Event-Study e adoção escalonada

IBM0288 - 2026.2

Prof. Raphael Gouvea

1 Aulas passadas: inferência causal

Queremos saber: a política, lei ou intervenção **causou** uma mudança no resultado?

O problema: nunca observamos a mesma unidade no estado tratado e no estado não tratado simultaneamente!

$$Y_i = D_i \cdot Y_i(1) + (1 - D_i) \cdot Y_i(0)$$

onde Y_i é o resultado observado; $Y_i(1)$ é o resultado potencial com tratamento e $Y_i(0)$ é o resultado potencial sem tratamento.

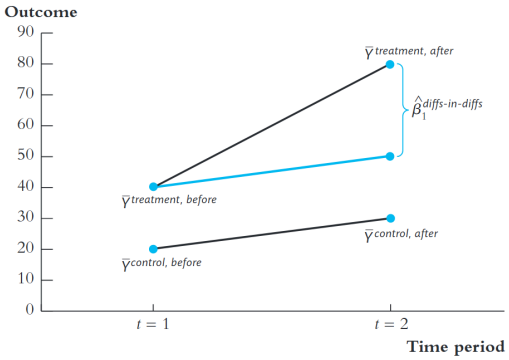
	$Y_i(1)$	$Y_i(0)$
$D_i = 1$ (tratado)	**observado**	✗ contrafactual
$D_i = 0$ (controle)	✗ contrafactual	**observado**

! O problema do contrafactual

O efeito causal individual é $Y_i(1) - Y_i(0)$. O contrafactual $Y_i(0)$ **nunca é observado** para os tratados. Portanto, precisamos estimá-lo com uma **estratégia de identificação**.

2 Aula passada: a solução DiD

A ideia-chave do DiD: usar a **evolução do grupo de controle** ao longo do tempo para estimar o contrafactual do grupo tratado.



$$\widehat{ATT} = \underbrace{(\bar{Y}_{trat,\,depois} - \bar{Y}_{trat,\,antes})}_{\text{mudança no grupo tratado}} - \underbrace{(\bar{Y}_{ctrl,\,depois} - \bar{Y}_{ctrl,\,antes})}_{\text{mudança no grupo de controle}}$$

3 Motivação: Medicaid e mortalidade

Pergunta causal: a expansão do Medicaid via ACA (*Affordable Care Act*) reduziu a mortalidade adulta nos EUA?

Contexto:

- Expandiu a elegibilidade do Medicaid para adultos de baixa renda em diversos estados;
- Objetivo principal: reduzir o número de pessoas sem seguro de saúde;
- Oferece cobertura mais acessível por meio de subsídios e programas públicos;
- Proíbe práticas como negar cobertura por condições preexistentes.

4 Motivação: ponto crucial

A expansão pelos estados foi **opcional**. Estados escolheram se e quando expandir!

Ano de expansão	Estados	% dos condados
2014	AR, AZ, CA, CO, CT, ...	36%
2015–2019	AK, IN, PA, LA, MT, ...	15%
Nunca	AL, FL, GA, TX, WY, ...	31%

Por que não basta comparar os dois grupos (controle e tratamento)?

Estados que decidiram expandir fizeram essa **escolha** por razões **econômicas e políticas**. Assim, uma comparação simples confundiria o efeito do Medicaid com as diferenças preexistentes entre os grupos (efeito seleção).

5 DiD 2x2

O caso mais simples: **2 grupos × 2 períodos**

- $t = 1$ (antes): 2013 — ambos os grupos ainda sem Medicaid expandido
- $t = 2$ (depois): 2014 — grupo tratado recebe a expansão

	2013 (antes)	2014 (depois)
Expansão 2014 ($D_i = 1$)	978 condados	**tratados**
Sem expansão ($D_i = 0$)	1.222 condados	controle

Resultado: $Y_{i,t}$ = taxa de mortalidade adulta (20–64) por 100.000 habitantes

Resultado potencial: $Y_{i,t}(0)$ = mortalidade que *teria ocorrido* sem a expansão do Medicaid

Resultado potencial: $Y_{i,t}(1)$ = mortalidade que *teria ocorrido* com a expansão do Medicaid

6 O parâmetro-alvo: ATT

O que queremos estimar é o **efeito médio do tratamento sobre os tratados**:

$$ATT = E[\, Y_{i,2}(1) - Y_{i,2}(0) \mid D_i = 1 \,]$$

O que cada parte significa:

- $Y_{i,2}(1)$: mortalidade *observada* nos condados que expandiram em 2014 → **observado!**
- $Y_{i,2}(0)$: mortalidade *contrafactual* — o que teria sido sem o Medicaid → **não observado!**

Hipótese de não antecipação: antes do tratamento, os grupos tratados ainda não reagem à expansão futura:

$$Y_{i,1}(1) = Y_{i,1}(0) \quad \forall \ \text{unidade tratada}$$

→ Em 2013, os condados que expandiram se comportam como se não fossem expandir.

Ibmec

7 A hipótese de tendências paralelas

Na ausência do tratamento, a mudança média no resultado teria sido a **mesma** nos dois grupos:

$$E[Y_{i,2}(0) - Y_{i,1}(0) \mid D_i = 1] = E[Y_{i,2}(0) - Y_{i,1}(0) \mid D_i = 0]$$

O que isso significa na prática?

- Se o Medicaid **não tivesse** expandido, a mortalidade no grupo de tratamento teria crescido **no mesmo ritmo** que no grupo de controle.
- Não exige que os **níveis** sejam iguais! Apenas que as **tendências** sejam paralelas!

8 Do contrafactual ao estimador

Sob a hipótese de tendências paralelas, podemos **identificar** o ATT a partir de médias observáveis:

Passo 1: O contrafactual dos tratados é:

$$E[Y_{i,2}(0) \mid D_i = 1] = \underbrace{E[Y_{i,1} \mid D_i = 1]}_{\text{média antes (tratados)}} + \underbrace{\big(E[Y_{i,2} \mid D_i = 0] - E[Y_{i,1} \mid D_i = 0] \big)}_{\text{tendência do controle}}$$

Passo 2: Substituindo na definição do ATT:

$$ATT = \big(E[Y_{i,2}|D_i = 1] - E[Y_{i,1}|D_i = 1] \big) - \big(E[Y_{i,2}|D_i = 0] - E[Y_{i,1}|D_i = 0] \big)$$

$\longrightarrow$ DiD = mudança nos tratados **menos** mudança nos controles.

9 O estimador DiD: quatro médias

Table 2: Simple 2 × 2 DiD

	Unweighted Averages			Weighted Averages		
	Expansion	No Expansion	Gap/DiD	Expansion	No Expansion	Gap/DiD
2013	419.2	474.0	*-54.8*	322.7	376.4	*-53.7*
2014	428.5	483.1	*-54.7*	326.5	382.7	*-56.2*
Trend/DiD	*9.3*	*9.1*	0.1	*3.7*	*6.3*	-2.6

This table reports average county-level mortality rates (deaths among adults aged 20-64 per 100,000 adults) in 2013 (top row) and 2014 (middle row) in states that expanded adult Medicaid eligibility in 2014 (columns 1 and 4, 978 counties) and states that have not expanded by 2019 (columns 2 and 5, 1,222 counties). The first three columns present unweighted averages and the second three columns present population-weighted averages. Columns 1, 2, 4, and 5 in the third row show time trends in mortality between 2013 and 2014 for each group of states. The first two rows of columns 3 and 6 show the cross-sectional gap in mortality between expansion and non-expansion states in 2013 and 2014. The entries in red text in the bottom row show the simple 2 × 2 difference-in-differences estimates without weights (column 3) and with them (column 6)

10 Como testar tendências paralelas?

A hipótese de tendências paralelas envolve $Y_{i,t}(0)$ **não observado** no período pós-tratamento! Não há teste definitivo!

O que podemos fazer:

1. **Balanceamento de covariáveis**: verificar se os grupos diferem em variáveis que determinam tendências do resultado
2. **Pré-tendências**: verificar se as tendências eram paralelas *antes* do tratamento
3. **Análise de sensibilidade**: quantificar o quão grande precisaria ser uma violação da hipótese para reverter os resultados

Ibmec

! Pré-tendências ≠ testar hipótese de tendências paralelas

Tendências paralelas antes do tratamento **apoiam** a hipótese de tendências paralelas, **mas não a provam**! A hipótese de tendências paralelas é uma hipótese sobre o comportamento do contrafactual não observado após o tratamento.

11 Balanceamento de covariáveis no Medicaid

Comparação dos grupos de controle e tratamento em 2013 (antes do tratamento):

Variável	Controle	Tratados	Dif. Normalizada
% Branca	81,6	90,5	**0,59** ⚠️
% Hispânica	9,6	8,2	−0,10
Taxa de desemprego	7,6	8,0	0,16
Taxa de pobreza	19,3	16,5	**−0,42** ⚠️
Renda mediana ($ mil)	43,0	48,0	**0,43** ⚠️

- Diferença normalizada acima de 0,25 em módulo indica desbalanceamento relevante.
- Grupos diferem em raça, pobreza e renda: variáveis que provavelmente afetam *tendências* de mortalidade. Isso coloca em dúvida a hipótese de tendências paralelas incondicional.

12 DiD com controles

Na ausência do tratamento, a evolução média do resultado **é a mesma entre tratados e não tratados** dentro de cada estrato de covariável X_i.

$$E[\Delta Y_t(0) \mid X_i, D_i = 1] = E[\Delta Y_t(0) \mid X_i, D_i = 0]$$

A hipótese de identificação é de tendências paralelas condicionais!

- Não exige que os grupos sejam iguais! Exige apenas que, **dado** X_i, os grupos caminhem em paralelo.

O caminho natural seria incluir covariáveis na regressão de efeito fixo:

$$Y_{it} = \theta_t + \eta_i + \beta_{\text{treat}} D_{it} + X_{it}' \beta_{\text{covs}} + e_{it}$$

13 Armadilhas TWFE com covariáveis

A estimação TWFE de DiD com covariáveis possui dois problemas:

- Controle errado: controlar pela pobreza *anual* (X_{it}) equivale a controlar pela *variação* da pobreza em TWFE, ou seja, áreas pobres ≠ áreas ficando mais pobres.
- Controles ruins: se X_{it} for afetado pelo próprio tratamento (ex.: Medicaid pode reduzir pobreza), incluí-lo *contamina* a estimativa do efeito causal.

Conclusão: mesmo que a hipótese de tendências paralelas condicionais seja válida, $\hat{\beta}_{\text{treat}}$ da regressão TWFE **não recupera o ATT**.

A solução é estimar o ATT diretamente por *regressão ajustada*, *IPW* ou métodos *duplamente robustos*.

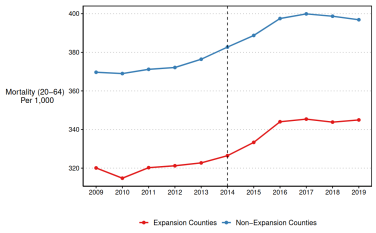
⚠️ DiD com covariáveis

Não iremos cobrir esses estimadores devido ao tempo e à complexidade. Porém, caso seja necessário utilizar covariáveis em seu DiD no futuro, é importante entender como esses estimadores funcionam — ver seção 4.4 de Baker et al. (2026). Implementá-los no R é super simples, pois os pacotes já trazem rotinas para isso.

14 *Event study*: DiD 2xT

- O termo "event study" refere-se à estimação e à apresentação de efeitos ao longo de diferentes períodos de tempo antes e depois do tratamento.
- Um desenho com um grupo tratado no mesmo período t e múltiplos períodos de tempo (2 × T) é o caso mais simples para discutir estudos de evento.

Figure 2: **County Mortality Trends by Expansion Decision**



Notes: This figure shows county population-weighted average mortality rates for adults ages 20-64 for 978 counties that expanded Medicaid in 2014 and 1,222 counties that did not expand Medicaid by 2019 from 2009 to 2019.

15 Estimação: *Event Study*

Com múltiplos períodos, estimamos o ATT para **cada período** t, sempre usando o período imediatamente anterior ao tratamento como base (g):

$$\widehat{ATT}(t) = \left(\bar{Y}_{D=1,\,t} - \bar{Y}_{D=1,\,g-1} \right) - \left(\bar{Y}_{D=0,\,t} - \bar{Y}_{D=0,\,g-1} \right)$$

Uma forma equivalente de obter todos os $ATT(t)$: estimar uma regressão TWFE com efeitos fixos de tempo, efeitos fixos das unidades e um conjunto de interações entre a dummy do grupo tratado e as dummies de tempo.

A omissão da interação de tratamento para $t = g - 1$ evita multicolinearidade e fixa $g - 1$ como o período de referência para todas as estimativas dos β.

$$Y_{i,t} = \theta_t + \eta_i + \sum_{k=1}^{g-2} \beta_k (\mathbf{1}\{G_i = g\} \cdot \mathbf{1}\{t = k\}) + \sum_{k=g}^{T} \beta_k (\mathbf{1}\{G_i = g\} \cdot \mathbf{1}\{t = k\}) + \varepsilon_{i,t}$$

Event Study

Event Study (DiD 2xT) nada mais é do que uma série de DiDs 2x2 (o caso mais simples), agora aplicados a cada t em uma janela de tempo ao redor do tratamento.

16 Tendências paralelas

Em *event study*, a hipótese de tendências paralelas é **mais forte**: tendências paralelas têm que valer para todos os períodos após o tratamento.

- Para identificar o ATT do ACA em 2019 é necessário que a hipótese de tendências paralelas seja válida para cada ano entre 2014 e 2019

Para os períodos pré-tratamento ($t < g$): sob a hipótese de não antecipação, $ATT(t)$ deve ser **zero**

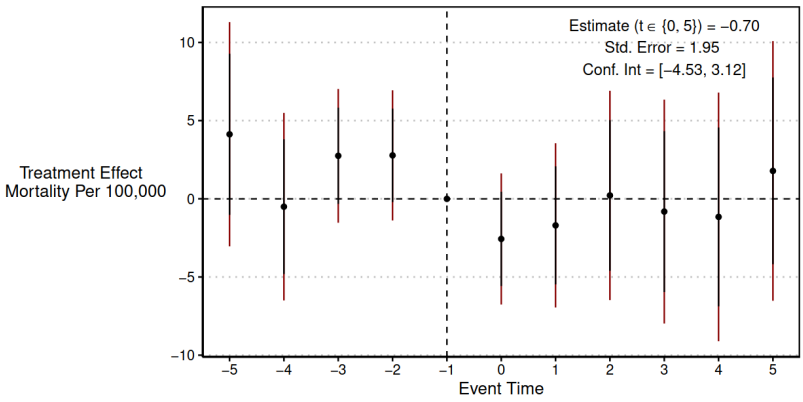
- Estimativas diferentes de zero antes do tratamento são evidência de que **a hipótese de tendências paralelas está violada**

Para os períodos pós-tratamento ($t \geq g$): estimamos o efeito causal ao longo do tempo

- Podemos ver se o efeito cresce, decresce ou se mantém estável

17 Resultados em *event study*

Figure 3: **2 × T Event Study**



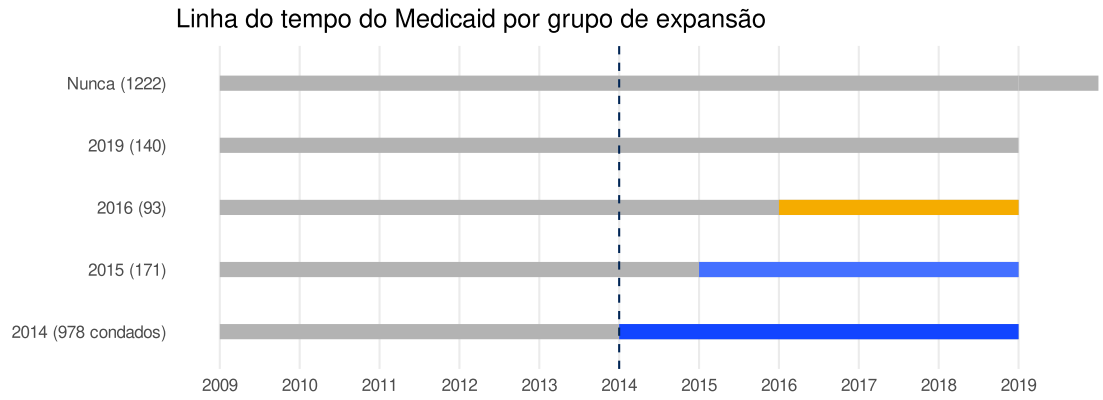
Notes: This figure shows the population-weighted event study estimates in the $2 \times T$ case, comparing 978 counties in states that expanded in 2014 to 1,222 counties in states that had not yet expanded by 2019. It uses the unconditional estimator from Callaway and Sant'Anna (2021). The outcome variable is the crude mortality rate for adults ages 20-64, and the standard errors are clustered at the county level. The point estimate is reported by the circles, and both point-wise (black) and simultaneous (red) confidence intervals are reported with the vertical lines (see section 5.1.3).

18 Quando o tratamento chega em momentos diferentes

No Medicaid, os estados expandiram em **anos diferentes** — isso é chamado de *staggered adoption* (adoção escalonada):

```
df_stag <- data.frame(
  estado = c("2014 (978 condados)", "2015 (171)", "2016 (93)", "2019 (140)", "Nunca (1222)"),
  inicio = c(2009, 2009, 2009, 2009, 2009),
  trat   = c(2014, 2015, 2016, 2019, 2020),
  fim    = c(2019, 2019, 2019, 2019, 2019)
)

ggplot(df_stag) +
  geom_segment(aes(x=inicio, xend=fim, y=estado, yend=estado),
               color="gray70", linewidth=4) +
  geom_segment(aes(x=trat, xend=fim, y=estado, yend=estado,
                   color=estado), linewidth=4, show.legend=FALSE) +
  geom_vline(xintercept=2014, linetype="dashed", color=cores$escuro) +
  scale_color_manual(values=c(
    "2014 (978 condados)"="#1245FF",
    "2015 (171)"="#4470FF",
    "2016 (93)"="#F5AC00",
    "2019 (140)"="#FF6B35",
    "Nunca (1222)"="gray70"
  )) +
  scale_x_continuous(breaks=2009:2019) +
  labs(x=NULL, y=NULL, title="Linha do tempo do Medicaid por grupo de expansão") +
  theme_minimal(base_size=12) +
  theme(panel.grid.minor=element_blank(), panel.grid.major.y=element_blank())
```



→ Cada grupo de tratamento define seus próprios parâmetros $ATT(g,t)$ — efeito para o grupo g no período t.

19 O problema com TWFE simples

A regressão TWFE convencional estima um coeficiente único:

$$Y_{i,t} = \theta_t + \eta_i + \beta^{twfe} D_{i,t} + e_{i,t}$$

O problema fundamental: com adoção escalonada, o TWFE usa grupos **já tratados** como controle uns dos outros.

→ Para estimar o efeito do grupo tratado em 2015, o TWFE compara com os condados tratados em 2014. Mas esses já foram afetados pelo Medicaid!

Consequência: com dois grupos de tratamento ($G_i = 1$ e $G_i = 2$) e nunca tratados:

$$\beta^{twfe} = ATT(2,2) + \underbrace{ATT(1,1) \times w_1 - ATT(1,2) \times w_1}_{\text{usa grupo 1 como controle para grupo 2}}$$

Se o efeito cresce com o tempo ($ATT(1,2) > ATT(1,1)$), $ATT(1,2)$ entra com **peso negativo**!

w_1 é o percentual de unidades tratadas

20 TWFE pode dar o sinal errado

> ! O problema do TWFE com adoção escalonada
>
> Mesmo que **todos** os efeitos $ATT(g,t)$ sejam positivos, o β^{twfe} pode ser **negativo**.
>
> Goodman-Bacon (2021) e de Chaisemartin & D'Haultfœuille (2020) demonstraram isso formalmente.

Por que isso acontece?

O TWFE compara grupos tratados com outros grupos tratados em momentos em que ambos já foram afetados pelo tratamento. A variação usada para identificação **não é mais o experimento que queremos**.

A solução: estimar cada $ATT(g,t)$ como um DiD 2×2 separado, usando apenas grupos não tratados (ou ainda não tratados) como controle, e depois **agregar**.

21 A solução: blocos 2×2 + agregação

Passo 1 — Estimar cada bloco 2×2:

$$\widehat{ATT}(g,t) = \left(\bar{Y}_{G_i=g,\,t} - \bar{Y}_{G_i=g,\,g-1} \right) - \left(\bar{Y}_{ctrl,\,t} - \bar{Y}_{ctrl,\,g-1} \right)$$

em que o grupo de controle é formado pelos **ainda não tratados** até o período t.

Passo 2 — Agregar com pesos:

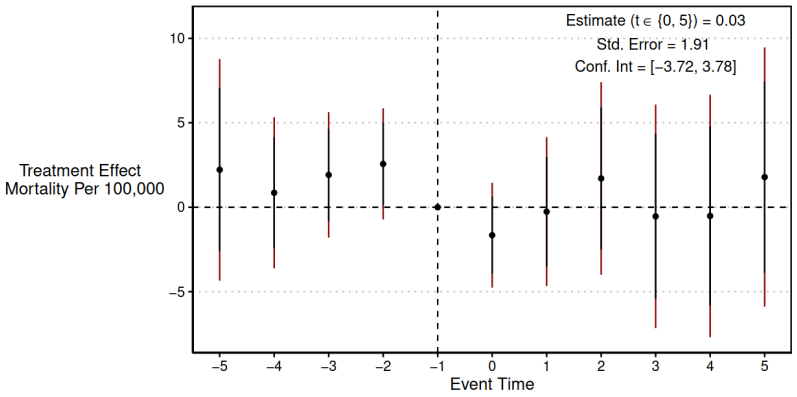
$$\widehat{ATT}_{agg} = \sum_{g,\, t \geq g} w_{g,t} \cdot \widehat{ATT}(g,t), \quad \sum w_{g,t} = 1$$

→ Pesos $w_{g,t}$ proporcionais ao tamanho do grupo g entre os tratados.

> Implementação em R
>
> O pacote `did` (Callaway & Sant'Anna, 2021) implementa exatamente esse procedimento: `att_gt()` para os blocos e `aggte()` para a agregação.

22 Resultado: adoção escalonada

Figure 8: **G × T Event Study without Covariates**



Notes: This figure shows the event study estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021), using 1,222 not-yet-treated units as the comparison group. The sample sizes used to estimate ATT parameters for each timing group are 2,200 for the 2014 group, 1,393 for the 2015 group, 1,315 for the 2016 group, and 1,362 for the 2019 group. The outcome variable is the crude mortality rate for adults ages 20-64. The point estimate is reported by the circles, and both 95% point-wise (black) and simultaneous (red) confidence intervals are reported with the vertical lines. We also report the simple average of all non-negative event times as a summary of the overall ATT (together with their standard errors and 95% confidence interval). All results use population weights.

23 Resumo

1. **Defina o parâmetro-alvo** — qual ATT você quer estimar? Para quais grupos e períodos?
2. **Enuncie a hipótese de identificação** — qual forma de tendências paralelas você está impondo? Justifique teoricamente.
3. **Avalie a plausibilidade** — balanceamento de covariáveis e pré-tendências no estudo de evento.
4. **Escolha o estimador adequado** — 2×2 simples, estudo de evento ou design com adoção escalonada.
5. **Estime e faça inferência** — erros-padrão clusterizados no nível de tratamento.
6. **Análise de sensibilidade** — quão robusto é o resultado a violações plausíveis das hipóteses?