Mestrado Profissional em Avaliação e
Monitoramento de Políticas Públicas

Métodos Quantitativos I

Aula 9: Modelos em Painel: tópicos avançados

Professores: Daniel Grimaldi e Arthur Bragança

3º Trimestre - 2025

Efeitos Aleatórios (RE)

RE vs FE

- Podemos pensar no FE como um modelo em que o intercepto varia livremente entre os indivíduos - conforme α_i na Equação (1).

$$y_{it} = \alpha_i + \beta_1 \, x^1_{it} + \epsilon_{it} \tag{1}$$

- Na prática, ao não impormos nenhuma restrição sobre o PGD de α_i, **ganhamos robustez ao custo de eficiência estatística**.
  - Dependendo do painel, cada α_i será estimado com algumas poucas observações, o que reduz a precisão das estimativas.

RE vs FE

- O modelo de efeitos aleatórios (RE) também assume que os interceptos são diferentes entre os indivíduos, mas assume que:
  - o PGD de α_i e equivale a uma distribuição aleatória conhecida - por exemplo, todos vêm de uma distribuição Normal; e
  - essa heterogeneidade não está correlacionada com as demais variáveis dependentes - $E[\alpha_i X_{it}] = 0$.
- Isso aumenta a precisão das estimativas, mas produz estimativas viesadas se o $E[\alpha_i X_{it}] \neq 0$

Modelagem de um RE

- Podemos pensar em um caso em que $\alpha_i = \alpha_0 + \upsilon_i$, sendo que $\upsilon_i \sim N(0,1)$, conforme Equação (2)

$$\begin{aligned} y_{it} =& \alpha_i + \beta_1 \; x^1_{it} + \epsilon_{it} \\ =& \alpha_0 + \beta_1 \; x^1_{it} + \upsilon_i + \epsilon_{it} \\ =& \alpha_0 + \beta_1 \; x^1_{it} + \eta_{it} \end{aligned} \tag{2}$$

- Neste caso, a omissão de υ_i do modelo de regressão leva a perda de eficiência por autocorrelação dos resíduos (existente entre distintas observações de um mesmo indivíduo), mas não produz viés.

RE como Mínimos Quadrados Ponderados

- Em um contexto de painel, um (pooled) MQO assume que:

$$E[\eta_{it}\eta_{js}] = \begin{cases} \sigma_\eta^2; \text{ \textit{se} } i = j \wedge t = s \\ 0; \text{ \textit{cc.}} \end{cases} \tag{3}$$

RE como Mínimos Quadrados Ponderados

- O modelo RE estima a Equação (2) por meio de Mínimos Quadrados Ponderados assumindo uma matriz de variância robusta à autocorrelação, de tal forma que:

$$E[\eta_{it}\eta_{js}] = \begin{cases} \sigma_\epsilon^2 + \sigma_\upsilon^2; \ \textit{se } i = j \wedge t = s \\ \sigma_\upsilon^2; \ \textit{se } i = j \wedge t \neq s \\ 0; \ \textit{cc.} \end{cases} \tag{4}$$

- Ao ajustarmos a estrutura imposta à matriz de covarância, conseguimos uma estimativa mais precisa.

Extensões de FE

Múltiplos efeitos fixos

- Vimos até agora apenas o caso em que incluímos um efeito fixo (indivíduos), mas um modelo em painel pode ser estimado com múltiplos efeitos fixos, desde exista grau de liberdade (variância) para múltiplas categorias:
  - Por exemplo, se no painel temos diferentes categorias de gênero em diferentes municípios, podemos incluir um efeito de gênero e outro de município
- Contudo, não é possível incluir simultaneamente um efeito fixo de município e outro de UF, pois não há variância de UFs dentro de um mesmo município.
  - Neste caso, a inclusão do efeito fixo de município já controla por efeitos de UFs.

Two-way Fixed Effects (TWFE)

- O TWFE é um caso particular de estimação em painel com múltiplos efeitos fixos. Mais especificamente, o TWFE inclui efeitos fixos para indivíduos (α_i) e para períodos de tempo (α_t)
- Podemos pensar nesse modelo como uma sequência de transformações within (a ordem não importa), onde controlamos por fatores não observados que afetaram os diferentes indivíduos (sem variar no tempo) e os diferentes períodos de tempo (sem variar entre indivíduos).

TWFE e DiD

O Diferenças-em-diferenças canônico (DiD) pode ser representado pela Equação (5), onde T_i e P_t representam dummies que, identificam respectivamente, indivíduos tratados e períodos após o tratamento.

$$y_{it} = \alpha + T_i + P_t + T_i P_t + \epsilon_{it} \tag{5}$$

- Note que esse modelo pode ser pensado como um caso particular de TWFE, com efeitos fixos para indivíduos tratados ($T_i = 1$) e períodos posteriores ao tratamento ($P_t = 1$)

Hands on

Caso Geral

- Vamos partir do mesmo exemplo da aula 8
- Ainda temos um painel de i municípios acompanhados por t anos e queremos estimar o efeito de uma política pública de facilitação de acesso a armas (P) sobre o nível de segurança pública - medido por um índice de morte violentas anuais por 100 mil habitantes (Y_{it})
- Mas faremos alguns ajustes para tornar o PGD um pouco mais flexível...
- ... e vamos **entender como mudanças nos parâmetros do PGD afetam estimativas de MQO, FE e RE**.

PGD: Formação de y_{it}

$$\alpha_i \sim N(0,2);\ \alpha_{it} = \alpha_i\ (1+\theta_1)^t \tag{6}$$

$$\gamma_0 \sim N(0,2);\ \gamma_t = \gamma_0\ (1+\theta_2)^t + \upsilon^3 t \tag{7}$$

$$x_i^1 \sim N(7, 0.5);\ \ x_{it}^1 = x_i^1 + \upsilon_{it}^1 \tag{8}$$

$$x_i^2 \sim N(40, 10);\ \ x_{it}^2 = x_i^2 + \upsilon_{it}^2 \tag{9}$$

$$y_{it} = \alpha_{it} + \gamma_t + \beta_1 x_{it}^1 + \beta_2 x_{it}^2 + \delta P_{it} \epsilon_{it} \tag{10}$$

$$\upsilon_{it}^1 \sim N(0,1);\ \upsilon_{it}^2 \sim N(0,5);\ \upsilon_t^3 \sim N(0,1); \epsilon_{it} \sim N(0,1) \tag{11}$$

PGD: Adesão a P

O acesso à política P é definido pelas equações (13) a (15).

$$S_i = \frac{1}{1 + e^{g(\alpha_{it})}} \tag{12}$$

$$g(\alpha_{it}) = 0 + \zeta(\sum_T \alpha_{it} - \mu_\alpha) \tag{13}$$

$$P_i \sim Bern(S_i) \tag{14}$$

P_{it} assume valores 0 ou 1, a dependender da realização de uma Bernoulli com probabilidade de sucesso S_i. **Por simplificação, todos os indivíduos que têm acesso à política, começam a receber o tratamento no mesmo período $t_p = \frac{T}{2}$.**

Caso 1: α_i não afeta acesso

- Neste primeiro caso $\alpha_{it} = \alpha_i \; \forall \; t$ e o acesso à política não é determinado por ele.
  - $\theta_1 = 0$ e $\zeta = 0$
- Os efeitos fixos de tempo são aleatórios:
  - $\theta_2 = 0$
- O impacto da política segue sendo negativo sobre os índices de criminalidade
  - $\delta = 1$
- Além disso: $\beta_1 = 0.5 \; e \; \beta_2 = 0.1$
- Qual é o modelo mais adequado?

Implementando o PGD - caso 1

Como de costume, escrevemos uma função (denominada pgd.R para simular esse PGD.

```
source("pgd.R")
set.seed(1)
data1 <- pgd(n.id=5570, n.t=6, beta=c(0.5, 0.1),
             delta=1, mu_x=c(7, 40),
             sigma_x=c(0.5, 10), mu_alpha=0,
             sigma_alpha=2, theta_1=0, theta_2=0,
             zeta=0)
```

Estimando os modelos - caso 1

```r
require(plm)
data.reg <- data1 %>%
  plm.data(index=c("id", "t"))

reg_ols <- plm(y ~ x1 + x2 + P,
               model="pooling", data=data.reg)
reg_fe <- plm(y ~ x1 + x2 + P,
              model="within", data=data.reg)
reg_twfe <- plm(y ~ x1 + x2 + P,
              model="within",
              effects="twoways", data=data.reg)
reg_re <- plm(y ~ x1 + x2 + P,
              model="random", data=data.reg)
```

Resultados - caso 1

	Dependent variable:			
	y			
	OLS	FE	TWFE	RE
	(1)	(2)	(3)	(4)
x1	0.4944***	0.4969***	0.4969***	0.4967***
	(0.0098)	(0.0059)	(0.0059)	(0.0059)
x2	0.1019***	0.0993***	0.0993***	0.0995***
	(0.0021)	(0.0012)	(0.0012)	(0.0012)
P	0.9648***	1.0036***	1.0036***	1.0013***
	(0.0283)	(0.0157)	(0.0157)	(0.0155)
Constant	2.4826***			2.5526***
	(0.1079)			(0.0686)
Observations	33,420	33,420	33,420	33,420
R^2	0.1557	0.3900	0.3900	0.3541
Note:				*p<0.1; **p<0.05; ***p<0.01

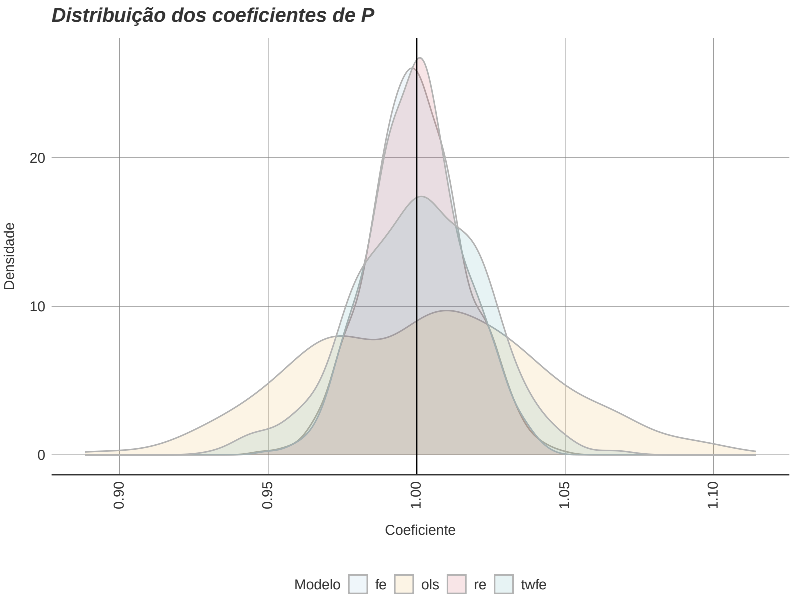
Usando LGN

```
get_models <- function(i){
  set.seed(i)
  data <- pgd(n.id=5570, n.t=6, beta=c(0.5, 0.1),
            delta=1, mu_x=c(7, 40),
            sigma_x=c(0.5, 10), mu_alpha=0,
            sigma_alpha=2, theta_1=0, theta_2=0,
            zeta=0)
  ols <- plm(y ~ x1+ x2 + P, data=data, model="pooling")
  fe <- plm(y ~ x1+ x2 + P, data=data, model="within")
  twfe <- plm(y ~ x1+ x2 + P, data=data, effect = "twoways",
              model="within")
  re <- plm(y ~ x1+ x2 + P, data=data, model="random")
  output <- data.frame(sample=i,
                       ols = ols$coefficients["P"],
                       fe = fe$coefficients["P"],
                       twfe = twfe$coefficients["P"],
                       re = re$coefficients["P"])}
```

Usando LGN

```
lgn.data <- lapply(1:500, get_models) %>%
  rbindlist()
head(lgn.data)
```

```
##     sample       ols        fe      twfe        re
##      <int>     <num>     <num>     <num>     <num>
## 1:       1 0.9647760 1.0036094 0.9871174 1.0012553
## 2:       2 1.0064966 1.0016655 1.0010830 1.0019593
## 3:       3 0.9942945 0.9947921 1.0225309 0.9947754
## 4:       4 0.9898379 0.9871762 0.9889057 0.9873263
## 5:       5 0.9654242 0.9901807 0.9758573 0.9887580
## 6:       6 0.9658109 0.9571945 0.9417144 0.9576901
```


Distribuição dos coeficientes de P
Densidade
20
10
0
0.90
0.95
1.00
1.05
1.10
Coeficiente
Modelo
fe
ols
re
twfe

Caso 2: α_i afeta acesso

- Neste segundo caso $\alpha_{it} = \alpha_i \; \forall \; t$ e o acesso à política é fortemente influenciado por ele.
  - $\theta_1 = 0$ e $\zeta = 1$
- Os demais parâmetros seguem os mesmos: qual é o modelo mais adequado?

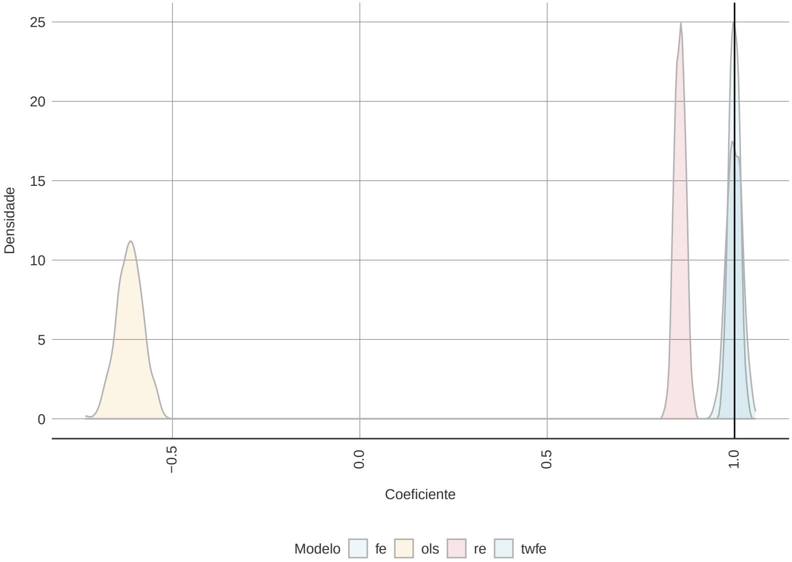
Implementando o PGD - caso 2

```
set.seed(13)
data2 <- pgd(n.id=5570, n.t=6, beta=c(0.5, 0.1),
            delta=1, mu_x=c(7, 40),
            sigma_x=c(0.5, 10), mu_alpha=0,
            sigma_alpha=2, theta_1=0, theta_2=0,
            zeta=1)
```

Resultados - caso 2

	Dependent variable:			
	y			
	OLS	FE	TWFE	RE
	(1)	(2)	(3)	(4)
x1	0.4878***	0.5013***	0.5013***	0.4993***
	(0.0096)	(0.0060)	(0.0060)	(0.0061)
x2	0.0994***	0.1002***	0.1002***	0.1001***
	(0.0020)	(0.0012)	(0.0012)	(0.0012)
P	−0.6895***	1.0039***	1.0039***	0.8548***
	(0.0272)	(0.0154)	(0.0154)	(0.0159)
Constant	2.9412***			2.4460***
	(0.1058)			(0.0694)
Observations	33,420	33,420	33,420	33,420
R^2	0.1437	0.3963	0.3963	0.3266
Note:				*p<0.1; **p<0.05; ***p<0.01

Distribuição dos coeficientes de P



Caso 3: Tendência temporal

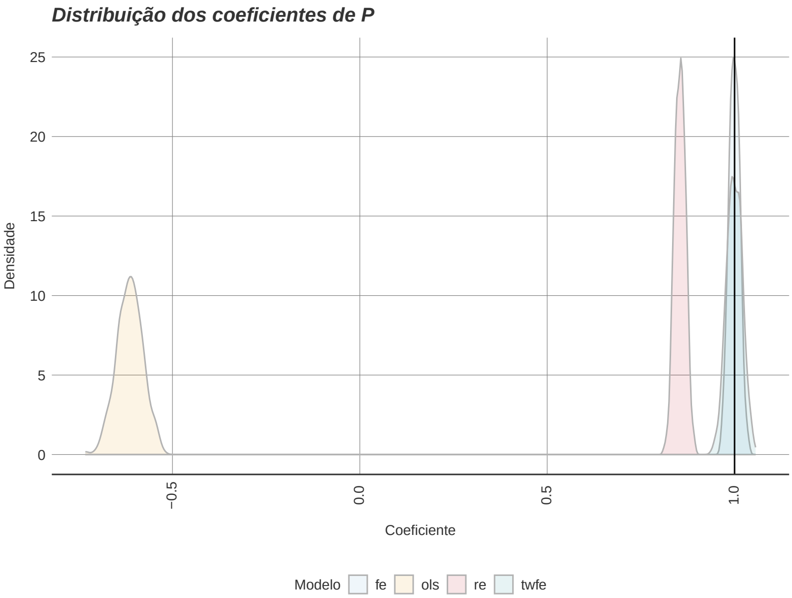
- Neste segundo caso $\alpha_{it} = \alpha_i \; \forall \; t$ e o acesso à política é fortemente influenciado por ele.
  - $\theta_1 = 0$ e $\zeta = 1$
- Além disso, há uma tendência temporal, de tal forma que γ_t cresce, em módulo, com o tempo.
  - $\theta_2 = 0.25$ e $\zeta = 1$
- Os demais parâmetros seguem os mesmos: qual é o modelo mais adequado?

Implementando o PGD - caso 3

```
set.seed(13)
data3 <- pgd(n.id=5570, n.t=6, beta=c(0.5, 0.1),
             delta=1, mu_x=c(7, 40),
             sigma_x=c(0.5, 10), mu_alpha=0,
             sigma_alpha=2, theta_1=0, theta_2=0.25,
             zeta=1)
```

Resultados - caso 3

	Dependent variable:			
	y			
	OLS	FE	TWFE	RE
	(1)	(2)	(3)	(4)
x1	0.4944***	0.4969***	0.4969***	0.4967***
	(0.0098)	(0.0059)	(0.0059)	(0.0059)
x2	0.1019***	0.0993***	0.0993***	0.0995***
	(0.0021)	(0.0012)	(0.0012)	(0.0012)
P	0.9648***	1.0036***	1.0036***	1.0013***
	(0.0283)	(0.0157)	(0.0157)	(0.0155)
Constant	2.4826***			2.5526***
	(0.1079)			(0.0686)
Observations	33,420	33,420	33,420	33,420
R^2	0.1557	0.3900	0.3900	0.3541
Note:				*p<0.1; **p<0.05; ***p<0.01



Distribuição dos coeficientes de P
25
20
15
10
5
0
Densidade
−0.5
0.0
0.5
1.0
Coeficiente
Modelo
fe
ols
re
twfe

Caso 4: Tendências distintas

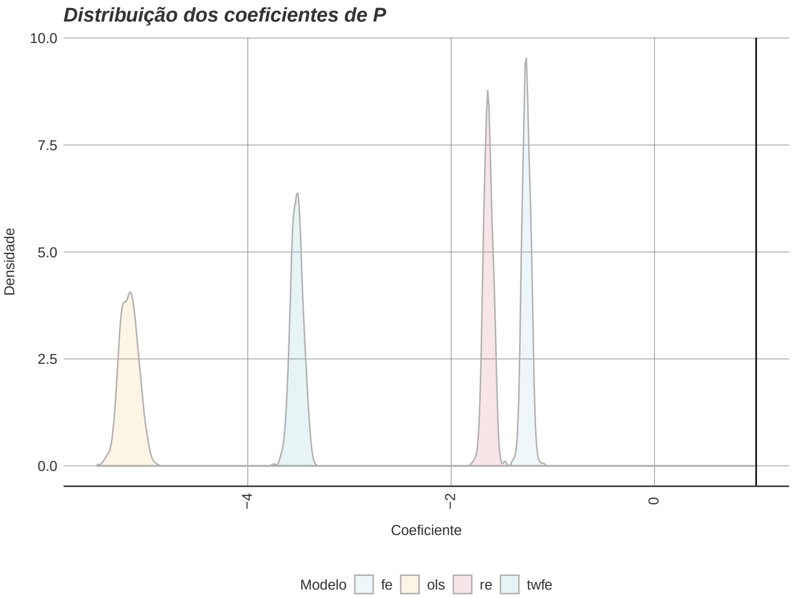
- Neste segundo caso $\alpha_{it} = \alpha_i \; \forall \; t$ e o acesso à política é fortemente influenciado por ele.
  - $\theta_1 = 0$ e $\zeta = 1$
- Além disso, há uma tendência temporal em $\alpha_i t$, de tal forma que α_t cresce, em módulo, com o tempo.
  - $\theta_1 = 0.25$ e $\zeta = 1$
- Os demais parâmetros seguem os mesmos do caso 3: qual é o modelo mais adequado?

Implementando o PGD - caso 4

```r
set.seed(13)
data4 <- pgd(n.id=5570, n.t=6, beta=c(0.5, 0.1),
            delta=1, mu_x=c(7, 40),
            sigma_x=c(0.5, 10), mu_alpha=0,
            sigma_alpha=2, theta_1=0.25, theta_2=0.25,
            zeta=1)
```

Resultados - caso 4

	Dependent variable:			
	y			
	OLS	FE	TWFE	RE
	(1)	(2)	(3)	(4)
x1	0.4852***	0.5081***	0.5081***	0.5034***
	(0.0198)	(0.0121)	(0.0121)	(0.0126)
x2	0.0974***	0.0982***	0.0982***	0.0981***
	(0.0041)	(0.0024)	(0.0024)	(0.0025)
P	−5.3723***	−1.3311***	−1.3311***	−1.7226***
	(0.0560)	(0.0311)	(0.0311)	(0.0328)
Constant	21.0874***			20.0135***
	(0.2175)			(0.1423)
Observations	33,420	33,420	33,420	33,420
R^2	0.2356	0.1584	0.1584	0.1488
Note:			*p<0.1; **p<0.05; ***p<0.01	



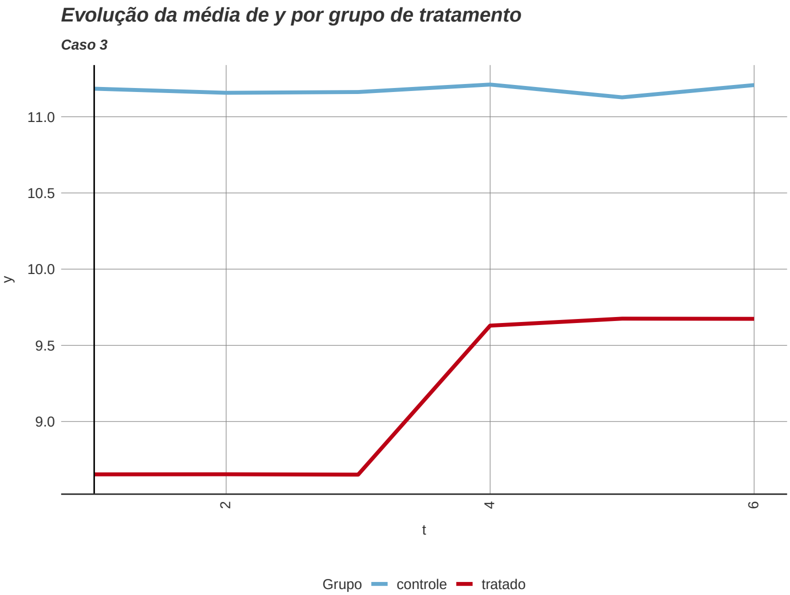
Distribuição dos coeficientes de P
10.0
7.5
5.0
2.5
0.0
Densidade
−4
−2
0
Coeficiente
Modelo
fe
ols
re
twfe

O que aconteceu?

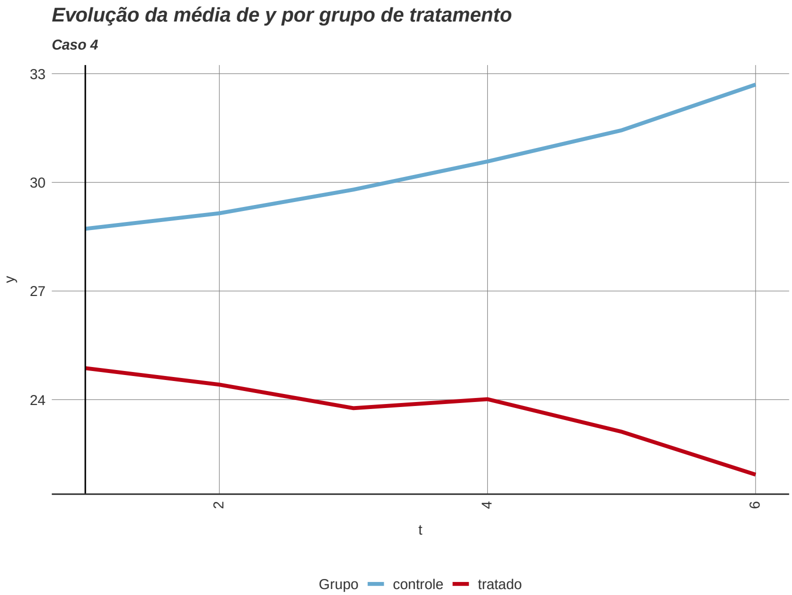
- No caso 4, α_{it} varia ao longo do tempo de tal forma que os indivíduos estão em trajetórias diferentes.
- Tal qual no caso do DiD, um painel de efeitos fixos depende da hipótese de que as variações within ("tendência") dos indivíduos controle são uma boa referência para a trajetória dos indivíduos tratados.
  - se α_{it} varia no tempo de maneira distinta entre tratados e controles, então isso não vale.
- Por isso, quando vamos usar TWFE, DiD ou mesmo painel de efeitos fixos para avaliar impacto de políticas, precisamos avaliar a validade da hipótese de tendências paralelas.

Tendências paralelas - caso 3

```r
data.temp <- data3 %>%
  group_by(D, t) %>%
  summarise(y = mean(y)) %>%
  mutate(grupo = case_when(D==1 ~ "tratado",
                           D==0 ~ "controle"))
```



Evolução da média de y por grupo de tratamento
Caso 3
11.0
10.5
10.0
9.5
9.0
y
2
4
6
t
Grupo
controle
tratado


Evolução da média de y por grupo de tratamento
Caso 4
33
30
27
24
y
2
4
6
t
Grupo
controle
tratado