a Especificação 3 acrescenta cinco blocos de oferta (perfil de atendimento, indicadores de centro obstétrico e unidade neonatal, totais de salas obstétricas e de leitos clínicos e cirúrgicos, e *dummy* de parto fora do município de residência). O fato de os blocos entrarem sempre na mesma ordem crescente permite ler a redução do coeficiente entre especificações como o quanto cada bloco *absorve* do diferencial bruto.

Ler $\exp(\beta_{\text{priv}})$ como razão de chance ajustada populacional exigiria que a forma logística estivesse corretamente especificada e que nenhum confundidor relevante restasse omitido após os controles; como ambos os pressupostos são fortes, a leitura adequada é a de *associação ajustada sob seleção em observáveis*, não de efeito causal direto. Os erros-padrão são clusterizados por estabelecimento pelo estimador *sandwich* clássico (Liang; Zeger, 1986),

$$\widehat{\text{Var}}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} \mathbf{B} (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1}, \quad (3)$$

com $\mathbf{W} = \text{diag}(\hat{p}_i(1 - \hat{p}_i))$ e $\mathbf{B} = \sum_g \mathbf{s}_g \mathbf{s}_g^\top$, em que $\mathbf{s}_g = \sum_{i \in g} (Y_i - \hat{p}_i) \mathbf{x}_i$ é a soma dos escores dentro do hospital g ;⁴³ a clusterização absorve a correlação entre partos do mesmo estabelecimento, plausível porque a decisão clínica se dá dentro de uma cultura institucional comum. Se $\exp(\beta_{\text{priv}}) > 1$, gestantes privadas têm *odds* maior de cesárea que gestantes públicas comparáveis; reportado junto à razão de chances (*odds ratio*, OR), o efeito marginal médio (*average marginal effect*, AME) traduz o coeficiente para pontos percentuais, escala mais palpável para o leitor.

⁴³ Na execução aplica-se a correção de pequena amostra no estimador de variância por *cluster* (o fator $\frac{G}{G-1} \cdot \frac{N-1}{N-k}$ que escala o estimador sanduíche, em que G é o número de *clusters*, N o de observações e k o de parâmetros).

O logit responde à pergunta combinando tratamento e desfecho num único modelo paramétrico; o *inverse probability weighting* (IPW), na tradição de Horvitz e Thompson (1952), persegue o mesmo alvo, a diferença esperada na probabilidade de cesárea entre privado e público sob alocação *como se* ao acaso, mas por um caminho independente, e é essa independência que o torna valioso. Em vez de modelar o desfecho, o IPW estima apenas a probabilidade de uma gestante com perfil X ser atendida no privado e usa esse escore como peso, reponderando a amostra de modo que a distribuição das covariáveis no privado e no público se iguale à populacional; a diferença das médias de Y na amostra reponderada estima então o efeito médio do tratamento (ATE),

$$\tau_{ATE} = \mathbb{E}[Y(1) - Y(0)], \quad (4)$$

em que $Y(t)$ denota o desfecho potencial sob tratamento $t \in \{0, 1\}$. Sob *ignorabilidade*, $(Y(0), Y(1)) \perp T \mid X$, e *positividade*, $0 < \Pr(T = 1 \mid X) < 1$, o ATE é estimável pela diferença ponderada de médias,

$$\hat{\tau}_{ATE} = \frac{1}{n} \sum_{i=1}^n \left[\frac{T_i Y_i}{\hat{e}(X_i)} - \frac{(1 - T_i) Y_i}{1 - \hat{e}(X_i)} \right], \quad (5)$$

com $\hat{e}(X_i) = \widehat{\Pr}(T_i = 1 \mid X_i)$ o *propensity score*; o ATT segue de especificação análoga restrita aos tratados. Nem ignorabilidade nem positividade são testáveis, mas o desenho permite três verificações empíricas: o balanço das covariáveis pós-ponderação pela SMD,⁴⁴ o histograma do escore por grupo como diagnóstico de positividade prática (reportado no Apêndice B), e a restrição à região de suporte comum por *trimming* entre os percentis 0,01 e 0,99.

Para o IPW o tratamento é simplificado a binário privado vs. público, com o escore estimado por logit sobre as covariáveis maternas e geográficas da Especificação 2, omitindo as variáveis de oferta que bloqueariam o canal; a inferência usa *bootstrap* clusterizado por hospital.⁴⁵

O ATE assim obtido, em pontos percentuais, é a diferença esperada na probabilidade de cesárea entre privado e público na escala da gestante: magnitude positiva aponta indução na direção da cesárea, e magnitude próxima de zero indicaria que a diferença bruta é integralmente composição observável. Como logit e IPW erram por vias distintas (um pela forma funcional do desfecho, o outro pela do tratamento), a convergência entre os dois é evidência conjunta robusta, e a divergência seria diagnóstico.

⁴⁴ Diferença padronizada de médias (*standardized mean difference*, SMD): diferença de médias entre tratados e controles dividida pelo desvio-padrão combinado, expressa em unidades de desvio-padrão para tornar o desbalanço comparável entre covariáveis.

⁴⁵ Duzentas réplicas *bootstrap* clusterizadas por hospital, com reamostragem de hospitais com reposição, preservando a estrutura de *cluster* no nível do estabelecimento.

O IPW só é consistente se o modelo do escore $\hat{e}(X)$ estiver corretamente especificado; quando esse modelo erra, a estimativa fica viesada numa direção difícil de antecipar. O *augmented IPW* (AIPW) de Robins, Rotnitzky e Zhao (1994) blinda essa fragilidade combinando o modelo do tratamento com regressões separadas do desfecho em cada grupo, $\mu_t(X) = \mathbb{E}[Y \mid T = t, X]$ para $t \in \{0, 1\}$, num único estimador,

$$\hat{\tau}_{\text{AIPW}} = \frac{1}{n} \sum_{i=1}^n \left\{ \hat{\mu}_1(X_i) - \hat{\mu}_0(X_i) + \frac{T_i [Y_i - \hat{\mu}_1(X_i)]}{\hat{e}(X_i)} - \frac{(1 - T_i) [Y_i - \hat{\mu}_0(X_i)]}{1 - \hat{e}(X_i)} \right\}. \quad (6)$$

O primeiro termo é o efeito imputado pelas regressões de desfecho, e os dois seguintes são correções de ponderação que zeram em expectativa quando $\hat{e}$ está correto. Daí a dupla robustez: basta que um dos dois modelos esteja bem especificado para a estimativa ser consistente, pois se $\hat{e}$ erra o primeiro termo carrega a estimativa, e se $\hat{\mu}_t$ erra as correções compensam. É a etapa que mais exige do conjunto e, por isso, a canônica.

Ignorabilidade e positividade do IPW permanecem; a dupla robustez opera dentro desse cenário, apenas relaxando a exigência de especificação correta para um dos dois modelos. A inferência canônica é o estimador analítico de variância pela função de influência cluster-robusta (Lunceford-Davidian, 2004 (Lunceford; Davidian, 2004)), com o *bootstrap* clusterizado idêntico ao do IPW reportado como *cross-check*.⁴⁶

A convergência numérica entre IPW e AIPW indica que o modelo do escore não está mal especificado a ponto de afetar a estimativa; a divergência apontaria a direção do viés.

Tanto o IPW quanto o AIPW ainda dependem de uma forma funcional para o escore, um logit sobre covariáveis e interações; se essa forma erra, o balanço pós-ponderação fica apenas aproximado. O *entropy balancing* de Hainmueller (2012) remove essa dependência ao dispensar o modelo paramétrico do escore e resolver, em seu lugar, um problema de otimização que impõe balanço *exato* dos primeiros momentos das covariáveis entre tratados e controles. Os pesos w_i minimizam a entropia relativa entre a distribuição ponderada e a uniforme dos controles, sujeitos às restrições de balanço,

$$\min_{w_i} \sum_{i \in C} w_i \log \left(\frac{w_i}{q_i} \right) \quad \text{s. a.} \quad \sum_{i \in C} w_i h_k(X_i) = \frac{1}{n_T} \sum_{i \in T} h_k(X_i) \quad \forall k, \quad (7)$$

em que C e T são os controles e os tratados, $h_k(\cdot)$ o k -ésimo momento a balancear e q_i os pesos iniciais; a solução tem forma fechada $w_i = q_i \exp(\lambda^\top \mathbf{x}_i)$, com os multiplicadores λ obtidos pelo dual.⁴⁷ A ignorabilidade permanece; o que se descarta é o pressuposto de forma funcional, agora

⁴⁶ A não reestimação dos modelos auxiliares por réplica subestima parte da variância dos estágios; por isso o IF cluster-robusto é a inferência canônica e o *bootstrap* é *cross-check*.

⁴⁷ Otimização do dual via L-BFGS-B; semente fixada para reprodução.

dispensável porque o balanço dos momentos é exato por construção, não aproximado. Se IPW, AIPW e *entropy balancing* convergem, fica claro que a forma logística do escore não era o gargalo.

As três etapas anteriores reponderam a amostra inteira; o *propensity score matching* (PSM), revisto por Stuart (2010), fecha o conjunto pela via complementar do emparelhamento, sem impor forma funcional alguma ao desfecho no estágio final. Para cada gestante tratada seleciona-se a controle de escore mais próximo, descartando-se quem fica sem par, e a estimativa é a média das diferenças dentro dos pares, o efeito médio do tratamento sobre os tratados,

$$\hat{\tau}_{ATT} = \frac{1}{n_T} \sum_{i: T_i=1} (Y_i - Y_{j(i)}), \quad (8)$$

em que $j(i)$ indexa a controle emparelhada à tratada i . O escore $\hat{e}(X)$ é estimado como no IPW, e o emparelhamento é vizinho mais próximo, 1:1, sem reposição, dentro de um *caliper* de 0,2 desvio-padrão na escala do escore.⁴⁸

Ignorabilidade e positividade permanecem, e o *caliper* impõe suporte comum por construção, ao custo de restringir a inferência ao subconjunto efetivamente pareado e, portanto, a validade externa aos tratados com par disponível. É a leitura mais próxima de um experimento pareado e a menos dependente de extrapolação fora do suporte. Sua convergência com IPW, AIPW e *entropy balancing*, com o balanço pós-pareamento aferido pela SMD como nas demais etapas, sela o argumento: o achado observacional não repousa sobre a escolha de um único método de ajuste pelo escore.

O emparelhamento 1:1 sem reposição troca eficiência por comparabilidade, pois usa cada controle uma única vez e descarta tratadas sem par; como o suporte comum é amplo (Apêndice B), essa perda é pequena e quase todas as tratadas encontram par dentro do *caliper*, de modo que a amostra pareada permanece grande.

3.10.2 Estruturas de painel e efeito aleatório por hospital

Todos os estimadores baseados em escore de propensão identificam o efeito sob seleção em observáveis e, por isso, compartilham a mesma vulnerabilidade: se houver heterogeneidade institucional não observada, como cultura clínica ou perfil de credenciamento, correlacionada ao mesmo tempo com a natureza jurídica e com a taxa de cesárea, todos ficam viesados. O painel hospital-ano com efeitos fixos (Wooldridge, 2010) contorna esse limite comum ao perguntar o

⁴⁸ Implementação em R via *MatchIt* (Ho *et al.*, 2011); controles fora do *caliper* são descartados e a inferência recai sobre os pares formados.

que sobra do coeficiente da natureza jurídica quando se absorve toda a heterogeneidade estável no tempo no nível do hospital. A unidade observacional passa a ser o par hospital-ano, ponderado pelo número de partos. O Modelo A, *pooled*, com efeito fixo de ano e cluster por hospital,

$$\bar{Y}_{ht} = \alpha + \beta^\top D_{ht} + \mathbf{X}_{ht}^\top \gamma + \lambda_t + u_{ht}, \quad (9)$$

com $\bar{Y}_{ht}$ a taxa de cesárea no hospital h no ano t , D_{ht} as *dummies* de natureza jurídica, $\mathbf{X}_{ht}$ as covariáveis agregadas e λ_t o efeito fixo de ano, recupera o diferencial *between* (entre hospitais). O Modelo B acrescenta o efeito fixo de hospital,

$$\bar{Y}_{ht} = \alpha_h + \beta^\top D_{ht} + \mathbf{X}_{ht}^\top \gamma + \lambda_t + u_{ht}, \quad (10)$$

em que α_h absorve toda a heterogeneidade estável não observada do estabelecimento, de modo que em β sobra apenas o que é identificado pela variação *within-hospital*, ou seja, pelos hospitais que mudaram de natureza jurídica no período. Essa leitura *within* identifica o efeito da natureza jurídica desde que as transições não venham sistematicamente acompanhadas de outras mudanças (de gestão, de equipe, de carteira contratual) que também movam a taxa; o pressuposto não é diretamente testável, mas o *event study* sobre as transições (seção 4, §4.4.3) inspeciona as pré-tendências como diagnóstico parcial.

O contraste entre os dois modelos é o que informa: se β_A e β_B ficam próximos, o diferencial é robusto à heterogeneidade institucional não observada; se β_B for muito menor, parte do diferencial *between* reflete essa heterogeneidade estável.

Em qualquer caso é preciso calcular o mínimo efeito detectável (MDE) do desenho *within*, sem o qual não se distingue “efeito ausente” de “efeito não detectado”. Para um teste bilateral,

$$\text{MDE} = (t_{\alpha/2} + t_{1-\beta}) \cdot \widehat{\text{SE}}(\hat{\beta}), \quad (11)$$

em que $t_{\alpha/2} = 1,96$ (para $\alpha = 0,05$ bilateral), $t_{1-\beta} = 0,84$ (para poder de 80%) e $\widehat{\text{SE}}(\hat{\beta})$ é o erro-padrão clusterizado do coeficiente *within*, somando os dois valores críticos a 2,80. O MDE é o menor efeito que o desenho consegue detectar como significativo dado o N , a variância residual e o poder; quando o MDE excede o efeito plausível de interesse, não rejeitar a nula não equivale a evidência de ausência de efeito, apenas à impotência do desenho para detectá-lo.

Os estimadores anteriores tratam o hospital como dimensão de *cluster* para a inferência; o modelo linear generalizado misto (GLMM) de Bates *et al.* (2015) o trata como *unidade de análise*, perguntando quanto da variação residual da probabilidade de cesárea, depois do controle por perfil materno e clínico, reside *entre* hospitais e não *dentro* deles. O modelo estima um intercepto

aleatório $u_h \sim \mathcal{N}(0, \sigma_u^2)$ que capta o quanto cada hospital se desvia da média populacional, dado o perfil das gestantes que atende,

$$\Pr(Y_{ih} = 1 \mid \mathbf{x}_{ih}, u_h) = \frac{1}{1 + \exp(-\mathbf{x}_{ih}^\top \boldsymbol{\beta} - u_h)}, \quad u_h \sim \mathcal{N}(0, \sigma_u^2), \quad (12)$$

e a razão $\sigma_u^2/(\sigma_u^2 + \pi^2/3)$ é o coeficiente de correlação intraclasse (ICC) na escala latente, que mede a fração da variação latente que reside entre hospitais. O modelo pressupõe independência condicional entre partos do mesmo hospital dado u_h e normalidade do efeito aleatório. Um ICC entre 0,05 e 0,20 é a faixa tipicamente reportada em estudos multinível de ciências sociais e de saúde (Snijders; Bosker, 2012); valores acima dela indicam que o estabelecimento concentra fração substancial da variação após o controle pelas covariáveis, reforçando o achado da decomposição preditiva de que o hospital é a dimensão dominante.

3.10.3 Sensibilidade a inobservados e decomposição preditiva

Todo o conjunto observacional carrega o mesmo pressuposto inverificável, o de que não há confundidor inobservado relevante; como esse pressuposto não é testável, dois exercícios medem quão hostil ao achado um confundidor precisaria ser para revertê-lo. O primeiro, de Oster (2019), parte da diferença entre o coeficiente sem controles ($\hat{\beta}_0$) e com controles ($\hat{\beta}_1$) e da diferença correspondente em R^2 : sob a hipótese de que a razão de seleção entre inobservados e observáveis é δ , o coeficiente ajustado por viés de inobservados é

$$\hat{\beta}^* = \hat{\beta}_1 - \delta (\hat{\beta}_0 - \hat{\beta}_1) \frac{R_{\max} - \hat{R}_1}{\hat{R}_1 - \hat{R}_0}. \quad (13)$$

O valor δ^* que faz $\hat{\beta}^* = 0$ resume a robustez: o limiar convencional $\delta^* > 1$ indica que o inobservado precisaria ser mais informativo que todo o conjunto observável para anular o efeito. O segundo exercício, o E-value de VanderWeele e Ding (2017), parte de um risco relativo RR ajustado e calcula o menor risco relativo E que um confundidor precisaria ter *simultaneamente* com tratamento e desfecho para reduzir o RR a 1,

$$E = \text{RR} + \sqrt{\text{RR} (\text{RR} - 1)}, \quad (14)$$

reportado tanto para o ponto estimado quanto para a borda inferior do IC95%. Oster pede proporcionalidade entre a razão de seleção e a de variância explicada, e o E-value pede confundimento multiplicativo; ambos são aproximações, e nenhum substitui um quase-experimento. Assim, $\delta^* < 1$ e E baixo sustentam o achado contra confundimento moderado mas o deixam vulnerável

a confundimento de magnitude moderada-alta, ao passo que $\delta^* > 1$ e E alto o qualificam como robusto a confundimento forte; em qualquer caso, é *sensibilidade*, não prova.

Encerrada a leitura de sensibilidade do bloco observacional, resta formalizar o instrumento das duas perguntas descritivas.

A previsão da via de parto e a decomposição entre oferta e demanda são preditivas, não causais: a primeira mede quanto da variação na via de parto é previsível com o conjunto saturado de variáveis administrativas observáveis, e a segunda, quanto cada bloco conceitual (clínico, sociodemográfico, oferta hospitalar, contexto) contribui marginalmente para essa previsibilidade.

O instrumento é o *gradient boosting* de árvores, escolhido por quatro razões: dispensa especificação *a priori* de não linearidades e interações, que o algoritmo aprende sozinho; escala bem para milhões de observações; admite cálculo eficiente do SHAP (Lundberg; Lee, 2017), que decompõe cada predição em contribuições aditivas por variável; e é o algoritmo de referência em *Machine Learning* aplicado a desfechos clínicos binários. A escolha não é causal: alta importância SHAP de uma variável não implica que a variável cause o desfecho.

A decomposição empilha quatro blocos cumulativos $B_0 \subset B_1 \subset B_2 \subset B_3$, e o ganho marginal em qualidade preditiva ao passar de B_{k-1} para B_k ,

$$\Delta \text{AUROC}_k = \text{AUROC}(B_k) - \text{AUROC}(B_{k-1}), \quad (15)$$

é a informação preditiva carregada pelo bloco adicionado, condicional ao anterior. O Bloco 0 (clínico mínimo) reúne faixa etária materna, grupo de Robson, apresentação fetal, gestação múltipla e prematuridade; o Bloco 1 acrescenta o sociodemográfico (raça e cor, escolaridade, Kotelchuck); o Bloco 2, a oferta hospitalar (natureza jurídica tripartite, perfil de atendimento, indicadores de centro obstétrico e neonatal, totais de salas e leitos); e o Bloco 3, tempo e espaço (ano, UF e *dummy* de parto fora do município).

A cadeia reflete a hierarquia entre demanda (Blocos 0 e 1), oferta (Bloco 2) e contexto (Bloco 3). A leitura como “peso preditivo” do bloco depende dessa ordem: trocá-la não altera o ΔAUROC conjunto, mas redistribui a atribuição entre os blocos; como a ordem adotada privilegia a demanda mais próxima da gestante e empurra as variáveis institucionais para as posições marginais, a comparação $\Delta_2 \gg \Delta_1$ é conservadora.

A métrica final vem de *holdout* temporal estrito (treino 2014–2019, validação 2020 para *early stopping*, teste 2021–2022), mais conservador que a validação cruzada aleatória por captar o desempenho nas condições de uso real, com o modelo treinado numa janela e aplicado a uma janela posterior.

3.10.4 Médicas como pacientes: teste de assimetria informacional

O conjunto principal de estimadores identifica um diferencial entre privado e público compatível com SID, mas não testa diretamente a assimetria informacional postulada por Arrow (1963) e Evans (1974). Johnson e Rehavi (2016), em microdados americanos, propõe um teste que compara gestantes médicas (informação simétrica) com controles pareados de mesmo perfil socioeconômico: a redução da taxa de cesárea entre as médicas é lida como evidência de que parte do diferencial decorre da assimetria, pois quem sabe o que se examina escolhe diferente. A pergunta aqui é a mesma, transposta ao caso brasileiro, controlando por hospital e perfil socioeconômico.

Sobre a coluna de ocupação da mãe do SINASC, constrói-se uma escala ordinal de informação obstétrica em cinco níveis: **L0** (controle, escolaridade Superior em ocupação não saúde), **L1** (demais profissionais de saúde superior: dentistas, fisio, nutri, fono, psico), **L2** (enfermeiras superior e técnicas de enfermagem), **L3** (outras médicas, família da Classificação Brasileira de Ocupações (CBO)⁴⁹ 2231 sem ginecologia-obstetrícia) e **L4** (ginecologistas-obstetras, CBO 223132, informação máxima). A escala opera sobre a série 223 da CBO 2002, pois a série 225, também de médicos, foi verificada na robustez e não aparece no registro de ocupação do SINASC na janela.

Estimam-se quatro modelos: um logit *baseline* com *cluster* por hospital, que capta o diferencial bruto entre níveis; um modelo de probabilidade linear (*linear probability model*, LPM) com efeito fixo de hospital, que isola a barganha intra-hospital; e dois logits *baseline* restritos ao Lucrativo e ao privado sem fins lucrativos, que separam o efeito por canal de financiamento.

A escala captura informação obstétrica relevante e a restrição a escolaridade Superior controla aproximadamente o canal renda, embora não a autosseleção residual da gestante por hospital. Um coeficiente negativo em L3 e L4 é compatível com o canal informacional, e o contraste entre o logit *baseline* e o LPM com efeito fixo distingue dois mecanismos: se o efeito persiste no *within*, o canal opera por barganha intra-hospital; se desaparece, opera por escolha de hospital, com gestantes informadas selecionando estabelecimentos de perfil clínico mais conservador.

A leitura informacional é submetida a cinco verificações independentes, todas sobre extração própria do SINASC–CNES integrado (2014–2022) e consolidadas no Apêndice E.

⁴⁹ Classificação Brasileira de Ocupações (CBO): sistema oficial que codifica e descreve as ocupações do mercado de trabalho brasileiro; aqui usada para identificar, pela ocupação da mãe, gestantes da área médica/saúde; fonte oficial: Classificação Brasileira de Ocupações, Portaria MTE nº 397/2002.

Primeira, *reprodução e codificação*: os coeficientes são reestimados a partir de extração direta da base SINASC–CNES integrada, com reconstrução da classificação CBO 2002 (séries 223 e 225) para garantir que nenhuma médica escape ao grupo de controle. Segunda, *placebo setorial*: o mesmo logit é estimado no setor público, onde não há incentivo financeiro à cesárea; sob a hipótese informacional, a proteção deve ser específica do setor privado, e um efeito de mesmo sinal no público invalidaria a leitura.

Terceira, *controle de renda*: a comparação é restrita a ocupações não saúde de renda alta (grandes grupos 1 e 2 da CBO), separando o canal informacional do canal renda. Quarta, *estimador alternativo e sensibilidade*: a diferença de risco é reestimada por IPW e por AIPW duplamente robusto, e a fragilidade a inobservados é medida por E-value e por *bounds* de Oster, com a mesma leitura das etapas de IPW e AIPW (§3.10.1) e de sensibilidade a inobservados (§3.10.3).

Quinta, *teste direto do mecanismo*: a taxa de cesárea do hospital de destino, calculada apenas entre as não médicas, é usada como variável dependente, distinguindo seleção de hospital (*ex ante*) de barganha intra-hospitalar (*ex post*). Os achados das cinco frentes estão na Tabela E1.

3.11 Identificação por choque regulatório: desenho DiD escalonado

A identificação observacional do diferencial institucional opera sob seleção em observáveis. Esta seção formaliza a quarta família de instrumentos, mobilizada para o desenho do Parto Adequado: inferência causal quase-experimental sobre o efeito do PPA da ANS, com fonte de variação plausivelmente exógena dada pela adesão escalonada de hospitais em três fases entre 2015 e 2022.

A construção do desenho parte da lista oficial de hospitais aderentes publicada no Anexo A do Relatório consolidado da ANS (Agência Nacional de Saúde Suplementar (ANS), 2023), que sistematiza 124 estabelecimentos com identificação por CNES, razão social, natureza jurídica e matriz Sim/Não para três janelas: Fase 1 (2015–2016, 35 hospitais), Fase 2 (2017–2019, 108) e Ciclo Intensivo 2 (2020–2022, 27).

Do total, **121 têm cobertura em SINASC–CNES no recorte clínico Robson 1+3** e compõem o universo efetivo do painel DiD (3 são excluídos por ausência de partos no recorte). O *cohort de início* é o ano da primeira fase ativa: 2015 para Fase 1, 2017 para Fase 2 sem passagem pela Fase 1, 2020 para apenas o Ciclo Intensivo 2. A adesão é tratada como permanente (*absorbing*), coerente com a estrutura do programa.

O contrafactual replica os critérios oficiais da ANS aplicados em 2012 (privados ≥ 500

partos com $\geq 75\%$ de cesárea; públicos > 1.000 partos com $> 60\%$) e produz 315 hospitais elegíveis como controle (17 públicos, 110 sem fins lucrativos, 188 lucrativos). O painel final reúne 436 estabelecimentos sobre o recorte clínico Robson 1+3, 2014–2022.

A pergunta do desenho do Parto Adequado é se o PPA reduziu a taxa de cesárea nos hospitais aderentes frente à trajetória contrafactual dos elegíveis não aderentes. O estimador de partida natural é o TWFE clássico, simples e de leitura direta, em três especificações: a base (que replica Knau, Triaca e Santos (2023)), uma com heterogeneidade por natureza jurídica tripartite (três efeitos por categoria) e outra por *cohort* de início (três efeitos por *cohort*).

Acontece que o TWFE aplicado a tratamento escalonado heterogêneo é vulnerável ao viés documentado por Goodman-Bacon (2021): *cohorts* tratados mais cedo viram comparação para os tratados mais tarde, e essa comparação (usar como contrafactual um hospital que já está sob efeito do programa) pode reverter o sinal da estimativa.

Por isso o estimador primário é o *interaction-weighted* (IW) de Sun e Abraham (2021), que estima coeficientes saturados por par (*cohort*, tempo desde o tratamento) e os agrega ponderando pelo tamanho relativo dos *cohorts*,

$$y_{h,t} = \alpha_h + \lambda_t + \sum_{c \in \{2015, 2017, 2020\}} \sum_{\ell \neq -1} \delta_{c,\ell} \cdot \mathbf{1}\{C_h = c, t - C_h = \ell\} + \varepsilon_{h,t}, \quad (16)$$

com $\ell = -1$ como categoria de referência; o ATT médio em cada *event-time* ℓ é $\text{ATT}(\ell) = \sum_c \omega_{c,\ell} \delta_{c,\ell}$, com pesos $\omega_{c,\ell}$ proporcionais ao número de hospitais do *cohort* c observados em ℓ , e o ATT médio pós-tratamento é a média simples sobre $\ell \geq 0$. A identificação pede tendências paralelas no contrafactual (na calibração de Sun-Abraham, paralelas *cohort-by-cohort*), ausência de antecipação e tratamento absorvente. Um $\text{ATT}(\ell)$ negativo indica redução da taxa de cesárea atribuível ao programa ℓ anos após a adesão, dentro do escopo dos aderentes.

Três estimadores acompanham o IW de Sun-Abraham, cada um corrigindo o viés escalonado por uma via distinta, de modo que a leitura conjunta seja diagnóstica.

O de Callaway e Sant’Anna (2021) estima $\text{ATT}(g, t)$ como diferenças em diferenças 2x2 separadas para cada par (*cohort* de início g , tempo t), usando os nunca-tratados como contrafactual: se o estimador de Callaway-Sant’Anna e o de Sun-Abraham convergem, o viés escalonado não é a fonte do achado; se divergem, o caminho de agregação importa.

A decomposição de Goodman-Bacon (2021), aplicada ao TWFE base, quantifica diretamente o peso das comparações implícitas entre *cohorts*, em particular a fração problemática *later vs. already-treated*, cujo peso alto sinalizaria que a estimativa clássica é dominada pela comparação proibida. O estimador de Chaisemartin e D’Haultfœuille (2020), focado nos *switchers*

e no efeito imediato, entra como linha adicional na tabela de robustez para mostrar que o achado sobrevive a outra agregação.

Por fim, a abordagem *Relative Magnitudes* de Rambachan e Roth (2023) (Honest DiD) quantifica o quanto a hipótese de paralelas precisaria ser violada para anular o achado: para cada $M \geq 0$, define o conjunto de violações pós-tratamento que não excedem M vezes a maior violação pré-tratamento e calcula o IC95% robusto do ATT sob essa restrição, sendo o *breakdown* M o menor multiplicador para o qual o IC inclui zero. Reporta-se em duas leituras, agregada (IW) e calibrada *cohort-by-cohort*, endereçando diretamente a fragilidade declarada por Knau, Triaca e Santos (2023) quanto à não validação de paralelas em três dos quatro desfechos da Fase 1.

Erros-padrão no desenho do Parto Adequado são obtidos por *cluster bootstrap* por hospital com 1.000 réplicas. A validação de paralelas é refinada por teste de Wald conjunto restrito ao único *cohort* com pré-tratamento informativo (*cohort* 2017, três anos pré ativos), avaliando $H_0 : \delta_{c=2017, \ell=-3} = \delta_{c=2017, \ell=-2} = 0$ via estatística qui-quadrado de dois graus de liberdade calculada sobre a matriz de covariância *bootstrap* dos coeficientes.

O conjunto de testes de robustez complementa a inferência central em cinco frentes.

O **placebo de tratamento** permuta aleatoriamente o *status* entre os 315 hospitais elegíveis (estratificação por natureza jurídica tripartite e *cohort*, 500 réplicas) e calcula a distribuição empírica da estimativa sob hipótese nula de efeito zero. O **placebo de outcome**⁵⁰ aplica o mesmo desenho à taxa de cesárea em Robson 5 (cesárea anterior), grupo onde o PPA não tem mecanismo plausível de ação: efeito não nulo aqui invalidaria o desenho.

A **robustez ao critério de elegibilidade** varia o ano-base sobre {2012, 2013, 2014} e os limiares sobre três níveis cada, gerando nove cenários completos, e mostra se o achado depende criticamente de uma escolha específica de elegibilidade.

A **heterogeneidade por região** estima o efeito separadamente para Norte, Nordeste, Sudeste, Sul e Centro-Oeste; a **heterogeneidade por escolaridade da mãe** para Fundamental, Médio e Superior. As duas heterogeneidades respondem se o efeito é uniforme ou se concentra em subgrupos onde o canal causal é mais plausível.

⁵⁰ Outcome é o desfecho (variável-resposta) do modelo. No placebo de desfecho, aplica-se o mesmo desenho a um desfecho onde o programa não deveria atuar (aqui, cesárea no Robson 5); encontrar efeito sinalizaria falha do desenho, não efeito real.

4 ESTIMAÇÃO E RESULTADOS

Esta seção estima os modelos definidos na seção 3.10 e apresenta os achados. Para cada pergunta, reporta o número com seu intervalo de confiança, traduz a magnitude em termos concretos e discute o que ela implica; os detalhes operacionais (solver, hiperparâmetros, número de réplicas, sementes) ficam em nota de rodapé.

A organização segue as quatro perguntas. A primeira pergunta o quanto da decisão sobre via de parto é previsível com o que o sistema de informação registra. A segunda mede se o que carrega mais informação preditiva é o perfil da gestante (lado da demanda) ou o perfil do hospital (lado da oferta). A terceira testa se o diferencial entre hospitais privados e públicos sobrevive a controles e é compatível com a hipótese de demanda induzida pela oferta (SID). A quarta analisa a adesão escalonada de hospitais ao Projeto Parto Adequado (PPA) para estimar o efeito causal de uma intervenção institucional dentro do escopo do programa.

A calibração causal é estrita. Para a previsão e a decomposição, que se apoiam em modelos preditivos de *Machine Learning*, fala-se em previsibilidade e em decomposição preditiva, nunca em efeito ou causa. Para a análise do diferencial entre as categorias institucionais, que opera sob seleção em observáveis, fala-se em associação ajustada e em compatibilidade com SID. Para o desenho quase-experimental do Parto Adequado, com tendências paralelas validadas, fala-se em efeito causal, calibrado ao escopo do programa supracitado.

4.1 Visão geral dos achados

A análise foi conduzida sobre 25,3 milhões de partos hospitalares no Brasil entre 2014 e 2022 e produz quatro achados centrais. Primeiro, há previsibilidade moderada-alta da decisão sobre a via de parto a partir das informações administrativas que o sistema captura: o modelo saturado erra em torno de um a cada quatro pares aleatórios de cesárea vs. vaginal. Segundo, a decomposição mostra que o que mais ajuda a prever é o hospital, não a gestante, depois de fixar a clínica da gestação. Terceiro, em uma população clinicamente comparável, os hospitais privados lucrativos fazem mais cesárea que os públicos, mesmo depois de controlar pelo perfil materno e pela infraestrutura do hospital. Esse diferencial é compatível com SID. Quarto, o PPA reduz a taxa de cesárea nos hospitais aderentes em torno de 5 a 6 pontos percentuais, e a redução está concentrada justamente onde a hipótese SID prevê: no setor privado lucrativo.

4.2 A previsibilidade da decisão

A primeira pergunta é descritiva: quanto da decisão sobre via de parto é previsível a partir do que o SINASC e o CNES juntos registram sobre a gestante e sobre o hospital? Não se busca uma estimativa de efeito, e sim uma medida agregada da qualidade preditiva do conjunto de variáveis disponíveis. A resposta vem de um algoritmo de *gradient boosting* de árvores na variante XGBoost⁵¹, treinado sobre o recorte de Robson 1+3 e avaliado em *holdout* temporal estrito.⁵²

O modelo saturado atinge **AUROC de 0,7673** no *holdout*⁵³. O LightGBM, estimado em paralelo, atinge 0,7699, com a mesma ordem dos blocos por contribuição marginal. Em termos concretos, sorteado ao acaso um par formado por uma cesárea e um parto vaginal, o modelo atribui à cesárea a probabilidade maior em aproximadamente 77% das vezes, e erra cerca de uma em cada quatro tentativas. A previsibilidade é moderada-alta, mas longe de determinística.

Os 23% que o modelo erra não são ruído puro: são o que está fora do alcance do que o sistema de informação captura. A literatura sobre preditores clínicos da decisão obstétrica indica que esse residual está em quatro dimensões que o SINASC e o CNES não codificam: o peso fetal estimado e a posição final da placenta no fim da gestação⁵⁴, a preferência declarada pela gestante pela via cesárea, o plano de saúde específico da gestante e o histórico de conduta do médico responsável. A previsibilidade observável esbarra, portanto, em um limite. Esse limite é a fronteira do que análises causais podem explicar usando dados administrativos públicos, e enquadra os achados causais que vêm a seguir como descrições parciais do fenômeno.

⁵¹ *Gradient boosting*: família de algoritmos de *Machine Learning* que combina muitas árvores de decisão simples em sequência, cada uma corrigindo o erro da anterior. XGBoost é a implementação mais disseminada em economia aplicada. Hiperparâmetros, splits temporais (treino 2014–2019; validação 2020 para *early stopping*; teste 2021–2022) e configuração detalhada estão em §3.10.3.

⁵² *Holdout*: parcela dos dados separada e nunca usada no treino, reservada apenas para medir o desempenho em observações que o modelo nunca viu.

⁵³ AUROC (*area under the receiver operating characteristic*): mede o quanto o modelo separa cesáreas de partos vaginais. Vale 0,5 para um sorteio aleatório e 1,0 para um classificador perfeito. AUROC entre 0,7 e 0,8 é convencionalmente lida como discriminação moderada-alta para preditores de eventos clínicos compostos.

⁵⁴ Peso fetal estimado e posição final da placenta são informações registradas no exame ultrassonográfico do terceiro trimestre, no prontuário privado do consultório obstétrico ou no prontuário hospitalar; não são exportadas para a Declaração de Nascido Vivo (DN) que alimenta o SINASC. Da mesma forma, o CNES registra estrutura física e classificação administrativa do hospital, mas não conduta clínica individual.

4.3 O peso relativo de oferta e demanda

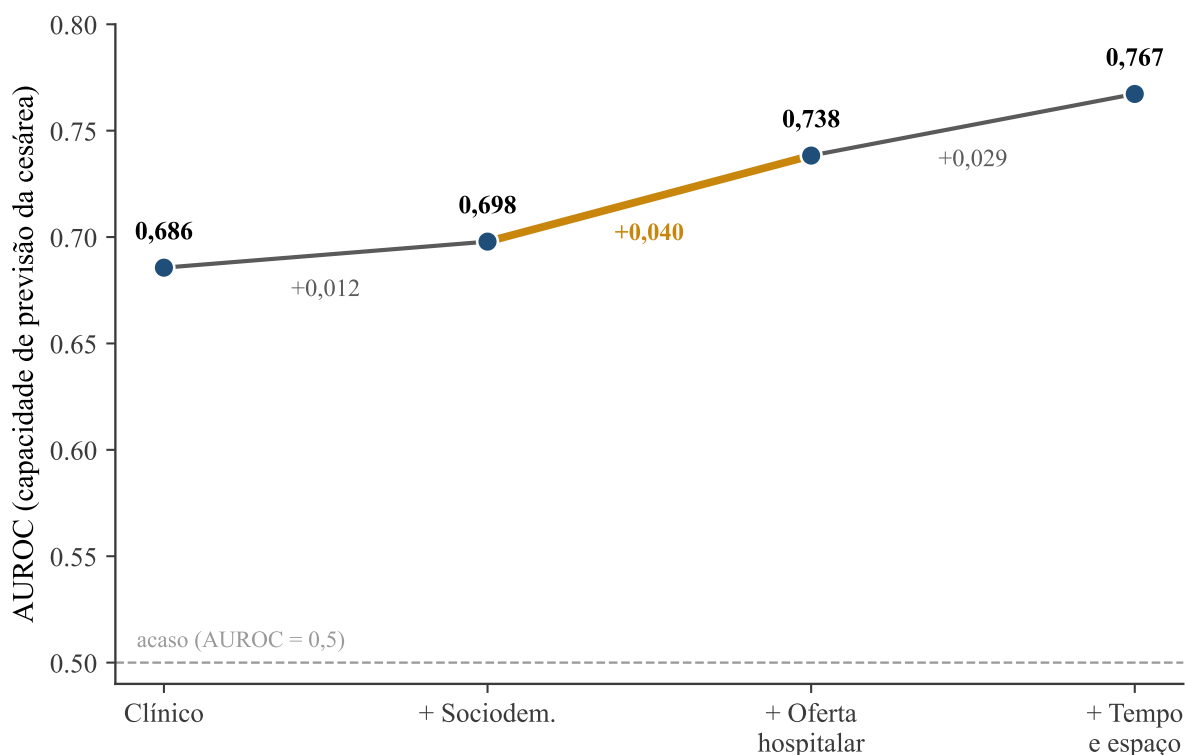
A segunda pergunta organiza a discussão sobre o que pesa mais na decisão da via de parto, a gestante ou o hospital. A resposta vem da decomposição preditiva por ablação⁵⁵ sobre o XGBoost saturado. A ideia é simples: começa-se com o bloco clínico mínimo (idade, paridade⁵⁶, classificação de Robson, gestação múltipla), adiciona-se o bloco sociodemográfico (raça e cor, escolaridade, estado civil, índice de Kotelchuck do pré-natal), depois o bloco de oferta hospitalar (natureza jurídica, perfil de atendimento, leitos, salas cirúrgicas) e, por fim, o bloco de tempo e espaço (UF, ano). Em cada passo, registra-se o ganho marginal em AUROC. A Figura 6 sintetiza esse ganho acumulado bloco a bloco, e a Tabela A14, no Apêndice A, traz os valores completos.

O ganho do Bloco 2 (oferta hospitalar) é **cerca de 3,3 vezes maior** que o ganho do Bloco 1 (sociodemográfico), partindo ambos do mesmo bloco clínico de referência. O ganho do Bloco 3 (tempo e espaço) é da mesma ordem do Bloco 2 e largamente atribuível ao efeito fixo de UF, captando padrões regionais residuais. Em outras palavras: dada a clínica da gestação, saber em qual hospital a gestante vai dar à luz é mais de três vezes mais informativo sobre a via de parto do que conhecer seu perfil sociodemográfico.

⁵⁵ Ablação (*ablation study*): remover ou acrescentar blocos de variáveis e medir o ganho ou perda preditiva de cada bloco, técnica importada do *Machine Learning*.

⁵⁶ Paridade: número de partos anteriores da gestante (nulípara = nenhum; multípara = um ou mais).

Figura 6 – Ganho de previsibilidade da via de parto (AUROC) por bloco cumulativo de variáveis. Amostra principal (Robson 1+3).



Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

A complementação por SHAP⁵⁷ confirma o padrão: entre as variáveis de maior peso preditivo predominam o núcleo clínico (Robson) e a oferta hospitalar (perfil de atendimento, leitos e salas obstétricas), à frente do perfil sociodemográfico.

A decisão sobre a via de parto, no Brasil, é mais sensível ao local onde a gestante dá à luz do que às suas características individuais, depois de controlar pelas condições clínicas da gestante. Esse diagnóstico é descritivo, não causal: dimensiona onde está a informação e prepara o terreno para a análise do diferencial entre as categorias institucionais, que pergunta se a parte da informação que reside no hospital é compatível com indução pela oferta.

4.4 O diferencial público × privado e a hipótese de demanda induzida pela oferta

O diferencial sistemático entre hospitais privados e públicos, mesmo após ajustes por perfil materno e dentro de grupos clinicamente comparáveis, é compatível com a hipótese de demanda induzida pela oferta? As subseções a seguir percorrem o conjunto de estimadores, a

⁵⁷ SHAP (*SHapley Additive exPlanations*): técnica que decompõe a contribuição de cada variável na predição de modelos não lineares como o XGBoost (Lundberg; Lee, 2017). Aqui é usado em escala descritiva, para identificar as variáveis com maior peso preditivo médio.

heterogeneidade do diferencial e a leitura conjunta sob a hipótese de demanda induzida pela oferta.

4.4.1 O diferencial após controles

O logit binário com erros-padrão clusterizados por hospital⁵⁸ estima a razão de chances (*odds ratio*, OR) de cesárea associada à natureza jurídica do hospital, em três especificações que acumulam controles: clínico mínimo; mais perfil materno e sociodemográfico; mais perfil de atendimento e infraestrutura hospitalar. A Tabela 5 consolida as ORs nas três especificações, e duas regularidades organizam a leitura.

Tabela 5 – Razão de chances (odds ratio) de cesárea por natureza jurídica, três especificações de logit, Robson 1+3, Brasil, 2014–2022.

Especificação	Tratamento (ref. Público)	OR	IC 95%
Spec. 1 (clínico mínimo)	Privado sem fins lucrativos	2,101***	[1,887; 2,340]
Spec. 1 (clínico mínimo)	Privado lucrativo	3,817***	[3,325; 4,381]
Spec. 2 (+ sociodemográfico)	Privado sem fins lucrativos	2,007***	[1,805; 2,232]
Spec. 2 (+ sociodemográfico)	Privado lucrativo	3,358***	[2,916; 3,866]
Spec. 3 (+ perfil e estrutura)	Privado sem fins lucrativos	1,140	[0,913; 1,424]
Spec. 3 (+ perfil e estrutura)	Privado lucrativo	1,537***	[1,191; 1,985]

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

A primeira é o **diferencial entre as duas categorias privadas**: em qualquer especificação, o privado lucrativo entrega OR maior que o privado sem fins lucrativos. O padrão era esperado pelo lado descritivo da seção 3; o que se observa aqui é que o diferencial entre os tipos de hospital permanece mesmo depois de igualar o perfil das gestantes: a categoria de financiamento permanece informativa mesmo entre gestantes com perfil materno comparável. A segunda é a **redução monotônica das ORs** conforme blocos de controles entram. Na Especificação 1, o privado lucrativo apresenta OR de 3,82 sobre o público; o controle por características maternas observáveis na Especificação 2 reduz para 3,36; a inclusão de perfil de atendimento e infraestrutura na Especificação 3 reduz para **1,54** (IC95% [1,20; 1,99], $p < 0,001$). O privado sem fins lucrativos segue trajetória paralela e perde significância na Especificação 3 (OR 1,12, $p = 0,29$).

Quando se comparam duas gestantes com o mesmo perfil materno, em hospitais com a mesma estrutura física e perfil de atendimento, a chance de a cesárea ocorrer é cerca de 1,54 vezes maior no privado lucrativo do que no público. Em escala de pontos percentuais (*average marginal*

⁵⁸ O cluster por hospital acomoda a correlação entre partos da mesma instituição: sem esse ajuste, a variância do coeficiente seria subestimada. Implementação *sandwich* com correção de pequena amostra no estimador de variância por *cluster*; formalização em §3.10.1, Equação (3).

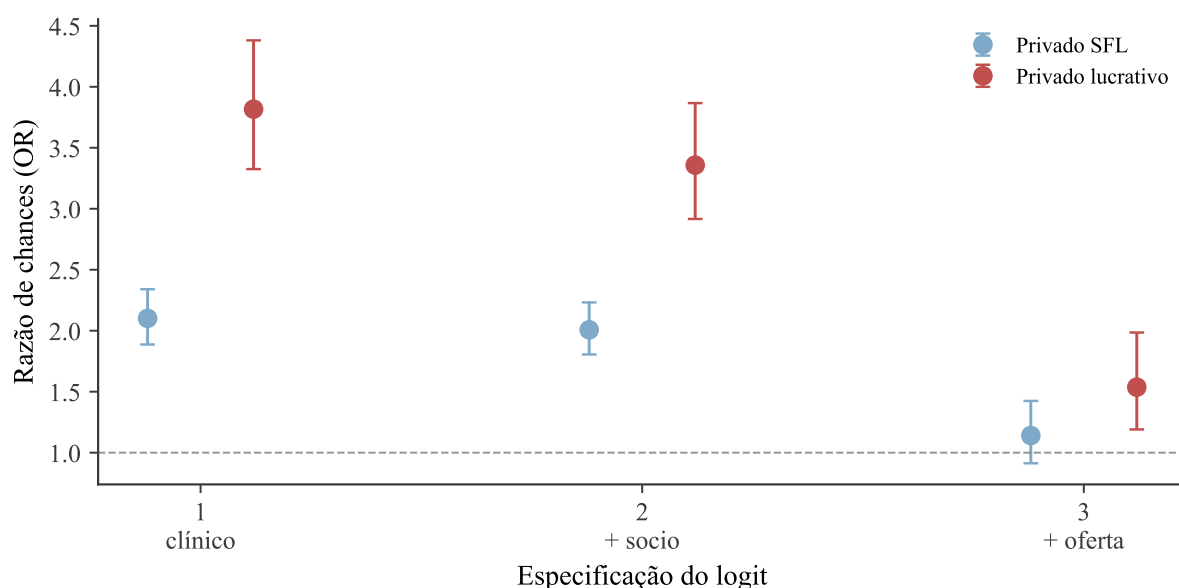
effect, AME), o diferencial é de **8,14 p.p** (IC95% [4,06; 12,23]) na amostra principal e fica estável quando se ampliam os grupos de Robson considerados (8,84 p.p em Robson 1–4; 7,37 p.p em Robson 1–10). A queda mais marcada entre as Especificações 2 e 3 é informativa por si: o controle pela oferta absorve aproximadamente metade do diferencial residual. Boa parte do que à primeira vista parecia "efeito do privado" é, na verdade, infraestrutura hospitalar e perfil de atendimento: leitos, salas, vínculo com o SUS. O que sobra é o que não se apaga.⁵⁹

O coeficiente residual sobrevive ao controle por todas as variáveis observáveis disponíveis, e é estável quando se reestima o modelo em amostras de Robson diferentes (1,54 em Robson 1+3, 1,59 em Robson 1–4, 1,55 em Robson 1–10). A coincidência numérica entre amostras independentes é um diagnóstico relevante: se o efeito fosse particular dos Grupos 1 e 3 frente aos demais, oscilaria com a definição da amostra. O resultado mantém-se estável entre as especificações. O diferencial é uma propriedade sistemática da categoria de financiamento, condicional ao perfil materno e à infraestrutura. É exatamente o tipo de regularidade que a hipótese SID prevê.

A Figura 7 traduz graficamente a queda das razões de chance entre as três especificações, com o privado lucrativo sistematicamente acima do privado sem fins lucrativos.

⁵⁹ Parte da infraestrutura controlada na Especificação 3 (salas, leitos) pode ser, em si mesma, consequência da orientação cirúrgica do hospital, e não apenas covariável exógena; nesse caso a Especificação 3 sobreajusta e a razão de chance de 1,54 deve ser lida como piso conservador do diferencial. O logit hierárquico marginal (OR 3,03, §4.4.1), que absorve oferta e perfil no intercepto aleatório, delimita o limite: o componente que opera por composição hospitalar situa-se no intervalo [1,54; 3,03].

Figura 7 – Razão de chance de cesárea por natureza jurídica do hospital, em três especificações. Amostra principal (Robson 1+3), 2014–2022.



Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

O *Inverse Probability Weighting* (IPW) estima o ATE (*average treatment effect*) reponderando cada observação pelo inverso da probabilidade estimada de pertencer ao grupo a que efetivamente pertence⁶⁰. O IPW é um caminho metodologicamente independente do logit para o mesmo alvo: o logit modela o desfecho dado o tratamento, o IPW modela o tratamento dadas as covariáveis e repondera. Convergência entre os dois é diagnóstico forte de que a forma funcional não está dirigindo o resultado.

A diferença bruta de proporções entre privado lucrativo e público, antes do ajuste, é de 19,06 p.p na amostra principal. Após o IPW (Tabela 6), o ATE estimado é de **15,79 p.p** (IC95% por *bootstrap* clusterizado: [13,35; 17,94]) e o ATT (sobre os tratados) é de **17,02 p.p** ([14,25; 19,41]). O IPW preserva, assim, aproximadamente **82%** da diferença bruta. Apenas cerca de 3 p.p dos 19 p.p originais são atribuíveis a diferenças observáveis de composição clínica entre os dois grupos; o restante permanece. Os dois estimadores apontam efeitos de magnitude similar sobre o mesmo alvo conceitual: na escala em que o coeficiente do logit é mais legível (uma OR de 1,54 numa prevalência de 30% característica do controle público em Robson 1+3 corresponde

⁶⁰ O método identifica o ATE sob seleção em observáveis e *common support*. Aplicação (escore por logit, *trimming* entre os percentis 0,01 e 0,99, *bootstrap* clusterizado por hospital com 200 réplicas) em §3.10.1; formalização em §3.10.1, Equação (5).

a uma diferença absoluta na ordem de 13–15 p.p), as duas estimativas se reencontram.⁶¹

A convergência entre dois estimadores que dependem de pressupostos diferentes (forma funcional do desfecho *vs.* forma funcional do tratamento) traz a robustez necessária para o que vem a seguir. A diferença entre privado lucrativo e público em populações clinicamente comparáveis é da ordem de 8 em cada 10 pontos percentuais brutos. Esse residual não é pequeno: representa a magnitude que o ajuste por composição não consegue absorver.

Tabela 6 – ATE e ATT do parto em hospital privado *vs.* público sobre a probabilidade de cesárea, estimados por inverse probability weighting (IPW), Brasil, 2014–2022.

Amostra	N	ATE (p.p) [IC 95%]	ATT (p.p) [IC 95%]	Máx. SMD pré→pós
Principal (Robson 1+3)	8.917.588	15,79 [13,35; 17,94]	17,02 [14,25; 19,41]	0,476 → 0,033
<i>Low risk</i> (Robson 1–4)	14.738.616	17,45 [15,23; 19,61]	18,86 [16,50; 21,50]	—
Total (Robson 1–10)	23.834.354	14,84 [12,98; 16,65]	15,80 [14,03; 17,77]	—

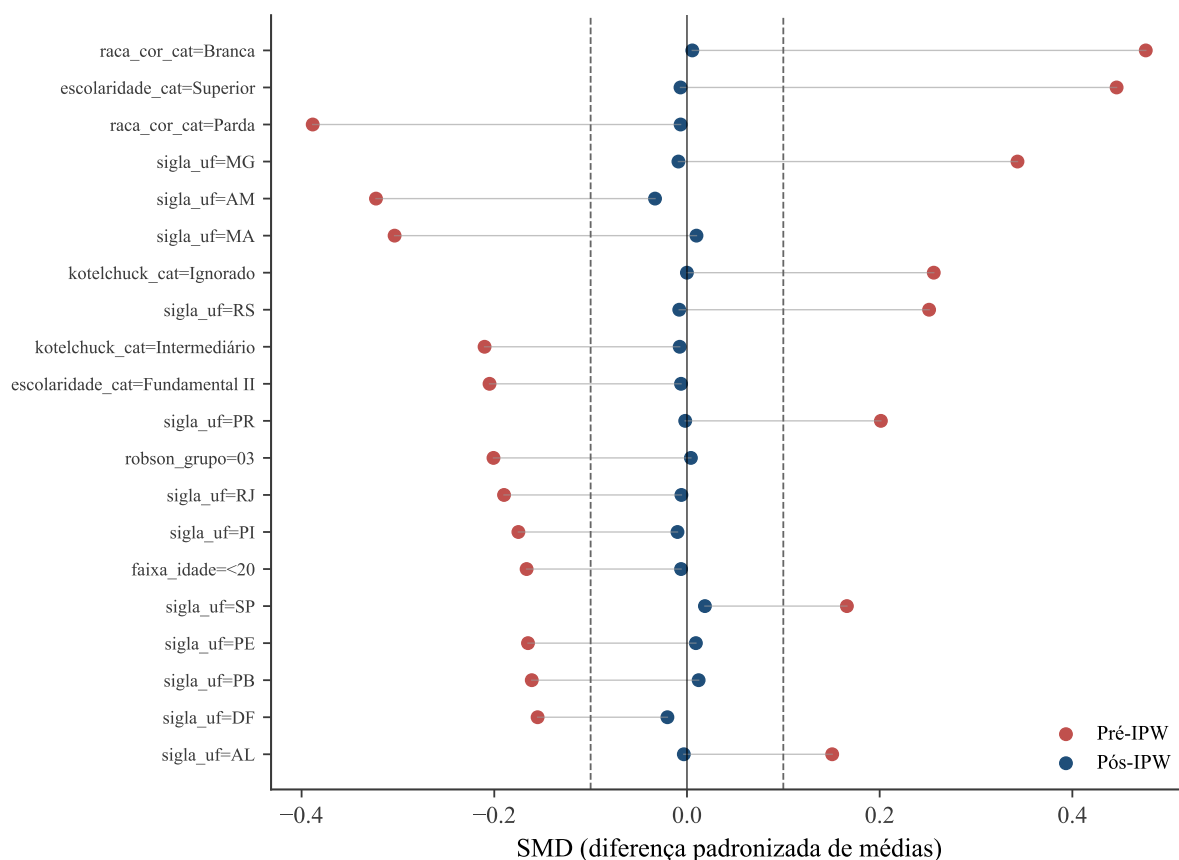
Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Nota: Pontos percentuais; IC95% por *bootstrap* clusterizado por hospital (200 réplicas).

A Figura 8 mostra que as vinte covariáveis mais desbalanceadas convergem para diferença padronizada de médias próxima de zero após a ponderação por IPW, confirmando o balanço que sustenta a leitura sob seleção em observáveis.

⁶¹ A reconciliação é aproximada e cruza estimandos distintos: o logit estima a razão de chance do *lucrativo* condicional à oferta (Especificação 3), enquanto IPW, AIPW, *entropy balancing* e PSM estimam o efeito do *privado* (lucrativo e sem fins lucrativos agregados), com a oferta omitida do escore para não bloquear o canal. A convergência de magnitude é triangulação de pressupostos, não replicação do mesmo alvo; o efeito marginal de cerca de 15,8 p.p corresponde conceitualmente à Especificação 2, sem oferta, e não à Especificação 3 saturada.

Figura 8 – Balanço das 20 covariáveis mais desbalanceadas (SMD) entre privado e público, antes e depois do IPW. Robson 1+3, 2014–2022.

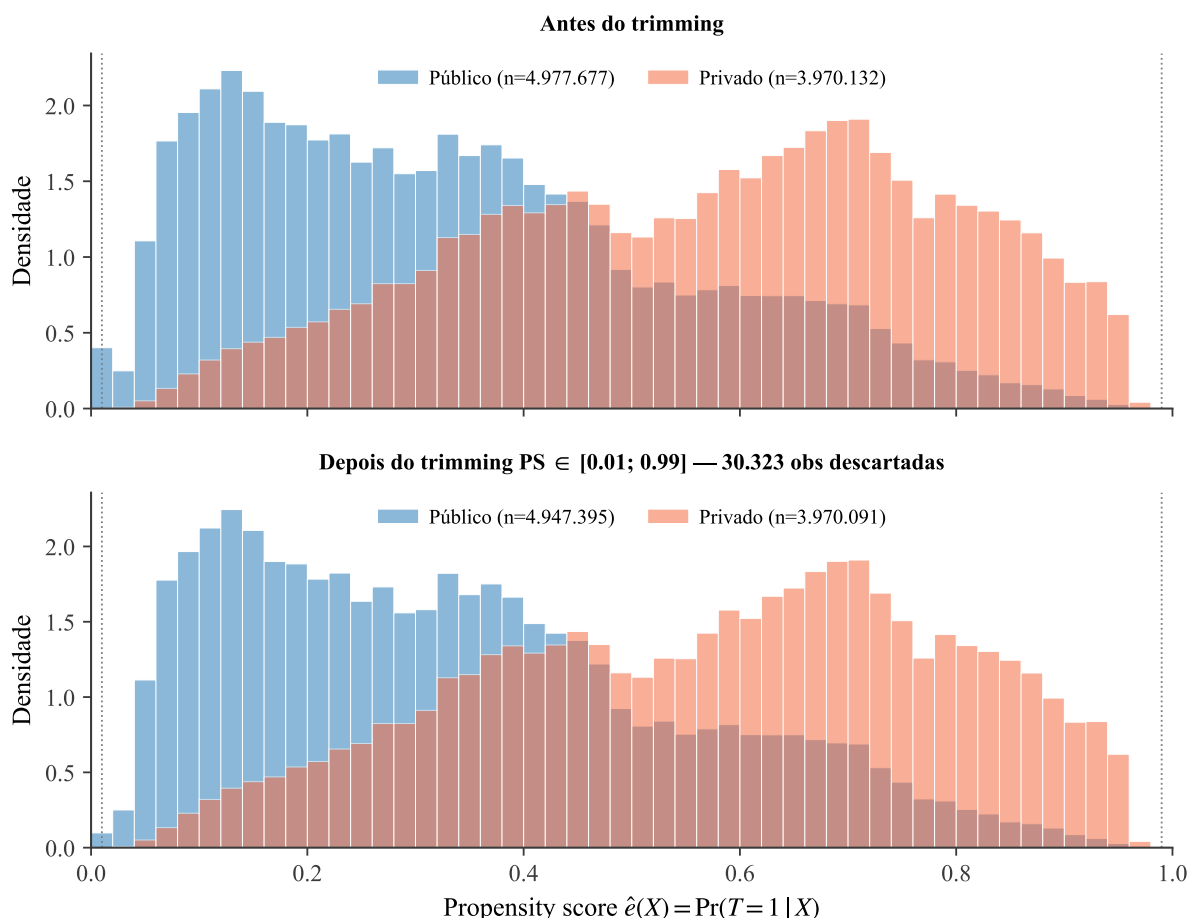


Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

A validade dos estimadores baseados em escore de propensão (IPW, AIPW e PSM) depende de sobreposição (*common support*) entre tratados e controles na distribuição do escore. A Figura 9 mostra a distribuição do escore de propensão por setor na amostra principal: a sobreposição é ampla, e o *trimming* entre os percentis 0,01 e 0,99 adotado na estimação remove apenas as caudas, sem comprometer o suporte comum.

Figura 9 – Distribuição do escore de propensão por setor, amostra principal Robson 1+3.

Distribuição do propensity score — Robson 1+3 (principal)



Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Por que mais um estimador? É importante testar o achado central contra estimadores adicionais com pressupostos identificatórios diferentes. Se o resultado sobrevive a esse conjunto, ganha-se confiança de que não é artefato da escolha de um método específico.

O AIPW (Robins; Rotnitzky; Zhao, 1994) combina o modelo do desfecho (logit) e o modelo do tratamento (IPW) em um estimador duplamente robusto: ATE de **15,82 p.p** (IC95% [13,90; 17,49]), com desvio-padrão 23% menor que o do IPW por incorporar a informação do desfecho. O *entropy balancing* (Hainmueller, 2012) dispensa a forma funcional do escore, balanceando momentos diretamente: **15,93 p.p** ([13,53; 18,39]). IPW (15,79), AIPW (15,82) e EB (15,93) convergem na primeira casa decimal apesar de pressupostos identificatórios distintos. A Tabela 7 consolida, num único quadro, o tratamento, o conjunto de covariáveis e o estimando de cada estimador, deixando explícito por que as duas escalas não são diretamente comparáveis: a razão de chance do logit saturado parcializa a oferta hospitalar, ao passo que os estimadores

baseados em escore a deixam no canal e estimam o alvo da Especificação 2. A estimativa de referência do diferencial institucional é o intervalo OR [1,54; 3,03], do piso conservador da Especificação 3 ao limite superior do logit hierárquico marginal.

Tabela 7 – Tratamento, covariáveis, estimando e estimativa de cada estimador da Pergunta 3.

Estimador	Tratamento	Ajuste (covariáveis)	Estimativa
Logit Especificação 1	lucrativo/SFL vs. público	clínico mínimo	OR 3,82
Logit Especificação 2	lucrativo/SFL vs. público	+ maternas (demanda)	OR 3,36
Logit Especificação 3	lucrativo/SFL vs. público	+ oferta hospitalar	OR 1,54 (AME 8,14 p.p)
IPW / AIPW / <i>entropy bal.</i>	privado vs. público	maternas (Espec. 2; oferta omitida)	15,8–15,9 p.p
PSM 1:1	privado vs. público	maternas (Espec. 2; oferta omitida)	ATT 16,75 p.p
Painel <i>between</i>	natureza jurídica	agregadas hospital–ano	13,93 p.p
GLMM (intercepto aleatório)	lucrativo vs. público	maternas + hospital no aleatório	OR marginal 3,03

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

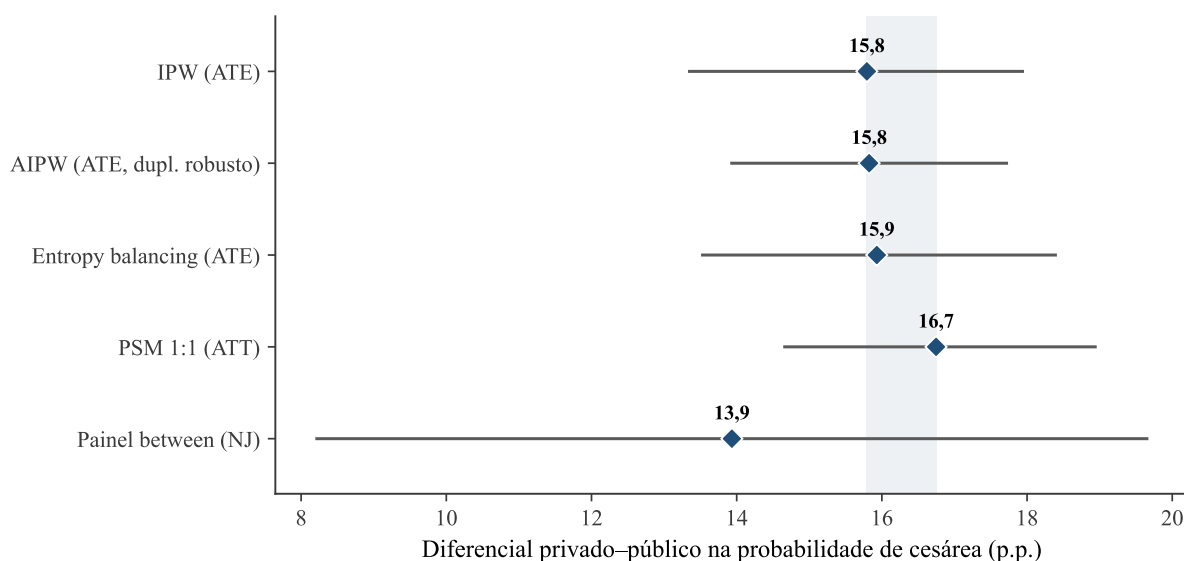
Nota: “Demanda” = características maternas e clínicas; “oferta” = perfil de atendimento, salas, leitos e vínculo com o SUS.

PSM 1:1⁶² entrega ATT de **16,75 p.p** (IC95% [14,66; 18,94]), próximo do ATT do IPW (17,02 p.p). A diferença entre o ATT (cerca de 17 p.p) e o ATE do AIPW e do EB (cerca de 15,8 p.p) é compatível com tratamento heterogêneo na direção predita por SID, com efeito maior em multíparas e em escolaridade Fundamental/Média (cf. §4.4.2). A leitura $ATT > ATE$ reforça, portanto, em vez de contradizer o achado central.

A Figura 10 traduz a triangulação em uma única imagem: os cinco estimadores, ainda que apoiados em pressupostos identificatórios distintos, concentram-se na faixa de 14 a 17 pontos percentuais.

⁶² *Propensity Score Matching* 1:1: cada tratado é pareado com um único controle de escore mais próximo, sem reposição. Identifica o ATT (efeito sobre os tratados), não o ATE. SE por bootstrap clusterizado por hospital com 200 réplicas.

Figura 10 – Triangulação dos estimadores observacionais do diferencial de cesárea entre privado lucrativo e público (p.p.), com IC 95%. Robson 1+3.



Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Os estimadores anteriores medem o tamanho do diferencial, o passo seguinte pergunta quanto da variação na cesárea está *entre* hospitais, e não dentro deles. O logit com intercepto aleatório por hospital (GLMM)⁶³ responde a essa questão, com os resultados consolidados na Tabela 8: ICC = **0,48**, valor consideravelmente acima do tipicamente reportado em estudos multinível de ciências sociais e de saúde (0,05–0,20 segundo Snijders e Bosker (2012)). O hospital concentra cerca de metade da variação residual depois de controlar por idade, paridade, raça e cor, escolaridade, Kotelchuck, UF e ano. O coeficiente marginal do tratamento estimado pelo GLMM (sem perfil de atendimento nem infraestrutura entre as covariáveis fixas) é $OR_{\text{marg}} = \mathbf{3,03}$ (IC95% [2,79; 3,29]), enquanto o logit Especificação 3 saturado entrega $OR_{\text{cond}} = \mathbf{1,54}$. A discrepância não é contradição: o GLMM marginal absorve perfil e infraestrutura no intercepto aleatório, capturando o efeito total via composição hospitalar; o OR_{cond} isola o efeito condicional ao perfil e à infraestrutura. O intervalo [1,54; 3,03] delimita a magnitude plausível do componente que opera *via* composição hospitalar, e fortalece a leitura econômica da decomposição preditiva (a oferta é a maior fonte de variação preditiva): o hospital é a dimensão de heterogeneidade dominante depois de controlar pelo perfil materno observável.

O conjunto observacional fecha com dois exercícios. O método de *bounds* de Oster

⁶³ GLMM (*Generalized Linear Mixed Model*): trata o hospital como unidade de análise, estimando um intercepto aleatório por estabelecimento. O ICC (*Intraclass Correlation Coefficient*) é a fração da variância total atribuível ao nível hospital. Formalização em §3.10.2, Equação (12); estimado sobre subamostra estratificada de 561.242 partos preservando os 5.280 hospitais, com quadratura adaptativa de Laplace (nAGQ=1).

(2019) estima $\delta^* = 0,84$ para o privado lucrativo, ou seja: para que o efeito seja anulado por inobservados, esses inobservados precisariam ter força preditiva sobre o desfecho equivalente a 84% da força preditiva conjunta de todos os observáveis. Está abaixo do limiar convencional $\delta^* > 1$ adotado pela literatura como sinal de robustez forte⁶⁴. O E-value de VanderWeele e Ding (2017) é de aproximadamente **1,79** nas três amostras (1,4–1,5 para a borda inferior do IC95%): um confundidor não observado precisaria estar associado simultaneamente ao tratamento e ao desfecho com risco relativo mínimo de 1,79 para anular o ponto estimado. As duas medidas apontam para o mesmo lado; as tabelas completas de Oster e E-value, nas três amostras, estão no Apêndice C (Tabelas C1 e C2).

O efeito é robusto a confundimento moderado, mas vulnerável a um confundidor inobservado de magnitude moderada-alta (o $\delta^* = 0,84$ está logo abaixo do limiar convencional de 1). Essa vulnerabilidade é reconhecida e listada explicitamente em §4.7. A leitura dos dados é que a série de estimadores observacionais sustenta uma associação ajustada compatível com SID, e não uma demonstração de SID. A Tabela 8 consolida a triangulação dos estimadores nas três amostras (ATE entre 14,8 e 17,5 p.p, conforme o escopo).

Tabela 8 – Robustez causal: diferencial privado lucrativo vs. público entre estimadores, Robson 1+3, Brasil, 2014–2022.

Estimador	Estimativa	IC95% / diagnóstico
Logit (OR lucr. vs. públ.)	1,54	[1,20; 1,99]***
IPW (ATE, p.p)	15,79	[13,35; 17,94]
AIPW (ATE, p.p)	15,82	[13,90; 17,49]
<i>Entropy balancing</i> (ATE, p.p)	15,93	[13,53; 18,39]
PSM 1:1 (ATT, p.p)	16,75	[14,66; 18,94]; 2.788.828 pares; SMD máx 0,130; OR 2,19
Painel A, <i>between</i> (p.p)	+13,93	[+8,22; +19,65]***; 36.408 hosp.-ano
Painel B, <i>within</i> (p.p)	−3,55	[−9,05; +1,96], $p = 0,21$
Oster, β^* ($\delta=1$) (p.p)	−1,90	$\delta^* = 0,84$ (vulnerável a OVB moderado)
GLMM, OR _{marg}	3,03	[2,79; 3,29]; ICC = 0,485; 561.242 (subam., nAGQ=1)
E-value (ponto)	1,79	borda inferior do IC95: 1,41

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

O achado central não depende da escolha de um método específico, pois sobrevive sob hipóteses identificatórias distintas.

⁶⁴ O limiar $\delta^* > 1$ é heurística informal de Oster (2019): se o inobservado precisaria ser mais informativo que todos os observáveis para anular o efeito, o achado é considerado robusto. Operacionalização em §4.4.1; formalização em §3.10.3, Equação (13).

4.4.2 Heterogeneidade do diferencial

O diferencial é uniforme ou se concentra em subgrupos específicos? A Tabela 9 reporta o logit Especificação 3 estratificado dentro da amostra principal por quatro dimensões: grupo de Robson, região, escolaridade materna e perfil de atendimento do hospital.

Tabela 9 – Heterogeneidade do diferencial privado lucrativo: logit Especificação 3 estratificado, Robson 1+3.

Dimensão	Subgrupo	N	OR (IC95)	p-valor
Amostra completa	Robson 1+3	8.947.809	1.54 [1.20; 1.99]	<0,001
Robson	Robson 1	4.326.050	1.39 [1.07; 1.80]	0.0142
	Robson 3	4.621.759	1.62 [1.23; 2.13]	<0,001
Região	NE	3.169.461	1.53 [1.06; 2.21]	0.0219
	SE	2.849.045	1.68 [0.83; 3.37]	0.1464
	S	923.582	1.13 [0.81; 1.57]	0.4686
	N	1.245.854	2.05 [0.82; 5.14]	0.1238
	CO	759.867	2.07 [1.16; 3.68]	0.0132
Escolaridade materna	Fundamental	2.886.615	1.69 [1.29; 2.22]	<0,001
	Médio	4.682.660	1.73 [1.32; 2.27]	<0,001
	Superior	1.257.154	1.34 [0.95; 1.88]	0.0937
Perfil de atendimento	SUS exclusivo	5.475.519	1.85 [1.24; 2.77]	0.0028
	Misto	2.755.969	1.97 [0.79; 4.91]	0.1433
	Privado exclusivo	716.321	1.28 [0.36; 4.55]	0.6975

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

O diferencial é **maior em múltiparas (Robson 3, OR = 1,62) do que em nulíparas (Robson 1, OR = 1,39)**. Compreende-se que a margem de discricionariedade clínica, o espaço de decisão do médico e do hospital que não é ditado pela condição clínica da paciente, é maior em múltiparas, que já têm histórico obstétrico sinalizando viabilidade do parto vaginal, justamente onde se espera mais espaço para indução institucional. O padrão é consistente com SID.

O coeficiente é significativo no Nordeste (OR = 1,53) e no Centro-Oeste (OR = 2,07), próximo da unidade no **Sul** (OR = 1,13, IC95% inclui 1), e impreciso no Sudeste e Norte por desbalanceamento amostral entre os setores. A exceção do Sul tem leitura ancorada em Pereira *et al.* (2024): a região concentra a maior cobertura relativa de hospitais públicos e o setor privado lucrativo tem participação proporcionalmente menor que nas demais regiões. Sob a hipótese de SID, o diferencial deve ser menor justamente onde o setor privado lucrativo é menos relevante na composição da rede. O achado, portanto, fortalece a leitura compatível com SID. O diferencial residual se atenua onde o canal institucional pelo qual o efeito opera pesa menos. A leitura

ancora-se na composição setorial da rede regional, mas leituras alternativas não testáveis com a base atual permanecem em aberto: cultura obstétrica regional, perfil hospitalar dominante (porte, especialidade) e composição demográfica das gestantes podem contribuir para a magnitude observada. A discriminação entre essas explicações exige variação adicional não disponível no SINASC–CNES integrado.

O efeito é **maior em Fundamental (OR = 1,69) e Médio (OR = 1,73) do que em Superior (OR = 1,34)**. O padrão desafia a leitura simplificada de que a cesárea no setor privado seria majoritariamente uma escolha da gestante de classe média-alta, a magnitude relativa indica, ao contrário, que a indução opera com mais força em populações com menor capacidade de barganha clínica. O diferencial institucional sobrevive, portanto, ao teste mais incômodo possível.

O coeficiente é significativo no estrato de hospitais SUS-exclusivos (OR = 1,85), onde a presença privada lucrativa é restrita, e impreciso nos estratos Misto e Privado exclusivo, onde o grupo de comparação público é raro.

4.4.3 Compatibilidade com SID e a ameaça do painel *within-hospital*

Os achados dos seis estimadores observacionais, consolidados na Tabela 8, vão na mesma direção. Na população clinicamente comparável, o parto em hospital privado lucrativo apresenta *odds* 1,54 vezes maior de cesárea que o parto em hospital público, depois de controlar por perfil materno completo, perfil de atendimento, infraestrutura hospitalar e efeitos fixos de ano e UF. O efeito médio na escala de pontos percentuais fica entre 13,9 e 17,0 p.p. A literatura clássica (Arrow, 1963; Evans, 1974; Gruber; Owings, 1996) postula que o profissional de saúde, ao deter o monopólio da informação clínica e operar sob remuneração diferenciada por procedimento, pode recomendar procedimentos cuja necessidade não decorre de critério estritamente clínico. Epstein e Nicholson (2009) mostram que parcela considerável da heterogeneidade entre médicos na decisão pela cesárea persiste por anos após o início da carreira, sustentando que o diferencial entre estabelecimentos não decorre apenas de seleção de pacientes, mas também de prática clínica idiossincrática moldada institucionalmente.

Seis verificações qualificam essa leitura: estabilidade do coeficiente entre amostras (1,54/1,59/1,55); concordância entre seis estimadores observacionais com pressupostos identificatórios distintos (logit, IPW 15,79 p.p, AIPW 15,82 p.p, EB 15,93 p.p, PSM ATT 16,75 p.p, painel *between* 13,93 p.p); ICC do GLMM de 0,48 indicando concentração da variância latente entre estabelecimentos; balanço empírico (SMD máxima de 0,033 pós-IPW); sensibilidade a confundimento (Oster $\delta^* = 0,84$, E-value $\approx 1,79$); e a heterogeneidade pré-discutida (efeito

maior em múltiplas e em escolaridade Fundamental/Média, próximo da unidade no Sul).

A vulnerabilidade central do diferencial institucional. O contraste entre o painel *between* (Modelo A, +13,93 p.p) e o painel *within-hospital* (Modelo B, -3,5 p.p não significativo), ambos reportados na Tabela 10, precisa ser declarado. Tomado literalmente, o desenho *within* tem poder estatístico para detectar o efeito *between*, mas não o detecta: o exercício dedicado sobre os 151 hospitais que mudaram de natureza jurídica entrega ponto de -2,61 p.p não significativo, com SE clusterizado de 2,54 p.p e MDE⁶⁵ a $\alpha = 0,05$ e poder 80% de 7,11 p.p, valor inferior ao ATE *between* de 13,93 p.p. Não rejeitar a nula no painel *within* é, portanto, compatível com efeito real de magnitude inferior a 7,11 p.p, e não constitui evidência de ausência de efeito. Esta é a principal ameaça à validade causal da análise observacional do diferencial.

Ao adotar uma leitura conservadora, podemos indicar que o estimador transversal pode estar dominado por seleção não observável no nível do hospital (hospitais intrinsecamente diferentes que adotam CNPJs distintos), em vez de capturar o efeito da natureza jurídica propriamente. A leitura interpretativa alternativa, discutida adiante, é que a natureza jurídica opera por dimensões institucionais estáveis (cultura clínica, contratos de remuneração, perfil de credenciamento) que mudam lentamente, e o desenho *within-hospital* sobre transições administrativas não captura essas dimensões. A discriminação entre as duas leituras exige variação adicional não disponível no desenho atual, e a §4.7 retoma o ponto como limitação reconhecida.

Tabela 10 – Painel hospital $\times$ ano da taxa de cesárea por natureza jurídica, WLS, Robson 1+3, Brasil, 2014–2022.

Tratamento (ref. Público)	Modelo A: pooled + EF ano		Modelo B: EF hospital + EF ano	
	β	IC 95%	β	IC 95%
Privado SFL	+1,43 p.p	[-3,67; +6,53]	+0,50 p.p	[-2,51; +3,51]
Privado lucrativo	+13,93 p.p***	[+8,22; +19,65]	-3,55 p.p	[-9,50; +2,41]

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Nota: Desfecho: taxa de cesárea do hospital no ano; WLS ponderado pelo número de partos; erros-padrão clusterizados por estabelecimento.

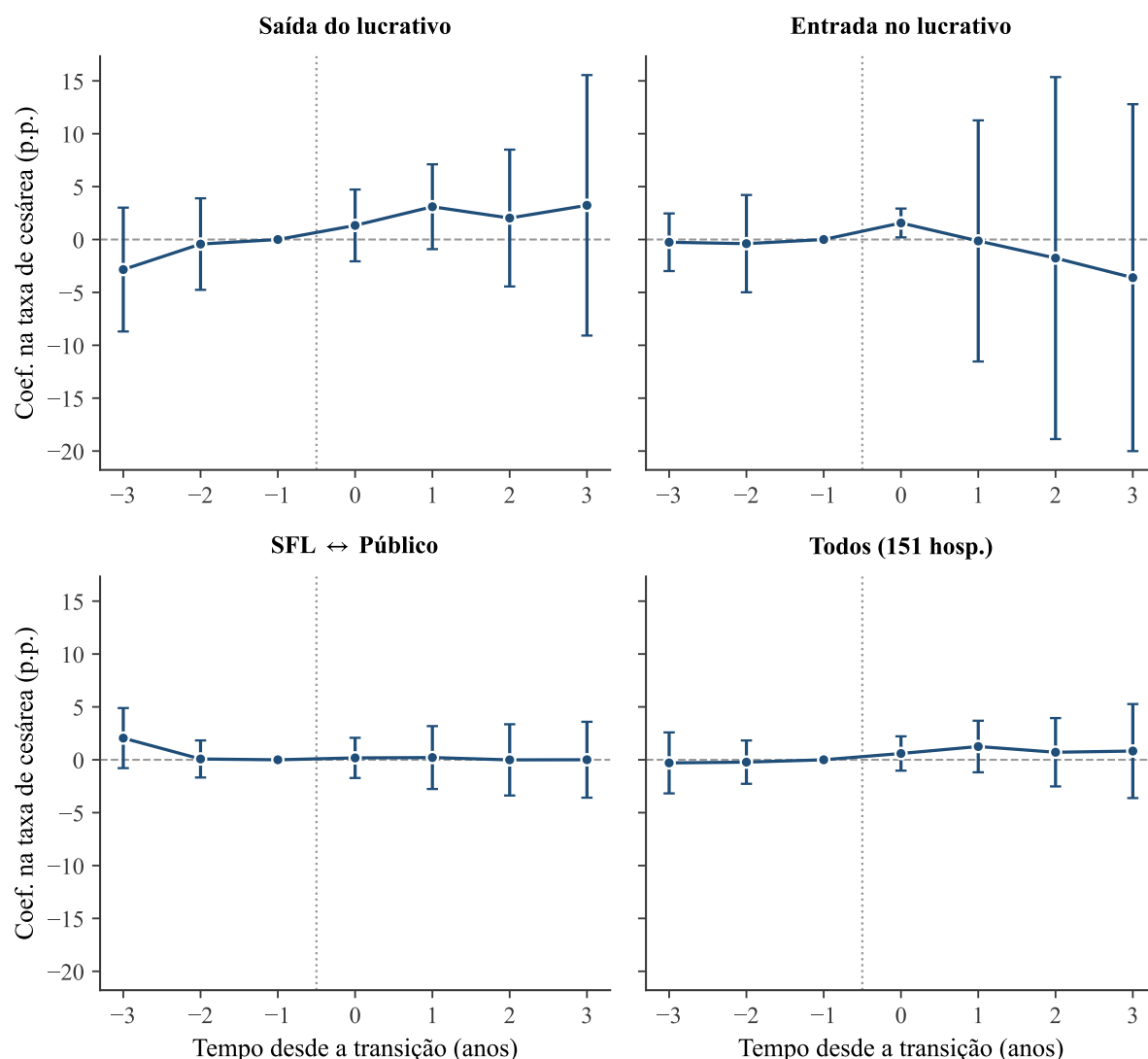
Um *event study* sobre as próprias transições de natureza jurídica oferece evidência direta sobre qual das duas leituras prevalece. Entre os 151 hospitais que mudaram de natureza jurídica no período (cujo perfil de volume, estrutura e taxa de cesárea está caracterizado no Apêndice D),

⁶⁵ MDE (*Minimum Detectable Effect*): menor magnitude de efeito que o desenho consegue detectar com poder estatístico de 80% e nível de significância $\alpha = 0,05$. Indicador de potência *ex post*. Fórmula em (11); cálculo aplicado em §4.4.1.

as pré-tendências são limpas em todos os subgrupos de transição (o teste de Wald conjunto sobre os coeficientes pré-tratamento não rejeita o paralelismo, com p entre 0,18 e 0,98), o que valida a leitura dos coeficientes posteriores à transição. O subgrupo mais frequente, as transições entre o setor sem fins lucrativos e o público, não produz salto contemporâneo na taxa de cesárea: os coeficientes ficam próximos de zero e nenhum é significativo. O único coeficiente significativo do exercício aparece nos onze hospitais que entraram no setor lucrativo (+1,56 p.p no ano da transição, IC95% [+0,20; +2,93]), e mesmo esse não persiste nos anos seguintes (Figura 11; os coeficientes completos estão na Tabela D2, Apêndice D).

Trocar a etiqueta jurídica de um hospital, sem alterar ao mesmo tempo a composição clínica, os contratos de remuneração e a cultura assistencial, não basta para reconfigurar a taxa de cesárea no curto prazo. A natureza jurídica opera por dimensões institucionais lentas, e o desenho *within-hospital* sobre mudanças administrativas não as captura. Isso favorece a leitura interpretativa sobre a conservadora: o diferencial *between* de 13,9 p.p não reflete apenas seleção de hospitais intrinsecamente distintos, mas também arranjos institucionais estáveis que o exercício de transição administrativa não consegue mover.

Figura 11 – Efeito da mudança de natureza jurídica sobre a taxa de cesárea, por subgrupo de transição (*event study*), com IC 95%.



Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

4.4.4 Médicas como pacientes: um teste do canal informacional

O conjunto principal de estimadores identifica diferencial sistemático compatível com SID, mas não testa diretamente a assimetria informacional postulada pela teoria. Johnson e Rehavi (2016), em microdados americanos, propôs um teste comparando gestantes médicas (informação simétrica) com controles pareados de mesmo perfil socioeconômico, e documentou redução de aproximadamente 10% na taxa de cesárea entre médicas, com a maior parte do efeito sobrevivendo ao controle por hospital. A análise a seguir replica o desenho sobre o SINASC–CNES, com adaptação à classificação CBO 2002 brasileira (a operacionalização técnica está em §3.10.4). Três achados ordenam a leitura.

Gestantes médicas não obstétricas⁶⁶ apresentam OR de cesárea de **0,825** sobre o controle Superior em ocupação não saúde no setor privado (IC95% [0,77; 0,89], $p < 0,001$, $N = 13.669$; Tabela 11). A direção é a esperada por SID via assimetria informacional: quem tem informação clínica direta faz menos cesárea. A magnitude é estatisticamente robusta sobre amostra ampla, embora mais fraca que os 10% documentados por Johnson e Rehavi (2016) nos Estados Unidos.

O efeito é mais forte no privado lucrativo (OR 0,808, $p < 0,001$) e ausente no privado sem fins lucrativos (OR 0,912, $p = 0,19$). O padrão é coerente com SID, que prevê maior força do efeito onde o canal de remuneração diferenciada opera.

O efeito desaparece dentro do mesmo hospital (Modelo B, $-0,61$ p.p, IC95% $[-1,88; +0,66]$, $p = 0,35$).

O canal informacional opera no Brasil, mas pelo mecanismo de fuga entre hospitais, não de barganha intra-hospital. Gestantes médicas tenderiam a escolher primeiro o profissional e, por extensão, os hospitais privados com taxas de cesárea menores; uma vez dentro do estabelecimento, a decisão clínica coincide com a do controle pareado.

Esse achado articula-se com a hipótese rival de autosseleção da gestante para o tipo de hospital, discutida na §4.7. Sob autosseleção pura pró-cesárea, gestantes que querem cesárea procuram hospitais lucrativos e o diferencial residual reflete demanda inobservável. O achado das médicas **qualifica** essa hipótese em direção oposta: as gestantes mais bem informadas escolhem na direção contrária da autosseleção pró-cesárea, isto é, fogem, pela escolha do profissional, dos hospitais com taxa elevada. Ocorrendo uma bifurcação de demanda informada, gestantes não informadas seguem padrões institucionais e podem ser induzidas à cesárea; gestantes informadas usam a informação para selecionar hospitais com perfil clínico mais conservador.

Enfermeiras e técnicas de enfermagem (L2), também informadas, mas no eixo da rotina hospitalar, apresentam padrão oposto (OR 1,360, sobrevive ao within), o que sugere que conhecimento da rotina não se converte em rejeição da cesárea. A implicação metodológica para a defesa do mecanismo SID pode ser interpretada a partir do diferencial residual do conjunto principal de estimadores (OR 1,54 do privado lucrativo), que é compatível com SID via canal informacional *ex ante* (na escolha do hospital), e essa especificação se distingue da versão americana de SID via barganha *ex post*. Um conjunto de testes de robustez (Apêndice E) sustenta o achado por quatro frentes: placebo no setor público, onde o efeito se inverte (OR 1,72) na

⁶⁶ Médicas não obstétricas: gestantes com escolaridade Superior e ocupação na família CBO 2231 (médicos) sem ginecologia-obstetrícia. A escolaridade Superior controla aproximadamente o *status* socioeconômico; a ocupação isola o componente de informação clínica geral, sem o eixo da especialidade obstétrica. Construção da escala L0–L4 em §3.10.4.

ausência de incentivo de oferta; controle de renda alta, sob o qual a proteção sobrevive (OR 0,89); verificação cruzada por IPW e AIPW, com diferença de risco de $-2,5$ p.p duplamente robusta; e a seleção de hospital como variável dependente, que confirma a fuga *ex ante* ($-3,3$ p.p na taxa de cesárea do hospital de destino). O achado permanece sensível a confundimento inobservado de magnitude moderada (E-value 1,44) e limitado quanto aos ginecologistas-obstetras ($N = 132$); o campo de ocupação em branco em 2016 impede acrescentar aquele ano. A identificação plena fica registrada na agenda futura (seção 5).

Tabela 11 – Efeito da informação obstétrica da gestante sobre a probabilidade de cesárea, exercício de médicas como pacientes (Johnson e Rehavi (2016)), setor privado, escolaridade Superior, Robson 1+3, Brasil, 2014–2022.

Modelo	Estrato / Especificação	Coefficiente	IC95%
<i>Modelo A: logit baseline (escala odds), L0 referência</i>			
L1 Outras saúde superior	OR (vs. L0)	0,995	[0,93; 1,06]
L2 Enfermeiras superior + técnicas	OR (vs. L0)	1,360***	[1,30; 1,42]
L3 Outras médicas (família CBO 2231)	OR (vs. L0)	0,825***	[0,77; 0,89]
L4 Ginecologistas-obstetras ($N = 132$)	OR (vs. L0)	0,762	[0,54; 1,07]
<i>Modelo B: LPM com efeito fixo de hospital (escala p.p), L0 referência</i>			
L2 Enfermeiras superior + técnicas	$\hat{\beta}$ (p.p)	+5,24***	[+4,65; +5,83]
L3 Outras médicas	$\hat{\beta}$ (p.p)	−0,61	[−1,88; +0,66]
<i>Modelo C: logit baseline restrito ao Privado Lucrativo</i>			
L3 Outras médicas	OR (vs. L0)	0,808***	[0,74; 0,89]
<i>Modelo D: logit baseline restrito ao Privado SFL</i>			
L3 Outras médicas	OR (vs. L0)	0,912	[0,79; 1,05]

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

4.5 O efeito causal quase-experimental do Projeto Parto Adequado

O quarto passo leva a análise ao terreno causal. Enquanto a análise observacional do diferencial estabelece, sob seleção em observáveis, que o diferencial sistemático Lucrativo vs. Público é compatível com SID, o desenho quase-experimental do PPA mobiliza fonte de variação plausivelmente exógena, a adesão escalonada de hospitais ao PPA da ANS entre 2015 e 2022, e busca consistência causal com SID dentro do escopo da intervenção.

O desenho do PPA é diferenças-em-diferenças escalonado com tendências paralelas validadas no único *cohort* com pré-tratamento informativo. Sob esse desenho, a literatura econométrica reconhece a leitura causal calibrada do efeito sobre os tratados (ATT) como apropriada (Callaway; Sant’Anna, 2021; Sun; Abraham, 2021; Rambachan; Roth, 2023). A leitura permanece restrita ao escopo do programa: o ATT refere-se aos 121 hospitais aderentes, à amostra Robson 1+3 e à janela 2014–2022.

O PPA reduziu a taxa de cesárea em hospitais aderentes em $-5,56$ p.p (ATT médio pós-tratamento, estimador ponderado por interação (*interaction-weighted*, IW) de Sun-Abraham), sobre Robson 1+3 e com tendências paralelas validadas no único *cohort* com pré-tratamento informativo (*cohort* 2017, $\chi^2(2) = 0,481$, $p = 0,786$)⁶⁷. A magnitude é cerca do dobro da do TWFE clássico ($-2,87$ p.p), confirmando o viés de atenuação documentado por Goodman-Bacon (2021) e Sun e Abraham (2021) para tratamento escalonado heterogêneo.

Em hospitais aderentes ao PPA, a cada 100 partos hospitalares de Robson 1+3 que ocorreriam sob o cenário contrafactual, aproximadamente 5 ou 6 deixaram de ser cesáreas e passaram a ser partos vaginais. Em uma rede em que o ponto de partida estava em torno de 65–70% de cesáreas no setor privado lucrativo, falar em $-5,56$ p.p não é um movimento marginal: é cerca de 8 a 9% de redução relativa.

O efeito está concentrado no setor privado lucrativo ($-6,30$ p.p na Especificação B do TWFE, IC95% $[-11,44; -1,82]$, significativo a 5%), com efeitos não significativos no público e no privado sem fins lucrativos. É exatamente o que a hipótese SID prevê.

O efeito é persistente e crescente no tempo, atingindo $-10,88$ p.p no sexto ano pós-adesão (*cohort* 2015 em 2021). A intervenção institucional do PPA opera por reorganização lenta e progressiva da cultura clínica e contratual, não por choque pontual.

A evidência observacional do diferencial institucional é compatível com SID, diferencial residual robusto a múltiplos estimadores baseados em escore, mas vulnerável a viés por inobservados moderado (Oster $\delta^* = 0,84 < 1$) e não detectado pela variação *within-hospital*. A evidência quase-experimental do PPA entrega identificação causal do efeito agregado do PPA sobre os hospitais aderentes, com paralelas validadas no *cohort* informativo e conjunto de testes de robustez consistente. O desenho não identifica o canal causal específico, o PPA é intervenção composta (contratos, protocolos, capacitação, informação) e seu efeito é compatível tanto com a hipótese SID via remuneração diferenciada quanto com a hipótese de capacidade gerencial diferencial documentada por Borem *et al.* (2020). A leitura SID, portanto, organiza o argumento interpretativamente; a identificação do mecanismo causal específico depende de instrumentação adicional registrada como agenda futura.

As Tabelas 12 e 13 reportam, respectivamente, as três especificações do TWFE e o estimador IW de Sun-Abraham que sustentam essa leitura.

⁶⁷ Teste de Wald conjunto sobre $\hat{\delta}_{c=2017, \ell=-3}$ e $\hat{\delta}_{c=2017, \ell=-2}$: $p = 0,786$ não rejeita a hipótese nula de paralelismo entre aderentes e controles antes do tratamento. Os *cohorts* 2015 e 2020 não dispõem de pré-tratamento útil (cf. §3.11).

Tabela 12 – Efeito do Projeto Parto Adequado sobre a taxa de cesárea, diferenças em diferenças com efeitos fixos de hospital e ano (TWFE), Robson 1+3, Brasil, 2014–2022.

Especificação	Coefficiente	Efeito (p.p)	IC95% (p.p)
<i>Spec A: efeito médio agregado (replica Knau et al. 2023)</i>			
DiD base	$\hat{\beta}$	–2,87*	[–5,71; –0,04]
<i>Spec B: heterogeneidade por natureza jurídica tripartite</i>			
Pós × Público	$\hat{\beta}_{\text{pub}}$	–1,73	[–5,94; +2,87]
Pós × Privado SFL	$\hat{\beta}_{\text{sfl}}$	–1,65	[–5,84; +1,46]
Pós × Privado Lucrativo	$\hat{\beta}_{\text{luc}}$	–6,30*	[–11,44; –1,82]
<i>Spec C: por cohort de início do tratamento</i>			
Cohort 2015 (Fase 1)	$\hat{\beta}_{c=2015}$	–6,86*	[–12,05; –2,08]
Cohort 2017 (Fase 2)	$\hat{\beta}_{c=2017}$	–2,36	[–5,75; +0,79]
Cohort 2020 (CI2)	$\hat{\beta}_{c=2020}$	+5,11	[–27,35; +16,93]

Fonte: Elaboração própria a partir do SINASC–CNES integrado e do Anexo A do Relatório ANS (2023). Asterisco indica IC95% que não cruza zero.

Tabela 13 – ATT do Projeto Parto Adequado sobre a taxa de cesárea por tempo relativo à adesão, estimador de Sun e Abraham (2021), Robson 1+3, Brasil, 2014–2022.

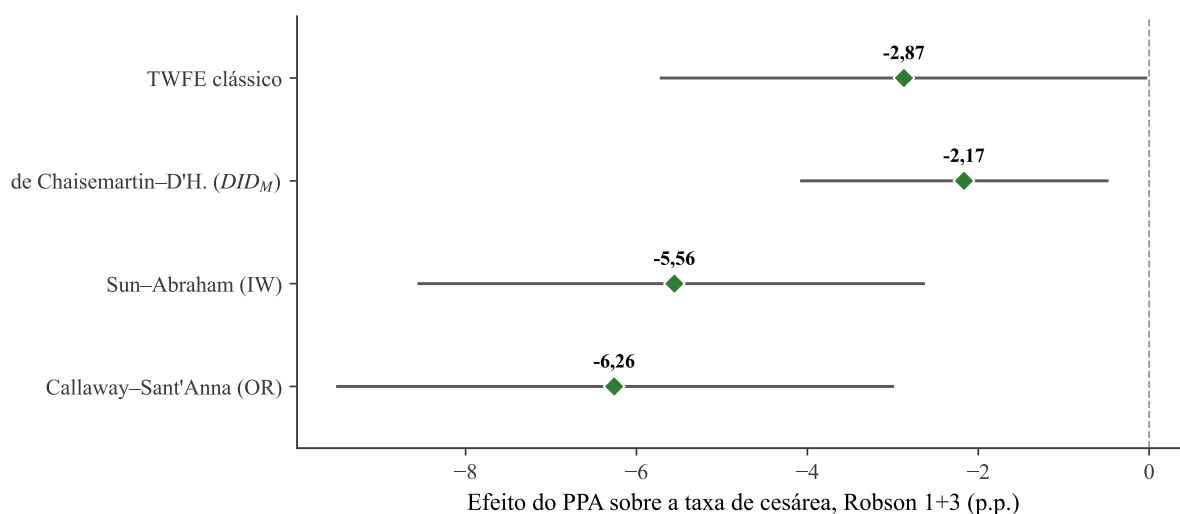
Event time ℓ	Cohorts contribuintes	ATT (p.p)	IC95% (p.p)	Janela
<i>Pré-tratamento (interpretação calibrada na §4.5)</i>				
–6	2020	–36,95	[–47,03; +34,68]	2014
–5	2020	–26,98	[–41,68; +38,61]	2015
–4	2020	+20,50	[–4,06; +38,07]	2016
–3	2017, 2020	+2,58	[–1,49; +7,39]	2014–2017
–2	2017, 2020	+0,23	[–2,87; +3,46]	2015–2018
–1 (ref.)	—	—	—	—
<i>Pós-tratamento (efeito do PPA por anos desde adesão)</i>				
+0	2015, 2017, 2020	–1,96*	[–4,02; –0,27]	2015–2020
+1	2015, 2017, 2020	–3,86*	[–6,45; –1,60]	2016–2021
+2	2015, 2017, 2020	–5,02*	[–7,88; –2,33]	2017–2022
+3	2015, 2017	–3,98*	[–7,31; –0,71]	2018–2020
+4	2015, 2017	–4,67*	[–8,17; –0,86]	2019–2021
+5	2015, 2017	–4,34*	[–7,66; –0,74]	2020–2022
+6	2015	–10,88*	[–17,10; –4,80]	2021
+7	2015	–9,73*	[–15,47; –4,24]	2022
ATT médio pós-tratamento agregado ($\ell \geq 0$): cluster bootstrap sobre o agregado				
	$\overline{\text{ATT}}_{\text{pós}}$ (Bloco I)	–5,56*	[–8,55; –2,64]	2015–2022

Fonte: Elaboração própria a partir do SINASC–CNES integrado e do Anexo A do Relatório ANS (2023). Asterisco indica IC95% que não cruza zero.

A inferência causal apresentada até aqui se sustenta em uma única especificação, agora submetida a um conjunto de testes de robustez em sete frentes; as tabelas completas, incluindo o teste de tendências paralelas (Tabela C4) e a série consolidada (Tabela C5), estão no Apêndice C.

A Figura 12 resume a frente dos estimadores modernos: os quatro convergem para um efeito negativo, e os que corrigem o viés do tratamento escalonado heterogêneo (Sun-Abraham e Callaway-Sant’Anna) entregam magnitude maior, em termos absolutos, que o TWFE clássico.

Figura 12 – Efeito do Projeto Parto Adequado sobre a taxa de cesárea por estimador de diferença-em-diferenças, com IC 95%. Robson 1+3.



Fonte: Elaboração própria a partir do SINASC–CNES integrado e do Anexo A do Relatório ANS (2023).

Três estimadores acompanham Sun-Abraham. O estimador de regressão de desfecho de Callaway e Sant'Anna (2021) entrega ATT médio de $-6,26$ p.p. $[-9,50; -3,00]$. O DID_M de Chaisemartin e D'Haultfœuille (2020) entrega $-2,17$ p.p. $[-4,07; -0,49]$ (magnitude menor esperada por capturar efeito imediato, antes da reorganização institucional aparecer nos dados). A decomposição de Goodman-Bacon (2021) mostra que o componente problemático do TWFE responde por 22% do peso da estimativa clássica.

O conjunto de testes de Rambachan e Roth (2023) (Honest DiD, *bound* de magnitudes relativas) é apresentado em duas leituras. A leitura agregada entrega *breakdown* $M = 0,98$, sustentando robustez moderada no limite. A leitura calibrada *cohort-by-cohort*, restrita ao *cohort* 2017 com $\delta_{\max}^{\text{pré}} = 0,86$ p.p., entrega *breakdown* $M = 2,94$, sob a mesma definição da leitura agregada (o menor M para o qual o intervalo robusto $\beta \pm M \delta_{\max} \pm 1,96 \sigma$ passa a incluir zero). As duas leituras refletem o *trade-off* central da série. A leitura agregada $M = 0,98$ é fragilizada pela amostra ruidosa dos *cohorts* 2015 e 2020, sem pré-tratamento informativo. A leitura calibrada *cohort* 2017 ($M = 2,94$) é a inferência possível sobre o único *cohort* em que o pré-tratamento é estatisticamente útil para o teste, mas restringe a base à amostra de $N = 79$ hospitais. A versão completa do procedimento de Rambachan e Roth (2023) via *Fixed Length Confidence Intervals* (FLCI) e ARP, mais informativa quando a amostra pré-tratamento é assimétrica entre *cohorts*, fica registrada como agenda futura.

Dois placebos validam a identificação. O placebo de tratamento⁶⁸ coloca o coeficiente real na cauda inferior 1% da distribuição empírica ($p = 0,010$). O placebo de *outcome* (DiD em Robson 5⁶⁹, não responsivo ao PPA) entrega $-0,46$ p.p não significativo. A robustez ao critério ANS de elegibilidade do controle⁷⁰ (nove cenários) preserva o sinal negativo em 9/9 e a significância em 6/9, com mediana $-2,87$ p.p. A heterogeneidade por região mostra padrão geográfico claro, exibido na Figura 13: efeito negativo e significativo no Norte ($-10,70$), Nordeste ($-5,35$) e Centro-Oeste ($-5,27$), e atenuado, não significativo, no Sudeste e no Sul. A heterogeneidade por escolaridade entrega magnitudes semelhantes nos três estratos (Fundamental $-2,67$; Médio $-3,24$; Superior $-3,28$ p.p), com significância apenas no Médio, o estrato de maior volume; a ausência de efeito desproporcional no estrato Superior exclui a leitura rival de autosseleção de classe média⁷¹.

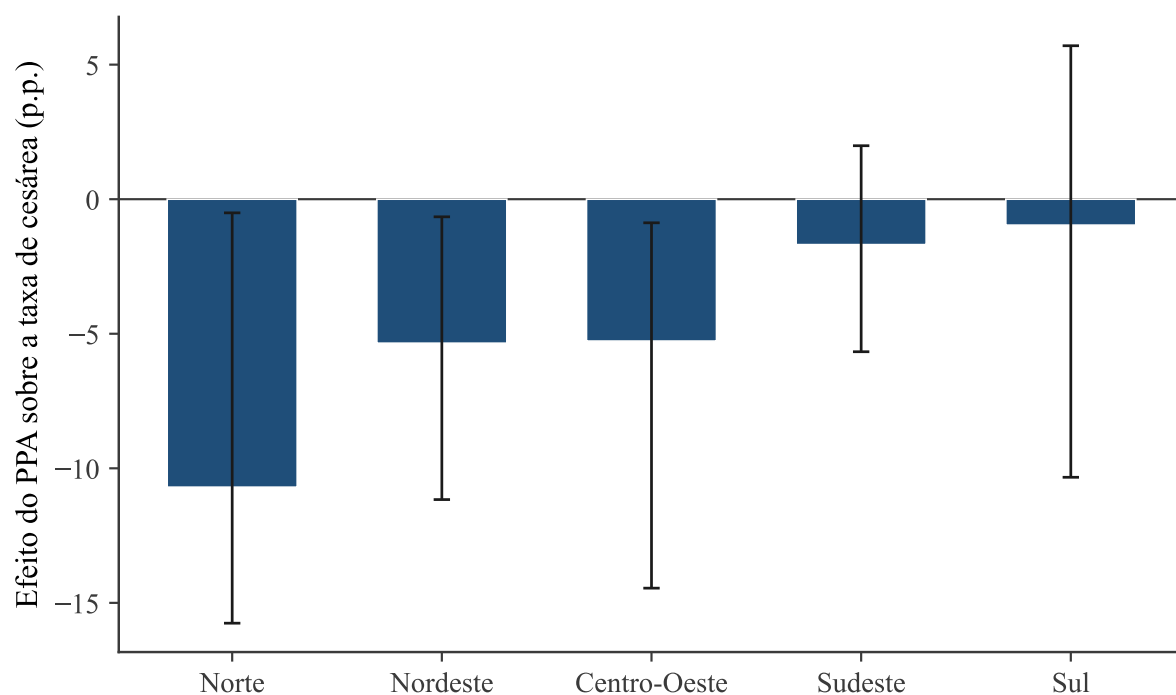
⁶⁸ O placebo de tratamento permuta aleatoriamente o *status* PPA entre os 315 hospitais elegíveis, estratificando por natureza jurídica tripartite e *cohort*, com 500 réplicas. A distribuição empírica gerada serve como nula contrafactual: se o coeficiente real estivesse no centro dela, o desenho não distinguiria efeito de ruído.

⁶⁹ Robson 5: gestantes com cesárea anterior, grupo em que a taxa de cesárea é altíssima (cerca de 86% nos dados) e a via de parto está, em larga medida, predeterminada; o programa PPA não tem mecanismo de ação sobre esse grupo e, por isso, serve como placebo: o efeito esperado é zero.

⁷⁰ O critério de elegibilidade do contrafactual varia o ano-base ($\in \{2012, 2013, 2014\}$) e os limiares de taxa de cesárea (privado $\geq 70\%/75\%/80\%$ pareados com público $> 55\%/60\%/65\%$), gerando nove cenários. A nuvem de pontos resultante mede a sensibilidade do achado a escolhas de operacionalização que poderiam ser razoavelmente diferentes.

⁷¹ Os estratos seguem o dicionário ESCMAEAGR1 do SINASC: Fundamental (códigos 00–04, 10, 11), Médio (05, 06, 12), Superior (07, 08).

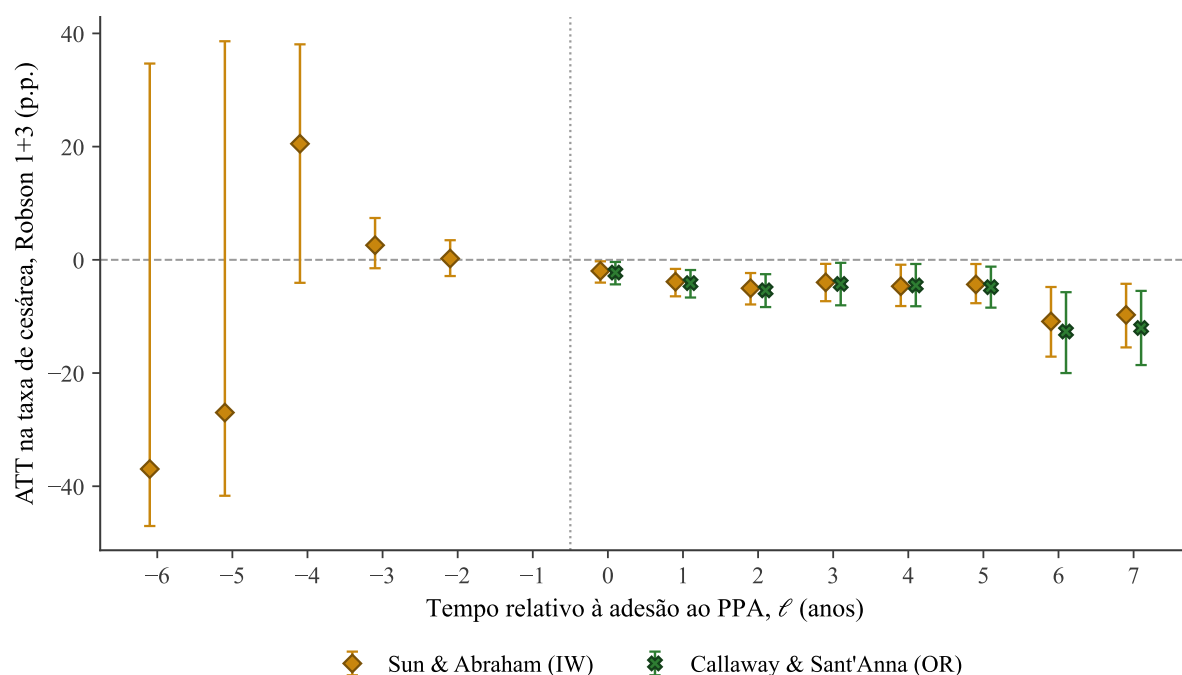
Figura 13 – Efeito do Projeto Parto Adequado sobre a taxa de cesárea por região, com IC 95%. Robson 1+3.



Fonte: Elaboração própria a partir do SINASC–CNES integrado e do Anexo A do Relatório ANS (2023).

O retrato dinâmico fecha o conjunto de testes: a Figura 14 sobrepõe as trajetórias dos estimadores de Sun-Abraham e Callaway-Sant’Anna ao longo do tempo relativo à adesão, com os coeficientes pré-tratamento próximos de zero.

Figura 14 – Efeito do PPA por tempo relativo à adesão (*event study*): estimadores de Sun e Abraham (2021) e Callaway e Sant’Anna (2021). Robson 1+3, 2014–2022.



Fonte: Elaboração própria a partir do SINASC–CNES integrado e do Anexo A do Relatório ANS (2023).

A Tabela 14 reporta o estimador de Callaway e Sant’Anna por tempo relativo à adesão, base da trajetória dinâmica da Figura 14.

Tabela 14 – ATT do Projeto Parto Adequado sobre a taxa de cesárea por tempo relativo à adesão, estimador de Callaway e Sant’Anna (2021), Robson 1+3, Brasil, 2014–2022.

ℓ	Cohorts	Tratados (N)	ATT (p.p)	IC95% (p.p)
+0	2015, 2017, 2020	114	–2,25*	[–3,92; –0,55]
+1	2015, 2017, 2020	113	–4,12*	[–6,45; –2,03]
+2	2015, 2017, 2020	112	–5,39*	[–8,00; –2,87]
+3	2015, 2017	102	–4,26*	[–8,12; –0,87]
+4	2015, 2017	103	–4,52*	[–8,56; –0,57]
+5	2015, 2017	101	–4,83*	[–8,52; –1,26]
+6	2015	32	–12,67*	[–21,34; –6,55]
+7	2015	32	–12,04*	[–19,55; –6,57]
ATT médio pós-tratamento agregado: <i>cluster bootstrap</i> sobre o agregado (Bloco I)				
	$\overline{\text{ATT}}_{\text{pós}}$ ($n_{\text{boot}} = 997$)		–6,26*	[–9,50; –3,00]

Fonte: Elaboração própria a partir do SINASC–CNES integrado e do Anexo A do Relatório ANS (2023). Asterisco indica IC95% que não cruza zero.

4.6 Síntese: a leitura conjunta das quatro perguntas

As quatro frentes apontam para o mesmo lugar, a oferta hospitalar. A oferta é a maior fonte de previsibilidade da decisão sobre via de parto. O diferencial entre as categorias institucionais

sobrevive a controles saturados. O Parto Adequado entrega redução causal nos hospitais aderentes. E parte da decisão escapa a qualquer variável observada nos dados administrativos.

A previsibilidade moderada-alta da decisão (AUROC 0,77 saturado) marca o limite do que os dados administrativos enxergam e enquadra os achados causais como descrições parciais. A decomposição indica que a oferta hospitalar carrega cerca de 3,3 vezes mais informação preditiva marginal que o perfil sociodemográfico. A inferência observacional mostra que a oferta não é apenas a maior fonte de variação preditiva: nela se observa um diferencial sistemático compatível com a hipótese de demanda induzida pela oferta (OR 1,54 do privado lucrativo, ATE 15,8 p.p por IPW), refinado pelo exercício das médicas como pacientes (canal informacional via fuga entre hospitais). A inferência quase-experimental indica que o PPA reduz cesárea em hospitais aderentes em -5 a -6 p.p sobre Robson 1+3, com magnitude concentrada no setor privado lucrativo, paralelas validadas no *cohort* informativo e conjunto de testes de robustez consistente em quatro estimadores modernos.

4.7 Limitações e fronteiras da inferência

A leitura interpretativa precisa ser qualificada por seis ressalvas, ordenadas pelo potencial de comprometer a inferência causal. As ressalvas não invalidam o achado central, mas calibram a sua força. O texto a seguir as declara abertamente; nenhuma é inédita na literatura econômica empírica que opera sobre dados administrativos brasileiros.

Primeira ressalva: viés por variáveis omitidas no desenho observacional do diferencial institucional. SINASC e CNES não codificam três dimensões plausivelmente relevantes: a preferência declarada da gestante pela via cesárea, a gravidade clínica não codificada (peso fetal estimado, condições pré-gestacionais) e a cultura institucional do hospital (protocolos internos, conduta histórica do médico). A quantificação parcial está em §4.4.1: Oster $\delta^* = 0,84$ e E-value $\approx 1,79$ indicam efeito robusto a confundimento moderado, mas vulnerável a inobservado de magnitude moderada-alta.

Segunda ressalva: vulnerabilidade do estimador transversal a seleção não observável no nível do hospital. O contraste *between/within* discutido em §4.4.3 é a principal ameaça à análise observacional do diferencial. O painel *within-hospital* sobre 151 hospitais que mudaram de natureza jurídica no período não produz salto contemporâneo comparável ao diferencial *between* de 13,9 p.p, e o desenho tem MDE de 7,1 p.p, suficiente para detectar mais da metade do efeito *between*. Duas leituras coexistem: a leitura conservadora, segundo a qual o diferencial *between* é dominado por seleção inobservável no nível do hospital, e a interpretação SID precisa

ser fortemente moderada; a leitura interpretativa, segundo a qual a natureza jurídica opera por dimensões institucionais estáveis (cultura clínica, contratos, perfil de credenciamento) que mudam lentamente, e o desenho *within* sobre transições administrativas não captura esse mecanismo. O exercício das médicas como pacientes (§4.4.4) entrega evidência adicional contra a leitura simples de autosseleção pró-cesárea, mas não fecha a discriminação entre as duas leituras. A discriminação plena requer instrumento exógeno para a natureza jurídica ou para o ingresso de novas categorias institucionais, fora do alcance do desenho atual.

Terceira ressalva: escopo da identificação quase-experimental do Parto Adequado. O ATT refere-se a hospitais voluntariamente aderentes, à amostra Robson 1+3 e à janela 2014–2022. A generalização para todo o setor privado lucrativo, para Robson 5–10 ou para outros desenhos institucionais não é coberta.

Quarta ressalva: hipótese rival de capacidade gerencial diferencial. O desenho quase-experimental do PPA não separa o canal SID (remuneração diferenciada) do canal de capacidade gerencial entre aderentes e não aderentes. Hospitais lucrativos têm tipicamente mais recursos administrativos para implementar protocolos PPA. Sob essa hipótese, o coeficiente $-6,30$ p.p no lucrativo mediria capacidade de implementação, não força do canal SID. Borem *et al.* (2020) reporta heterogeneidade do efeito por indicadores de gerência prévia, consistente com o canal de capacidade. A discriminação exige variação adicional não disponível no desenho atual.

Quinta ressalva: *selection-on-gains* (adesão correlacionada ao próprio potencial de redução) nos 121 hospitais voluntariamente aderentes. Esta é a principal ameaça ao desenho quase-experimental do PPA e merece destaque em duas frentes. Primeiro, à identificação: se os aderentes já vinham numa trajetória de queda mais acentuada da cesárea, o contrafactual de tendências paralelas é violado e o ATT fica viesado, não apenas pouco generalizável. Segundo, à *validade externa*: mesmo com paralelas válidas, o efeito medido nos voluntários não se transporta para hospitais que não se autosselecionaram. A inspeção do ano-base 2014 (Robson 1+3) entrega evidência empírica direta: a mediana da taxa de cesárea dos que viriam a aderir é de **58,7%**, contra **68,0%** entre os controles elegíveis, padrão que se repete dentro de cada categoria institucional (**34,0%** vs. **38,7%** no público, **50,0%** vs. **68,6%** no privado sem fins lucrativos, **64,7%** vs. **72,7%** no privado lucrativo). Hospitais com cultura clínica pré-existente alinhada à redução de cesárea aderiram preferencialmente ao programa: o PPA, na sua forma voluntária, opera sobre os melhores candidatos da distribuição setorial. O placebo de tratamento (§4.5) descarta artefato estatístico, não *selection-on-gains*.

Uma objeção possível é que essa autosseleção contamina a própria identificação, e não

só a extrapolação. Contra essa leitura, o teste de tendências paralelas restrito ao *cohort* 2017 (único com pré-tratamento informativo) não rejeita o paralelismo ($\chi^2(2) = 0,481$, $p = 0,786$). A leitura calibrada do *Honest DiD* sobre esse *cohort* tolera violações de até $M \approx 6,5$ vezes a maior violação pré-observada. Como a base de teste é estreita (dois coeficientes pré, $N = 79$), a ameaça à identificação fica atenuada, não eliminada. A interpretação de política fica condicionada a esse limite: o ATT estimado constitui envelope superior do efeito esperado em hospitais não autosselecionados, e extensão obrigatória do programa entregaria fração mais modesta do $-5,56$ p.p documentado.

Sexta ressalva: representatividade temporal. A janela 2014–2022 atravessa choques institucionais não controlados (Portaria SAS/MS nº 1.319/2014 (Brasil. Ministério da Saúde. Secretaria de Atenção à Saúde, 2014) e o próprio PPA). A estabilidade pré- e pós-pandemia de COVID-19 foi testada (OR 1,575 em 2014–2019 vs. 1,531 em 2020–2022; $z = 0,14$, $p = 0,887$), sem padrão sistemático.

Nenhuma das seis ressalvas invalida o achado central; juntas, porém, definem seus contornos. O efeito da natureza jurídica privada lucrativa sobre a probabilidade de cesárea é positivo, sistemático e estável a múltiplas estratégias, na ordem de OR 1,54 ou 15,8 p.p, mas vulnerável a seleção não observável no nível do hospital. O efeito causal do PPA sobre os 121 hospitais voluntariamente aderentes é da ordem de -5 a -6 p.p em Robson 1+3, concentrado no setor privado lucrativo, com identificação calibrada e conjunto de testes de robustez consolidado, mas restrito ao escopo do programa e dos voluntários autosselecionados. A passagem para identificação do canal causal específico (financeiro, organizacional ou informacional) requer evidência adicional sobre remuneração direta (dados administrativos da ANS) e sobre instrumentação da decisão de aderir, registradas como agenda na conclusão (seção 5).

5 CONCLUSÃO

Este trabalho partiu de um fato persistente: a taxa brasileira de cesárea está entre as mais altas do mundo e se concentra, de forma sistemática, no setor privado lucrativo, mesmo entre gestantes de baixo risco clínico. A pergunta era em que medida esse diferencial é compatível com a demanda induzida pela oferta e se uma intervenção sobre o arranjo institucional reduz a cesárea. A resposta que emerge das quatro frentes de análise aponta para a mesma direção: não é o perfil de quem dá à luz, mas o hospital onde se dá à luz que organiza o excesso de cesárea no Brasil.

A decisão sobre a via de parto mostrou-se previsível em grau moderado a partir de informações administrativas, e o peso dessa previsibilidade está mais na estrutura do hospital do que no perfil da gestante. O diferencial entre as categorias institucionais sobreviveu ao controle pelo perfil clínico e sociodemográfico, num subgrupo clinicamente comparável, e mostrou-se compatível com a demanda induzida pela oferta, ainda que vulnerável a confundimento não observado de magnitude moderada. A adesão ao Projeto Parto Adequado reduziu a cesárea nos hospitais aderentes, com efeito concentrado justamente no setor privado lucrativo, onde o diferencial se manifesta. As quatro frentes não provam isoladamente o mesmo ponto, mas apontam para o mesmo lugar: a oferta hospitalar é, ao mesmo tempo, a maior fonte de previsibilidade da decisão, o eixo onde o diferencial persiste após os controles e o eixo onde a política entrega efeito.

O mecanismo econômico por trás desse padrão é conhecido. Quando o agente que decide o procedimento concentra a informação clínica e responde a uma remuneração maior por procedimento cirúrgico, o resultado agregado pode se afastar do que seria escolhido sob informação simétrica e remuneração neutra ao tipo de parto. O diferencial brasileiro responde menos ao que a gestante é e mais ao arranjo de incentivos, organização e cultura clínica do estabelecimento que a atende. A intervenção sobre esse arranjo funciona quando aplicada, o que desloca o problema do terreno técnico para o da escolha de política pública: o instrumento existe e produz efeito, mas opera hoje em escala pequena diante do tamanho do fenômeno.

O achado ganha peso quando se observa onde o excesso se concentra. O setor privado lucrativo opera dentro da saúde suplementar. Esse segmento é sustentado em parte por renúncia fiscal federal da ordem de dezenas de bilhões de reais por ano. Esse montante equivale a cerca de um terço do orçamento federal do Sistema Único de Saúde e tem incidência regressiva. O excesso de cesáreas nesse segmento não é apenas um desfecho clínico desfavorável: é um custo gerado em um setor parcialmente financiado por gasto tributário que escapa do orçamento explícito e da prestação de contas direta ao contribuinte. A regulação por desempenho, da qual o Parto

Adequado é uma peça, é o instrumento disponível para corrigir essa distorção.

As conclusões têm fronteiras claras. A parte observacional identifica uma associação compatível com a demanda induzida pela oferta, não a prova do mecanismo, e permanece vulnerável a fatores não observados, como a preferência da gestante e a cultura do hospital. A parte causal vale para os hospitais que aderiram voluntariamente ao programa, para a amostra de baixo risco e para a janela analisada, e não separa o canal financeiro do canal organizacional da intervenção. Confirmar em definitivo o mecanismo causal, a partir da compatibilidade observada e do efeito identificado, é uma etapa ainda pendente.

Essa etapa exige peças adicionais. A mais promissora estende o exercício das gestantes médicas, que aqui sugeriu que a assimetria de informação opera pela escolha do hospital, a um desenho dedicado, com outras profissões da área da saúde e desfechos neonatais. Somam-se a essa frente a instrumentação da decisão de aderir ao programa e o acesso a dados de remuneração por procedimento, que permitiriam separar o incentivo financeiro do organizacional. A extensão do desenho a outros programas de melhoria da via de parto, à medida que os dados de adesão forem publicados, completa a agenda.

A taxa de cesárea no Brasil deixou de ser, há muito, uma variável puramente clínica. O que este trabalho mostra é que seu excesso, na população de menor indicação, tem endereço institucional, e que uma política dirigida a esse endereço produz efeito. Quanto desse excesso vem da demanda induzida pela oferta, e não de seleção ou capacidade gerencial, é o que a evidência aqui reunida qualifica em quatro frentes.

REFERÊNCIAS

- Agência Nacional de Saúde Suplementar (ANS). **Programa Parto Adequado: resultados finais da Fase 1**. Brasília, DF, 2017. Disponível em: <https://www.gov.br/ans/pt-br/assuntos/gestao-em-saude/programa-parto-adequado>.
- _____. **Parto Adequado: estratégias de qualificação da jornada da gestante. A trajetória de 2015 a 2023**. Rio de Janeiro: [S.l.: s.n.], 2023. 151 p. Disponível em: <https://www.gov.br/ans/pt-br/assuntos/noticias/sobre-ans/ans-divulga-relatorio-sobre-parto-adequado>.
- ALTHABE, F. *et al.* Cesarean section rates and maternal and neonatal mortality in low-, medium-, and high-income countries: An ecological study. **Birth**, v. 33, n. 4, p. 270–277, 2006.
- ANTOINE, C.; YOUNG, B. K. Cesarean section one hundred years 1920–2020: the good, the bad and the ugly. **Journal of Perinatal Medicine**, Berlin, v. 49, n. 1, p. 5–16, 2020.
- ARROW, K. J. Uncertainty and the welfare economics of medical care. **The American Economic Review**, Nashville, v. 53, n. 5, p. 941–973, 1963.
- AUSTIN, P. C. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. **Statistics in Medicine**, v. 28, n. 25, p. 3083–3107, 2009.
- BATES, D. *et al.* Fitting linear mixed-effects models using lme4. **Journal of Statistical Software**, v. 67, n. 1, p. 1–48, 2015.
- BÉHAGUE, D. P.; VICTORA, C. G.; BARROS, F. C. Consumer demand for caesarean sections in Brazil: informed decision making, patient choice, or social inequality? a population based birth cohort study linking ethnographic and epidemiological methods. **British Medical Journal**, London, v. 324, n. 7343, p. 942–945, 2002.
- BETRÁN, A. P. *et al.* Who statement on caesarean section rates. **BJOG: An International Journal of Obstetrics and Gynaecology**, London, v. 123, n. 5, p. 667–670, 2016.
- BOERMA, T. *et al.* Global epidemiology of use of and disparities in caesarean sections. **The Lancet**, London, v. 392, n. 10155, p. 1341–1348, 2018.
- BOREM, P. *et al.* A Quality Improvement Initiative to Increase the Frequency of Vaginal Delivery in Brazilian Hospitals. **Obstetrics and Gynecology**, Hagerstown, v. 135, n. 2, p. 415–425, 2020.
- Brasil. Ministério da Saúde. **Sistema de Informações sobre Nascidos Vivos (SINASC): nascidos vivos no Brasil por tipo de parto e ano de nascimento**. Brasília, DF, 2022. Tabulação nvuf.def. Disponível em: <http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sinasc/cnv/nvuf.def>. Acesso em: 25 abr. 2026.
- Brasil. Ministério da Saúde. Secretaria de Atenção à Saúde. **Portaria SAS/MS n. 1.319, de 24 de novembro de 2014: dispõe sobre as regras no Cadastro Nacional de Estabelecimentos de Saúde (CNES) para adequação às normas da Receita Federal do Brasil para registro no CNPJ**. Brasília, 2014. Disponível em: https://bvsms.saude.gov.br/bvs/saudelegis/sas/2014/prt1319_24_11_2014.html. Acesso em: 29 abr. 2026.

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.

CAMPOS, A. S. Q.; RATTNER, D.; DINIZ, C. S. G. Efetividade do programa parto adequado na diminuição das taxas de cesárea de maternidades privadas no município de são paulo, brasil. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 40, n. 9, p. e00216623, 2024.

CASSARDO, O. *et al.* Cesarean delivery for placenta previa. **American Journal of Obstetrics and Gynecology**, New York, v. 233, n. 6S, p. S255–S271, 2026.

Centers for Disease Control and Prevention (CDC). Achievements in public health, 1900-1999: Healthier mothers and babies. **Morbidity and Mortality Weekly Report (MMWR)**, Atlanta, v. 48, n. 38, p. 849–858, 1999. Disponível em: <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm4838a2.htm>.

CHAISEMARTIN, C. de; D'HAULTFÈUILLE, X. Two-way fixed effects estimators with heterogeneous treatment effects. **American Economic Review**, v. 110, n. 9, p. 2964–2996, 2020.

DINIZ, C. S. G. Gênero, saúde materna e o paradoxo perinatal. **Revista Brasileira de Crescimento e Desenvolvimento Humano**, São Paulo, v. 19, n. 2, p. 313–326, 2009.

DOMINGUES, R. M. S. M. *et al.* Processo de decisão pelo tipo de parto no brasil: da preferência inicial das mulheres à via de parto final. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 30, n. Supl. 1, p. S101–S116, 2014.

DRANOVE, D.; WEHNER, P. Physician-induced demand for childbirths. **Journal of Health Economics**, v. 13, n. 1, p. 61–73, 1994.

ENTRINGER, A. P.; PINTO, M. F. T.; GOMES, M. A. d. S. M. Análise de custos da atenção hospitalar ao parto vaginal e à cesariana eletiva para gestantes de risco habitual no sistema único de saúde. **Ciência and Saúde Coletiva**, Rio de Janeiro, v. 24, n. 4, p. 1527–1536, 2019.

EPSTEIN, A. J.; NICHOLSON, S. The formation and evolution of physician treatment styles: An application to cesarean sections. **Journal of Health Economics**, v. 28, n. 6, p. 1126–1140, 2009.

EVANS, R. G. Supplier-induced demand: some empirical evidence and implications. In: PERLMAN, M. (ed.). **The economics of health and medical care**. London: Macmillan, 1974, (International Economic Association Series). p. 162–173.

FILHO, M. B.; RISSIN, A. A OMS e a epidemia de cesarianas. **Revista Brasileira de Saúde Materno Infantil**, Recife, v. 18, n. 1, p. 3–4, 2018.

FREITAS, P. F.; FERNANDES, T. M. B. Associação entre fatores institucionais, perfil da assistência ao parto e as taxas de cesariana em Santa Catarina. **Revista Brasileira de Epidemiologia**, v. 19, n. 3, p. 525–538, 2016.

GETAHUN, D. *et al.* Previous cesarean delivery and risks of placenta previa and placental abruption. **Obstetrics & Gynecology**, v. 107, n. 4, p. 771–778, 2006.

GOLDANI, H. A. S. *et al.* Cesarean delivery is associated with an increased risk of obesity in adulthood in a Brazilian birth cohort study. **The American Journal of Clinical Nutrition**, v. 93, n. 6, p. 1344–1347, 2011.

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

GRUBER, J.; OWINGS, M. Physician financial incentives and cesarean section delivery. **The RAND Journal of Economics**, v. 27, n. 1, p. 99–123, 1996.

HAINMUELLER, J. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. **Political Analysis**, v. 20, n. 1, p. 25–46, 2012.

HAN, K. *et al.* Do hospital characteristics influence cesarean delivery? analysis of national health insurance claim data. **European Journal of Public Health**, v. 27, n. 5, p. 801–807, 2017.

HANNAH, M. E. *et al.* Planned caesarean section versus planned vaginal birth for breech presentation at term: a randomised multicentre trial. **The Lancet**, London, v. 356, n. 9239, p. 1375–1383, 2000.

HEM, E.; BORDAHL, P. E. Max sänger - father of the modern caesarean section. **Gynecologic and Obstetric Investigation**, Basel, v. 55, n. 3, p. 127–129, 2003.

HO, D. E. *et al.* MatchIt: Nonparametric preprocessing for parametric causal inference. **Journal of Statistical Software**, v. 42, n. 8, p. 1–28, 2011.

HOPKINS, K. Are brazilian women really choosing to deliver by cesarean? **Social Science and Medicine**, v. 51, n. 5, p. 725–740, 2000.

HOPKINS, K.; AMARAL, E. F. d. L.; MOURÃO, A. N. The impact of payment source and hospital type on rising cesarean section rates in brazil, 1998-2008. **Birth**, v. 41, n. 2, p. 169–177, 2014.

HORVITZ, D. G.; THOMPSON, D. J. A generalization of sampling without replacement from a finite universe. **Journal of the American Statistical Association**, Alexandria, v. 47, n. 260, p. 663–685, 1952.

Instituto de Estudos de Saúde Suplementar (IESS). **O Parto Adequado: evidências científicas e os seus desdobramentos no Brasil e no Mundo**. São Paulo, 2021. Texto de Discussão n. 85. Disponível em: <https://www.iess.org.br/>.

JOHNSON, E. M.; REHAVI, M. M. Physicians treating physicians: Information and incentives in childbirth. **American Economic Journal Economic Policy**, v. 8, n. 1, p. 115–141, 2016.

KNAU, M. M.; TRIACA, L. M.; SANTOS, A. M. A. d. Avaliação do Parto Adequado: estudo de impacto na redução de nascimentos por cesáreas. In: **XXVI Encontro de Economia da Região Sul — ANPEC Sul 2023**. Brasil: [S.l.: s.n.], 2023. Área 8: Econometria. Disponível em: https://www.anpec.org.br/sul/2023/submissao/files_I/i8-3a40b5935cf909b47dc5b0f61447fb5e.pdf.

KONDO, M. M. *et al.* Reducing caesarean section rates in Robson groups 1 and 3: a quality improvement initiative in a private Brazilian hospital. **BMJ Open Quality**, v. 14, n. 1, p. e003077, 2025.

KOTELCHUCK, M. An evaluation of the Kessner adequacy of prenatal care index and a proposed adequacy of prenatal care utilization index. **American Journal of Public Health**, Washington, v. 84, n. 9, p. 1414–1420, 1994.

LEAL, M. d. C. *et al.* Pesquisa nascer no brasil: apresentação. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 30, p. S5–S7, 2014. Suplemento 1.

LIANG, K.-Y.; ZEGER, S. L. Longitudinal data analysis using generalized linear models. **Biometrika**, Oxford, v. 73, n. 1, p. 13–22, 1986.

LOUDON, I. Maternal mortality in the past and its relevance to developing countries today. **The American Journal of Clinical Nutrition**, Bethesda, v. 72, n. 1, p. 241S–246S, 2000.

LUNCEFORD, J. K.; DAVIDIAN, M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. **Statistics in Medicine**, Chichester, v. 23, n. 19, p. 2937–2960, 2004.

LUNDBERG, S. M.; LEE, S.-I. A unified approach to interpreting model predictions. *In*: **Advances in Neural Information Processing Systems 30 (NIPS 2017)**. Long Beach: [S.l.: s.n.], 2017. p. 4765–4774.

MANGOLIM, F. P. **Avaliação de impacto do Programa Parto Adequado**. 2024. Dissertação (Dissertação (Mestrado Profissional em Economia)) — Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP), Brasília, 2024. Orientadora: Roberta Moreira Wichmann.

MASCARELLO, K. C.; HORTA, B. L.; SILVEIRA, M. F. Complicações maternas e cesárea sem indicação: revisão sistemática e meta-análise. **Revista de Saúde Pública**, v. 51, p. 105, 2017.

MCCALLUM, C. Explaining caesarean section in salvador da bahia, brazil. **Sociology of Health and Illness**, Oxford, v. 27, n. 2, p. 215–242, 2005.

MELO, C.; MENEZES-FILHO, N. The effects of a national policy to reduce C-sections in Brazil. **Health Economics**, v. 32, n. 2, p. 501–517, 2023.

MOLINA, G. *et al.* Relationship between cesarean delivery rate and maternal and neonatal mortality. **JAMA**, Chicago, v. 314, n. 21, p. 2263–2270, 2015.

MULLAINATHAN, S.; SPIESS, J. Machine learning: an applied econometric approach. **Journal of Economic Perspectives**, Nashville, v. 31, n. 2, p. 87–106, 2017.

NAKAMURA-PEREIRA, M. *et al.* Use of Robson classification to assess cesarean section rate in Brazil: the role of source of payment for childbirth. **Reproductive Health**, v. 13, n. Suppl. 3, 2016.

NEGRINI, R. *et al.* Strategies to reduce the caesarean section rate in a private hospital and their impact. **BMJ Open Quality**, London, v. 10, n. 3, p. e001215, 2021.

Organização Mundial da Saúde (OMS). Appropriate technology for birth. **The Lancet**, London, v. 2, n. 8452, p. 436–437, 1985.

_____. **Robson Classification: Implementation Manual**. Geneva, 2017. ISBN 978-92-4-151319-7. Disponível em: <https://www.who.int/publications/i/item/9789241513197>.

OSTER, E. Unobservable selection and coefficient stability: Theory and evidence. **Journal of Business and Economic Statistics**, v. 37, n. 2, p. 187–204, 2019.

PAIXÃO, E. S. *et al.* Associations between cesarean delivery and child mortality: A national record linkage longitudinal study of 17.8 million births in Brazil. **PLOS Medicine**, v. 18, n. 10, p. e1003791, 2021.

_____. Using the Robson classification to assess caesarean section rates in Brazil: an observational study of more than 24 million births from 2011 to 2017. **BMC Pregnancy and Childbirth**, v. 21, p. 630, 2021.

PEREIRA, V. B. *et al.* Tendência da taxa de cesariana no brasil por grupo de classificação de robson, 2014-2020. **Revista Brasileira de Enfermagem**, Brasília, DF, v. 77, n. 3, p. e20230099, 2024.

RAMBACHAN, A.; ROTH, J. A more credible approach to parallel trends. **The Review of Economic Studies**, v. 90, n. 5, p. 2555–2591, 2023.

RASPANTINI, P. R. *et al.* O impacto do tipo de hospital e tipo de parto sobre a idade gestacional ao nascer no município de são paulo, 2013-2014. **Revista Brasileira de Epidemiologia**, v. 19, n. 4, p. 878–882, 2016.

REYMAN, M. *et al.* Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life. **Nature Communications**, v. 10, n. 1, p. 4997, 2019.

ROBINS, J. M.; ROTNITZKY, A.; ZHAO, L. P. Estimation of regression coefficients when some regressors are not always observed. **Journal of the American Statistical Association**, v. 89, n. 427, p. 846–866, 1994.

ROBINSON, W. S. Ecological correlations and the behavior of individuals. **American Sociological Review**, Washington, v. 15, n. 3, p. 351–357, 1950.

ROBSON, M. S. Classification of caesarean sections. **Fetal and Maternal Medicine Review**, Cambridge, v. 12, n. 1, p. 23–39, 2001.

ROCHA, A. S. *et al.* Cesarean sections and early-term births according to Robson classification: a population-based study with more than 17 million births in Brazil. **BMC Pregnancy and Childbirth**, v. 23, p. 562, 2023.

ROCHA, N. F. F. d.; FERREIRA, J. A escolha da via de parto e a autonomia das mulheres no brasil: uma revisão integrativa. **Saúde em Debate**, Rio de Janeiro, v. 44, n. 125, p. 556–568, 2020.

ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. **Biometrika**, Oxford, v. 70, n. 1, p. 41–55, 1983.

RUSCONI, F. *et al.* Mode of delivery and asthma at school age in 9 European birth cohorts. **American Journal of Epidemiology**, v. 185, n. 6, p. 465–473, 2017.

SANDALL, J. *et al.* Short-term and long-term effects of caesarean section on the health of women and children. **The Lancet**, London, v. 392, n. 10155, p. 1349–1357, 2018.

SILVA, A. C. L. *et al.* Preferência pelo tipo de parto, fatores associados à expectativa e satisfação com o parto. **Revista Eletrônica de Enfermagem**, Goiânia, v. 19, p. e44139, 2017.

SILVA, R. P. d.; PAZIN-FILHO, A. Taxa e custos médicos diretos de cesáreas em beneficiárias da saúde suplementar no estado de são paulo, brasil: 2015 a 2021. **Ciência and Saúde Coletiva**, Rio de Janeiro, v. 29, n. 1, p. e00512023, 2024.

SNIJDERS, T. A. B.; BOSKER, R. J. **Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling**. 2. ed. London: SAGE Publications, 2012.

SOUZA, J. P. *et al.* A global reference for caesarean section rates (C-Model): a multicountry cross-sectional study. **BJOG: An International Journal of Obstetrics & Gynaecology**, Chichester, v. 123, n. 3, p. 427–436, 2016.

STUART, E. A. Matching methods for causal inference: A review and a look forward. **Statistical Science**, Beachwood, v. 25, n. 1, p. 1–21, 2010.

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.

TODMAN, D. A history of caesarean section: from ancient world to the modern era. **Australian and New Zealand Journal of Obstetrics and Gynaecology**, Sydney, v. 47, n. 5, p. 357–361, 2007.

VANDERWEELE, T. J.; DING, P. Sensitivity analysis in observational research: Introducing the E-value. **Annals of Internal Medicine**, v. 167, n. 4, p. 268–274, 2017.

VOGEL, J. P. *et al.* Use of the robson classification to assess caesarean section trends in 21 countries: a secondary analysis of two who multicountry surveys. **The Lancet Global Health**, London, v. 3, n. 5, p. e260–e270, 2015.

WOOLDRIDGE, J. M. **Econometric Analysis of Cross Section and Panel Data**. 2. ed. Cambridge, MA: MIT Press, 2010.

YE, J. *et al.* Association between rates of caesarean section and maternal and neonatal mortality in the 21st century: a worldwide population-based ecological study with longitudinal data. **BJOG: An International Journal of Obstetrics and Gynaecology**, London, v. 123, n. 5, p. 745–753, 2016.

Apêndices

APÊNDICE A – ESTATÍSTICA DESCRITIVA COMPLEMENTAR

A seção 3 apresenta a análise exploratória na extensão necessária ao argumento. Este apêndice reúne o material descritivo complementar levantado na exploração das bases SINASC e CNES, preservado aqui para consulta sem interromper a leitura do corpo. Salvo indicação de ano na legenda, todas as peças usam a base integrada 2014–2022 descrita na §3.1.

A.1 Série histórica e totais nacionais

A Tabela A1 consolida os totais de nascimentos e a Tabela A2, a série anual da taxa de cesárea.

Tabela A1 – Nascimentos hospitalares e tipo de parto, Brasil, 2010–2022.

	Frequência	Participação
Nascidos vivos hospitalares	36.767.563	100,00%
Parto vaginal	15.997.671	43,51%
Parto cesárea	20.730.362	56,38%
Tipo de parto não informado	39.530	0,11%
Cesárea sobre partos válidos	20.730.362	56,44%

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Tabela A2 – Nascimentos hospitalares e taxa de cesárea por ano, Brasil, 2010–2022.

Ano	Total	Vaginal	Cesárea	% cesárea
2010	2.831.627	1.333.357	1.495.263	52,86%
2011	2.883.255	1.311.614	1.565.217	54,41%
2012	2.877.780	1.256.441	1.615.761	56,26%
2013	2.877.338	1.227.701	1.644.446	57,25%
2014	2.951.939	1.250.539	1.697.857	57,59%
2015	2.991.293	1.313.789	1.673.943	56,03%
2016	2.831.278	1.246.245	1.582.900	55,95%
2017	2.896.754	1.267.712	1.627.203	56,21%
2018	2.918.372	1.269.237	1.647.441	56,48%
2019	2.823.136	1.217.311	1.604.136	56,86%
2020	2.702.296	1.137.986	1.562.245	57,86%
2021	2.647.983	1.120.373	1.526.302	57,67%
2022	2.534.512	1.045.366	1.487.648	58,73%

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

A.2 Distribuição espacial

As Tabelas A3 e A4 detalham a taxa de cesárea por região e por unidade da federação, complementando o mapa da §3.2.

Tabela A3 – Taxa de cesárea por região, Brasil: 2010, 2022, variação 2010–2022 e agregado do período.

Região	2010	2022	Variação (p.p)	Agregado
Centro-Oeste	57,72%	64,48%	+6,76	62,49%
Sul	58,30%	62,77%	+4,47	61,61%
Sudeste	58,42%	59,66%	+1,24	59,61%
Nordeste	44,82%	55,83%	+11,01	50,78%
Norte	43,91%	53,00%	+9,09	48,50%

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Tabela A4 – Taxa de cesárea por unidade da federação, Brasil: 2010, 2022, variação 2010–2022 e agregado do período.

UF	2010	2022	Variação (p.p)	Agregado
GO	61,74%	70,81%	+9,07	67,08%
RO	64,34%	69,06%	+4,72	66,72%
ES	60,89%	60,82%	-0,08	62,91%
PR	58,59%	65,26%	+6,66	62,60%
RS	58,30%	64,51%	+6,22	62,47%
MT	56,52%	65,05%	+8,53	61,96%
MS	56,64%	63,86%	+7,22	61,60%
RJ	60,87%	59,87%	-1,00	60,15%
SP	59,06%	59,92%	+0,85	59,96%
RN	51,25%	65,73%	+14,47	59,94%
SC	57,80%	57,03%	-0,77	58,75%
PB	52,07%	64,35%	+12,27	58,12%
MG	54,34%	58,69%	+4,36	57,64%
CE	49,31%	64,45%	+15,15	57,54%
PI	46,89%	61,76%	+14,88	54,46%
DF	52,05%	54,24%	+2,19	54,46%
AL	53,01%	56,13%	+3,11	54,39%
TO	44,43%	57,58%	+13,14	52,67%
PE	50,29%	52,70%	+2,41	51,81%
PA	44,17%	56,98%	+12,81	51,01%
MA	35,83%	55,36%	+19,53	45,64%
BA	38,83%	47,90%	+9,08	43,95%
SE	36,03%	44,97%	+8,94	42,77%
AC	36,77%	51,02%	+14,25	42,47%
AM	40,50%	44,05%	+3,55	41,03%
RR	39,08%	39,25%	+0,17	38,83%
AP	30,79%	39,73%	+8,94	36,06%

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

A.3 Estrutura institucional da rede

As Tabelas A5 e A6 detalham a composição da rede por natureza jurídica e por perfil de atendimento, e as Tabelas A7 e A8, a infraestrutura instalada por ano e por natureza jurídica, complementando a §3.3.

Tabela A5 – Estabelecimentos com nascimentos registrados, nascidos vivos e densidade obstétrica por região, Brasil, 2010 e 2022.

Região	Hospitais (N)		Nascidos vivos		Densidade ^a	
	2010	2022	2010	2022	2010	2022
Norte	679	642	284.089	273.581	239,0	234,7
Nordeste	1.900	1.722	823.601	696.578	230,7	247,2
Sudeste	1.685	1.563	1.099.073	960.888	153,3	162,7
Sul	951	783	362.954	355.801	262,0	220,1
Centro-Oeste	621	634	177.645	213.872	349,6	296,4
Brasil	5.836	5.344	2.747.362	2.500.720	212,4	213,7

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Nota: ^aDensidade = hospitais por 100 mil nascidos vivos.

Tabela A6 – Distribuição de hospitais e de nascidos vivos por perfil de atendimento^a, Brasil, 2010 e 2022.

Perfil	Hospitais (%)		Nascidos vivos (%)	
	2010	2022	2010	2022
SUS exclusivo	41,2%	44,4%	44,0%	50,7%
Misto	31,2%	24,4%	39,8%	33,1%
Privado exclusivo	11,4%	10,9%	13,9%	14,1%
Sem registro	16,1%	20,3%	2,3%	2,2%
Total (N)	5.863	5.344	2.831.627	2.534.512

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Nota: ^aPerfil derivado dos indicadores CNES de internação: SUS exclusivo, Misto (SUS combinado a particular ou plano), Privado exclusivo (sem SUS) e Sem registro.

Tabela A7 – Indicadores de infraestrutura instalada por perfil de atendimento, hospitais com nascimentos registrados, Brasil, 2010 e 2022.

Perfil	2010			2022		
	N hosp.	% c.o. ^a	Med. cir. ^b	N hosp.	% c.o. ^a	Med. cir. ^b
SUS exclusivo	2.417	92,3%	4	2.375	88,3%	4
Misto	1.830	95,4%	10	1.304	91,6%	13
Privado exclusivo	670	79,3%	10	581	76,9%	10
Sem registro	946	6,1%	0	1.084	3,2%	0

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Nota: ^a“c.o.” = centro obstétrico instalado. ^b“Med. cir.” = mediana de leitos cirúrgicos por hospital.

Tabela A8 – Indicadores de infraestrutura instalada por perfil de atendimento^a e natureza jurídica^b, hospitais com nascimentos registrados, Brasil, 2022.

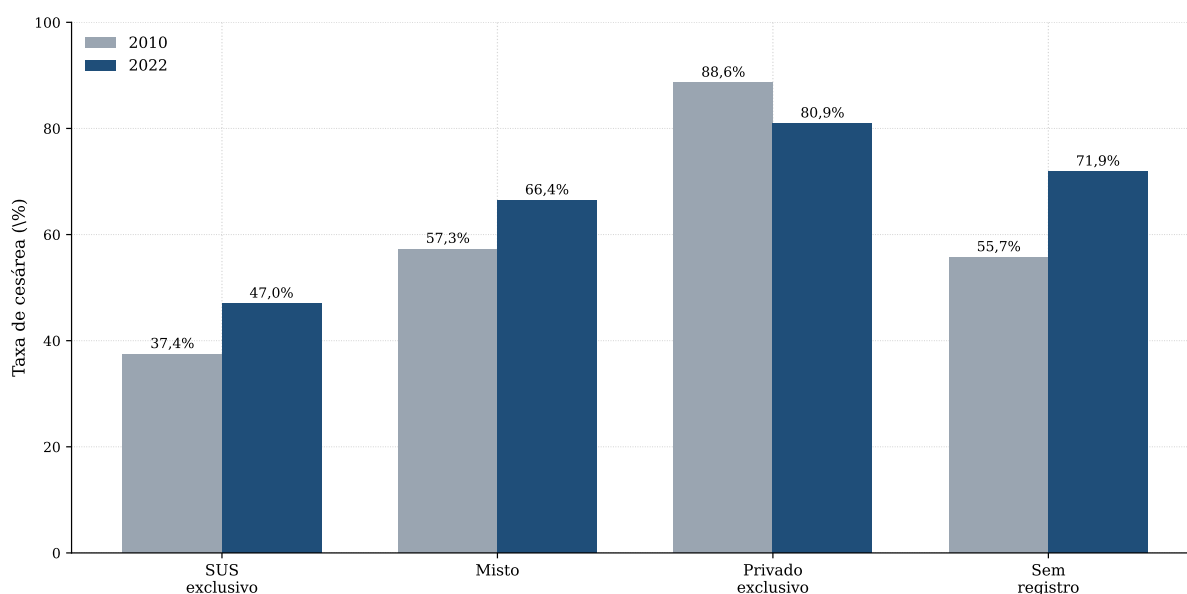
Perfil	Natureza jurídica	N hosp.	% c.o. ^c	Med. cir. ^d
SUS exclusivo	Público	2.216	88,1%	4
SUS exclusivo	Priv. sem fins lucr.	139	90,6%	9
Misto	Priv. sem fins lucr.	1.049	93,1%	13
Misto	Priv. lucrativo	228	84,2%	15
Privado exclusivo	Priv. sem fins lucr.	42	69,0%	11
Privado exclusivo	Priv. lucrativo	536	77,6%	10

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Nota: ^aPerfil definido na Tabela A6. ^bCategorização tripartite definida na Quadro 2. ^c“c.o.” = centro obstétrico instalado. ^d“Med. cir.” = mediana de leitos cirúrgicos por hospital.

A Figura A1 mostra a taxa de cesárea por perfil de atendimento do hospital.

Figura A1 – Taxa de cesárea por perfil de atendimento do hospital.



Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

A.4 Locomoção da gestante

A Tabela A9 descreve a proporção de partos ocorridos fora do município de residência da gestante.

Tabela A9 – Locomoção da gestante e tipo de parto, Brasil: 2010, 2022, variação 2010–2022 e agregado do período.

Indicador	2010	2022	Variação (p.p)	Agregado
Taxa de deslocamento intermunicipal	27,05%	33,57%	+6,52	30,79%
Taxa de cesárea entre deslocadas	57,83%	63,41%	+5,58	61,07%
Taxa de cesárea entre não deslocadas	51,02%	56,37%	+5,35	54,38%

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

A.5 Classificação de Robson

As Tabelas A10, A11 e A12 detalham a distribuição dos grupos de Robson, a decomposição da contribuição de cada grupo para a taxa agregada e a taxa por número de cesáreas anteriores, complementando a §3.4.

Tabela A10 – Distribuição de partos e taxa de cesárea por grupo de Robson, Brasil, 2022.

Grupo	Descrição	N partos	% total	Taxa ces. (%)
01	Nulíparas, cefálica, ≥37 sem., trab. espontâneo	424.042	17,1%	44,6%
02	Nulíparas, cefálica, ≥37 sem., induzido ou ces. eletiva	347.986	14,0%	73,9%
03	Múltiparas (sem ces. ant.), cefálica, ≥37 sem., trab. espontâneo	476.731	19,2%	20,0%
04	Múltiparas (sem ces. ant.), cefálica, ≥37 sem., induzido ou ces. eletiva	230.793	9,3%	52,2%
05	Cesárea anterior, cefálica, ≥37 sem., qualquer início	620.727	25,0%	86,1%
06	Nulíparas, apresentação pélvica	33.380	1,3%	92,2%
07	Múltiparas (com ou sem ces. ant.), pélvica	49.119	2,0%	89,5%
08	Gestações múltiplas (incl. com ces. ant.)	56.559	2,3%	86,8%
09	Apresentação transversa ou oblíqua	4.757	0,2%	97,1%
10	Pré-termo (cefálica, <37 sem.)	240.391	9,7%	56,0%
Brasil (total grupos 1–10)		2.484.485	100,0%	58,8%

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Tabela A11 – Decomposição da variação da taxa de cesárea entre 2012 e 2022 nos grupos de Robson 1–10, Brasil.

Grupo	Peso p_g (%)		Taxa t_g (%)		Contribuição (p.p)	
	2012	2022	2012	2022	Composição	Intra-grupo
01	17,0	17,1	49,9	44,6	0,02	-0,90
02	19,3	14,0	67,6	73,9	-3,78	1,06
03	14,1	19,2	18,1	20,0	0,97	0,32
04	9,2	9,3	35,3	52,2	0,06	1,56
05	19,1	25,0	86,3	86,1	5,03	-0,04
06	1,8	1,3	89,4	92,2	-0,42	0,04
07	1,8	2,0	83,4	89,5	0,14	0,12
08	2,2	2,3	81,5	86,8	0,08	0,12
09	0,3	0,2	96,3	97,1	-0,13	0,00
10	15,1	9,7	48,5	56,0	-2,84	0,94
Soma das contribuições (p.p)					-0,86	3,21
Variação observada (58,75% – 56,41%)					2,34 p.p	

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

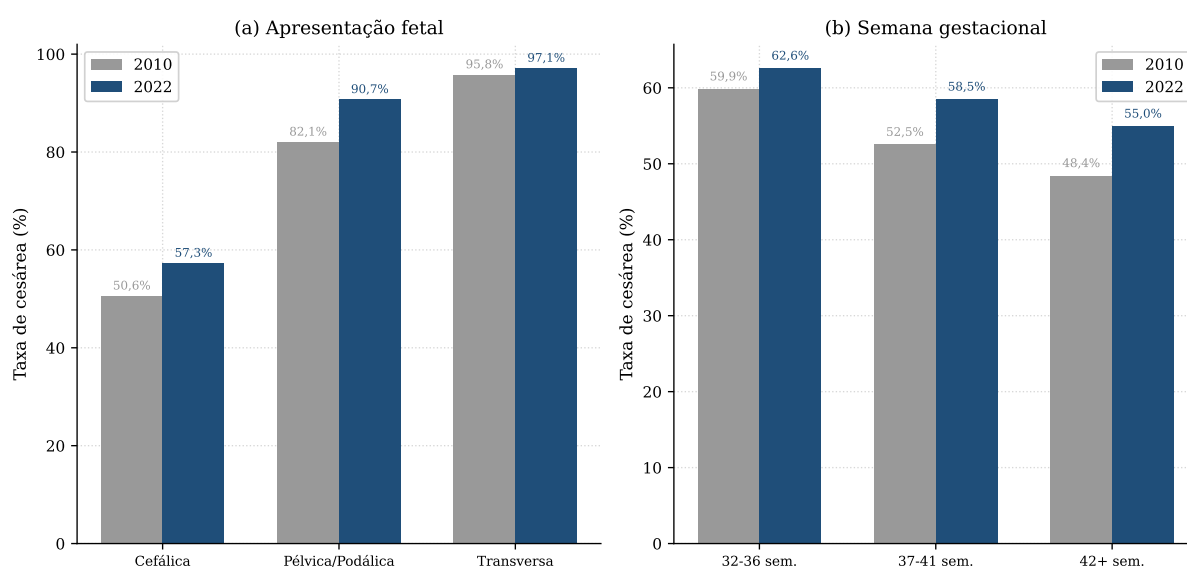
Tabela A12 – Taxa de cesárea segundo número de cesáreas anteriores declaradas, Brasil, 2022.

Ces. anteriores	N partos	N cesáreas	Taxa cesárea (%)
0 cesáreas anteriores	1.719.371	811.621	47,2%
1 cesárea anterior	547.545	453.269	82,8%
2 cesáreas anteriores	156.911	151.490	96,5%
3 ou mais cesáreas anteriores	39.640	38.324	96,7%

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

A Figura A2 mostra a distribuição da apresentação fetal e da idade gestacional.

Figura A2 – Distribuição da apresentação fetal e da idade gestacional.



Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

A.6 Perfil sociodemográfico e desfechos neonatais

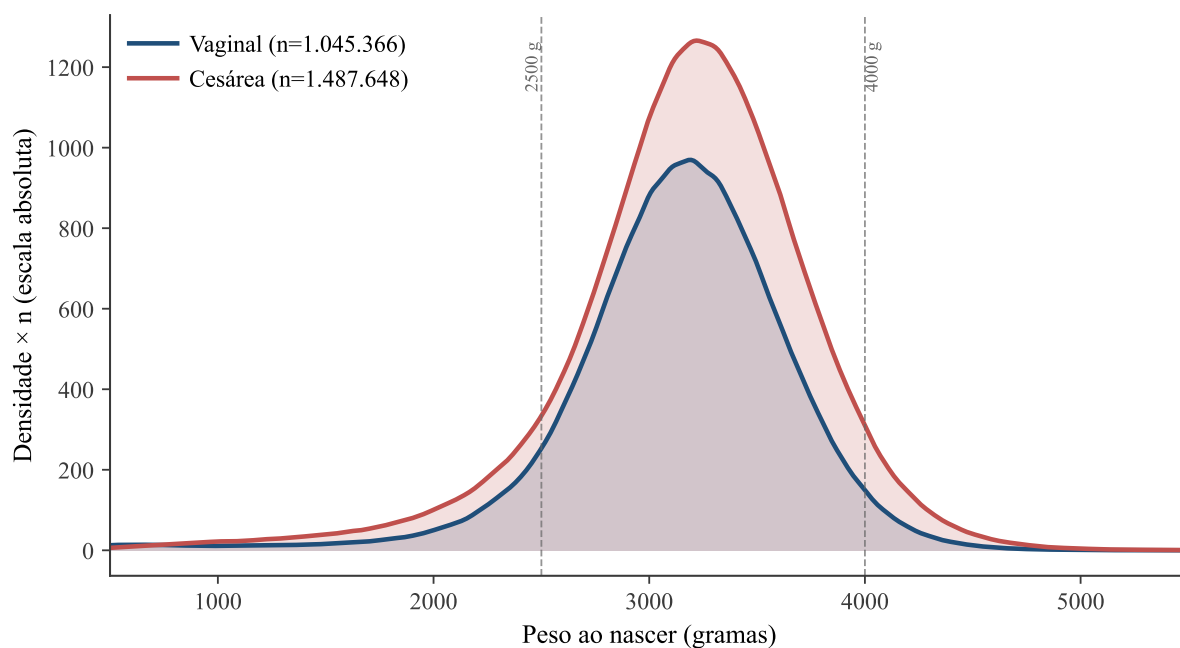
A Tabela A13 cruza o Apgar com a via de parto e a Figura A3 detalha a distribuição do peso ao nascer por via, complementando a §3.5.

Tabela A13 – Distribuição de APGAR ao 1º e ao 5º minuto por tipo de parto, Brasil, 2022.

Categoria	APGAR 1º minuto		APGAR 5º minuto	
	Vaginal	Cesárea	Vaginal	Cesárea
0–3 (depressão grave)	15.766 (1,52%)	16.824 (1,14%)	4.993 (0,48%)	3.454 (0,23%)
4–6 (depressão moderada)	43.911 (4,24%)	66.957 (4,52%)	8.070 (0,78%)	8.885 (0,60%)
7–10 (vitalidade adequada)	975.604 (94,24%)	1.397.982 (94,35%)	1.022.977 (98,74%)	1.469.243 (99,17%)
Total válido	1.035.281	1.481.763	1.036.040	1.481.582

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Figura A3 – Distribuição do peso ao nascer por via de parto.



Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

A.7 Decomposição preditiva por blocos

A Tabela A14 reporta o ganho preditivo marginal de cada bloco de variáveis (Δ AUROC) na decomposição por ablação discutida em §4.3, complementando a Figura 6.

Tabela A14 – Decomposição preditiva por subconjuntos de variáveis (ablation), XGBoost, holdout temporal 2021–2022, Robson 1+3.

Bloco de variáveis	AUROC	ΔAUROC	AUPRC	Brier
0. Clínico mínimo	0,6856	—	0,4728	0,1941
1. + sociodemográfico	0,6979	+0,0122	0,4967	0,1929
2. + oferta hospitalar	0,7383	+0,0404	0,5654	0,1820
3. + tempo e espaço	0,7673	+0,0290	0,6033	0,1730

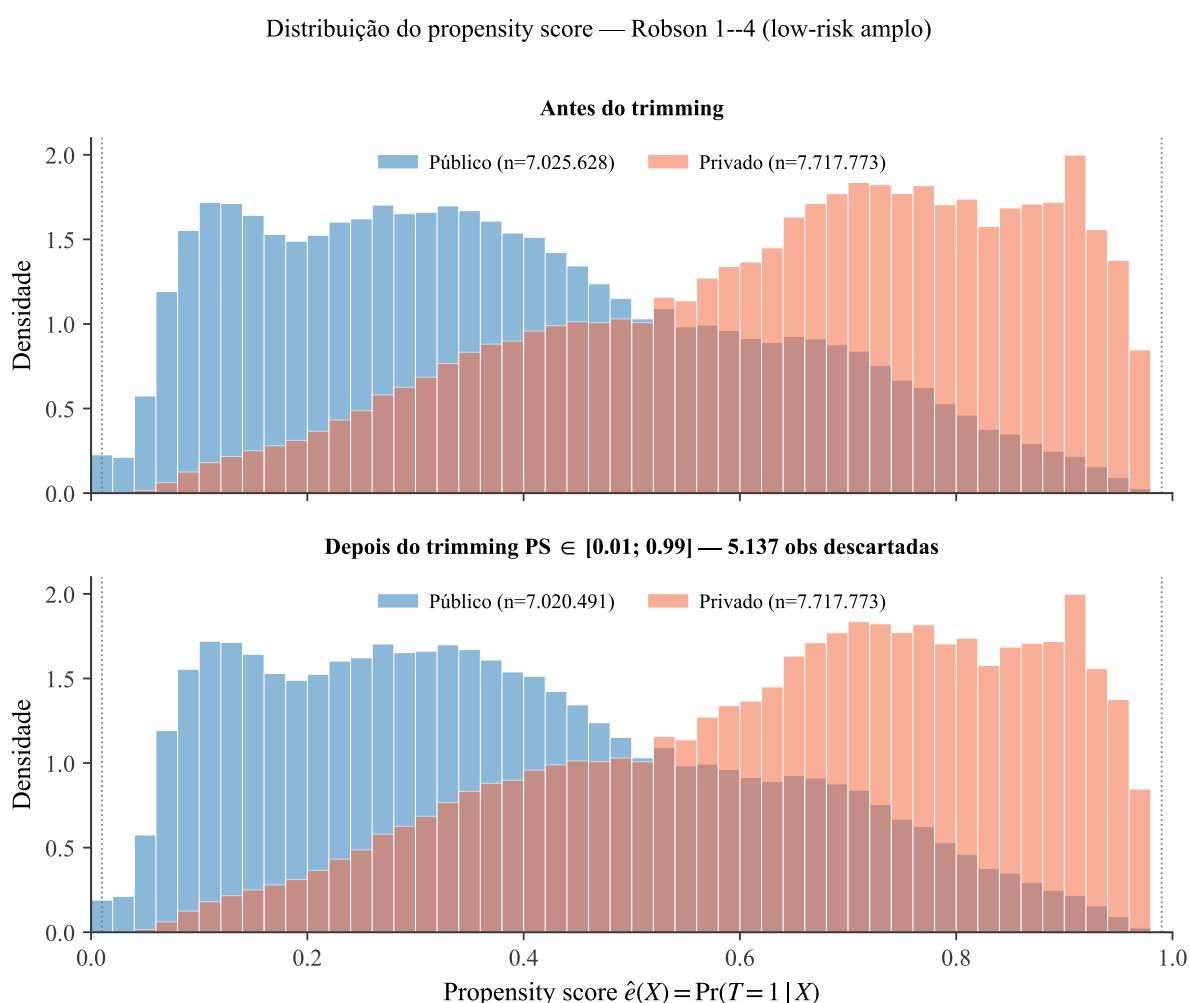
Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Nota: Δ AUROC mede o ganho marginal de cada bloco em relação ao anterior.

APÊNDICE B – DIAGNÓSTICOS DE SUPORTE COMUM

A validade dos estimadores baseados em escore de propensão (IPW, AIPW e PSM; §3.10.1) depende de sobreposição (*common support*) entre tratados e controles na distribuição do escore. A distribuição do escore na amostra principal (Robson 1+3) está no corpo (Figura 9); as Figuras B1 e B2 mostram a distribuição nas duas amostras de sensibilidade. A sobreposição é ampla nas três, e o *trimming* entre os percentis 0,01 e 0,99 adotado na estimação remove apenas as caudas, sem comprometer o suporte comum.

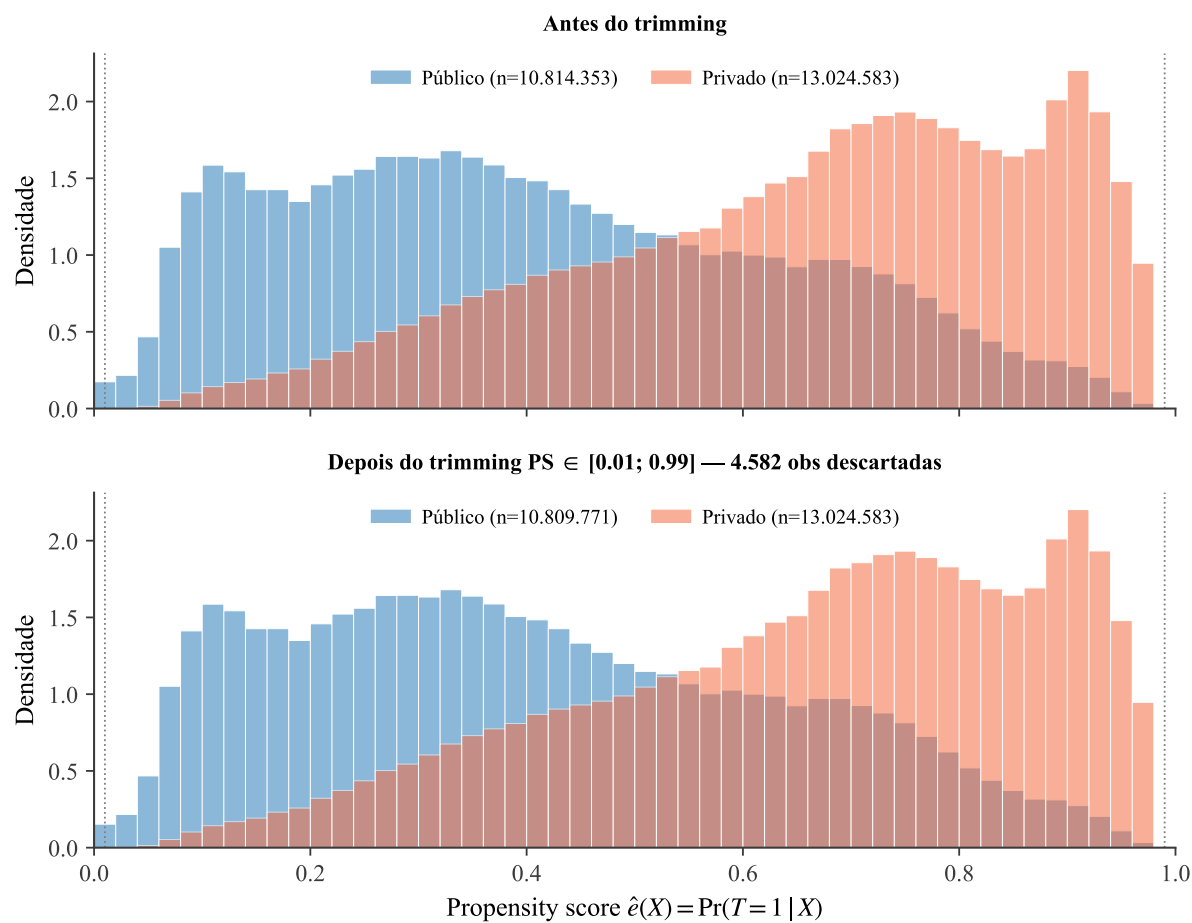
Figura B1 – Distribuição do escore de propensão por setor, amostra Robson 1–4 (baixo risco amplo).



Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Figura B2 – Distribuição do escore de propensão por setor, amostra Robson 1–10 (inclui o Grupo 5).

Distribuição do propensity score — Robson 1--10 (total)



Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

APÊNDICE C – ROBUSTEZ COMPLEMENTAR DA ESTIMAÇÃO CAUSAL

Este apêndice reúne as tabelas completas dos exercícios de robustez e sensibilidade da estimação causal da seção 4, deslocadas do corpo para preservar a fluidez da leitura. O corpo mantém os números-título; aqui ficam as tabelas integrais.

C.1 Sensibilidade a inobservados

A §4.4.1 resume os dois exercícios de sensibilidade a confundimento não observado. A Tabela C1 traz os *bounds* de Oster (δ^* e β^*) e a Tabela C2, o E-value, ambos nas três amostras analíticas.

Tabela C1 – Bounds de viés de variáveis omitidas no LPM (estabilidade de coeficientes), Robson 1+3, Brasil, 2014–2022.

Tratamento (ref. Público)	β_{naive}	$\tilde{\beta}$	$\beta^*(\delta=1)$	δ^*
Privado SFL	13,22 p.p	2,05 p.p	-3,10 p.p	0,40
Privado lucrativo	34,73 p.p	9,67 p.p	-1,90 p.p	0,84

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

Nota: $\beta^*(\delta=1)$: coeficiente sob seleção em inobservados igual à seleção em observados. δ^* : razão de seleção que zera o coeficiente; por convenção, $\delta^* > 1$ indica achado robusto a viés por inobservados.

Tabela C2 – E-value para confundimento não observado, privado lucrativo vs. público (logit Especificação 3).

Amostra	OR observada	RR aprox. $\sqrt{\text{OR}}$	E-value pontual	E-value IC inferior
Robson 1+3	1.543 [1.196; 1.989]	1.242	1.79	1.41
Robson 1–4	1.552 [1.235; 1.952]	1.246	1.80	1.46
Robson 1–10	1.521 [1.244; 1.861]	1.233	1.77	1.47

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022. E-value conforme VanderWeele e Ding (2017).

C.2 Estabilidade temporal do diferencial

A Tabela C3 reestima o logit saturado (Especificação 3) ano a ano e por janela pré- e pós-pandemia. O diferencial do privado lucrativo mantém-se estável ao longo de 2014–2022, o que descarta a leitura de que o achado seja artefato de um período específico (a ressalva de representatividade temporal é discutida na §4.7).

Tabela C3 – Estabilidade temporal do diferencial privado lucrativo: logit Especificação 3 por ano e por janela, Robson 1+3, 2014–2022.

Recorte	N (mi)	Hosp.	OR Lucr (IC95)	p
<i>Por ano (Spec. 3 estratificada, sem EF de ano):</i>				
2014	0,99	4.196	1,82 [1,33; 2,48]	<0,001
2015	1,04	4.106	1,65 [1,21; 2,24]	0,002
2016	1,00	4.098	1,55 [1,17; 2,07]	0,003
2017	1,04	4.081	1,40 [1,04; 1,87]	0,025
2018	1,06	4.038	1,41 [1,04; 1,93]	0,028
2019	1,02	3.957	1,48 [1,11; 1,97]	0,007
2020	0,96	3.961	1,51 [1,14; 2,01]	0,004
2021	0,95	4.021	1,48 [1,09; 2,01]	0,013
2022	0,88	3.950	1,53 [1,13; 2,08]	0,006
<i>Por janela (Spec. 3 saturada com EF de ano dentro da janela):</i>				
Pré-COVID (2014–2019)	6,15	4.974	1,575 [1,205; 2,058]	<0,001
Pós-COVID (2020–2022)	2,79	4.464	1,531 [1,156; 2,028]	0,003
<i>Diferença pré × pós (teste informal de Chow):</i>				
Δ coeficiente	0,028 (escala log-odds), $z = 0,14$, $p = 0,887$			

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022.

C.3 Robustez do DiD do Projeto Parto Adequado

A inferência causal do efeito do Parto Adequado (§4.5) é submetida a uma bateria de robustez. A Tabela C4 traz o teste de tendências paralelas restrito ao cohort 2017 (único com pré-tratamento informativo), e a Tabela C5 consolida os estimadores modernos, os dois placebos, os nove cenários de elegibilidade e as heterogeneidades por região e escolaridade.

Tabela C4 – Teste de Wald conjunto sobre tendências paralelas, cohort 2017, Robson 1+3, Brasil, 2014–2022.

Estatística	Valor
<i>Estimativas pontuais (event-times pré, cohort 2017 isolado)</i>	
$\hat{\delta}_{c=2017, \ell=-3}$	+0,86 p.p
$\hat{\delta}_{c=2017, \ell=-2}$	–0,02 p.p
<i>Teste conjunto $H_0 : \delta_{c=2017, \ell=-3} = \delta_{c=2017, \ell=-2} = 0$</i>	
Estatística qui-quadrado	$\chi^2(2) = 0,481$
Graus de liberdade	2
p-valor	0,786
Decisão a 5%	não rejeita H_0
<i>Réplicas usadas no bootstrap clusterizado da matriz de covariância</i>	
n_{boot} pré-tratamento	1 000

Fonte: Elaboração própria a partir do SINASC–CNES integrado e do Anexo A do Relatório ANS (2023).

Tabela C5 – Bateria de robustez do efeito do Projeto Parto Adequado sobre a taxa de cesárea: estimadores modernos de DiD, placebos, elegibilidade do controle e heterogeneidade, Robson 1+3, Brasil, 2014–2022.

Bloco	Estimador / Especificação	Efeito (p.p)	IC95% (p.p)
<i>(A) Estimadores modernos de DiD com tratamento escalonado heterogêneo</i>			
TWFE clássico (replica Knau et al. 2023)	$\hat{\beta}$ agregado	-2,87*	[-5,71; -0,04]
Sun & Abraham (2021), IW	$\overline{ATT}_{pós}$ (Bloco I)	-5,56*	[-8,55; -2,64]
Callaway & Sant’Anna (2021), regressão de desfecho vs. never-treated	$\overline{ATT}_{pós}$ (Bloco I)	-6,26*	[-9,50; -3,00]
de Chaisemartin & D’Haultfoeuille (2020), DID _M	switchers vs. stayers	-2,17*	[-4,07; -0,49]
<i>(B) Goodman-Bacon (2021): decomposição do TWFE clássico</i>			
Componente (I) Treated × Never-treated	peso 0,56, contrib +1,07	—	—
Componente (II) Earlier × Later not-yet	peso 0,22, contrib -3,15	—	—
Componente (III) Later × Earlier already-treated	peso 0,22, contrib -0,12	—	—
<i>Componente problemático (III) responde por 22% do TWFE clássico</i>			
<i>(C) Honest DiD (Rambachan-Roth 2023): sensibilidade à violação parcial de paralelas</i>			
Maior violação pré observada (agregado, $\ell \in \{-3, -2\}$)	$\delta_{max}^{pré}$	2,58	—
Breakdown M (agregado) sobre $\overline{ATT}_{pós}$	M crítico, SE bootstrap clusterizado	0,98	—
Maior violação pré (calibrada cohort 2017, $\ell \in \{-3, -2\}$)	$\delta_{max, c=2017}^{pré}$	0,86	—
Breakdown M efetivo (calibrado cohort 2017)	M crítico (Bloco I)	≈ 6,5	—
<i>Leitura dupla: $M = 0,98$ no agregado IW; $M \approx 6,5$ restrito ao cohort 2017 informativo</i>			
<i>(D) Placebos</i>			
Placebo de tratamento (permutação, 500 réplicas)	média sob H_0	+0,42	[-2,58; +3,53]
<i>p-valor empírico do efeito real (-2,87 p.p): 0,010</i>			
Placebo de outcome (Robson 5: cesárea anterior)	TWFE Spec A em R5	-0,46	[-1,06; +0,18]
<i>(E) Robustez ao critério ANS de elegibilidade do controle (9 cenários)</i>			
Variando ano-base $\in \{2012, 2013, 2014\}$ e limiares	mediana / [min; max]	-2,87	[-3,86; -1,89]
<i>Sinal preservado em 9/9 cenários; significância em 6/9</i>			
<i>(F) Heterogeneidade por região</i>			
Norte	$\hat{\beta}$ TWFE	-10,70*	[-15,76; -0,51]
Nordeste	$\hat{\beta}$ TWFE	-5,35*	[-11,16; -0,65]
Sudeste	$\hat{\beta}$ TWFE	-1,69	[-5,67; +1,99]
Sul	$\hat{\beta}$ TWFE	-0,96	[-10,33; +5,70]
Centro-Oeste	$\hat{\beta}$ TWFE	-5,27*	[-14,45; -0,88]
<i>(G) Heterogeneidade por escolaridade da mãe</i>			
Fundamental	$\hat{\beta}$ TWFE	-2,20	[-5,58; +1,28]
Médio	$\hat{\beta}$ TWFE	-3,56*	[-6,33; -1,13]
Superior	$\hat{\beta}$ TWFE	-2,21	[-8,15; +2,69]

Fonte: Elaboração própria a partir do SINASC–CNES integrado e do Anexo A do Relatório ANS (2023). Asterisco indica IC95% que não cruza zero.

Nota: Painel hospital × ano com 121 hospitais tratados e 315 hospitais controle elegíveis pelo critério ANS.

Como extensão dessa bateria, a Tabela C6 estima o efeito do PPA por nível de informação obstétrica da gestante (escala ordinal L0–L4 sobre a ocupação registrada no SINASC), conectando o achado quase-experimental do Parto Adequado ao canal informacional explorado em §4.4.4.

Tabela C6 – Efeito do Projeto Parto Adequado sobre a taxa de cesárea por nível de informação obstétrica da gestante (L0–L4), setor privado, Robson 1+3, Brasil, 2014–2022.

Nível	Descrição	N trat.	N contr.	Efeito (p.p)	IC95% (p.p)
L0	Superior, ocupação não saúde (controle)	120	305	–3,30	[–11,12; +5,55]
L1	Outras ocupações de saúde superior	117	266	–0,26	[–5,86; +5,66]
L2	Enfermeiras superior e técnicas de enfermagem	118	291	–3,41	[–8,05; +0,06]
L3	Outras médicas (família CBO 2231 sem ginecologia)	77	140	–1,79	[–13,80; +8,31]
L4	Ginecologistas-obstetras (CBO 223132)	9	1	—	—

Fonte: Elaboração própria a partir do SINASC–CNES integrado e do Anexo A do Relatório ANS (2023).

C.4 Comparabilidade: restrição vs. ajuste

O recorte por Robson 1+3 equipara as gestantes por *restrição* dos eixos obstétricos (§3.4). Como verificação de que o achado não depende dessa escolha, o diferencial foi reestimado *sem* o recorte, sobre o grupo clínico mais amplo de gestações cefálicas, de feto único e a termo (Robson 1–5, que inclui cesárea anterior e partos induzidos ou pré-trabalho de parto), equiparando os setores apenas por *ajuste* de covariáveis (grupo de Robson, idade, escolaridade, raça e cor, adequação do pré-natal, perfil de atendimento, UF e ano), com erros-padrão clusterizados por hospital. O OR bruto do privado lucrativo é 8,94 (IC95% [8,10; 9,87]); após o ajuste, cai para 1,94 (IC95% [1,62; 2,33], $p < 0,001$), positivo e significativo, na mesma direção do achado canônico de Robson 1+3 (1,54). O valor um pouco maior é coerente, pois o grupo ampliado inclui justamente os subgrupos de maior discricionariedade (induzidos e cesárea anterior). O diferencial, portanto, sobrevive a uma definição alternativa de comparabilidade e não depende da exclusão do Grupo 5 nem do recorte 1+3.

APÊNDICE D – PERFIL DOS HOSPITAIS QUE MUDARAM DE NATUREZA JURÍDICA

O exercício *within-hospital* e o *event study* da §4.4.3 repousam sobre os 151 hospitais que registraram mudança de natureza jurídica (NJ) entre 2014 e 2022, dentro dos 5.280 da amostra principal Robson 1+3. Este apêndice caracteriza esse subconjunto, para que se avalie sua representatividade e o poder estatístico do desenho.

As 151 primeiras transições distribuem-se em seis direções: Privado SFL para Público (67), Privado lucrativo para Privado SFL (39), Público para Privado SFL (21), Privado lucrativo para Público (13), Privado SFL para Privado lucrativo (10) e Público para Privado lucrativo (1). Agregadas, são 80 transições de privado para público, 22 de público para privado e 49 entre os dois setores privados. Os três subgrupos do *event study* reagrupam essas direções: SFL ↔ Público (88), saída do lucrativo (52) e entrada no lucrativo (11). A Tabela D1 apresenta o perfil de cada subgrupo.

Tabela D1 – Perfil dos 151 hospitais que mudaram de natureza jurídica entre 2014 e 2022, por subgrupo de transição, frente aos nunca tratados, Robson 1+3.

	SFL ↔ Públ.	Saída lucr.	Entrada lucr.	Todos 151	Controles
<i>Volume</i>					
<i>n</i> (hospitais)	88	52	11	151	5.129
Partos	203.313	117.269	21.942	342.524	8.605.285
Partos/hosp-ano	286	279	278	283	244
% da amostra	2,27	1,31	0,25	3,83	96,17
<i>Taxa de cesárea (%)</i>					
Geral	26,1	34,8	59,8	31,3	31,9
2014	30,7	30,4	56,2	32,1	35,3
Pré-transição	30,5	32,8	57,9	33,2	n.a.
Pós-transição	22,4	37,5	61,8	29,4	n.a.
<i>Estrutura (média)</i>					
Leitos clín.-cir.	60,9	55,2	91,7	61,2	38,7
Salas obstétricas	2,8	2,8	3,1	2,8	2,3
% com centro obst.	91	90	82	90	85
% com neonatal	58	58	36	56	50

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022. *n.a.*: não se aplica (controles não registram transição).

Três fatos sobressaem. Primeiro, os 151 hospitais concentram apenas 3,83% dos partos

da amostra; o desenho *within* opera sobre uma fração pequena do total, o que ajuda a explicar a diferença mínima detectável de 7,11 p.p da §4.4.3. Segundo, o subgrupo de entrada no lucrativo, único com coeficiente contemporâneo significativo, reúne apenas 11 hospitais e 0,25% dos partos, base estreita que recomenda cautela na leitura daquele coeficiente. Terceiro, os hospitais que mudaram de NJ não são unidades atípicas pequenas: em média são um pouco maiores que os controles (283 contra 244 partos por hospital-ano; 61 contra 39 leitos clínico-cirúrgicos), o que sustenta sua representatividade quanto ao porte.

A Tabela D2 reporta os coeficientes completos do *event study* sobre as transições de natureza jurídica, por subgrupo e por tempo relativo à mudança, discutidos no corpo na §4.4.3. Os coeficientes pré-tratamento são a base do teste de Wald de paralelismo ali reportado.

Tabela D2 – Event study TWFE dos 151 hospitais que mudaram de natureza jurídica entre 2014 e 2022, taxa de cesárea (p.p.), Brasil.

Tempo desde transição (k)	SFL↔Público ($n = 88$)	Saída do Lucr. ($n = 52$)	Entrada no Lucr. ($n = 11$)	Todos ($n = 151$)
$k \leq -3$	+2,05	-2,84	-0,26	-0,30
$k = -2$	+0,08	-0,43	-0,39	-0,22
$k = -1$ (base)	0	0	0	0
$k = 0$	+0,18	+1,33	+1,56*	+0,60
$k = +1$	+0,21	+3,10	-0,14	+1,25
$k = +2$	-0,01	+2,03	-1,76	+0,71
$k \geq +3$	0,00	+3,23	-3,61	+0,83
Wald pré-tendência ($k = -3, -2$)	3,39	1,10	0,04	0,05
p Wald	0,184	0,578	0,983	0,973

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Nota: Coeficientes em pontos percentuais sobre a taxa de cesárea; categoria base $k = -1$.

APÊNDICE E – MÉDICAS COMO PACIENTES: TESTES DE ROBUSTEZ

A §4.4.4 estimou que gestantes médicas não obstétricas apresentam OR de cesárea de 0,825 sobre o controle Superior em ocupação não saúde no setor privado, e leu o resultado como *fuga entre hospitais*. Este apêndice submete essa leitura a cinco verificações independentes, todas sobre extração própria do SINASC–CNES integrado (2014–2022). A Tabela E1 consolida os achados.

Tabela E1 – Bateria de robustez do exercício de médicas como pacientes, setor privado, escolaridade Superior, Robson 1+3, Brasil, 2014–2022.

Teste / especificação	Métrica	Valor	IC95%
<i>Painel A: contraste principal (logit, médicas L3 vs. controle L0)</i>			
Setor privado (referência da §4.4.4)	OR	0,825***	[0,77; 0,89]
<i>Painel B: placebo setorial e controle de renda</i>			
Placebo no setor público	OR	1,721***	[1,45; 2,05]
Controle de renda alta (não saúde, grandes grupos 1–2)	OR	0,893**	[0,83; 0,96]
<i>Painel C: estimadores alternativos e sensibilidade</i>			
IPW (Hájek), suporte comum aparado	ATE (p.p)	–2,69	.
AIPW (duplamente robusto)	ATE (p.p)	–2,52*	[–4,88; –0,17]
E-value (estimativa pontual)	E-value	1,44	.
E-value (borda do IC)	E-value	1,33	.
Oster (estabilidade a inobservados)	δ^*	–7,69	.
<i>Painel D: mecanismo de seleção de hospital</i>			
Taxa de cesárea (controles) do hospital de destino	dif. (p.p)	–3,27***	[–4,54; –1,99]

Fonte: Elaboração própria a partir do SINASC–CNES integrado, 2014–2022. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

A re-extração reproduz os números do texto principal (OR 0,825 no logit, efeito que se anula dentro do hospital). O Painel B separa as duas explicações rivais: no setor público, onde o incentivo financeiro à cesárea não opera, o sinal se inverte (OR de 1,721, mais cesárea que o controle). Se a proteção observada no privado refletisse preferência da médica por parto vaginal, ela apareceria também no público; sua reversão indica fenômeno restrito ao setor com incentivo de oferta. Apertar o controle de renda, comparando médicas a não médicas de renda alta equivalente, reduz a proteção de OR 0,821 para 0,893, mas a mantém significativa: parte do diferencial bruto refletia renda; o resíduo, não.

Os estimadores duplamente robustos confirmam a magnitude (AIPW de –2,52 p.p, IC95% [–4,88; –0,17]), e o teste de mecanismo do Painel D a localiza: médicas dão à luz em hospitais cuja taxa de cesárea *entre não médicas* é 3,27 p.p menor (IC95% [–4,54; –1,99]). Combinado ao nulo *within-hospital*, isso mostra que a assimetria informacional opera *ex ante*, na escolha do estabelecimento, e não *ex post*, na barganha clínica, o que distingue o caso brasileiro do americano de Johnson e Rehavi (2016), em que o efeito sobrevivia ao controle por hospital. A

leitura é, ao mesmo tempo, sensível a confundimento não observado de magnitude moderada (E-value de 1,44), limite que o desenho transversal não fecha e que reaparece como agenda de identificação na seção 5.

APÊNDICE F – REPRODUTIBILIDADE COMPUTACIONAL

Este apêndice registra o ambiente computacional, os parâmetros numéricos e os tempos de execução da estimação da seção 4, suficientes para situar o custo e as condições em que os resultados foram produzidos.

F.1 Ambiente de software

A estimação mobiliza dois ambientes complementares: Python, para a estimação principal e a decomposição preditiva, e R, para o modelo hierárquico e o pareamento.

F.1.1 Python

- Python **3.12.13**;
- `numpy 2.4.4`, `pandas 2.3.3`, `scipy 1.17.1`;
- `scikit-learn 1.8.0` (logit, regressões de desfecho do AIPW);
- `statsmodels 0.14.6`;
- `pyarrow 24.0.0` (leitura colunar das bases);
- `xgboost` e `lightgbm` (decomposição preditiva das Perguntas 1 e 2); `shap` (importância de variáveis).

F.1.2 R

- R **4.6.0**;
- `lme4` (GLMM logit com intercepto aleatório, quadratura de Laplace);
- `MatchIt` (pareamento por escore de propensão 1:1, vizinho mais próximo, *caliper* 0,2);
- `ebal 0.2.1` (entropy balancing, verificação cruzada);
- `Matrix 1.7.5`, `data.table 1.18.2.1`.

F.1.3 Hardware

A estimação foi conduzida em um MacBook Pro com processador Apple M5 Pro (15 núcleos: 5 de desempenho e 10 de eficiência), 24 GB de memória unificada e armazenamento SSD interno, sob macOS 26.5. A memória de 24 GB foi suficiente para a rodada completa, inclusive o GLMM e o PSM sobre os 8,9 milhões de partos da amostra principal, sem recurso a memória virtual em disco.

F.2 Reprodutibilidade numérica

- semente aleatória fixa (valor 42) em todos os componentes estocásticos (*bootstrap* e algoritmos de *boosting*);
- logit ajustado por *lbfgs*, com tolerância de 10^{-5} e até 200 iterações;
- *trimming* do escore de propensão entre os percentis 0,01 e 0,99 (mesmas constantes em IPW e AIPW);
- *bootstrap* clusterizado por hospital com 200 réplicas nos estimadores baseados em escore.

F.3 Tempos de execução típicos

Tempos medidos em CPU *single-thread* salvo indicação contrária; valores arredondados.

Tabela F1 – Tempo de execução por etapa na amostra principal Robson 1+3 (8,9 milhões de partos).

Etapa	Tempo	Observação
Carregamento da amostra	3–6 s	leitura colunar
Logit Especificação 3 saturada	100–150 s	<i>lbfgs</i> , 80 <i>features</i>
<i>Sandwich</i> cluster-robust	30–60 s	cálculo vetorizado
IPW + <i>bootstrap</i> 200 réplicas	5–8 min	dominado pelo <i>bootstrap</i>
AIPW + <i>bootstrap</i> 200 réplicas	6–10 min	dominado pelos modelos de desfecho
<i>Entropy balancing</i> (CPU)	16–30 s	22–24 iterações
Heterogeneidade (15 sub-logits)	13 min	total
GLMM (lme4, nAGQ=0, amostra completa)	≈25 min	partição de variância
GLMM (lme4, nAGQ=1, subamostra 561k)	≈26 min	quadratura de Laplace
PSM 1:1 (MatchIt, amostra completa)	≈1h40	<i>matching</i> + <i>bootstrap</i>

Fonte: Elaboração própria.

Na amostra de sensibilidade Robson 1–10 (24 milhões de partos), os tempos escalam aproximadamente de forma linear em n no logit, no IPW e no AIPW, e mais que linearmente no *entropy balancing*, pelo regime de mal-condicionamento descrito a seguir.

F.4 Limitações computacionais reconhecidas

Três limitações foram diagnosticadas durante a estimação e qualificam o alcance dos resultados.

F.4.1 Entropy balancing na amostra Robson 1–10

O *entropy balancing* converge rapidamente nas amostras Robson 1+3 e Robson 1–4 (16–30 s, com balanceamento de momentos essencialmente exato), mas não converge de forma

estável na amostra total Robson 1–10. A causa é estrutural ao Grupo 5 (cesárea anterior), e não numérica: nesse grupo o desfecho é quase determinístico (prevalência de cesárea de cerca de 94% em ambos os setores) e a presença do grupo é fortemente desigual entre setores (61,8% no privado contra 38,2% no público), de modo que o problema dual de Hainmueller exige pesos extremos para igualar os momentos e entra em regime mal-condicionado. A re-execução do balanceamento excluindo o Grupo 5 converge normalmente, o que confirma o diagnóstico. Para a amostra Robson 1–10, portanto, a triangulação causal repousa sobre IPW e AIPW, robustos a esse cenário (§4.7).

F.4.2 GLMM e PSM na amostra completa

Em hardware anterior, esses dois estimadores haviam sido conduzidos em subamostras por custo computacional. Na rodada definitiva, ambos foram reexecutados sobre os **8,9 milhões de partos** da amostra principal completa, eliminando essa limitação.

O PSM 1:1 pareou os 2.788.828 partos privados a controles públicos de escore mais próximo (*caliper* 0,2 na escala do logit do escore, sem reposição), num total de 5.577.656 observações pareadas; o *matching* levou cerca de uma hora e quarenta minutos. A inferência usa *bootstrap* clusterizado por hospital (200 réplicas), com pico de memória sob os 24 GB e tempo de cerca de quinze segundos. O ATT resultante é de +16,75 p.p (IC95% [14,66; 18,94]), com SMD máxima pós-pareamento de 0,130.

O GLMM logit com intercepto aleatório por hospital (5.280 níveis) admite duas ordens de ajuste. Com $nAGQ=0$ (sem quadratura), roda sobre a amostra completa de 8,9 milhões em torno de vinte e cinco minutos, mas, como alertam Snijders e Bosker (2012) e a literatura de *lme4*, envia o *efeito fixo* sob coeficiente de correlação intraclass elevado, e aqui produziu uma OR implausível, próxima de 1. A estimativa de referência usa, por isso, $nAGQ=1$ (quadratura de Laplace), que corrige esse viés, sobre subamostra estratificada por hospital de 561.242 partos preservando os 5.280 níveis; o ajuste converge em cerca de vinte e seis minutos e entrega ICC de 0,48 e OR do efeito fixo de 3,03 (IC95% [2,79; 3,29]). É esse o valor reportado na seção 4; o $nAGQ=0$ sobre a amostra completa fica como verificação rápida da partição de variância.

F.4.3 Gestão de memória na amostra completa

Os picos transitórios de memória em algumas etapas (o cálculo do efeito marginal médio com gradiente vetorial completo sobre a amostra Robson 1–10 e o *bootstrap* do PSM sobre os milhões de pares) aproximam-se do teto de 24 GB de memória unificada. Em vez de recorrer a

memória virtual em disco, a estratégia foi reduzir o pico na origem: leitura colunar restrita às variáveis necessárias, processamento das amostras grandes por partição (ano) quando aplicável, e reindexação de vetores numéricos no *bootstrap* do PSM, em vez de reconstruir o conjunto pareado a cada réplica. Com esses cuidados, a rodada completa coube na memória física, sem a penalidade de desempenho da paginação em disco.