

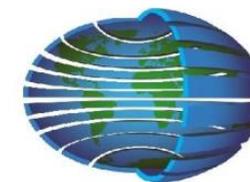
Uma Introdução à Regressão com Dados em Painel

Rafael Mesquita, Antônio Fernandes, Dalson B. Figueiredo Filho
Universidade Federal de Pernambuco (UFPE)

2021



UNIVERSIDADE
FEDERAL
DE PERNAMBUCO



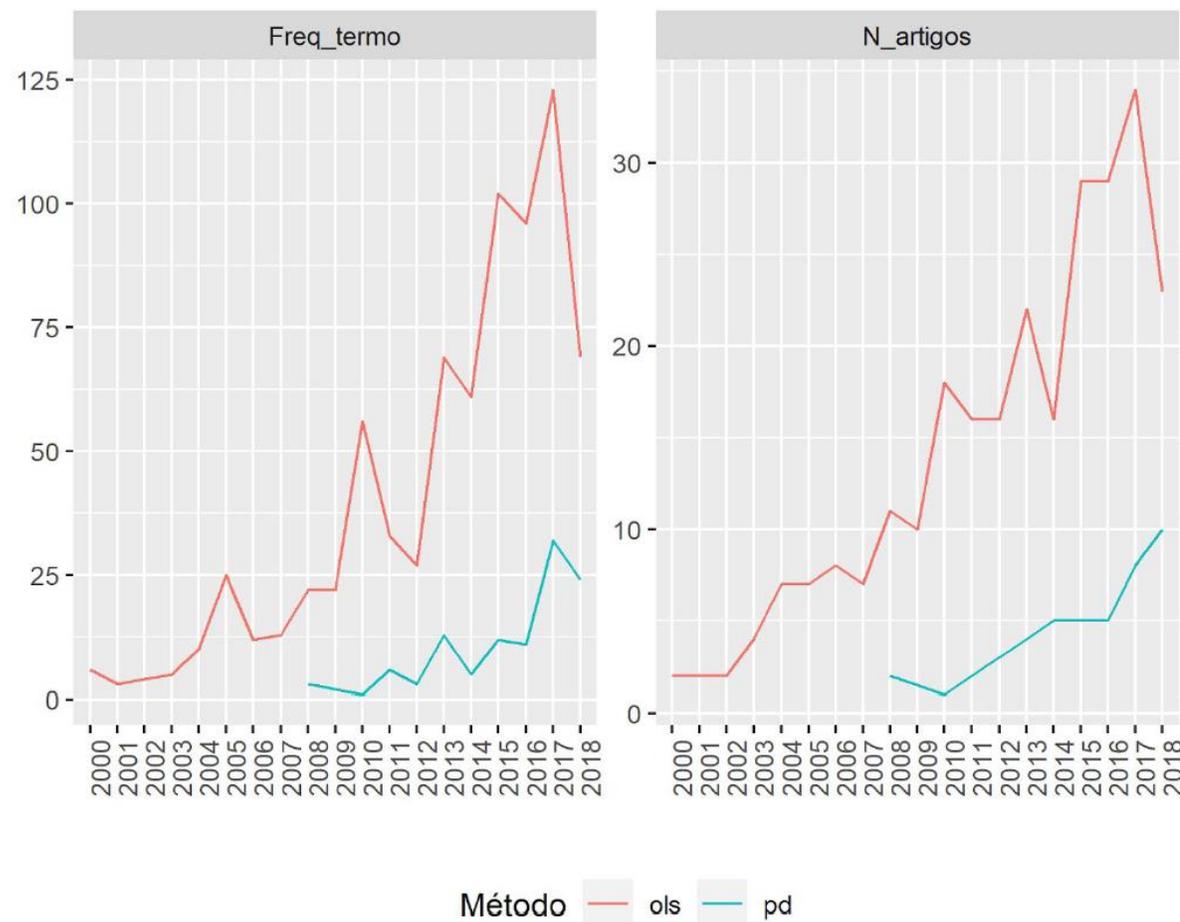
PPGCP
Programa de
Pós-Graduação em
Ciência Política

Materiais de suporte aos slides

- Artigo na *Política Hoje*
 - <https://periodicos.ufpe.br/revistas/politica hoje/article/view/246522/38645>
- Repositório OSF
 - https://osf.io/5yx7g/?view_only=ac1691cced8549238d6d6e0a9d2b7f7b

A aplicabilidade de dados em painel para CPRI

- Diversos fenômenos da CPRI adequam-se naturalmente ao formato de painel de dados
 - Ex.: Eleições em municípios; Comércio entre países a cada ano
- Porém ainda pouco usados na CPRI brasileira
 - Apenas 0,6% em amostra de 7,7 mil artigos



Fonte: Mesquita et al. (2021)

A aplicabilidade de dados em painel para CPRI

- Dados em painel oferecem diversos benefícios como:
 - 1. Facilitar a detecção de causalidade no tempo
 - 2. Monitoramento de variações individuais dos casos de interesse
 - 3. Redução de erros de mensuração
 - 4. Amostras maiores
 - 5. Formas de eliminar viés de variável omitida (que não poderiam ser implementadas com dados transversais estáticos)

A estrutura de dados longitudinais

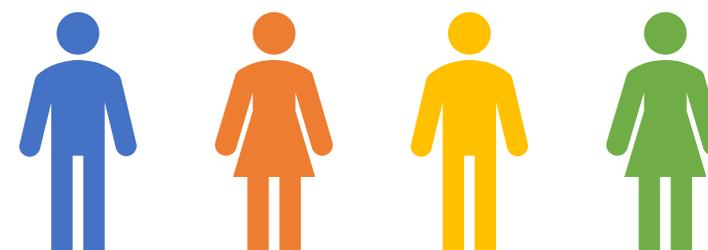
A estrutura de dados longitudinais

- Dados transversais (*cross section*) possuem observações de 'i' sujeitos em um único período no tempo (ou desprezando a passagem do tempo)
- Séries temporais (*time series*) são observações de um único sujeito repetidas ao longo de 't' períodos de tempo
- Dados em painel combinam ambas: observam 'i' sujeitos ao longo de 't' períodos de tempo

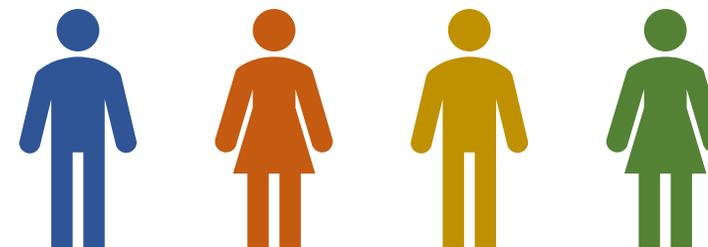
Dados transversais (*cross section*)

Série temporal (*time series*)

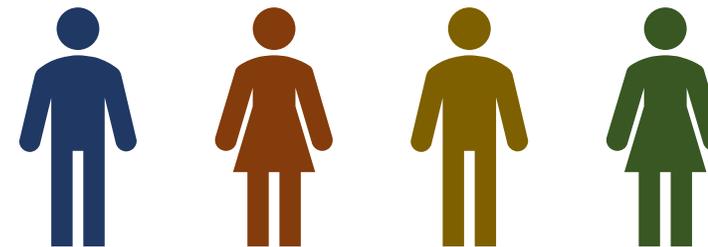
Dados em painel (*panel data*)



t_2



t_3



O layout de dados em painel

Transversal / *Cross-Section*

UF	Tempo	Taxa de homicídios
i	t	X_i
Acre	2011	22,0
Bahia	2011	39,4
Santa Catarina	2011	12,8

Série Temporal / *Time-Series*

UF	Tempo	Taxa de homicídios
i	t	X_t
Bahia	2011	39,4
Bahia	2012	43,4
Bahia	2013	37,8

Dados em Painel / *Time-Series Cross-Section*

UF	Tempo	Taxa de homicídios
i	t	X_{it}
Acre	2011	22
Acre	2012	27,4
Acre	2013	30,1
Bahia	2011	39,4
Bahia	2012	43,4
Bahia	2013	37,8
Santa Catarina	2011	12,8
Santa Catarina	2012	12,9
Santa Catarina	2013	11,9

Fonte:
Mesquita et al. (2021), com
dados de Cerqueira (2017)

O layout de dados em painel

UF	2011	2012	2013
Acre	22	27,4	30,1
Bahia	39,4	43,4	37,8
Santa Catarina	12,8	12,9	11,9

UF	Tempo	Taxa de homicídios	Média Geral
i	t	X_{it}	\bar{X}
Acre	2011	22	26,4
Acre	2012	27,4	26,4
Acre	2013	30,1	26,4
Bahia	2011	39,4	26,4
Bahia	2012	43,4	26,4
Bahia	2013	37,8	26,4
Santa Catarina	2011	12,8	26,4
Santa Catarina	2012	12,9	26,4
Santa Catarina	2013	11,9	26,4

Formato *wide*

- Cada linha, uma UF

Formato *long*

- Cada linha, uma observação de UF x Ano

Medidas agregadas por tempo e unidade

- A repetição permite agregar valores de diferentes maneiras
 - Ex.: Média de uma única UF ao longo do tempo; Média de todas as UFs em certo ano.
- Essas medidas são úteis para estatística descritiva e algumas formas de estimação

Transversal / Cross-Section

UF	Tempo	Taxa de homicídios	Média Geral	Desvio Entre Casos
i	t	X_i	\bar{X}	$X_i - \bar{X}$
Acre	2011	22,0	24,7	-2,7
Bahia	2011	39,4	24,7	14,7
Santa Catarina	2011	12,8	24,7	-11,9

Série Temporal / Time-Series

UF	Tempo	Taxa de homicídios	Média Individual	Desvio Intra-Caso
i	t	X_t	\bar{X}	$X_t - \bar{X}$
Bahia	2011	39,4	40,2	-0,8
Bahia	2012	43,4	40,2	3,2
Bahia	2013	37,8	40,2	-2,4

Dados em Painel / Time-Series Cross-Section

UF	Tempo	Taxa de homicídios	Média Geral	Média Individual	Desvio Geral	Desvio Intra-Caso	Desvio Entre Casos
i	t	X_{it}	\bar{X}	\bar{X}_i	$X_{it} - \bar{X}$	$X_{it} - \bar{X}_i$	$\bar{X}_i - \bar{X}$
Acre	2011	22	26,4	26,5	-4,4	-4,5	0,1
Acre	2012	27,4	26,4	26,5	1,0	0,9	0,1
Acre	2013	30,1	26,4	26,5	3,7	3,6	0,1
Bahia	2011	39,4	26,4	40,2	13,0	-0,8	13,8
Bahia	2012	43,4	26,4	40,2	17,0	3,2	13,8
Bahia	2013	37,8	26,4	40,2	11,4	-2,4	13,8
Santa Catarina	2011	12,8	26,4	12,5	-13,6	0,3	-13,9
Santa Catarina	2012	12,9	26,4	12,5	-13,5	0,4	-13,9
Santa Catarina	2013	11,9	26,4	12,5	-14,5	-0,6	-13,9

Tipos de painel

Tipos de painel: quanto à proporção N/T

$N > T$

- Curto ou *cross-section dominant*
- Exemplo típico: censo

$T > N$

- Longo ou *temporally-dominant*
- *Time-Series Cross-Section* (TSCS) em alguns autores (Beck)
- Exemplo típico: dados de PIB país-ano

Tipos de painel: quanto ao *missingness*

Balanced

- Todos os sujeitos são observados em todos os períodos
- Total de observações = $N \cdot T$

UF	Tempo	Taxa de homicídios
i	t	X_{it}
Acre	2011	22
Acre	2012	27,4
Acre	2013	30,1
Bahia	2011	39,4
Bahia	2012	43,4
Bahia	2013	37,8
Santa Catarina	2011	12,8
Santa Catarina	2012	12,9
Santa Catarina	2013	11,9

Unbalanced

- Alguns sujeitos não são observados em alguns períodos de tempo
- Total de observações $< N \cdot T$

UF	Tempo	Taxa de homicídios
i	t	X_{it}
Acre	2011	22
Acre	2012	27,4
Acre	2013	
Bahia	2011	39,4
Bahia	2012	43,4
Bahia	2013	37,8
Santa Catarina	2011	
Santa Catarina	2012	
Santa Catarina	2013	11,9

As 4 principais abordagens

Pooled
OLS

Primeiras
Diferenças

Efeitos
Fixos

Efeitos
Aleatórios

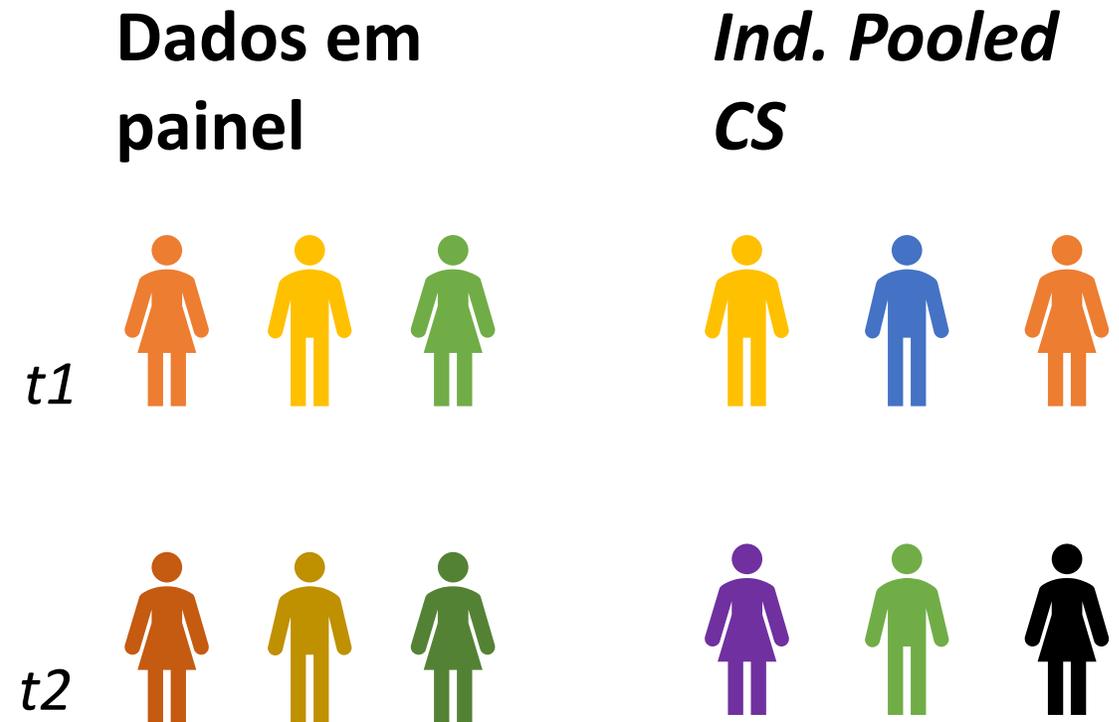
Pooled OLS

Pooled OLS

- Se tenho observações de várias entidades ao longo do tempo, é apropriado aplicar um modelo MQO/OLS normal, da mesma forma que eu faria para dados transversais?
- Distinção entre **dados em painel** e ***independently pooled cross sections***

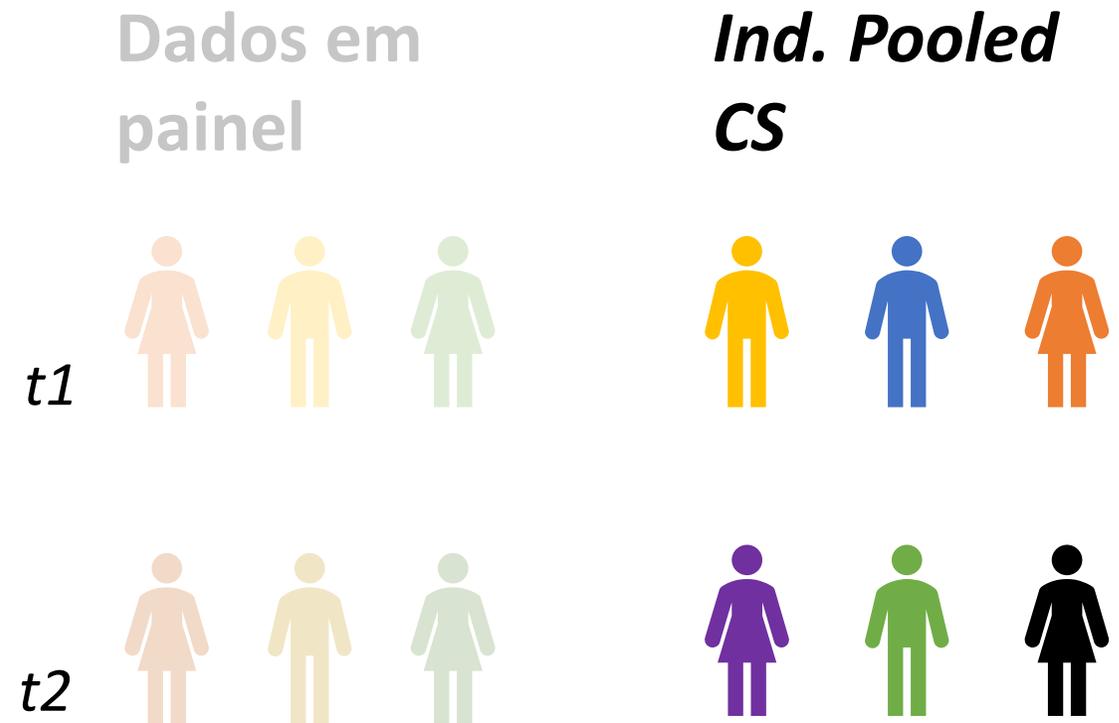
Independently pooled cross sections vs. Dados em painel

- Amostras aleatórias de casos distintos em t_1 e t_2 podem ser acumuladas (“empilhadas”, “*pooled*”) para aumentar o N
- O *pooling* de casos não ameaça os pressupostos da regressão linear se houver aleatoriedade
- *Ind. pooled cross sections* não são propriamente dados em painel/longitudinais, porque não se observam as mesmas unidades ao longo do tempo



Independently pooled cross sections

- Normalmente, pode-se usar regressão OLS habitual
- As diferenças entre $t1$ e $t2$ podem ser satisfatoriamente capturadas por uma *dummy*



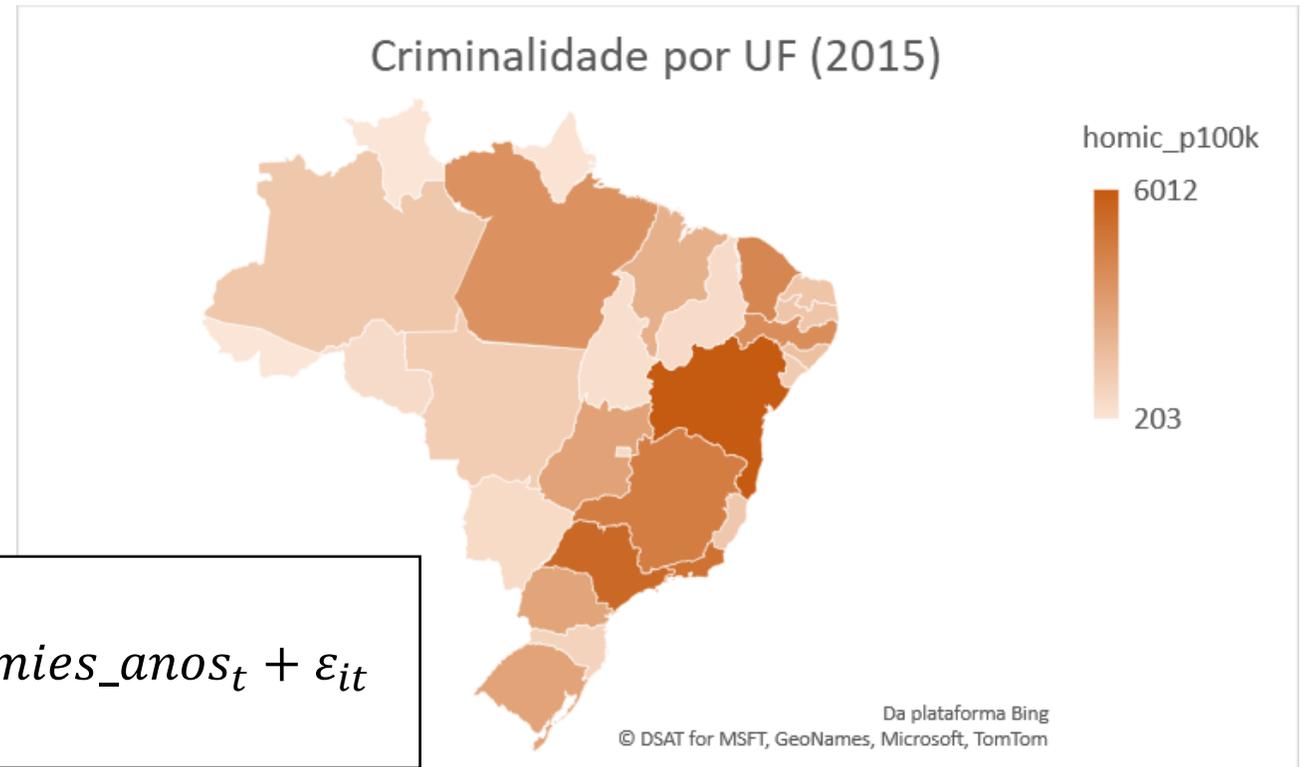
Pooled OLS

- Mas e se foram os mesmos sujeitos ao longo do tempo? É possível aplicar uma regressão OLS tradicional?
- Sim, se o conjunto de VIs for apropriado e assegurar que os **erros permanecem aleatórios e não-correlacionados**



Exemplo: Criminalidade nas UFs

Explicar a taxa de criminalidade nas 27 UFs de 2015 a 2017, a partir do nível de desemprego e tamanho da população, utilizando OLS



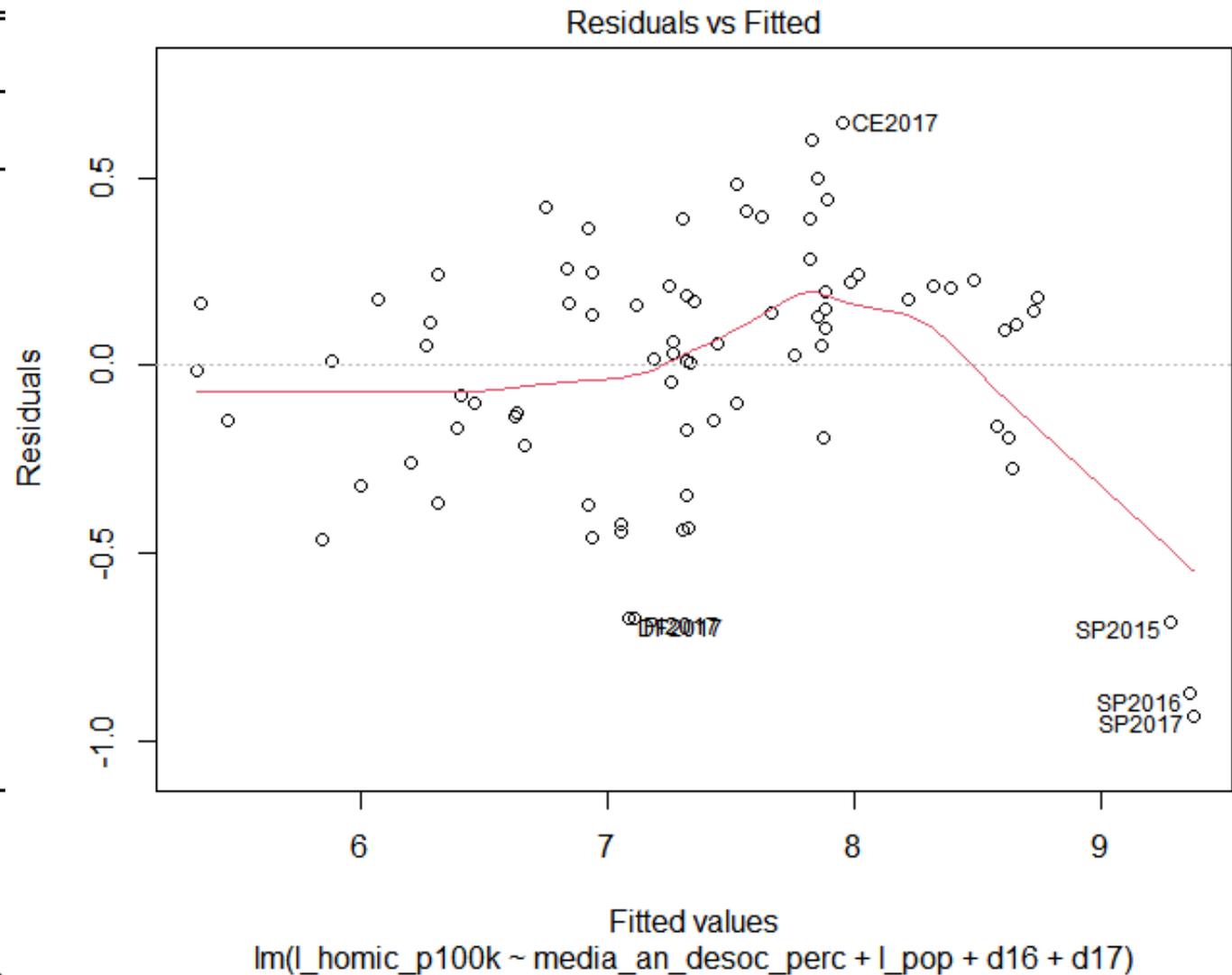
Modelo:

$$crime_{it} = \alpha + desemp_{it} + pop_{it} + dummies_anos_t + \varepsilon_{it}$$

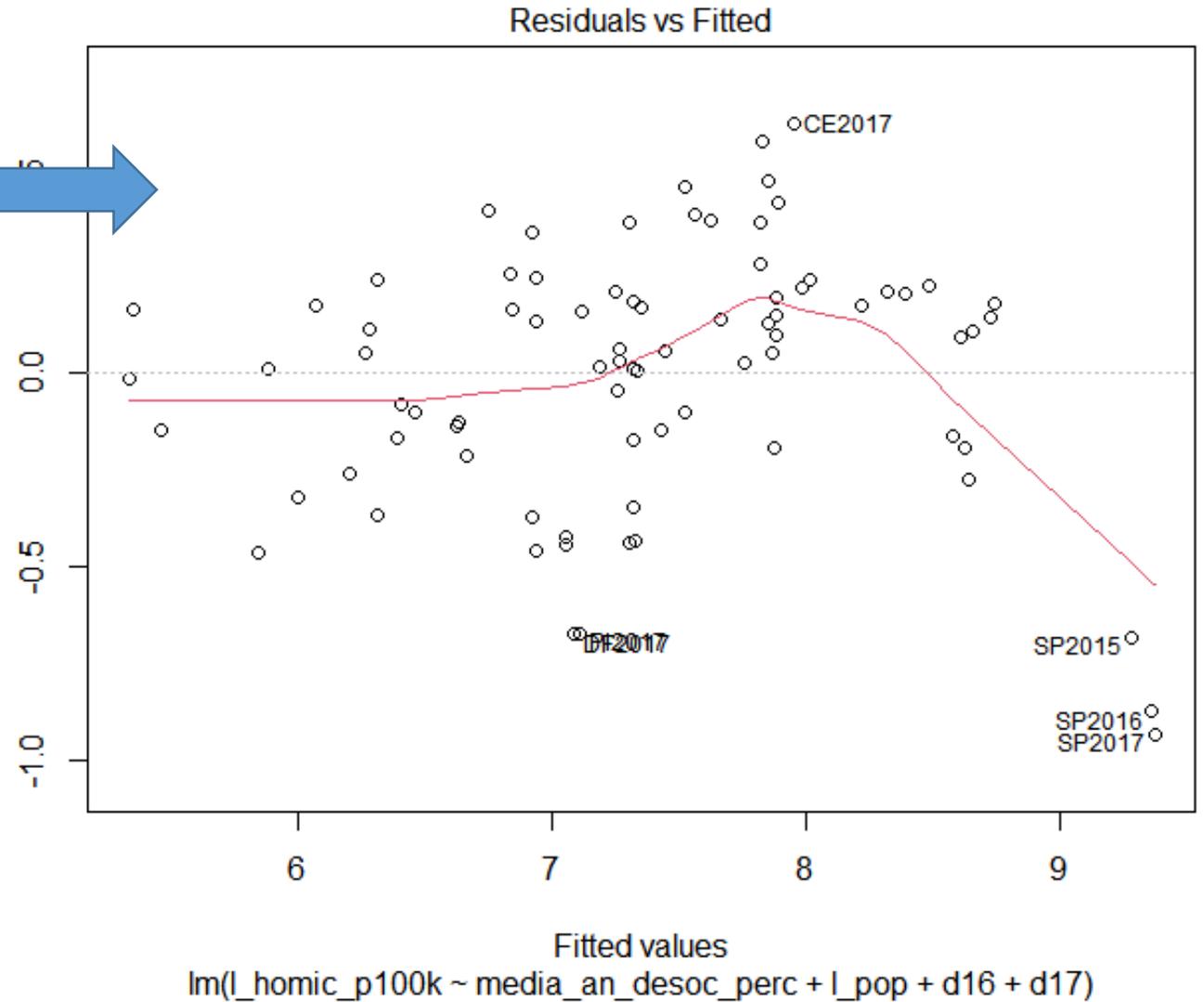
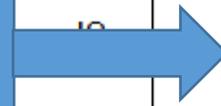
Dados no OSF: https://osf.io/5yx7g/?view_only=ac1691cced8549238d6d6e0a9d2b7f7b

Dependent variable:	
l_homic_p100k	
media_an_desoc_perc	0.077*** (0.015)
l_pop	0.841*** (0.036)
d16	-0.160 (0.101)
d17	-0.239** (0.110)
Constant	-6.247*** (0.571)
Observations	81
R2	0.883
Adjusted R2	0.877
Residual Std. Error	0.335 (df = 76)
F Statistic	143.023*** (df = 4; 76)

Note: *p<0.1; **p<0.05; ***p<0.01

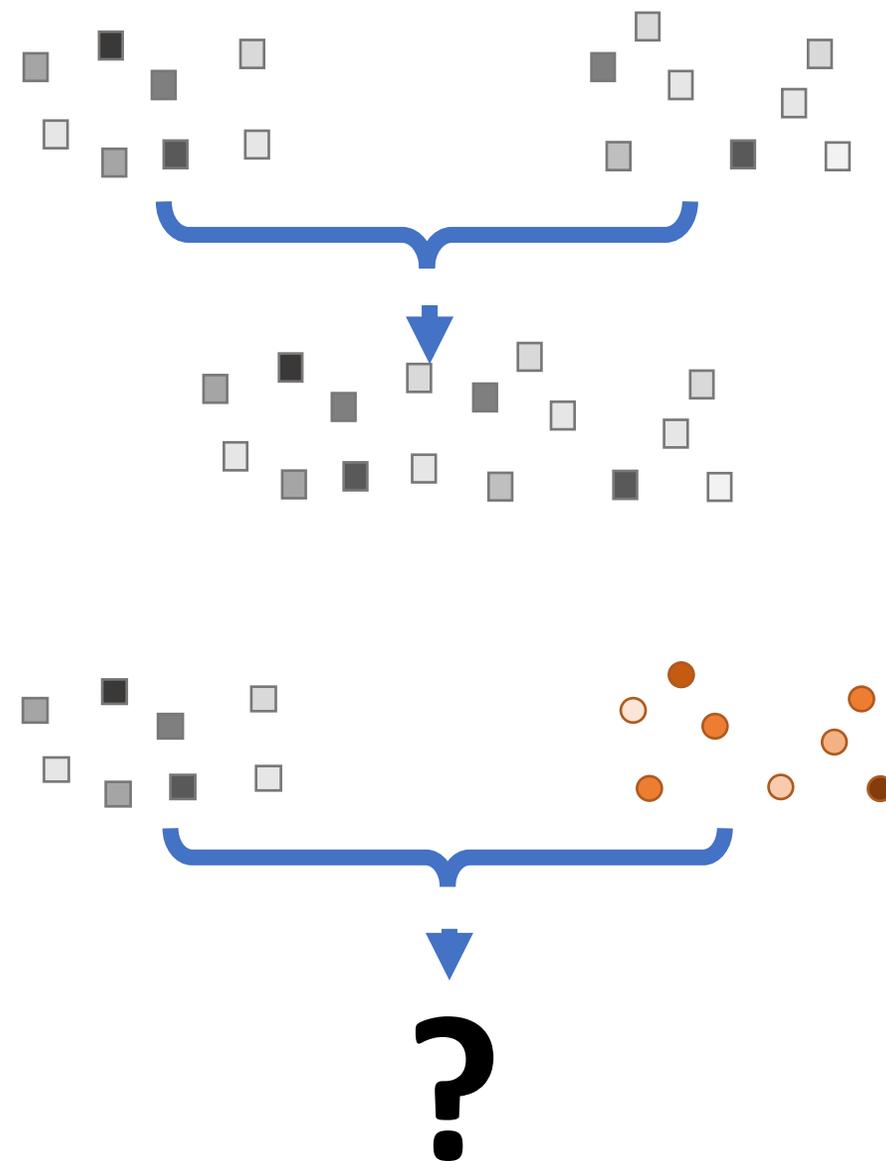


- O *plot* diagnóstico de resíduos vs. preditos sugere que pode haver alguma coisa peculiar, não-modelada, ocorrendo em DF, CE e SP, que **sistematicamente** leva seus resultados a divergirem das previsões do modelo
- Isso sugere que o *pooling* não é recomendado para esses dados



Quando se pode fazer *pooling*?

- Supõe que duas ou mais amostras podem ser combinadas em um único *pool* de observações
- O modelo linear original ainda é bom o bastante para descrever toda a variação sistemática
- **Ou seja, a introdução de mais observações não traz uma parcela extra importante de variação não-sistemática**
- Normalmente isso resulta de amostragens aleatórias (*independently pooled cross sections*) ou de modelos bem especificados, cujas VIs esgotem a variação sistemática adequadamente
- **Porém, um novo conjunto de observações que traz consigo heterogeneidade importante, não capturada pelas VIs, inviabiliza o *pooling***



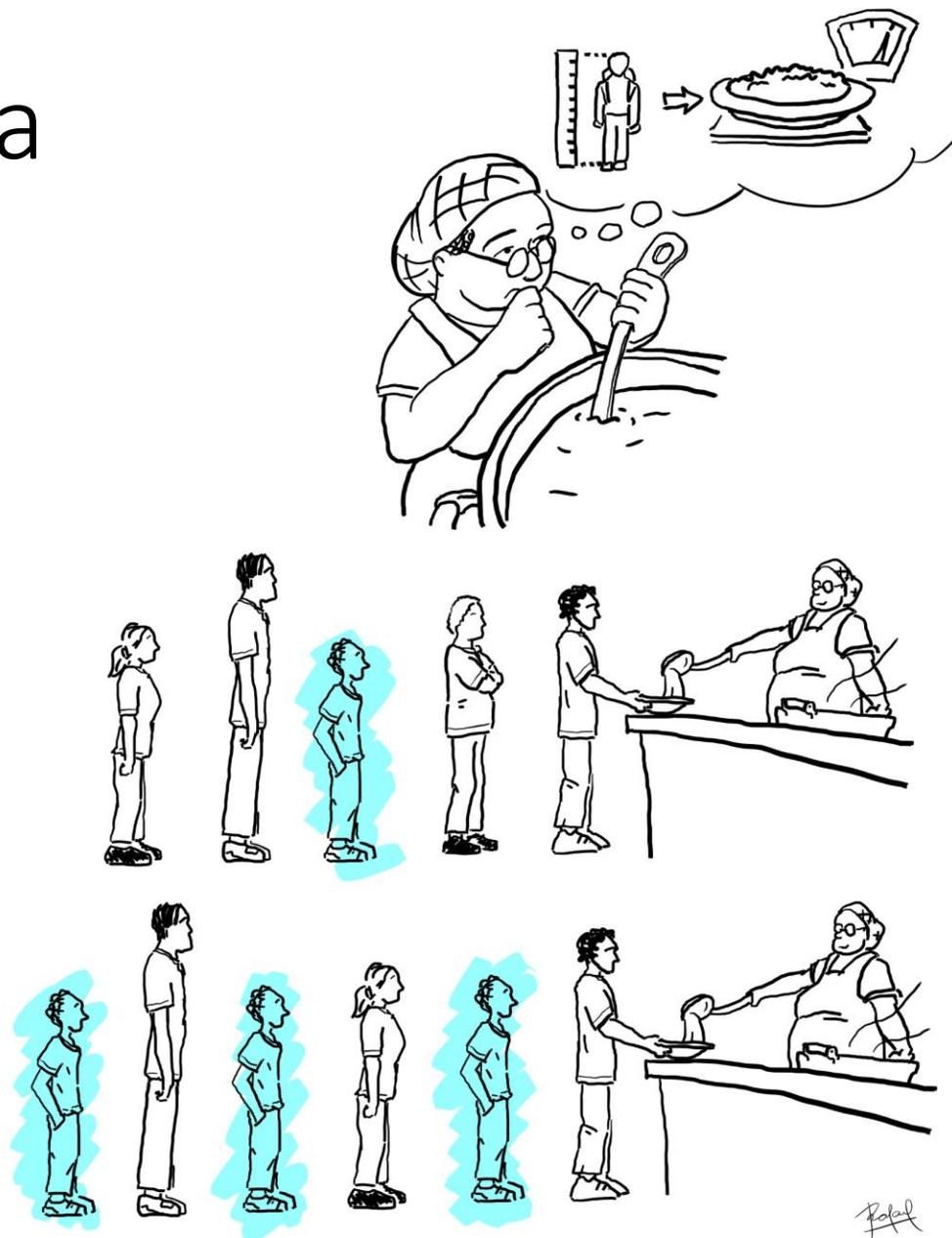
Quando se pode fazer *pooling*?

- Intuitivamente, nós imaginamos que observar repetidamente **o mesmo sujeito** nos leva a ter uma base de dados **mais homogênea**
- Mas, na verdade, a repetição pode levar a **mais heterogeneidade** e por isso desencorajar o *pooling*
- Tudo vai depender do quão satisfatoriamente as VIs explicam a variação dos dados
- Exemplo: “a fila da merenda”



Exemplo: a fila da merenda

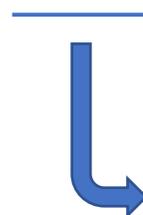
- Uma merendeira pode ter um “modelo mental” prevendo a **quantidade de comida que uma criança irá comer (Y)** como função de uma variável observada, a **altura da criança (X)**, mais **fatores aleatórios (ϵ)**
- Mas pode haver **uma criança** na fila que, mesmo baixinha, come muito mais do que seu tamanho sugere, por força de algum fator não-observado (ex.: gula)
- Se essa criança só entrar 1 vez na fila, o “modelo mental” da cozinheira não falha em servir a turma, porque o apetite “não modelado” deste aluno guloso será compensado pela moderação de algum outro estudante mais adiante na fila (ex.: um aluno que, mesmo alto, comeu pouco).
- Mas se a criança gulosa entrar **repetidamente** na fila, a comida “prevista” para a turma irá acabar mais cedo! Isso porque a heterogeneidade não-modelada que antes se dissipava na população e podia ser tratada como “ruído” se torna um fator mais persistente



Heterogeneidade não-observada

Decompondo o “erro composto”

$$Y_{it} = \alpha_i + \beta X_{it} + \varepsilon_{it}$$



$$\varepsilon_{it} = \mu_i + \nu_{it}$$



Efeito individual, fixo, ou heterogeneidade não-observada

Fixo para cada unidade espacial, invariável com o tempo

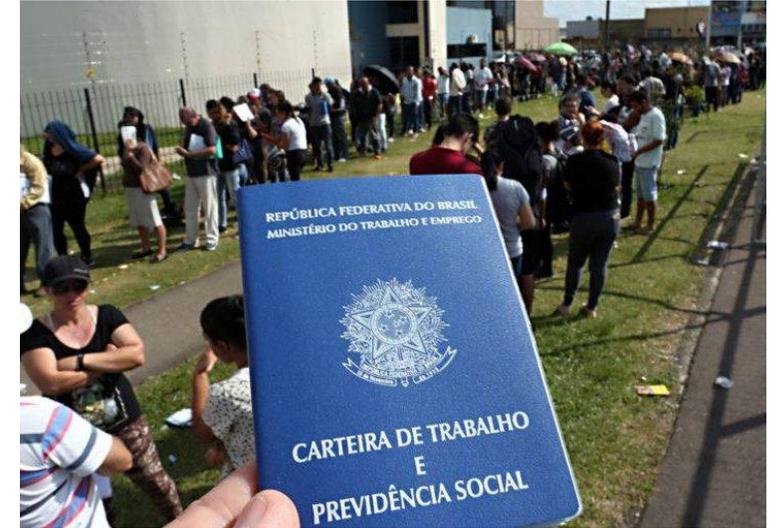


Efeito variável, ou idiossincrático

Variável para o tempo e unidades

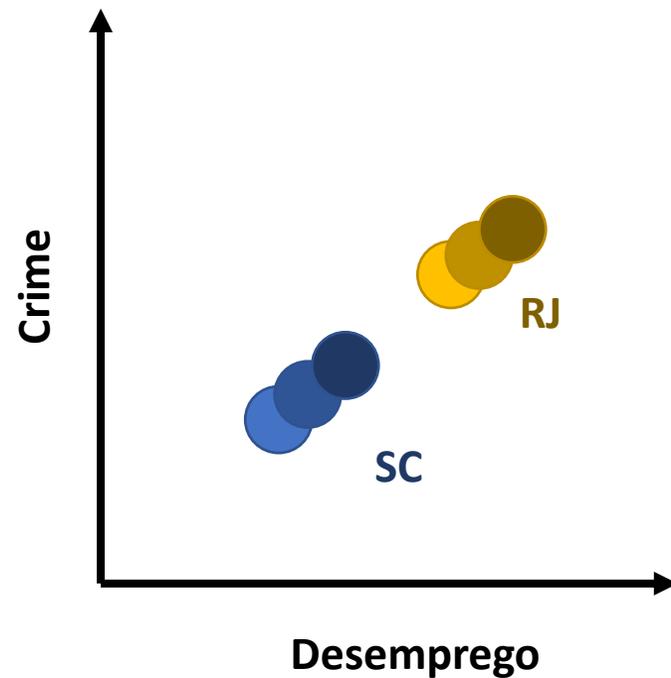
Heterogeneidade não-observada

- A heterogeneidade individual não-observada (μ_i) encapsula todos os fatores peculiares à unidade espacial, não capturados por X , que influem sobre o resultado Y
 - Ex.: Por que esses índices de criminalidade, se o desemprego é pequeno? “É porque é o *Rio de Janeiro*”
 - Criminalidade = desemprego + força do crime organizado
- Por que as UFs diferem na taxa de criminalidade? Se for apenas devido a níveis distintos de desemprego, as unidades não são heterogêneas.
 - Ex.: Se o RJ tivesse o desemprego de SC, teria a mesma taxa de criminalidade desta



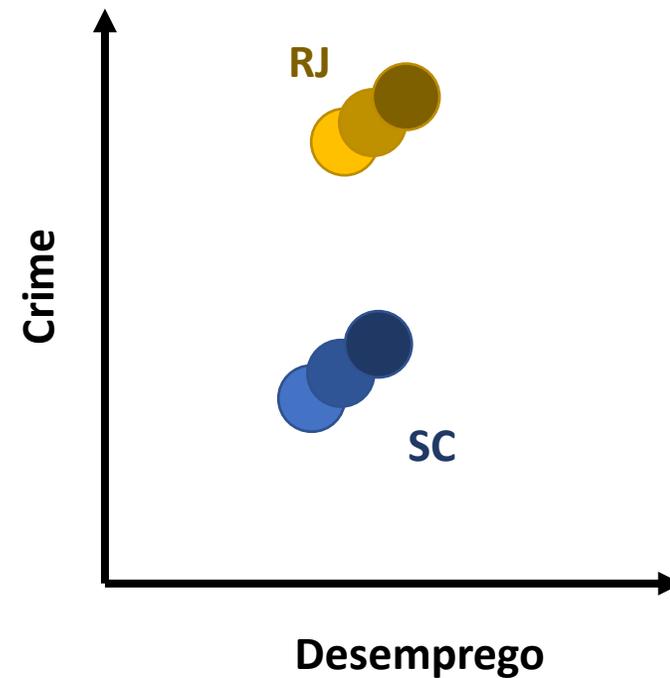
Heterogeneidade não-observada

Sem heterogeneidade



Mais desemprego,
mais crime

Com heterogeneidade



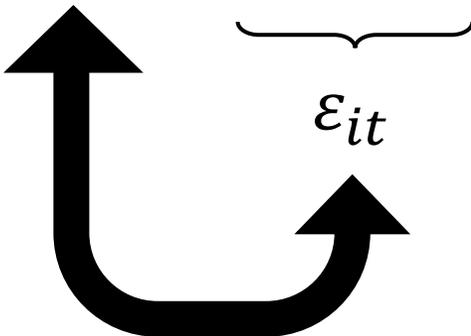
Mesmo nível de
desemprego,
mas crime
distinto

Heterogeneidade não-observada

- Se a heterogeneidade não-observada for desprezível, POLS será apropriado
- Se houver heterogeneidade, sua presença violará pressupostos clássicos da regressão linear

Como heterogeneidade viola pressupostos

- **Pressuposto da exogeneidade/média condicional zero:**
- **$E(\varepsilon | x) = 0$**
- Se μ_i tiver correlação com X_{it} então teremos um caso de viés de variável omitida

$$Y_{it} = \alpha_i + \beta X_{it} + \underbrace{\mu_i + v_{it}}_{\varepsilon_{it}}$$


Como heterogeneidade viola pressupostos

- **Pressuposto da amostra aleatória/ausência de autocorrelação:**
- **$\text{Corr}(\varepsilon_s, \varepsilon_j) = 0$ para $s \neq j$**
- Se μ_i for constante para cada unidade i , os erros ε em $t1$ serão correlacionados com os erros em $t2$, já que certamente μ_i estará presente nas duas observações

$$Y_{it} = \alpha_i + \beta X_{it} + \mu_i + v_{it}$$

$$t1 \quad Y_{i1} = \alpha_i + \beta X_{i1} + \underline{\mu_i} + v_{i1}$$

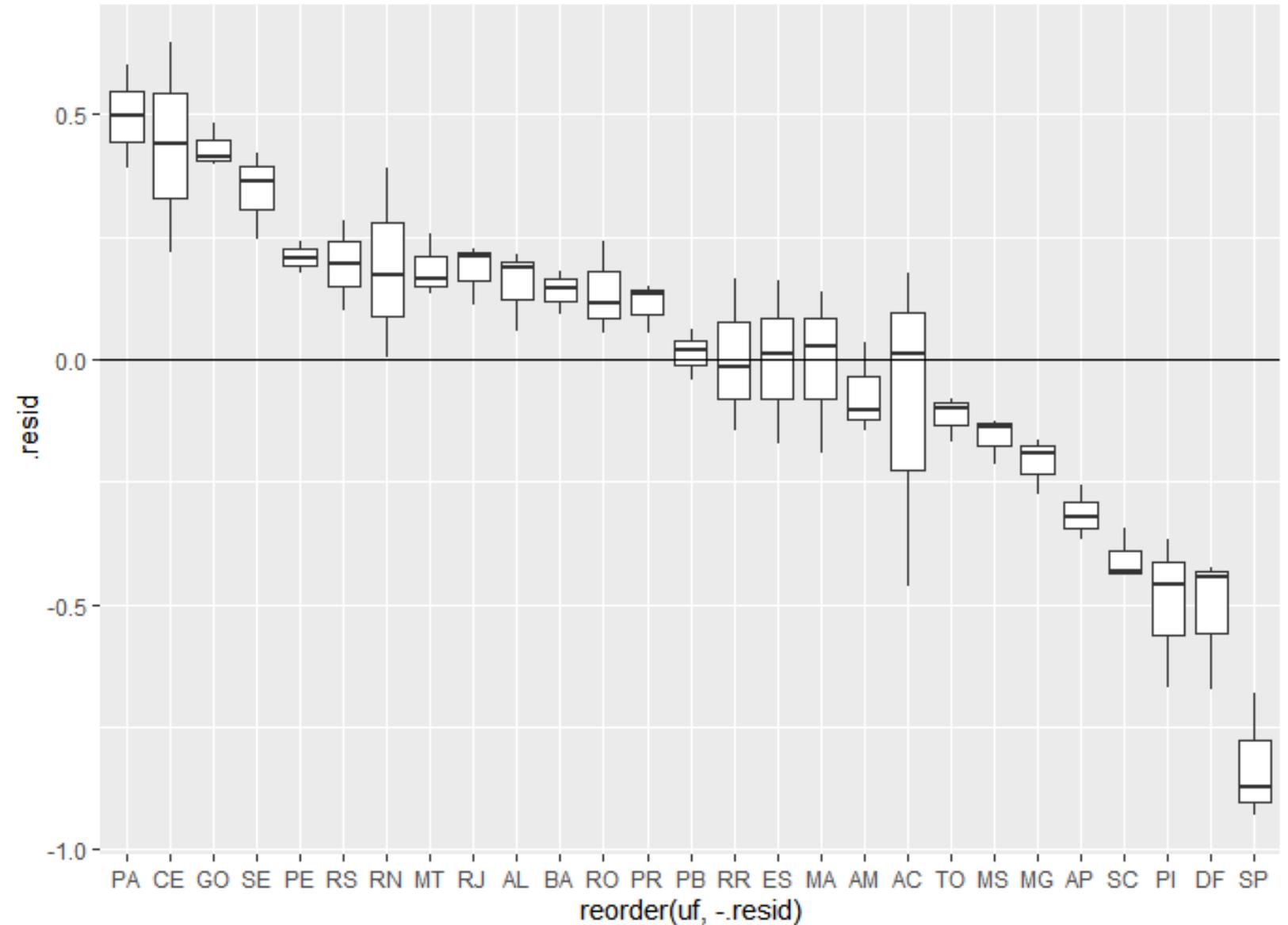
$$t2 \quad Y_{i2} = \alpha_i + \beta X_{i2} + \underline{\mu_i} + v_{i2}$$

Diagnosticando heterogeneidade (I)

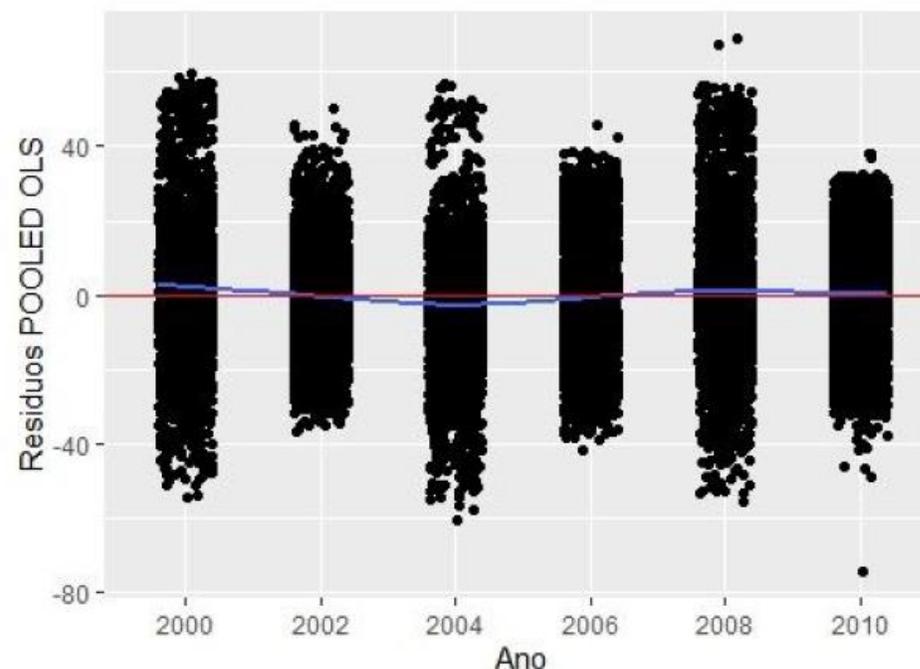
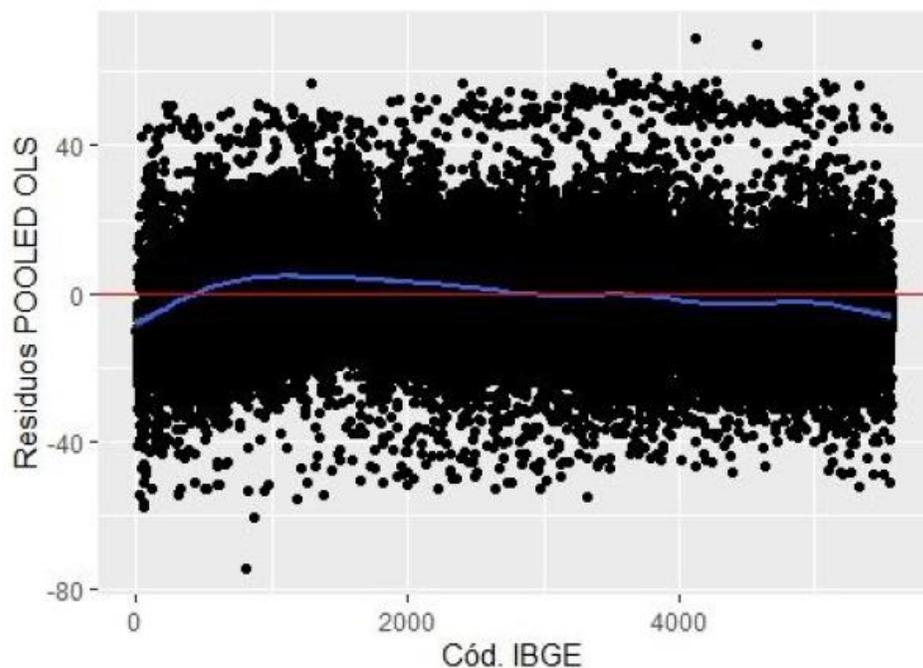
Diagnosticando heterogeneidade (I)

- (1) Conhecimento substantivo
 - Ex.: Podemos esperar que a criminalidade no RJ seja causada pelos mesmos determinantes que na SC?
- (2) Análise gráfica
 - Agrupar resíduos por unidade espacial ou por ano

Resíduos do modelo
Pooled OLS para
criminalidade,
agrupados por UF



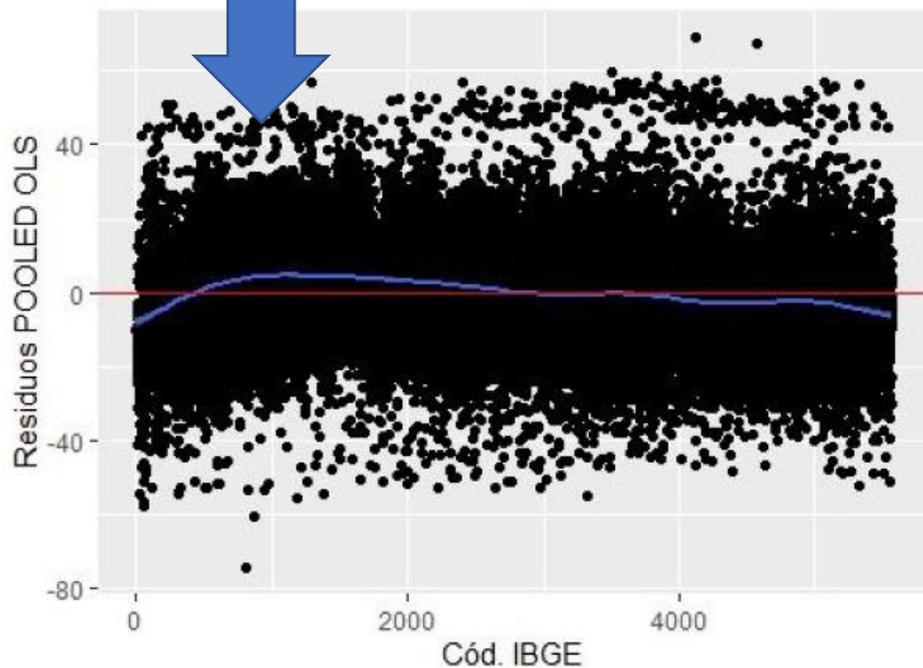
Dados de Fernandes e Fernandes (2017) para “votos no incumbente” em eleições locais e nacionais de 2000 a 2010, para +5mil municípios. Resíduos agrupados por município e por ano/eleição



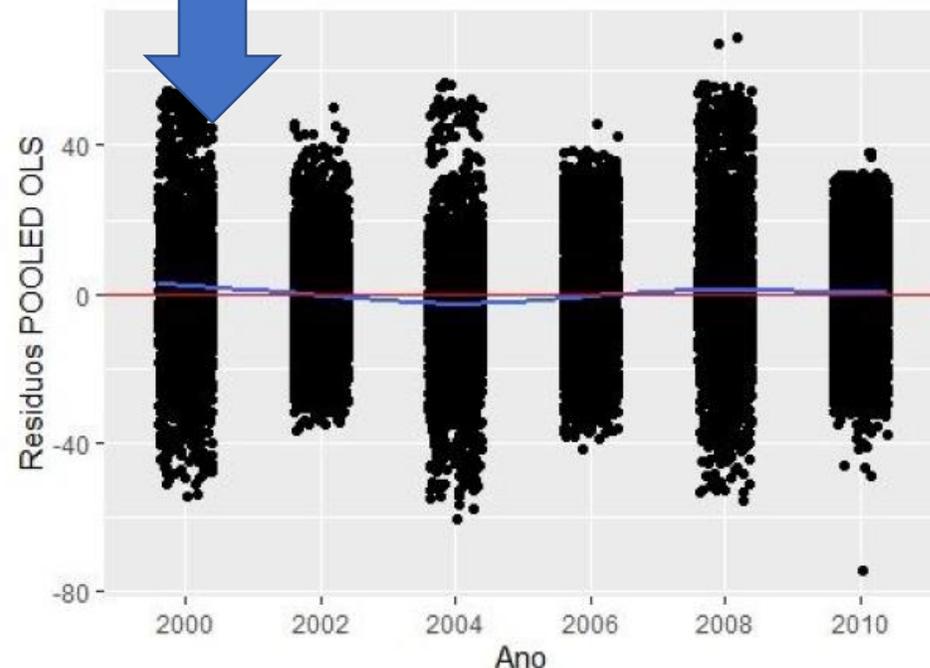
Fonte: Mesquita et al. (2021), a partir dos dados de Fernandes e Fernandes (2017)

Dados no OSF: https://osf.io/5yx7g/?view_only=ac1691cced8549238d6d6e0a9d2b7f7b

Há algo que sistematicamente diferencia as previsões para municípios de código IBGE 500 a 3000 daqueles de código >4000



Há algo que sistematicamente diferencia as previsões para eleições majoritárias em 2002, 2006 e 2010 (nacionais?)



Fonte: Mesquita et al. (2021), a partir dos dados de Fernandes e Fernandes (2017)

Dados no OSF: https://osf.io/5yx7g/?view_only=ac1691cced8549238d6d6e0a9d2b7f7b

Eliminando heterogeneidade: Primeiras Diferenças

Primeiras Diferenças

- Primeiras diferenças (ou *First differences*, FD) são uma forma de remover os efeitos fixos individuais (μ_i) e permitir uma regressão MQO livre de violações
- Como μ_i é o mesmo para cada i ao longo do tempo, é possível subtrair (“diferenciar”), para cada i , uma observação da sua anterior, e com isso elementos invariantes no tempo são removidos e apenas os variáveis se preservam

Primeiras Diferenças

$$Y_{it} = \alpha_i + \beta X_{it} + \mu_i + \nu_{it}$$

$$Y_{i1} = \alpha_i + \beta X_{i1} + \mu_i + \nu_{i1} \quad t1$$

$$- \left[Y_{i2} = \alpha_i + \beta X_{i2} + \mu_i + \nu_{i2} \right] \quad t2$$

$$\Delta Y_i = \beta \Delta X_i + \Delta \nu_i$$

ID	Tempo	Taxa de homicídios	Taxa de desemprego
i	t	Y_{it}	X_{it}
Acre	2011	22	5,4
Acre	2012	27,4	8
Acre	2013	30,1	9,6
Acre	2014	29,4	9,8
Bahia	2011	39,4	10,5
Bahia	2012	43,4	10,2
Bahia	2013	37,8	9,9
Bahia	2014	40	10
Santa Catarina	2011	12,8	3,6
Santa Catarina	2012	12,9	3,1
Santa Catarina	2013	11,9	3,4
Santa Catarina	2014	13,5	3,1



ID	Tempo	Taxa de homicídios	Taxa de desemprego
i	t	ΔY_{it}	ΔX_{it}
Acre	d1	5,4	2,6
Acre	d2	-0,7	0,2
Bahia	d1	4	-0,3
Bahia	d2	2,2	0,1
Santa Catarina	d1	0,1	-0,5
Santa Catarina	d2	1,6	-0,3

Logo, idealmente T é par. Se T for ímpar, 1 período transversal é desperdiçado

Obs.: concretamente, só diferenciamos elementos observáveis da planilha (X e Y). Porém, os não observados (μ_i) também são removidos quando da regressão

Fonte:
Mesquita et al. (2021)

Primeiras Diferenças

- **Benefícios de FD**
- Espera-se que a heterogeneidade não-observada seja um problema em diversos tipos de dados. FD se beneficia das mensurações repetidas para livrar-se dessa fonte de viés
- É uma forma de sanar casos de autocorrelação

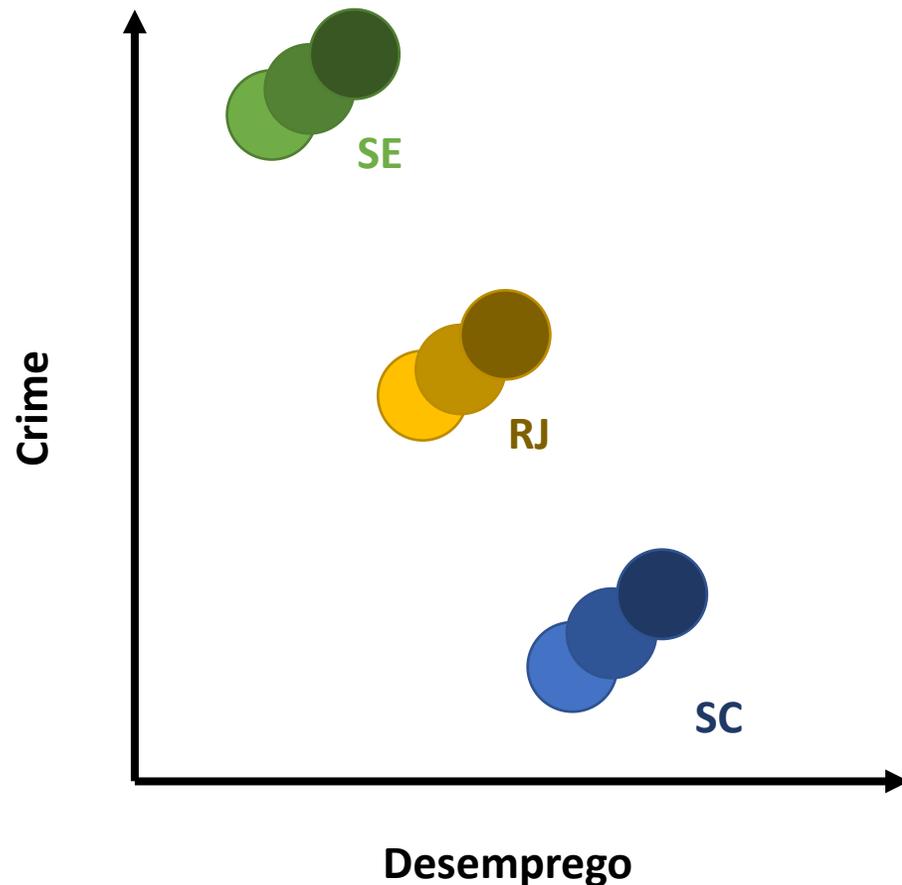
Primeiras Diferenças

- **Limitações de FD**
- Tudo o que for invariável no tempo (ou variar monotonicamente) é diferenciado (ex.: geografia, gênero)
- Diminuição do número de observações pela metade
- Aumento dos erros-padrão

- Por isso, recomenda-se FD como uma solução drástica para casos de autocorrelação severa (Beck 2008) e quando se tem um número razoavelmente alto de observações

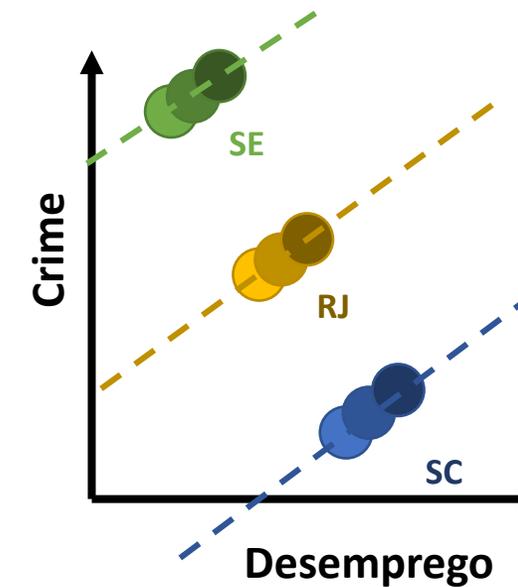
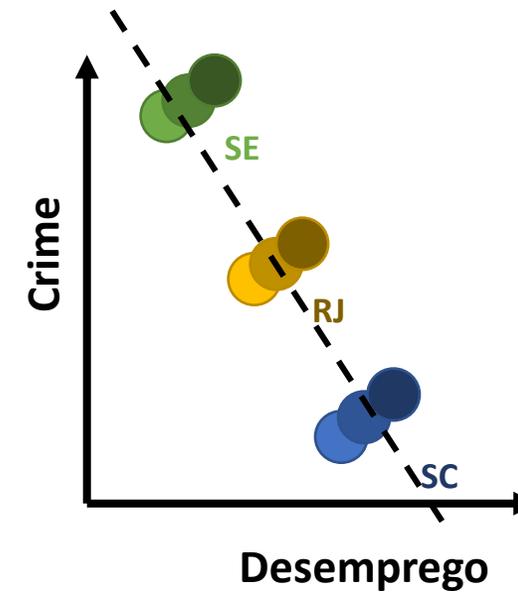
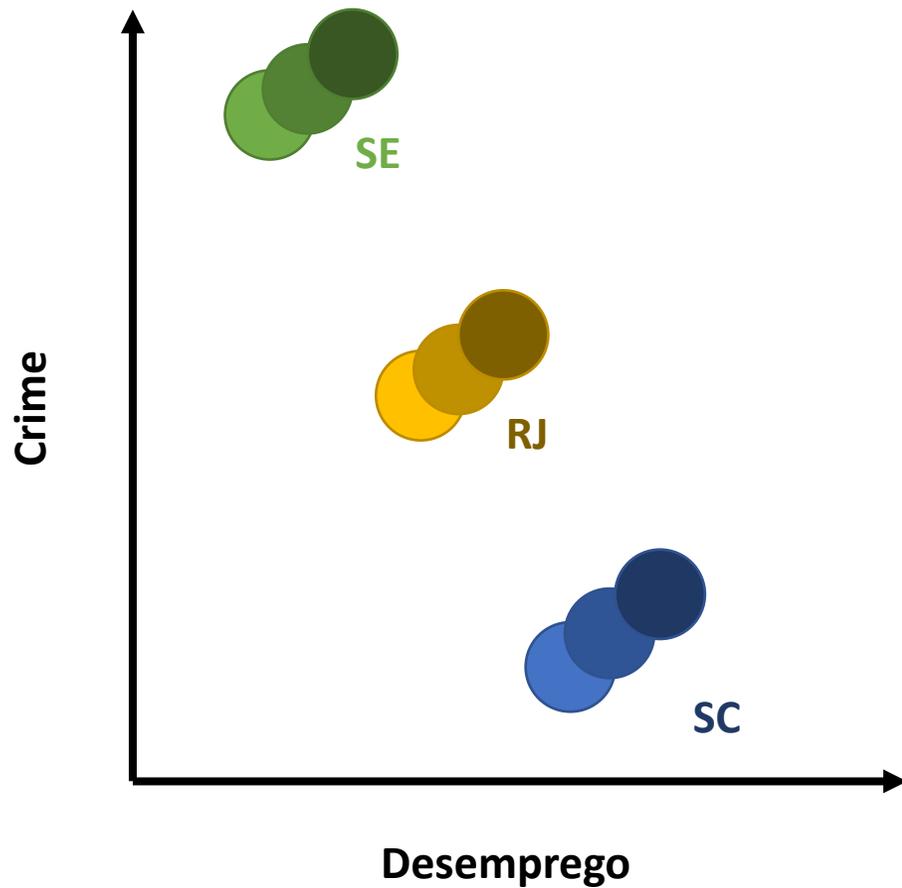
Efeitos Fixos

Uma ou várias retas de regressão?



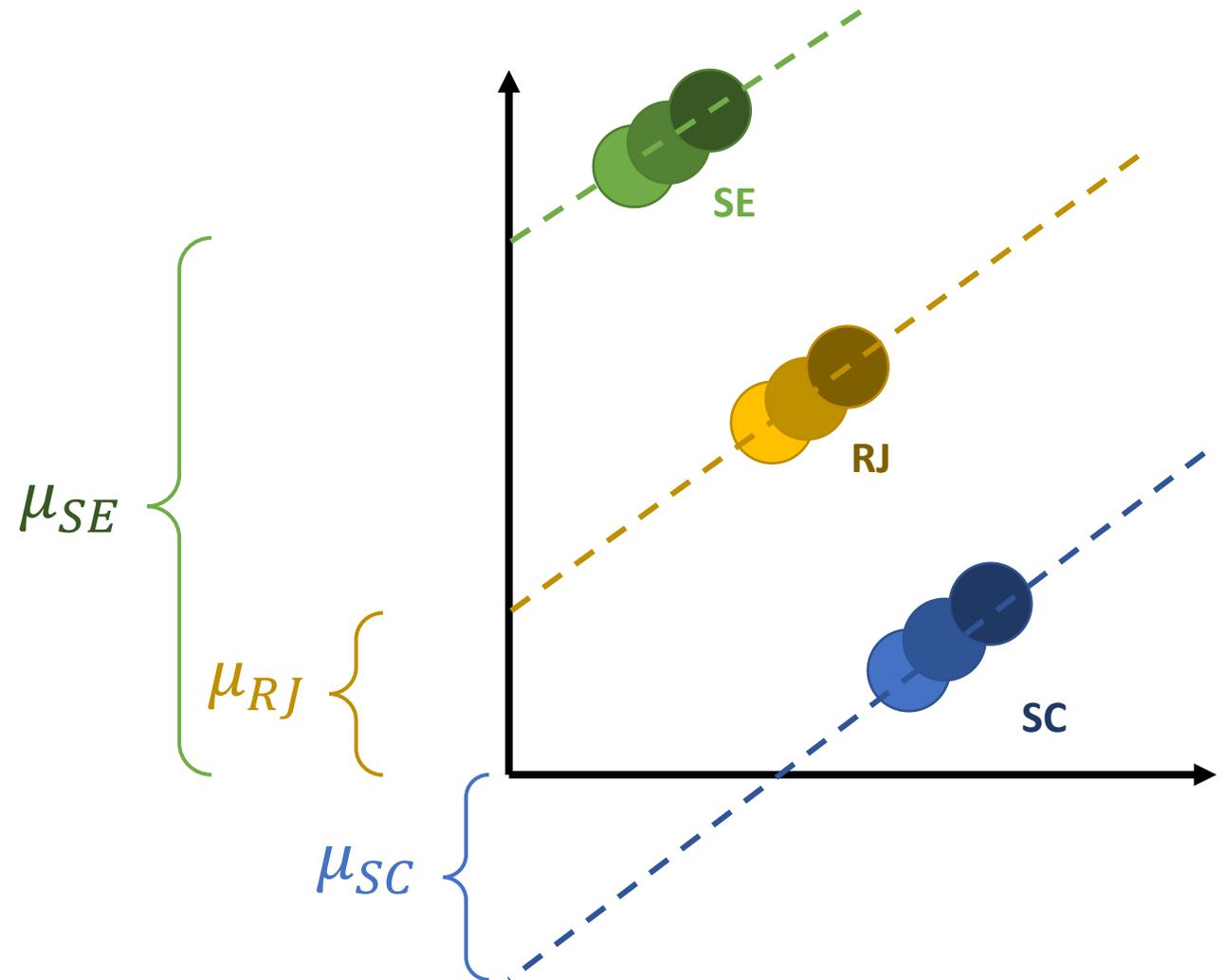
- Uma outra forma de entender a relação entre heterogeneidade não-modelada e *pooling* é pensar se é possível traçar uma única linha de regressão que cubra adequadamente todos os casos
- No exemplo ao lado, essa única linha para toda a população de casos iria prever uma relação **negativa** entre desemprego e crime

Uma ou várias retas de regressão?



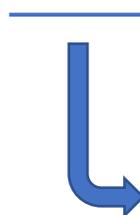
Uma ou várias retas de regressão?

- Cada UF deveria ter seu próprio intercepto
- Mas o ângulo é o mesmo, de modo o β é o mesmo para todas as unidades



Recap.: decompondo o “erro composto”

$$Y_{it} = \alpha_i + \beta X_{it} + \varepsilon_{it}$$



$$\varepsilon_{it} = \mu_i + \nu_{it}$$

Efeito fixo, ou
heterogeneidade não-
observada



Erro variável, ou
idiossincrático

$$Y_{it} = \alpha_i + \beta X_{it} + \mu_i + \nu_{it}$$

Modelo padrão

$$Y_{it} = \alpha_i + \beta X_{it} + \mu_i + v_{it}$$

- Semelhante ao FD, FE também busca um meio de apagar a heterogeneidade individual (μ_i)

Subtração da média individual

$$Y_{it} = \alpha_i + \beta X_{it} + \mu_i + v_{it}$$
$$- \left[\bar{Y}_i = \alpha_i + \beta \bar{X}_i + \mu_i + \bar{v}_i \right]$$

$$\ddot{Y}_i = \beta \ddot{X}_i + \ddot{v}_i$$

- *Time demeaning, group-mean centering* ou *within transformation*: subtração cada observação it da média de um indivíduo i para todo o t

Efeitos Fixos / Fixed Effects

ID	Tempo	Taxa de homicídios	Taxa de desemprego	Média Individual (homic.)	Média Individual (desemp.)
i	t	Y_{it}	X_{it}	\bar{Y}_i	\bar{X}_i
Acre	2011	22	5,4	27,2	8,2
Acre	2012	27,4	8	27,2	8,2
Acre	2013	30,1	9,6	27,2	8,2
Acre	2014	29,4	9,8	27,2	8,2
Bahia	2011	39,4	10,5	40,2	10,2
Bahia	2012	43,4	10,2	40,2	10,2
Bahia	2013	37,8	9,9	40,2	10,2
Bahia	2014	40	10	40,2	10,2
Santa Catarina	2011	12,8	3,6	12,8	3,3
Santa Catarina	2012	12,9	3,1	12,8	3,3
Santa Catarina	2013	11,9	3,4	12,8	3,3
Santa Catarina	2014	13,5	3,1	12,8	3,3

ID	Tempo	Desvio Intra-Caso (homic.)	Desvio Intra-Caso (desemp.)
i	t	$Y_{it} - \bar{Y}_i$	$X_{it} - \bar{X}_i$
Acre	2011	-5,2	-2,8
Acre	2012	0,2	-0,2
Acre	2013	2,9	1,4
Acre	2014	2,2	1,6
Bahia	2011	-0,8	0,4
Bahia	2012	3,3	0,0
Bahia	2013	-2,4	-0,3
Bahia	2014	-0,1	-0,2
Santa Catarina	2011	0,0	0,3
Santa Catarina	2012	0,1	-0,2
Santa Catarina	2013	-0,9	0,1
Santa Catarina	2014	0,7	-0,2



Within transformation

Efeitos fixos do tempo e *two ways*

- Ao invés de efeitos para cada indivíduo i , é também possível supor efeitos fixos para cada período temporal t (ex.: choque exógeno, como uma crise financeira)

$$Y_{it} = \alpha_i + \beta X_{it} + \eta_t + v_{it}$$

- Combinando efeitos fixos para indivíduos e para períodos temporais, tem-se o *two ways* FE

$$Y_{it} = \alpha_i + \beta X_{it} + \mu_i + \eta_t + v_{it}$$

Vantagens e desvantagens do FE

- **Vantagens**

- Intuitivo (idêntico a POLS com *dummies* para todas as unidades i : *Least Squares Dummy Variable, LSDV*)
- Livra-se da heterogeneidade individual não-modelada
- Abordagem mais popular para dados observacionais
- Permite visualizar o intercepto individual das unidades i

- **Desvantagens**

- Similar a FD, variáveis que não variam com o tempo não são computadas (ex.: geografia, gênero)
- Menos parcimonioso
- Menor eficiência e menor precisão, pois, para cada nova unidade i , calcula-se novo intercepto, e usa-se apenas a variabilidade intra-indivíduo (que, normalmente é menor que a entre-indivíduos)
- Sensibilidade à amostra

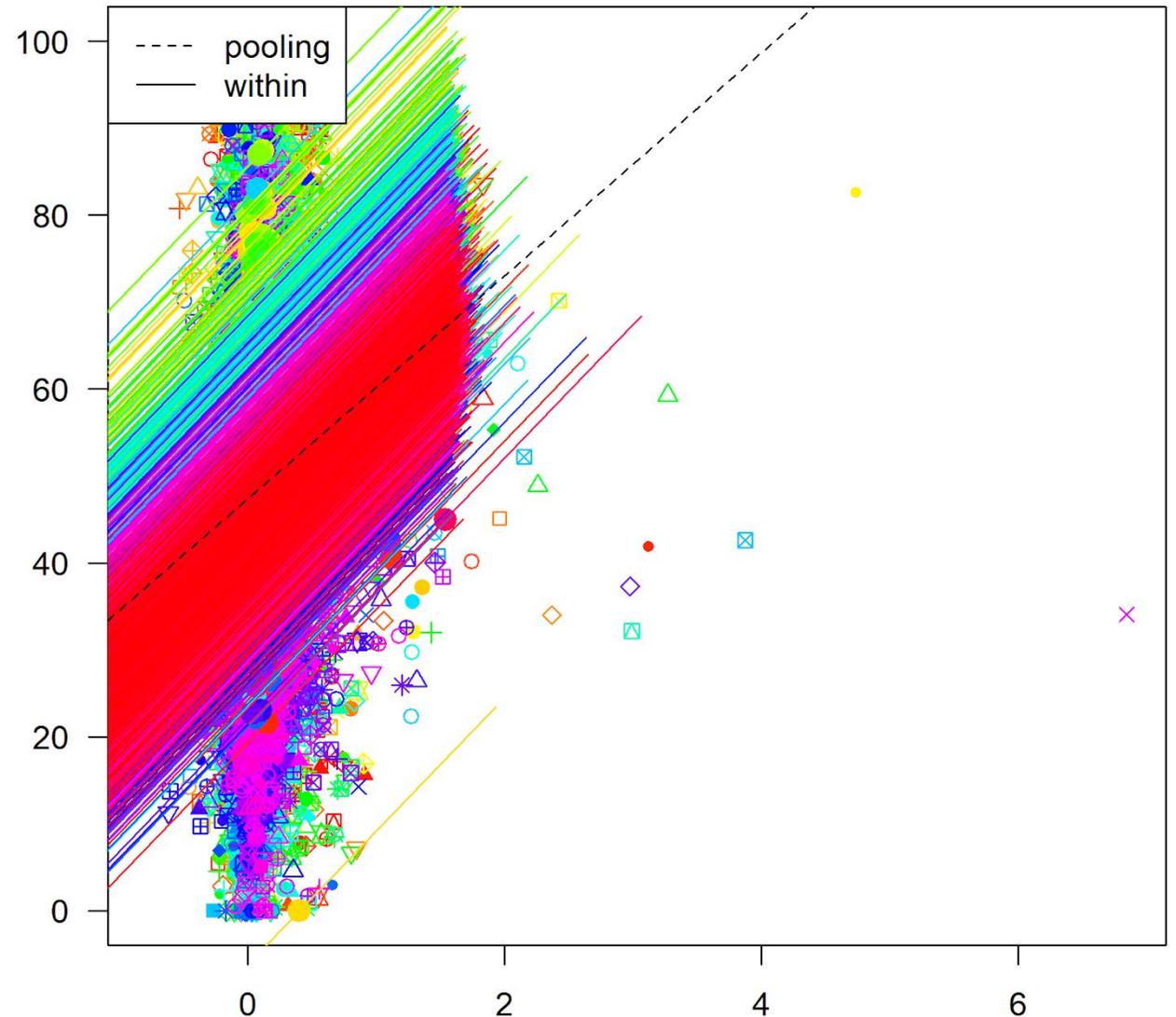
Como sei se devo usar FE?

Diagnosticando heterogeneidade (II)

- (1) Conhecimento substantivo
- (2) Análise gráfica
- (3) Testes

Análise gráfica

- Objetivo é averiguar se é preferível cobertura por única reta de regressão e intercepto (POLS) vs. várias retas e interceptos individuais (FE)
- Plots *pooling* vs. *within* são úteis para relações bivariadas $X \sim Y$



Testes

Há dois testes para recomendar o abandono do POLS:

F-Test

- POLS vs. FE
- Teste F para verificar se efeitos individuais são conjuntamente significativos
- Hipótese nula:
 - $\mu_i = \dots \mu_N = 0$
(os efeitos individuais são iguais a zero)
- Hipótese alt.:
 - $\mu_i \neq \dots \mu_N \neq 0$

Breusch-Pagan LM

(não confundir com teste Breusch-Pagan de heteroscedasticidade!)

- POLS vs. RE
- Teste para verificar se a variância dos efeitos individuais é zero
- Hipótese nula:
 - $\sigma_{\mu_i}^2 = 0$
(variância dos efeitos indiv. é igual a zero)
- Hipótese alt.:
 - $\sigma_{\mu_i}^2 \neq 0$

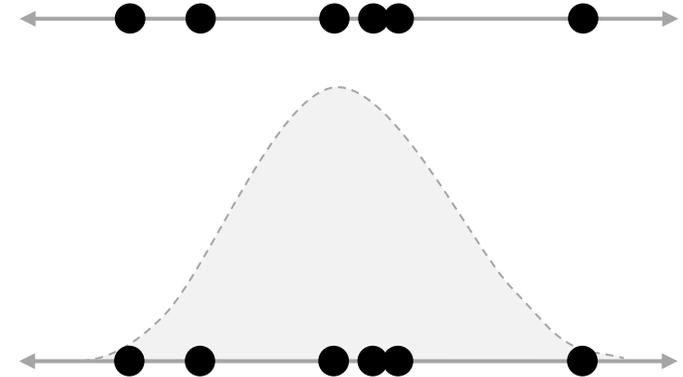
Testes

- Rejeição das H_0 sugere que há sim efeitos individuais fixos significativos
- Podem testar também efeitos temporais ou *two ways*

Efeitos Aleatórios

Aleatoriedade

- A heterogeneidade individual (μ_i) pode ser tratada como um parâmetro isolado (“cada caso é um caso”), ou como **realizações aleatórias** pertencentes a uma mesma distribuição
- Essa singularidade vs. generalidade dos efeitos individuais depende da natureza dos dados
 - Ex.: Países exportadores na economia mundial (casos muito singulares e heterogêneos) vs. Respondentes de um *survey* (casos mais similares e homogêneos)



Aleatoriedade

- A abordagem RE supõe que os efeitos individuais μ_i vêm de uma mesma **distribuição normal**, cuja média e variância são computadas a partir dos dados coletados.

$$\mu_i \sim N(0, \sigma_\mu^2)$$

Recap.: Viés de variável omitida

- Eliminamos o efeito fixo (μ_i) por meio de FD ou FE por ser uma forma de viés de variável omitida
- Relembrando: a omissão de uma variável “Z” gera viés se:
 - a. Z tiver correlação com Y; ou
 - b. Z tiver correlação com X.
- Então **não** haverá risco de viés se:
 - a. **A variável omitida Z não estiver correlacionada com Y; ou**
 - b. **A variável omitida Z não estiver correlacionada com X**

RE

- RE supõe que não há relação entre o efeito individual μ e X , de modo que aquele não precisa ser removido via FE ou FD
- *Mas se pressupomos que μ e X são independentes, por que então não usar simplesmente Pooled OLS?*
 - Resposta: O efeito fixo μ não é considerado fonte de viés na abordagem RE, mas μ ainda pode provocar correlação serial nos erros e portanto ainda é preciso lidar com ele (rever slides “Como a heterogeneidade viola pressupostos”)

Quasi-demeaning

- RE propõe o *quasi-demeaning* dos dados: subtraem-se as observações de uma *parcela* da média intra-grupo

$$Y_{it} - \theta \bar{Y}_i = \alpha_i(1 - \theta) + \beta(X_{it} - \theta \bar{X}_i) + (\varepsilon_{it} - \theta \bar{\varepsilon}_i)$$

- O tamanho da parcela é dado por θ , que é estimado:

$$\hat{\theta} = 1 - \sqrt{\frac{1}{1 + T \left(\frac{\sigma_{\mu}^2}{\hat{\sigma}_{\nu}^2} \right)}}$$

(em Wooldrige 2013, p.493)

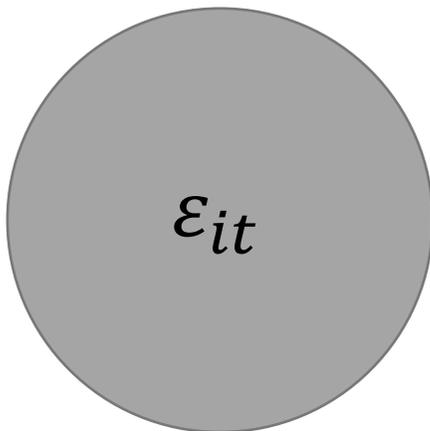
Theta

- θ varia de 0 a 1
- Pode ser interpretado como um diagnóstico de qual componente mais contribui com a variância do “erro composto” (ε_{it}): se o efeito fixo (μ_i) ou idiossincrático (ν_{it})
- RE como “meio termo” entre POLS e FE
 - Quando há pouca variância nos efeitos fixos (μ_i), então $\theta \rightarrow 0$ e as estimativas de RE se aproximam de POLS
 - Quando há muita variância nos efeitos fixos, então $\theta \rightarrow 1$ e RE terá resultados próximos a FE

Então as quatro abordagens diferem basicamente sobre o que fazer quanto à heterogeneidade não observada

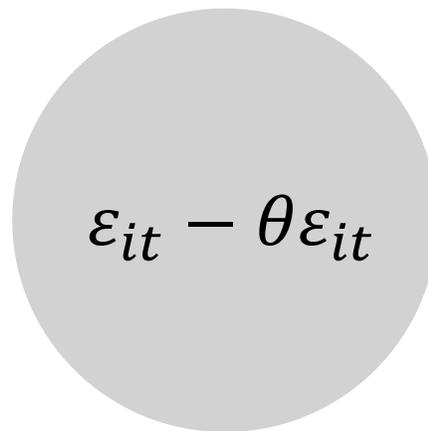
POLS

- Nada é feito quando à heterogeneidade


$$\epsilon_{it}$$

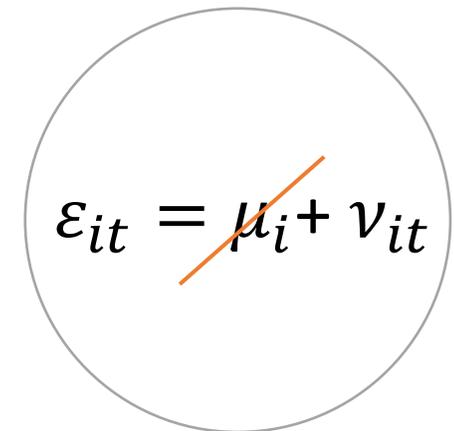
RE

- Elimina-se parcela da heterogeneidade


$$\epsilon_{it} - \theta\epsilon_{it}$$

FE/FD

- Elimina-se toda a heterogeneidade


$$\epsilon_{it} = \cancel{\mu_i} + \nu_{it}$$

Vantagens e desvantagens RE

- **Vantagens**

- Como $X_i - \theta X_i \neq 0$ (se $\theta \neq 1$), é possível inserir X de valor constante no tempo (ex.: geografia, gênero, etc.)
- Mais eficiente (erros padrão menores)

- **Desvantagens**

- Sempre terá viés, tanto mais grave quanto maior a $cov(X, \mu_i)$

FE ou RE?

- Escolha deve se pautar na:
- **(1) Natureza dos dados**
 - Intercambialidade: o *nome* dos casos importa? Isso pode revelar se é devido considerar os casos como realizações aleatórias ou não
 - Observacional (FE) vs. experimental (RE)
- **(2) Interesse substantivo por variáveis fixas no tempo**
- **(3) Viés [β] vs. Ineficiência [erro padrão(β), p-valor]**
 - FE: -viés, +ineficiência
 - RE: +viés, -ineficiência
- **(4) Teste de Hausman**

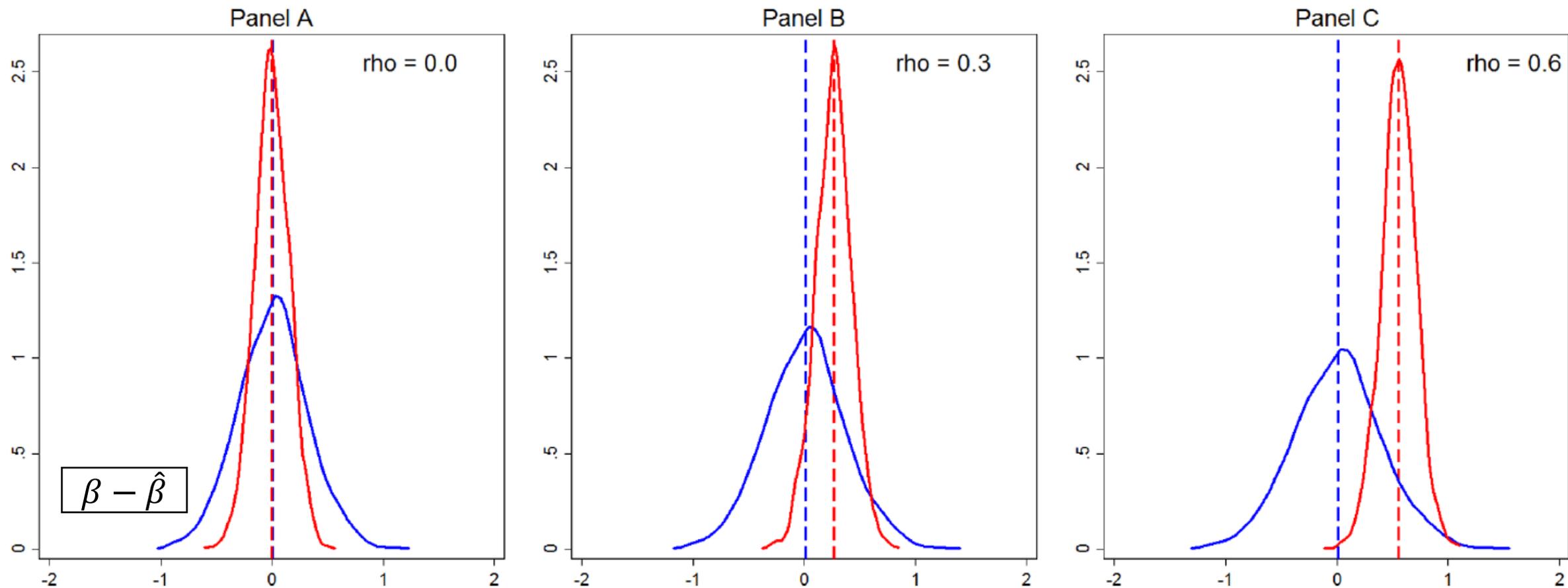


Figure 2. Distribution of errors. Red lines show the distribution of errors from RE estimation, while the blue lines show the distribution of errors from FE estimation. Each panel shows the correlation between the explanatory variable and the group-level effect set to a different value of ρ (0.0, 0.3, 0.6), increasing from left to right. Simulation based on the correct specification of a model with 50 groups and 10 observations per group. 50% of the variation of the outcome variable is explained by residuals, while only 10% of the variation in the explanatory variable is within-groups. doi:10.1371/journal.pone.0110257.g002

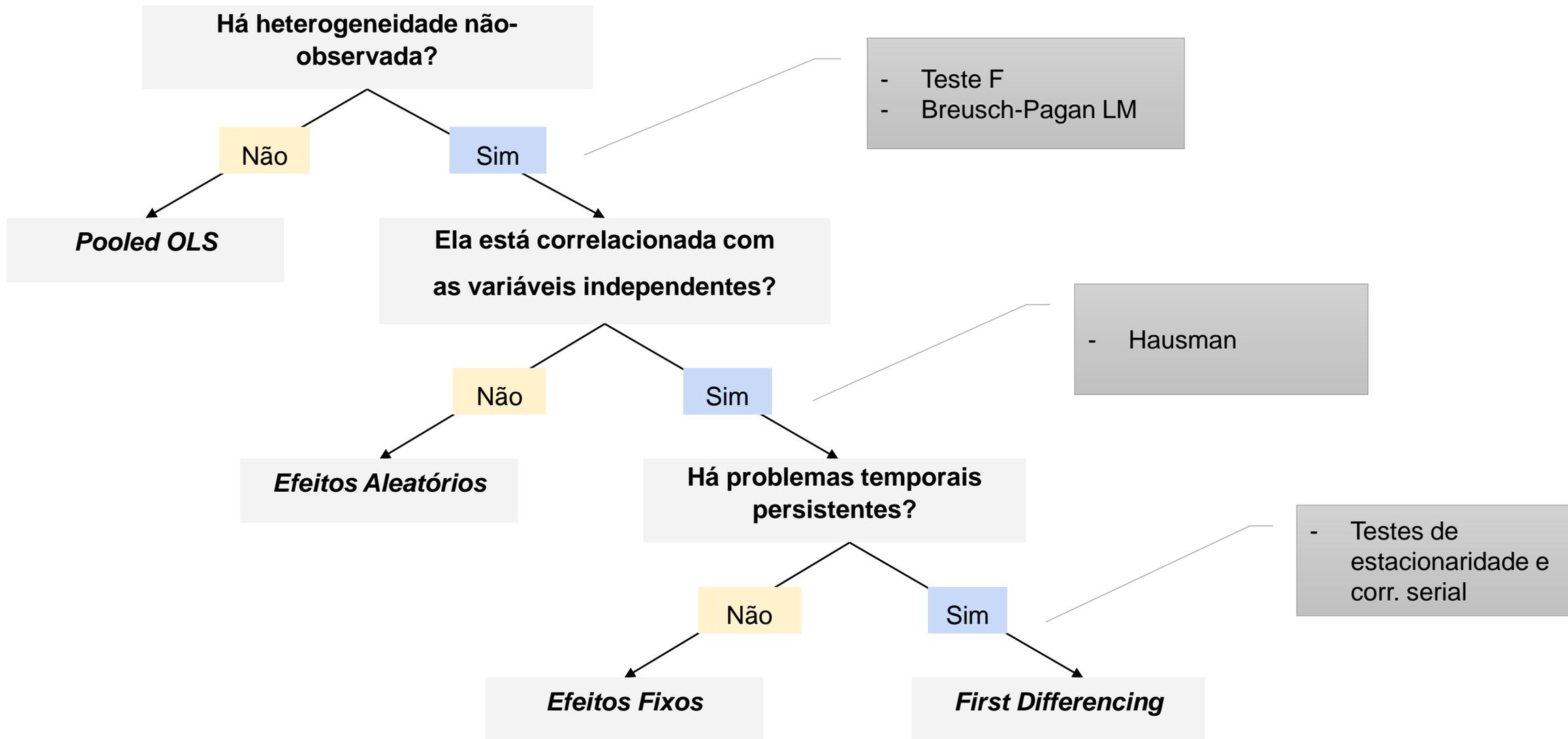
Fonte: Dieleman e Templin (2014, p. 5)

Teste de Hausman (1978)

- **H0: $\hat{\beta}_{fe} = \hat{\beta}_{re}$** (os 2 estimadores são consistentes)
 - Sob H0, prefere-se então RE por ser mais eficiente
- É um teste feito para modelos bem especificados. Supõe-se:
 - 1. Não há *misspecification*
 - 2. O erro idiossincrático v_{it} é independente de X_{it} e μ_i
 - 3. O erro idiossincrático não tem correlação serial nem heteroscedasticidade
 - 4. Normalmente, amostras grandes
- Em havendo violações: há versão robusta do teste

Escolhendo um estimador

POLS, RE, FE, FD



Fonte: Mesquita et al. (2021)

Aplicação:

Qual o impacto do crescimento econômico local sobre percentual de votos recebidos pelo candidato incumbente nas eleições presidenciais e municipais brasileiras entre 2000 e 2010?

Fernandes e Fernandes (2017)

Dados https://osf.io/5yx7g/?view_only=ac1691cced8549238d6d6e0a9d2b7f7b

Pacote R *plm* (Croissant e Milo 2008)

Base e modelo

```
##### ABRINDO BANCO DE DADOS #####

library(haven) # pacote para ler dados do Stata
library(plm) # pacote para executar os modelos de dados em painel

BANCO <- read_dta("fernandes_2017.dta") # Arquivo da base de dados no
diretório

BANCO <- pdata.frame(BANCO, index = c("codibge", "ano")) # aqui o banco é
convertido ao formato pdata.frame
# para executar os modelos. No 'index', São acrescentadas as dimensões
espacial e temporal dos dados.
# Nesse caso, 'codibge' representa a dimensão espacial (cód. do munic.) e
'ano' o ano da eleição.

##### FÓRMULA FERNANDES 2017 #####

form <- as.numeric(fracaovotos) ~ cresc+crescuf+crescbr+lpibreal+lpibuf+
  lpibbrasil+prefeitobasepresidente+persaude+lpop+leec+
  lheu+lse+laseps+ldesporc+ldespcor+linvest+ldespes
```

```
##### POOLED OLS #####
```

```
POLS <- plm(form, data = BANCO, model = "pooling")
```

```
##### EFEITOS FIXOS #####
```

```
mode_fe <- plm(form, data = BANCO, model = "within") # O modelo de EF é executado com o model = 'within'
```

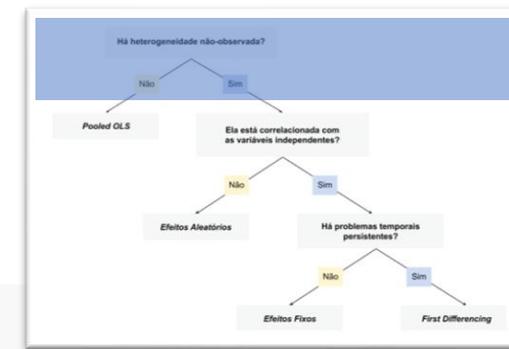
```
##### EFEITOS ALEATÓRIOS #####
```

```
mode_re <- plm(form, data = BANCO, model = "random") # o modelo de EA é executado com o model = "random".
```

```
##### PRIMEIRAS DIFERENÇAS #####
```

```
mode_fd <- plm(form, data = BANCO, model = "fd") # o modelo de FD é executado com o model = "fd".
```

Criar os quatro modelos:
POLS, FE, RE e FD



Há heterogeneidade não-observada?

F-TEST POLS

- **Teste F**
- FE é preferível a POLS

```
pFtest(mode_fe, POLS) # F-test com o modelo de efeitos fixos e Pooled
```

```
##
## F test for individual effects
##
## data: form
## F = 1.4214, df1 = 5548, df2 = 20786, p-value < 2.2e-16
## alternative hypothesis: significant effects
```

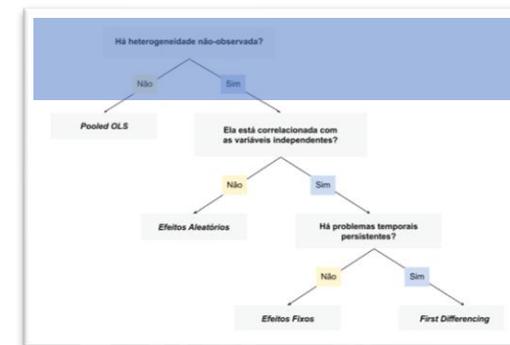
- **BP**
- RE é preferível a POLS

BREUSCH-PAGAN TEST

```
plmtest(POLS, type="bp", effect = "individual") # teste de Breusch-Pagan + efeitos individuais
```

```
##
## Lagrange Multiplier Test - (Breusch-Pagan) for unbalanced panels
##
## data: form
## chisq = 107.15, df = 1, p-value < 2.2e-16
## alternative hypothesis: significant effects
```

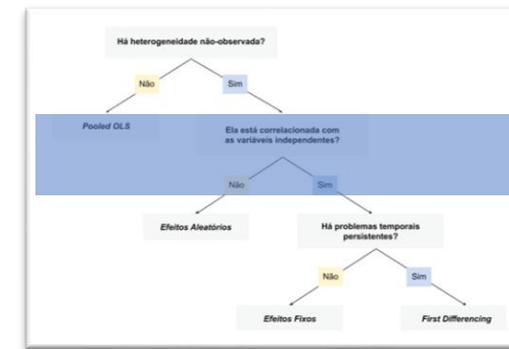
Há heterogeneidade não-observada?



- **BP**
- Há heterogeneidade não só individual mas também temporal
- Deve-se usar efeitos “two ways”

```
plmtest(POLS, type="bp", effect = "twoways") # teste de Breusch-Pagan +  
efeitos individuais e temporais
```

```
##  
## Lagrange Multiplier Test - two-ways effects (Breusch-Pagan) for  
## unbalanced panels  
##  
## data: form  
## chisq = 2762.9, df = 2, p-value < 2.2e-16  
## alternative hypothesis: significant effects
```



Ela está correlacionada com as VIs?

- Hausman
- O modelo FE é preferível ao RE

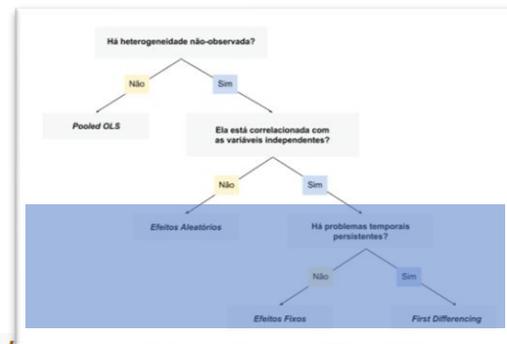
HAUSMAN TEST

```
phptest(mode_fe_2w, mode_re_2w) # teste de Hausman (modelo fixo e aleat.)
```

```
##
## Hausman Test
##
## data: form_d
## chisq = 851.98, df = 15, p-value < 2.2e-16
## alternative hypothesis: one model is inconsistent
```

```
phptest(form_d, data=BANCO, method="aux", vcov=vcovHC) # versão robusta do teste
```

```
##
## Regression-based Hausman test, vcov: vcovHC
##
## data: form_d
## chisq = 810.83, df = 18, p-value < 2.2e-16
## alternative hypothesis: one model is inconsistent
```



Há problemas de correlação serial?

- Há correlação serial tanto no modelo FE quanto FD

TESTAR CORR. SERIAL PARA FE

```
pwartest(mode_fe_2w) #pwartest para o modelo FE
```

```
##
## Wooldridge's test for serial correlation in FE panels
##
## data: mode_fe_2w
## F = 71.527, df1 = 1, df2 = 20801, p-value < 2.2e-16
## alternative hypothesis: serial correlation
```

- Decisão final: adotar FE ou FD, com robustificação

TESTAR CORR. SERIAL PARA FD

```
pwfdtest(mode_fd_dic) #pwfdtest para o modelo FD
```

```
##
## Wooldridge's first-difference test for serial correlation in panels
##
## data: mode_fd_dic
## F = 6004.9, df1 = 1, df2 = 15307, p-value < 2.2e-16
## alternative hypothesis: serial correlation in differenced errors
```

(Excerto da tabela completa)

	Fração Votos				
	POLS	RE (2W)	FE (2W)	FE (2W, Robust.)	FD
Cresc. Munic.	1.683 ^{***} (0.608)	1.098 ^{***} (0.040)	1.148 [*] (0.652)	1.148 (1.517)	0.190 (0.669)
Cresc. UF	0.095 ^{***} (0.027)	0.086 ^{***} (0.002)	-0.058 [*] (0.030)	-0.058 (0.163)	-0.144 ^{***} (0.033)
Cresc. BR	3.660 ^{***} (0.090)	3.709 ^{***} (0.088)			4.231 ^{***} (0.091)
Prefeito Base Presid.	1.211 ^{***} (0.208)	0.742 ^{***} (0.014)	1.289 ^{***} (0.235)	1.289 ^{***} (0.409)	1.136 ^{***} (0.254)
Perc. Orç. Saúde	0.040 ^{***} (0.014)	0.121 ^{***} (0.001)	0.044 [*] (0.023)	0.044 (0.035)	0.058 ^{***} (0.019)
Desp. Educ. Cult. (log)	6.917 ^{***} (0.358)	6.461 ^{***} (0.025)	1.543 ^{***} (0.525)	1.543 (1.284)	0.162 (0.553)
Eleição Municip.	1.074 ^{***} (0.277)	0.828 ^{***} (0.246)			-0.915 ^{***} (0.264)
Intercepto	102.026 ^{**} (43.701)	152.778 ^{***} (37.792)			-4.350 ^{***} (0.373)
N Obs.	26,352	26,352	26,352		20,803
R ²	0.321	0.319	0.039		0.254
Estat. F	691.129^{***} (df = 18; 26333)	12,309.190^{***}	55.646^{***} (df = 15; 20783)		392.979^{***} (df = 18; 20784)

*** p < 0,001; ** p < 0,01; *p < 0,05

- Coeficientes POLS próximos a RE, e os FE próximos a FD
- Mudança de sinal de algumas variáveis (“Cresc. UF”) mostram que fatores não-modelados provocavam viés
- Mudança de magnitude e significância para outras
- Variáveis fixas no ano para todos os municip. (“Cresc. BR”, *dummy* “Eleição Municip.”) não são estimadas no FE 2W
- Da esquerda para direita da tabela, vê-se o resultado de ignorar a heterogeneidade (POLS) e eliminá-la totalmente (FE/FD)

Aprofundamento:
Lidando com problemas espaciais e
temporais dos modelos de dados em
painel

Relembrando os pressupostos da regressão linear multivariada

A partir de Wooldridge (2013)

	Pressuposto	Formal	Problema	Causas comuns	Solução possível
P 1	Relação linear	$Y = \alpha + \beta X + \varepsilon$	Viés	Má especificação do modelo	Rever modelo
P 2	Amostra aleatória (ausência de autocorrelação)	$\text{Corr}(\varepsilon_i, \varepsilon_j) = 0$ $i \neq j$	Viés	Forma de coleta Natureza dos dados	Aleatorização; Rever modelo para eliminar correlação
P 3	Ausência de colinearidade perfeita	$\text{Corr}(X_1, X_2) \neq \pm 1$	Impossível estimar $\hat{\beta}$; Viés; Infla $EP(\hat{\beta})$, p-valor	X_1 e X_2 não ocorrem de forma indep. N pequeno	Rever modelo; Aumentar N
P 4	Exogeneidade	$E(\varepsilon x) = 0$	Viés	Variável omitida; Erro sistemático de mensuração; Truncagem	Rever modelo
P 5	Homoscedasticidade	$\text{Var}(\varepsilon x) = \sigma^2$	Invalida $EP(\hat{\beta})$, p-valor	Má especificação do modelo; Erro sistemático de mensuração; N pequeno	Rever modelo; Aumentar N; Robustificação
P 6	Distribuição normal dos erros	$\varepsilon \sim \text{Normal}(0, \sigma^2)$	Invalida $EP(\hat{\beta})$, p-valor	Má especificação do modelo; Erro sistemático de mensuração	(menos importância em N grande)

	Pressuposto
P 1	Relação linear
P 2	Amostra aleatória (ausência de autocorrelação)
P 3	Ausência de colinearidade perfeita
P 4	Exogeneidade
P 5	Homoscedasticidade
P 6	Distribuição normal dos erros

$\hat{\beta}$ é livre de viés

$\hat{\beta}$ tem a menor variância

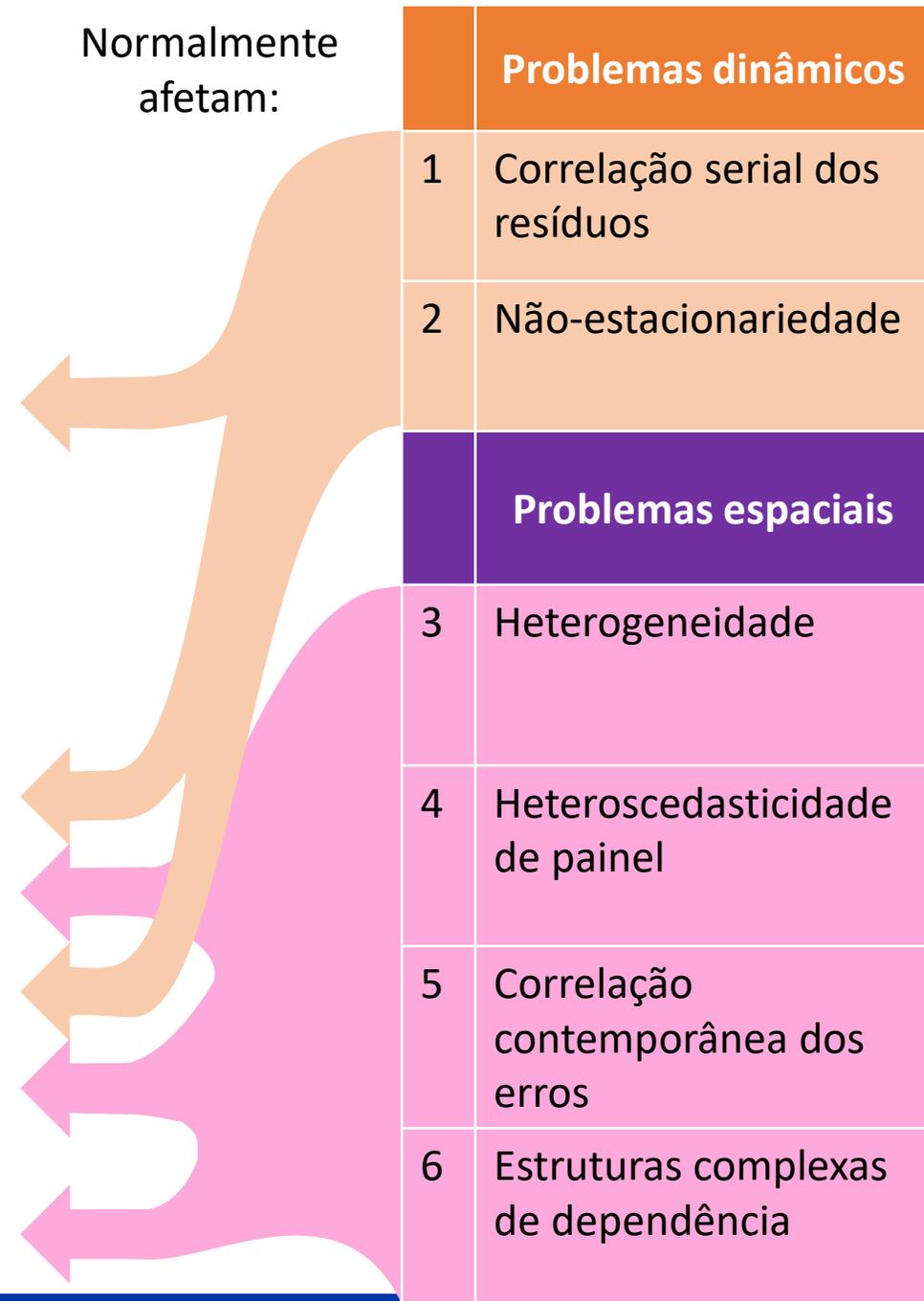
Premissas **Gauss-Markov**

B est
L inear
U nbiased
E stimator

$\hat{\beta}$ tem a menor variância de qualquer estimador não-viesado

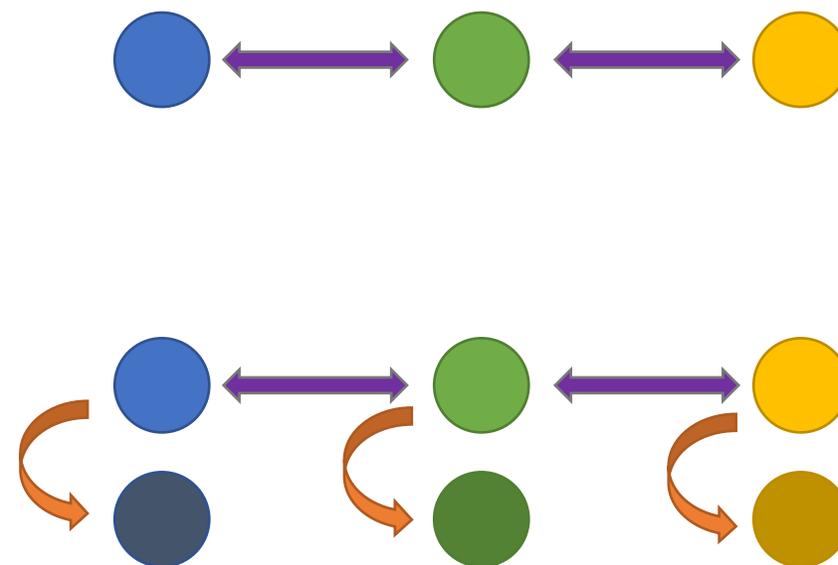
Premissas clássicas da regressão linear

	Pressuposto	Formalização	Por que violá-lo é um problema?
P.1	Relação linear	$Y = \alpha + \beta X + \varepsilon$	Viés
P.2	Ausência de autocorrelação, Amostra aleatória	$\text{Corr}(\varepsilon_i, \varepsilon_j) = 0$ $i \neq j$	Viés
P.3	Ausência de colinearidade perfeita	$\text{Corr}(X_1, X_2) \neq \pm 1$	Impossível estimar $\hat{\beta}$; Viés; Infla Erro-Padrão($\hat{\beta}$), p-valor
P.4	Exogeneidade	$E(\varepsilon x) = 0$	Viés
P.5	Homoscedasticidade	$\text{Var}(\varepsilon x) = \sigma^2$	Invalida Erro-Padrão($\hat{\beta}$), p-valor
P.6	Distribuição normal dos erros	$\varepsilon \sim \text{Normal}(0, \sigma^2)$	Invalida Erro-Padrão($\hat{\beta}$), p-valor



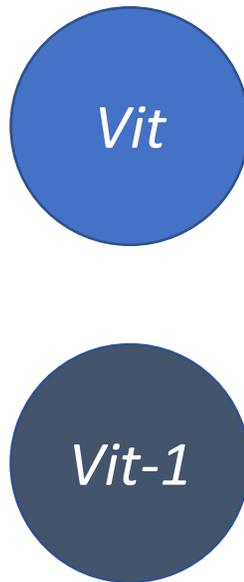
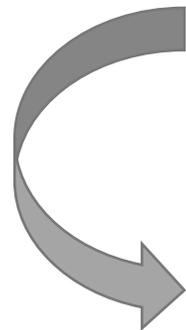
O problema da dependência

- Numa regressão transversal simples, só precisamos monitorar 1 dimensão para verificar independência entre observações: **espacial**
- Em dados em painel, há 2 dimensões: **espacial** e **temporal**



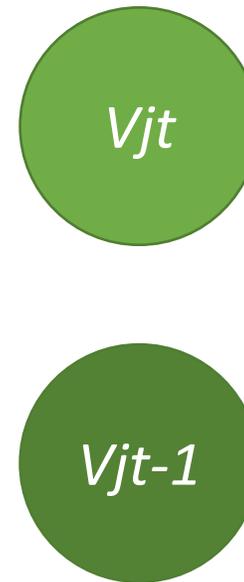
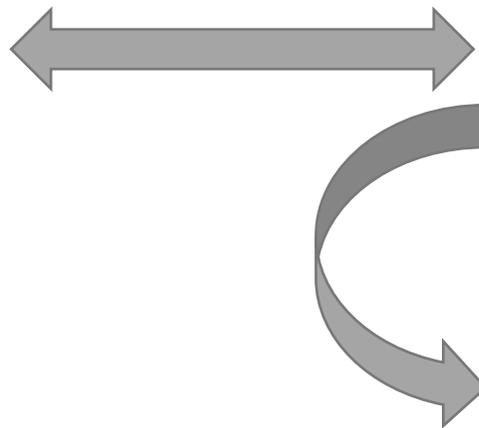
Dependências temporais

Os resíd. de uma unidade i em t estão relacionados com os da mesma unidade em $t-1$?



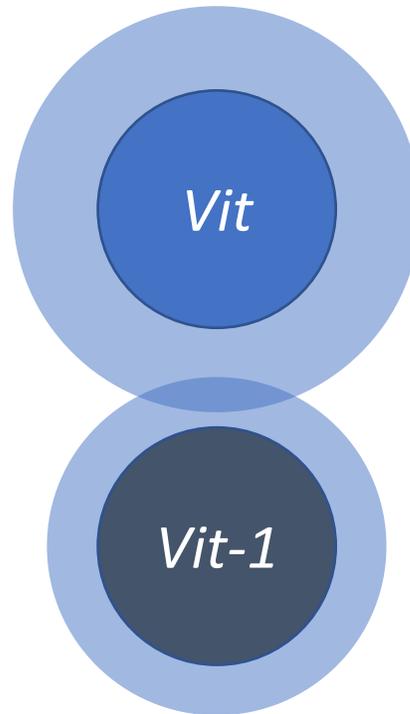
Dependências espaciais

Os resíd. da unidade i em t estão relacionados com os de outra unidade j em t ?



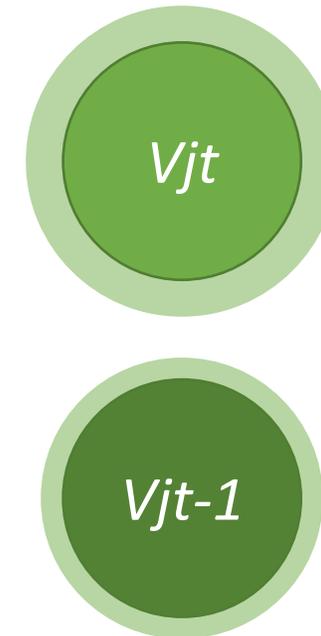
Dependências temporais

A variância do momento t é diferente do momento $t-1$?



Dependências espaciais

A variância dos resíd. da unidade i é diferente da variância da unidade j ?



Problemas temporais

1. Não
estacionariedade

2.
Autocorrelação

**Normalmente
violam:**

P 2 Ausência de
autocorrelação

P 4 Exogeneidade

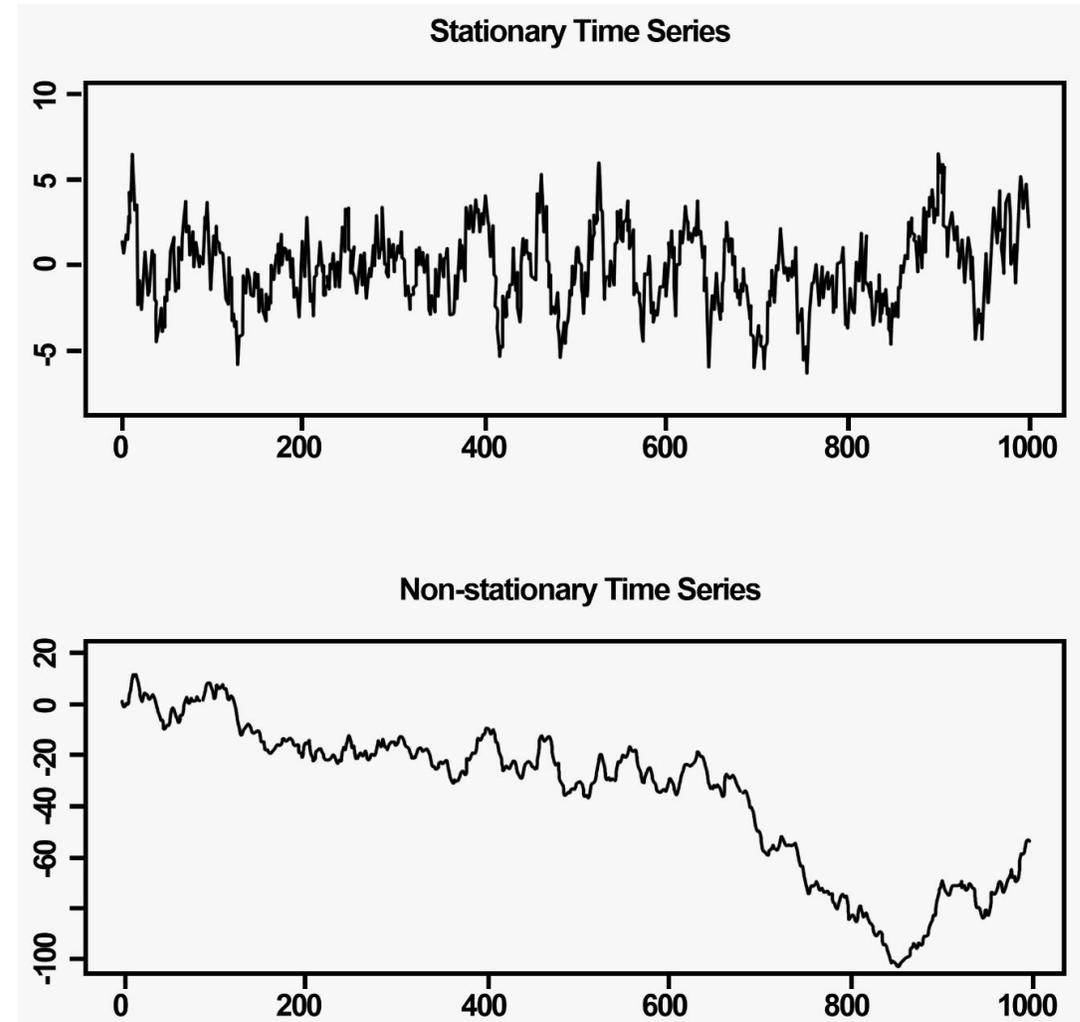
P 5
Homocedasticidade

Problemas temporais

- **Considerações iniciais:**
- Recomenda-se lidar primeiro com os problemas temporais antes dos espaciais
- Serão tanto mais sérios quanto maior for T em relação a N
- Técnicas de séries temporais aplicáveis para $T > \sim 20$

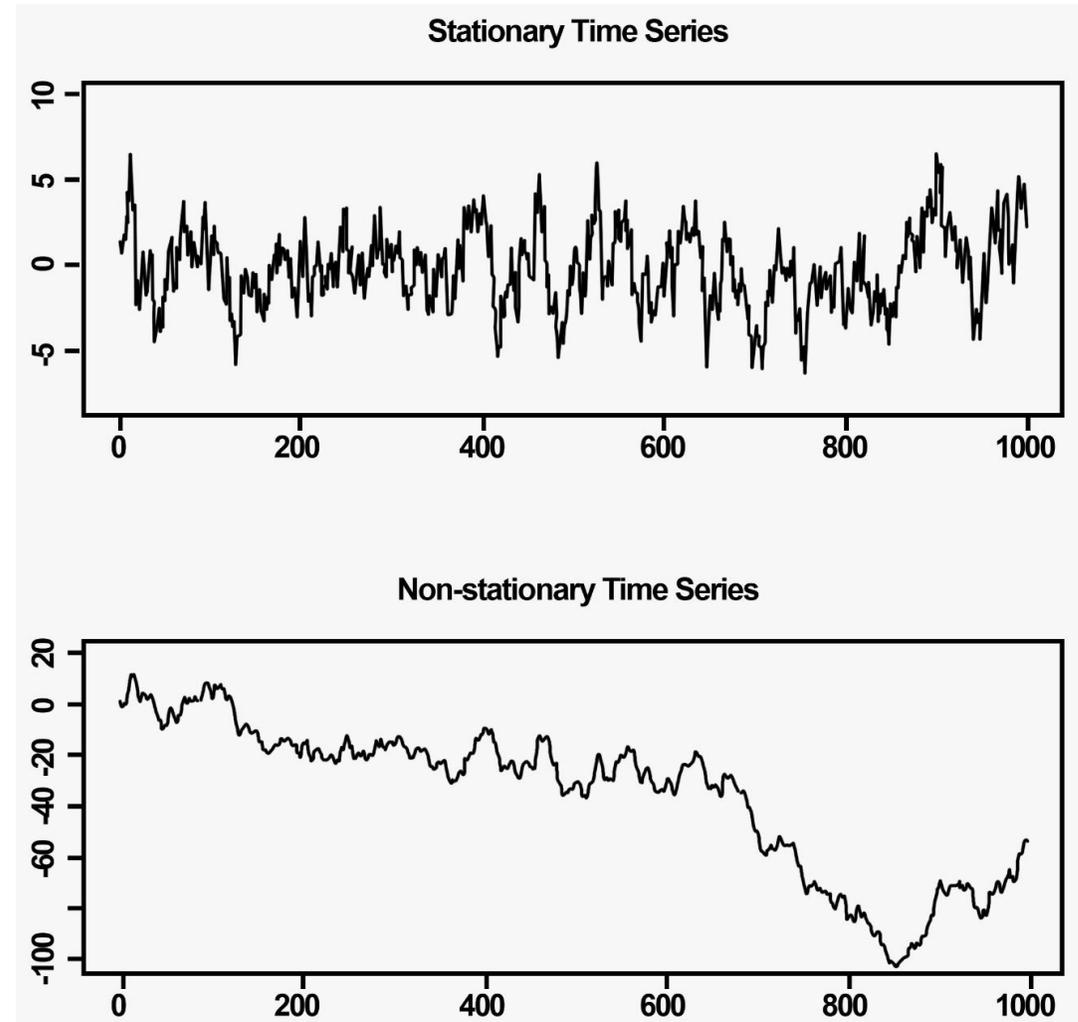
1. Não-estacionariedade: definição

- Série temporal estacionária: média e variância constantes ao longo do tempo
 - Ainda que os valores oscilem, retornarão à média
- Não-estacionária: média e variância não são constantes
 - A série não tende a retornar a uma média anterior após desvios
 - Efeitos de choques persistentes
- **Pergunte-se:**
 - *“A série retorna à origem?”*
 - *“É semelhante em todo momento?”*



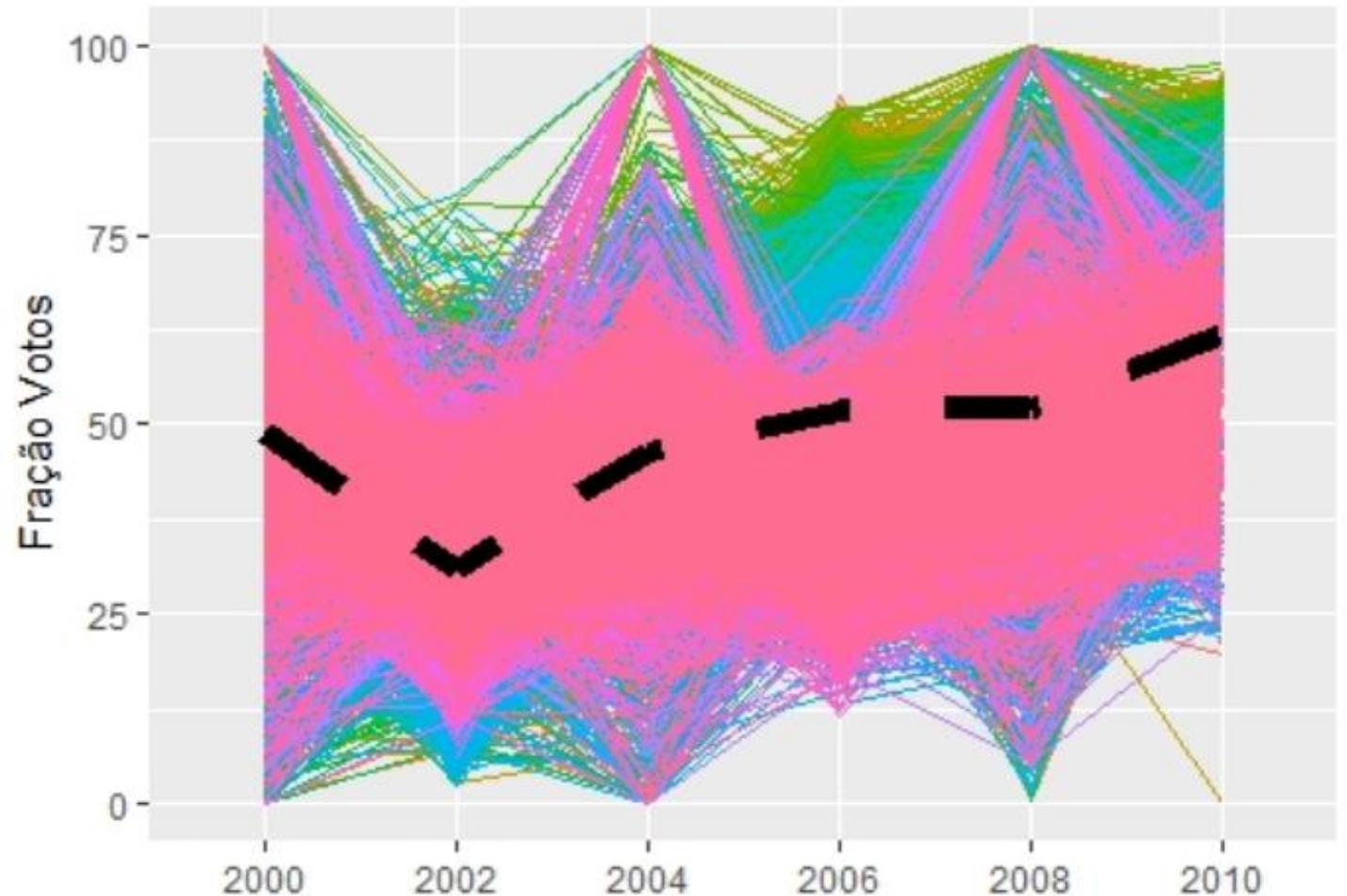
1. Não-estacionariedade: definição

- Por que é um problema?
 - Ausência de autocorrelação
 - Exogeneidade
 - Homocedasticidade
- Causas comuns de não-estacionariedade
 - Tendência
 - Sazonalidade



1. Não-estacionariedade: diagnóstico

- Diagnósticos
 - 1. Inspeção visual (plots de séries temporais)
 - 2. Teste de raiz unitária para séries mais longas
 - 3. Plots ACF e PACF, também para séries mais longas



1. Não-estacionariedade: soluções

- Alterar a operacionalização da variável (ex.: corrigir pela inflação ao longo do tempo)
- Adicionar *time trend* (especialmente para tendência)
 - Captura o efeito da “passagem do tempo”
 - Mas não são estimadas para modelos FE ou FD (cresc. monotônico)
- Aplicar FD

2. Correlação serial dos resíduos: definição

- Ocorre quando o resíduo v_{it} correlaciona-se com v_{it-1}
- Comum em dados de painel já que as observações são repetidas
- Tipicamente, processos autorregressivos de 1ª ordem, ou “AR(1)”
- Por que é um problema?
 - Autocorrelação
 - Exogeneidade
 - Distribuição dos erros

2. Correlação serial dos resíduos: diagnóstico

- Diagnóstico
 - 1. Regressão dos resíduos (para AR1)
 - 2. Testes

Diagnóstico da correlação serial tipo AR(1)

- Regressão de v_{it} em v_{it-1} (POLS) (Wooldridge)
- $\varepsilon_{it} = \rho\varepsilon_{it-1} + e_{it}$
- Ver se ρ é significativa

Testes do plm para correlação serial (tipicamente: $H_0 =$ não há corr.)

POLS	<ul style="list-style-type: none">• Regressão manual $\varepsilon_{it} \sim \varepsilon_{it-1}, \rho$• pwtest(pols) # Teste genérico de Wooldridge para ausência de efeitos não-observados nos resíd.• pbsytest(pols) # Teste para AR(1) ou RE(1) (Bera, Sosa e Yoon), AR(1) e RE(1) (Batalgi e Li)<ul style="list-style-type: none">• obs: <code>pbsytest(pols, test=c("ar", "re", "j"))</code>
RE	<ul style="list-style-type: none">• pbltest(formula, data=.) # Teste de Batalgi e Li para AR(1)/MA(1)
FE	<ul style="list-style-type: none">• pwartest(fe) # Teste de Wooldridge para AR(1) em FE
FD	<ul style="list-style-type: none">• pwfdtest(fd) # Teste de Wooldridge para AR(1) em FD
Geral	<ul style="list-style-type: none">• pbgtest(modelo) # Teste de Breusch-Godfrey/Wooldridge para corr. serial• pdwtest(modelo) # Teste de Durbin Watson<ul style="list-style-type: none">• obs: BG e DW para FE exigem T longo• pbnftest(modelo) # Generalização do DW (Bhargava, Narendranathan e Franzini)<ul style="list-style-type: none">• obs: não fornece p-valor. Habitualmente:• DW: 0 – 2 autocorr. positiva; 2 não há autocorr.; 2 – 4 autocorr negativa.

2. Correlação serial dos resíduos: soluções

- Efeitos fixos do tempo
- Introduzir *lagged dependente variable* (LDV)
 - Porém, pode diminuir efeito dos demais coeficientes
- Robustificação dos erros-padrão
- Aplicar FD

Problemas espaciais

3. Heterogeneidade
(rever slides 'heterogeneidade não-observada')

4.
Heteroscedasticidade
de painel

5. Correlação
contemporânea de
erros

6. Estruturas
complexas de
dependência

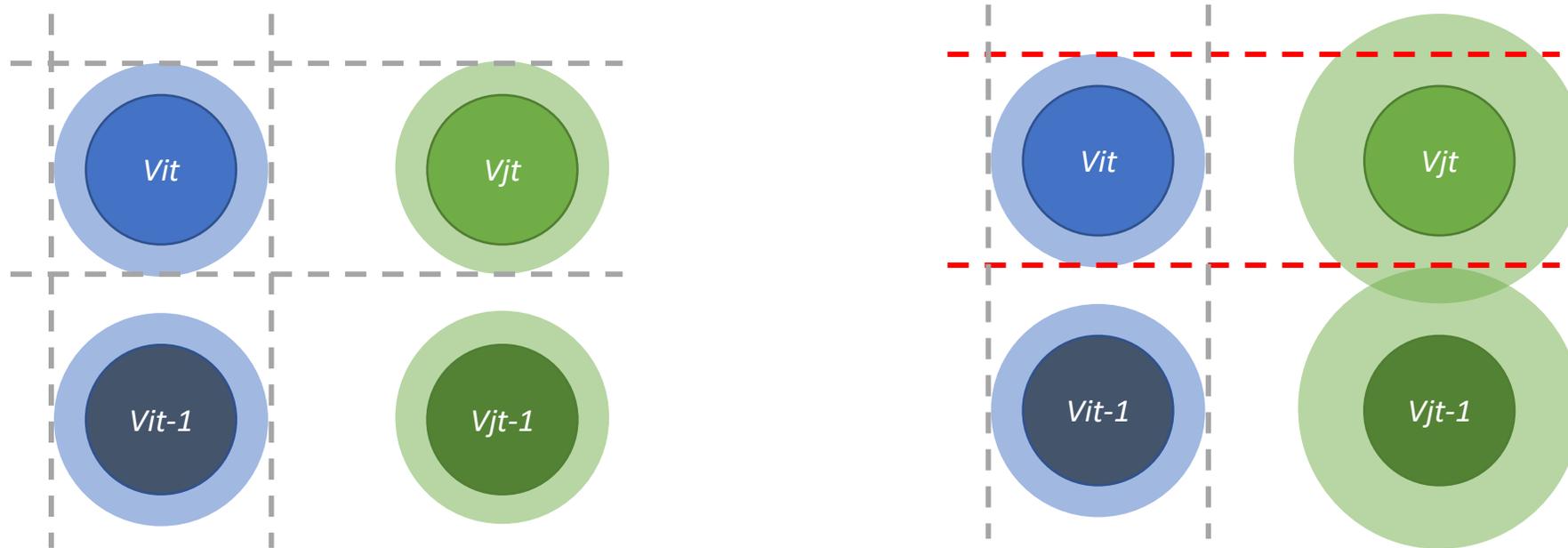
**Normalmente
violam:**

P 4 Exogeneidade

P 5
Homocedasticidade

P 6 Distribuição dos
erros

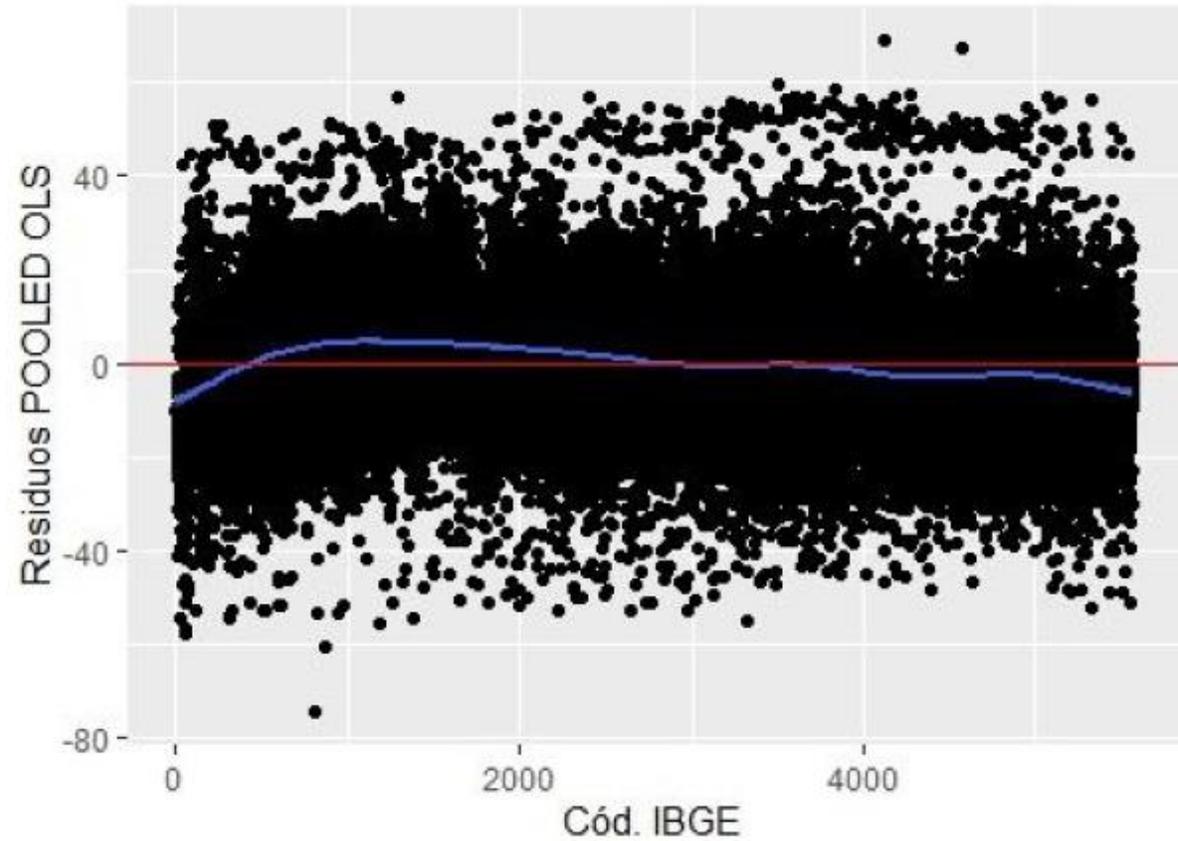
4. Heteroscedasticidade de painel: definição



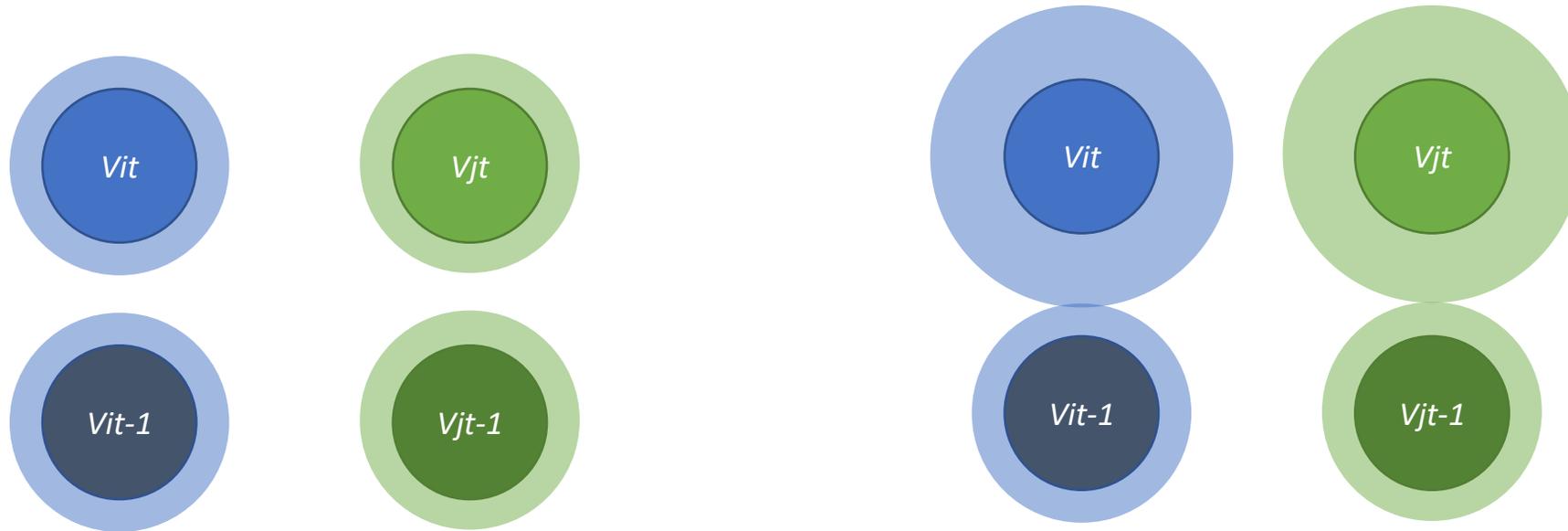
- Homoscedasticidade requer que a variância dos erros seja a mesma para toda a amostra
- Ocorre heteroscedasticidade de painel quando resíduos têm variância constante ao longo do tempo para cada unidade (*within unit*), mas inconstante entre as unidades (*across units*)

4. Heteroscedasticidade de painel: diagnóstico

- Inspeção visual:
 - Agrupar resíduos por unidade espacial
- Testes



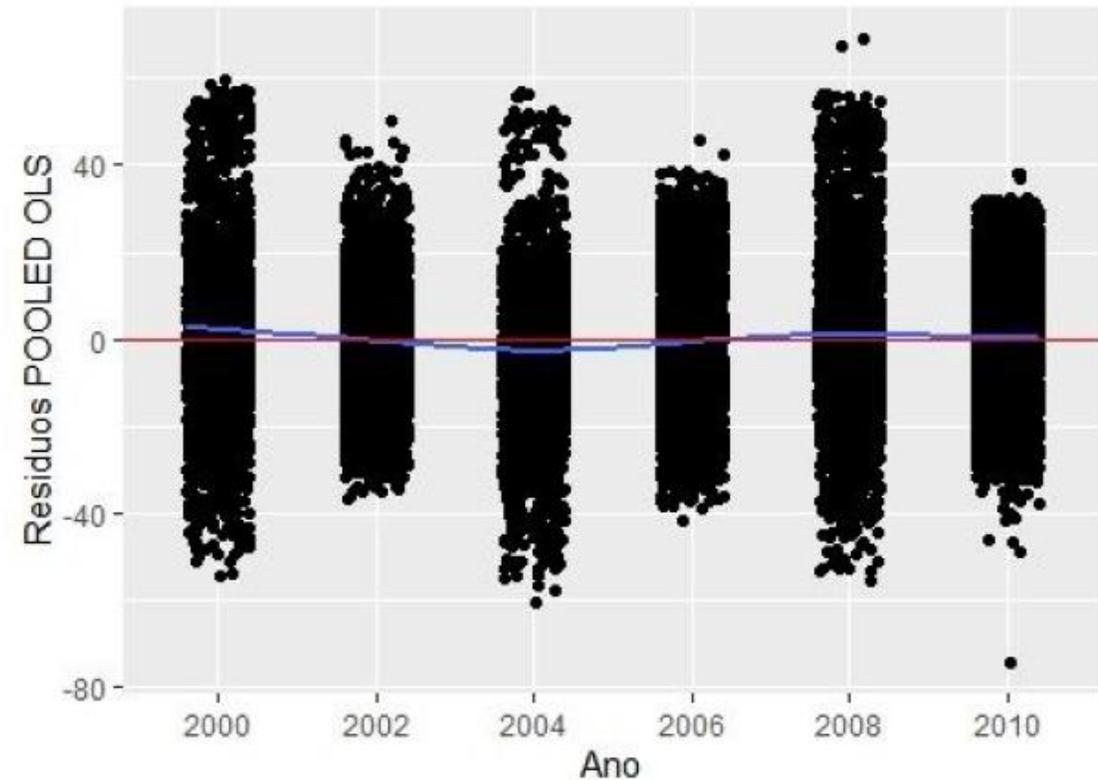
5. Correlação contemporânea dos erros: definição



- O resíduo de um caso está correlacionado com o resíduo de outros casos para o mesmo momento no tempo
- Choques exógenos afetam todas as unidades de forma comum no mesmo instante t

5. Correlação contemporânea dos erros: diagnóstico

- Plot resíduos agrupados por ano



5. Correlação contemporânea dos erros: soluções

- Adicionar *dummies* para os “anos de choque”
- Efeitos temporais ou *two ways*
- Robustificação

6. Estruturas complexas de dependência

- Há algum processo de difusão espacial (ex.: unidades geograficamente próximas sofrerão impactos parecidos)

Testes do plm para dependência transversal (cross sectional dependence) (tipicamente: H_0 = não há dependência transversal)

T longo, N curto	<ul style="list-style-type: none">• <code>pcdtest(modelo, test="lm")</code> # teste Breusch-Pagan LM
T longo, N longo	<ul style="list-style-type: none">• <code>pcdtest(modelo, test="sclm")</code> # scaled Breusch-Pagan LM
Geral	<ul style="list-style-type: none">• <code>pcdtest(modelo, test="cd")</code> # teste Pesaran de cross-sectional dependence<ul style="list-style-type: none">• obs.: teste padrão do <code>pcdtest()</code>• <code>pcdtest(modelo, test="rho")</code> <code>pcdtest(modelo, test="absrho")</code> # valor dos coeficientes de correlação ρ entre pares de observações

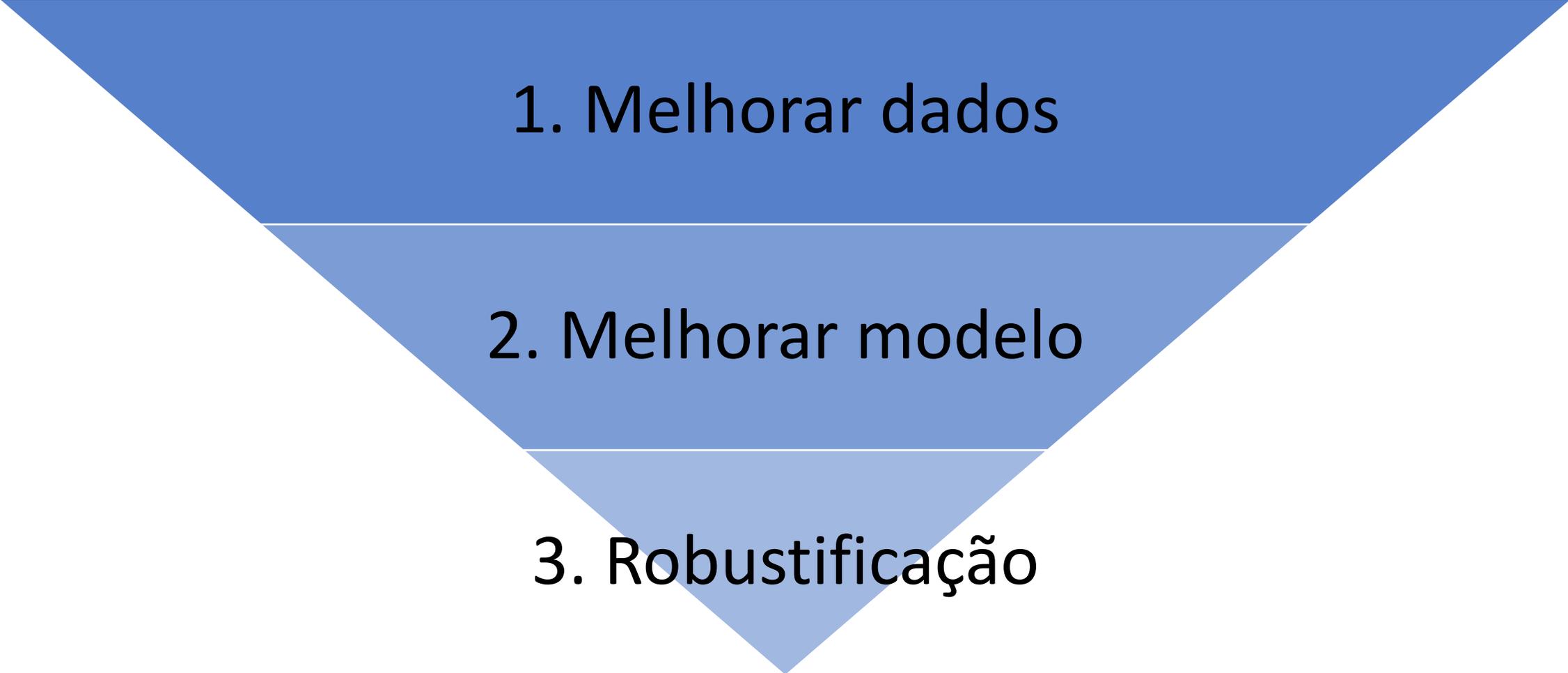
Robustificação

- Erros-padrão e p-valor precisam ser feitos robustos contra
 - (1) problemas espaciais (heteroscedasticidade, dependência transversal)
 - (2) problemas temporais (autocorrelação)

Matrizes de variância x covariância dos pacotes sandwich, plm para robustificação

	Erro-padrão robusto contra...
vcovHC	<ul style="list-style-type: none">• Heteroscedasticidade (igual a dados transversais)
vcovNW	<ul style="list-style-type: none">• Heteroscedasticidade e correlação serial
vcovSCC	<ul style="list-style-type: none">• Heteroscedasticidade e correlação serial (para $T > 20$)
vcovDC	<ul style="list-style-type: none">• Double-clustering
vcovBK	<ul style="list-style-type: none">• Panel Corrected Standard Erros (PCSE)<ul style="list-style-type: none">• obs.: geralmente para POLS com LDV1

O que tentar primeiro?



1. Melhorar dados

2. Melhorar modelo

3. Robustificação

Referência destes slides

MESQUITA, Rafael; FERNANDES, Antônio; FIGUEIREDO FILHO, Dalson B.;. Uma introdução à regressão com dados de painel. **Política Hoje**, Vol. 30, N. 1, 2021.