

Efeitos fixos e aleatórios

Avaliação de Políticas Públicas B

16, 18 e 30 de junho de 2025

Leitura básica:

FACURE, Matheus. **Causal inference for the brave and true**. 2022. [Capítulo 14: Panel data and fixed effects.] Disponível em: <https://matheusfacure.github.io/python-causality-handbook/landing-page.html>

Leitura complementar:

HUNTINGTON-KLEIN, Nick. **The effect: an introduction to research design and causality**. 2023. [Capítulo 16: Fixed effects.] Disponível em: <https://theeffectbook.net/>

STOCK, James H.; WATSON, Mark W. **Introduction to Econometrics**. 4. ed. New York: Pearson, 2019. [Capítulo 10: Regression with panel data.] Edições anteriores também contém este capítulo.

TOWNSEND, Zac et al. The choice between fixed and random effects. In: SCOTT, Marc A.; SIMONOFF, Jeffrey S.; MARX, Brian D. (org.). **The SAGE handbook of multilevel modeling**. Londres: SAGE, 2013, p. 73-88. Disponível em: http://agfda.userweb.mwn.de/ALD_2016/downloads/Uebung2/Townsend_et_al_2013.pdf

Estudos observacionais: estratégias e abordagens para enfrentar heterogeneidade não observada

Problema

Em estudos observacionais, não há presunção de comparabilidade entre grupo tratado e não tratado (antes da exposição ao tratamento). De fato, esses grupos podem diferir substancialmente, tanto nas características observáveis como nas não observáveis. Por isso, a associação observada entre T (tratamento) e Y (variável de resultado) não pode ser considerada uma estimativa do efeito causal de T em Y. Para a estimação de efeito, teremos de usar alguma estratégia para enfrentar a não comparabilidade entre os grupos.

Estratégia

1

Modelar o mecanismo de tratamento (T)

2

Modelar a variável de resultado (Y)

3

Uma combinação das estratégias anteriores

Exemplo de abordagem

- **Inverse Probability of Treatment Weighting (IPTW)** com probabilidade de tratamento definida como **propensity score (PS)**
 - Matching (pode envolver PS)
 - Regressão c/ descontinuidade
-
- Controle estatístico (pode envolver PS)
 - Diferença em diferenças
 - Efeitos fixos
-
- Estimação do tipo double robust

Estudos observacionais: estratégias e abordagens para enfrentar heterogeneidade não observada

Problema

Em estudos observacionais, não há presunção de comparabilidade entre grupo tratado e não tratado (antes da exposição ao tratamento). De fato, esses grupos podem diferir substancialmente, tanto nas características observáveis como nas não observáveis. Por isso, a associação observada entre T (tratamento) e Y (variável de resultado) não pode ser considerada uma estimativa do efeito causal de T em Y. Para a estimação de efeito, teremos de usar alguma estratégia para enfrentar a não comparabilidade entre os grupos.

Estratégia

1

Modelar o mecanismo de tratamento (T)

2

Modelar a variável de resultado (Y)

3

Uma combinação das estratégias anteriores

Exemplo de abordagem

- **Inverse Probability of Treatment Weighting (IPTW)** com probabilidade de tratamento definida como **propensity score (PS)**
- Matching (pode envolver PS)
- Regressão c/ descontinuidade

- Controle estatístico (pode envolver PS)
- Diferença em diferenças
- **Efeitos fixos**

- Estimação do tipo double robust

Reflexões iniciais

- Em DiD, combinamos 2 comparações (depois vs. antes e tratamento vs. comparação) para neutralizar dois tipos de confounders (variáveis omitidas correlacionadas com o tratamento):
 - a_i : fatores constantes no tempo mas diferentes entre grupos (i);
 - d_t : fatores variáveis no tempo (t) que incidem sobre os dois grupos.
- Podemos aplicar essa lógica a dados longitudinais (e.g., de painel, de corte transversal empilhado) mesmo quando não temos dois grupos (tratamento e comparação) definidos, desde que tenhamos variação do tratamento no tempo para as unidades observadas.

Pooled OLS: não costuma ser uma boa ideia

- Imagine que tenhamos dados sobre desemprego e criminalidade coletados para diferentes cidades em diferentes momentos no tempo.
 - Para simplificação, trabalharemos inicialmente com $T = 2$ (i.e., apenas duas rodadas de observação) e consideraremos apenas heterogeneidades não observadas entre cidades (i.e., a_i).

$$\text{Crime}_{it} = B_0 + B_1 \text{Desemprego}_{it} + a_i + u_{it}$$

Este é o processo
gerador dos dados.

Pooled OLS: não costuma ser uma boa ideia

- Se consideramos que $\text{cov}(a_i, \text{Desemprego}_{it}) = 0$, então podemos estimar o que se chama de Pooled OLS.

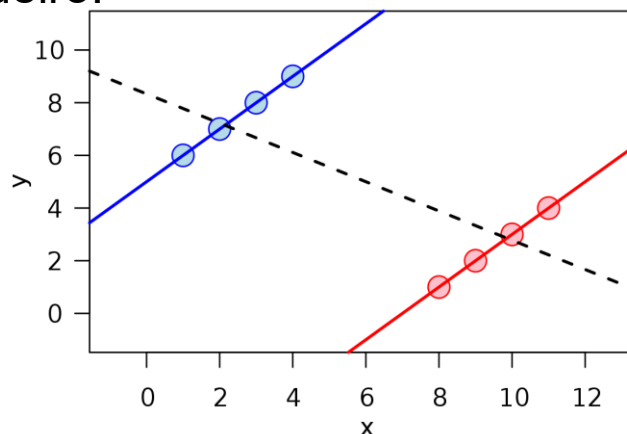
Nota: ignoramos o agrupamento das observações em cidades, pois partimos do pressuposto de que as heterogeneidades não observadas específicas às cidades (i.e., a_i) não se correlacionam com Desemprego_{it} .

$$\text{Crime}_{it} = B_0 + B_1 \text{Desemprego}_{it} + u_{it}$$

Este é o modelo pooled OLS.

- Todavia, se o pressuposto acima não se sustentar, teremos \hat{B}_1 enviesado.

Exemplo: imaginemos que Desemprego_{it} tende a ser mais elevado em cidades que têm forças de segurança mais consolidadas (o que contribui para menor Crime_{it}). Neste caso, pooled OLS seria um estimador enviesado: produziria viés negativo no efeito estimado de Desemprego_{it} em Crime_{it} , com o \hat{B}_1 Pooled tendendo a ser menor que o B_1 verdadeiro.



Simpson's paradox: a positive trend (— , —) appears for two separate groups, whereas a negative trend (- - -) appears when the groups are combined.

https://en.wikipedia.org/wiki/Simpson's_paradox

Efeitos fixos: requisitos e intuição

Requisitos

- Dados **longitudinais** para **múltiplas unidades** observadas (**pelo menos duas medições por unidade**).
 - Idealmente, dados de painel; no mínimo, dados de corte transversal empilhados.
- **Fatores explicativos variantes no tempo** e entre **unidades**.
 - Fatores explicativos constantes no tempo serão absorvidos por efeitos fixos de entidade.
 - Fatores explicativos invariantes entre unidades serão absorvidos pelos efeitos fixos de tempo.

Intuição

- Modela as **singularidades não observadas** de cada agrupamento através de **interceptos próprios** aos grupos (nota: estimação nem sempre envolve dummies por grupos, mas a intuição permanece válida).
- Agrupamentos podem ser por **entidade** (grupo não temporal – e.g., municípios, estados) **e/ou por tempo** (momento da medida).

Intuição: efeitos fixos de entidade vs. tempo

Efeitos fixos de entidade

Interceptos próprios a cada entidade captam efeito combinado de variáveis não observadas que afetam o Y, variam entre entidades (mas não dentro delas), sendo constantes no tempo (pelo menos, no intervalo temporal observado).

Efeitos fixos de tempo

Interceptos próprios a cada momento de observação captam efeito combinado de variáveis não observadas que afetam o Y, não variam entre entidades (i.e., afetam todas as entidades da mesma forma num certo momento do tempo), e que são variáveis no tempo.

Consideremos o seguinte banco de dados de corte transversal empilhado

Observation	State	Year	Tax (tax per beer case in dollars)	Fatality (# of traffic fatalities/10,000)
1	AL	1982	1,54	2,13
2	AL	1988	1,50	2,49
3	AZ	1982	0,21	2,50
4	AZ	1988	0,34	2,71
...				

- i = entidades (estados contíguos); $i \in \mathbb{N}^* \mid 1 \leq i \leq 48$
- t = momentos no tempo; $t \in \mathbb{N}^* \mid 1 \leq t \leq 2$
- Número total de observações no banco de dados = $48 * 2 = 96$

Efeitos fixos: Entidade

- Imagine que você tenha calculado a seguinte regressão:

$$Fatality_{it} = \beta_0 + \beta_1 Tax_{it} + \beta_2 AZ_i + \dots + \beta_{48} WY_i + \varepsilon_{it}$$

- Seu foco é o coeficiente da variável **tax** (i.e., β_1).
- Tipicamente, os coeficientes das variáveis dummy para as entidades não são substantivamente interessantes (i.e., são apenas controles): **controlam por todos os fatores que influenciam fatality e que variam entre estados mas que não variam ao longo do período de tempo analisado** (e.g., aceitação cultural da embriaguês ao volante).
- Esse tipo de especificação chama-se **regressão de efeitos fixos de entidade**.

Efeitos fixos: entidade e tempo

Suponha que você adicionou um conjunto de variáveis dummy para cada ano na amostra:

$$Fatality_{it} = \beta_0 + \beta_1 Tax_{it} + a_i + d_t + \varepsilon_{it}$$

a_i

Efeitos fixos de entidade (no caso, estados): Controlam por fatores que afetam o indicador de fatalidade e que diferem entre estados, mas que são constantes ao longo do período de tempo examinado (e.g., aceitação cultural da embriaguês ao volante, frequência de batidas policiais).

d_t

Efeitos fixos de tempo: Controlam por fatores que afetam a fatalidade, que não variam entre estados, mas que variam ao longo do período de tempo examinado (e.g., uma nova norma federal sobre padrões de segurança em automóveis, uma campanha institucional em nível nacional).

Estimação de efeitos fixos: LSDV vs. demeaning

- FE podem ser estimados pela inclusão de dummies para (n-1) agrupamentos (neste exemplo, estados) – especificação também conhecida como **Least Squares Dummy Variables (LSDV)** estimator.

$$Fatality_{it} = \beta_0 + \beta_1 Tax_{it} + a_i + d_t + \varepsilon_{it}$$

- Alternativamente, FE podem ser estimados pela diferença em relação à média (**demeaning**): todas as variáveis são apresentadas como desvios em relação a sua respectiva média em i , t ou it .

Efeitos fixos de entidade:

$$(Fatality_{it} - média_i(Fatality_{it})) = \beta_0 + \beta_1(Tax_{it} - média_i(Tax_{it})) + \varepsilon_{it}$$

Efeitos fixos de tempo:

$$(Fatality_{it} - média_t(Fatality_{it})) = \beta_0 + \beta_1(Tax_{it} - média_t(Tax_{it})) + \varepsilon_{it}$$

Efeitos fixos de entidade e tempo:

Considere a seguinte transformação de uma variável v_{it} :

$$\ddot{v}_{it} = v_{it} - \underbrace{média_i(v_{it})}_{\text{Desvio em relação à média do grupo } i} - \underbrace{média_t(v_{it} - média_i(v_{it}))}_{\text{Desvio em relação à média do grupo } i}$$

Desvio em relação à
média do grupo i

Desvio em relação à
média do grupo i

$$\ddot{Fatality}_{it} = \beta_0 + \beta_1 \ddot{Tax}_{it} + \varepsilon_{it}$$

Todos os três β_0 (em azul) são dispensáveis: variáveis demeaned têm média zero. Em regressões simples, reta estimada passa por (\bar{X}, \bar{Y}) . Em regressões múltiplas, reta estimada passa por (\bar{X}, \bar{Y}) desde que demais variáveis explicativas encontrem-se em seus valores médios.

Exemplo: efeitos fixos de entidade e de tempo

A RELAÇÃO ENTRE O DESEMPENHO DA CARREIRA NO MERCADO DE TRABALHO E A ESCOLHA PROFISSIONAL DOS JOVENS

Otávio Bartalotti[§]

Naércio Menezes-Filho[□]

ECON. APLIC., SÃO PAULO, V. 11, N. 4, P. 487-505, OUTUBRO-DEZEMBRO 2007

<https://doi.org/10.1590/S1413-80502007000400002>

RESUMO

Este artigo examina como o desempenho relativo no mercado de trabalho de cada profissão afeta a escolha profissional dos futuros universitários. O desempenho no mercado de trabalho é medido pela média e pelo desvio padrão dos salários recebidos por profissão e pela sua taxa de desemprego no censo demográfico nos anos próximos ao vestibular. O número de pleiteantes a ingresso na carreira é medido pelo número de inscritos no exame da Fuvest. Utiliza-se dados em painel para os anos de 1991 e 2000 para controlar pelo efeito específico de cada profissão. Os resultados apontam para um efeito positivo e robusto do salário médio da profissão sobre a escolha profissional, que persiste na análise de painel, e para efeitos negativos da dispersão salarial da renda e do desemprego que, no entanto, não se mostraram significantes.

Palavras-chave: carreiras, profissões, mercado de trabalho, renda, Fuvest.

Exemplo: efeitos fixos de entidade e de tempo

Tabela 3 – Determinantes do número de inscritos nos cursos

Variável Dependente	Logaritmo da média do número de inscritos apurada nos processos seletivos para ingresso em 1992-1993 e 2001-2002.	Efeitos Fixos	
		Coeficientes	Coeficientes
Constante		-6.810 (6.670)	-7.693 (6.602)
Ln (Renda)		2.086** (0.970)	2.041** (0.964)
Desvio padrão Ln (Renda)		-0.705 (1.018)	-0.654 (1.010)
Ln (Desemprego)		-0.006 (0.085)	0.010 (0.086)
<i>Dummy ano</i>		-0.336 (0.329)	-0.321 (0.329)
Ln (Vagas)			0.240 (0.220)
R ²		0.120	0.258

Notas: Estatística *t* entre parênteses. ***= coeficiente significativo a 1%; **= coeficiente significativo a 5%; *= coeficiente significativo a 10%.

<https://doi.org/10.1590/S1413-80502007000400002>

Efeitos fixos estimam uma transformação do ATE

Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects

Clément de Chaisemartin and Xavier D'Haultfoeuille

NBER Working Paper No. 25904 <https://www.nber.org/papers/w25904>

May 2019

JEL No. C21,C23

ABSTRACT

Linear regressions with period and group fixed effects are widely used to estimate treatment effects. We show that they identify weighted sums of the average treatment effects (ATE) in each group and period, with weights that may be negative. Due to the negative weights, the linear regression estimand may for instance be negative while all the ATEs are positive. In two articles that have used those regressions, half of the weights are negative. We propose another estimator that solves this issue. In one of the articles we revisit, it is of a different sign than the linear regression estimator.

Clément de Chaisemartin
Department of Economics
University of California at Santa Barbara
Santa Barbara, CA 93106
and NBER
clementdechaisemartin@ucsb.edu

Xavier D'Haultfoeuille
CREST
5 avenue Henry Le Chatelier
91764 Palaiseau cedex
FRANCE
xavier.dhaultfoeuille@ensae.fr

O artigo baseado nesse *working paper* foi publicado como:
DE CHAISEMARTIN, Clément; D'HAULTFOEUILLE, Xavier. Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. American Economic Review, v. 110, n. 9, p. 2964-96, 2020.
DOI: 10.1257/aer.20181169

Gestão local e seu impacto no desempenho fiscal: análise dos municípios mineiros

Ana Paula Karruz

Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil

André Machado Moraes

Banco de Desenvolvimento de Minas Gerais (BDMG), Belo Horizonte, MG, Brasil

Karruz e Moraes (2020)

<https://doi.org/10.21874/rsp.v71i1.3777>

Resumo

O estudo busca elucidar e articular a gama de condicionantes do desempenho fiscal local, dimensionar a influência da gestão fiscal municipal nesse desempenho, e descrever como a qualidade da gestão fiscal se distribui entre municípios com diferentes características. Propõe um modelo conceitual e de regressão para o desempenho fiscal, tendo o Índice Firjan de Gestão Fiscal (IFGF) dos municípios de Minas Gerais como variável dependente. A identificação de efeitos apoia-se na natureza longitudinal dos dados (2007-2016) e na modelagem em painel. Transferências voluntárias da União e do governo estadual parecem favorecer o desempenho fiscal local, especialmente via investimentos. Um terço da variação do IFGF não pode ser atribuída às covariáveis e é tomada, por eliminação, como *proxy* para a qualidade da gestão fiscal. Esta apresenta moderada correlação com a população e mostra sinais de dependência espacial, com concentração de valores altos na região da capital.

Palavras-chave: gestão fiscal, municípios, Índice Firjan de Desempenho Fiscal (IFGF)

Efeitos fixos: exemplo

Quadro 04 | Principais influências sobre o desempenho fiscal municipal, por natureza de variação

Varia entre municípios?	Varia frequentemente?			
	Não (fatores estruturais)		Sim (fatores conjunturais)	
Não	Fatores (institucionais): <ul style="list-style-type: none">- Limites constitucionais da tributação municipal- Regras para transferências não discricionárias- Centralização do policy decision-making (e.g., percentuais mínimos de gastos em saúde e educação)- Determinações da Lei de Responsabilidade Fiscal (LRF)	Referências: <ul style="list-style-type: none">- Arretche (2004);- Baião, Cunha e Souza (2017); Veloso et al. (2011);- Arretche (2012);- Sistema Firjan (2017a).	Fatores: <ul style="list-style-type: none">- Contexto fiscal (federal, estadual)- Contexto econômico (nacional, estadual)	Referências: <ul style="list-style-type: none">- Sistema Firjan (2017a);- Sistema Firjan (2017a).
Sim	Fatores (estruturais locais): <ul style="list-style-type: none">- Ano de criação- Outros fatores estruturais locais, tais como:<ul style="list-style-type: none">° Recursos naturais° Principais atividades econômicas estabelecidas° Infraestrutura instalada° Extensão territorial° Status de capital° Proximidade com cidades-polo	Referências: <ul style="list-style-type: none">- Araújo Júnior (2016)- (Autores do presente trabalho reputam estes fatores como influentes sobre a capacidade de arrecadação própria e a demanda por serviços públicos)	Fatores: <ul style="list-style-type: none">- Contexto fiscal (municipal)- Potencial de arrecadação própria° Contexto econômico (municipal)° Tamanho populacional (base de arrecadação)- Demanda por serviços públicos (escala e complexidade)- Capacidade de atração de recursos públicos* (alinhamento partidário)- Transferências intergovernamentais (realizadas)° Compulsórias redistributivas (influxo de baixo custo)° Voluntárias- Ciclos políticos° Alternância ideológica° Ano eleitoral° Probabilidade de manutenção do incumbente ou seu grupo no poder- Qualidade da gestão fiscal	Referências: <ul style="list-style-type: none">- Sistema Firjan (2017a);- Soares e Melo (2016); Veloso et al. (2011)- Gobetti e Klering (2007);- Bueno (2018); Puttomatti (2013); Stokes et al. (2013)- Baião, Cunha e Souza (2017); Brollo e Nannicini (2012); Bueno (2018); Costa e Castelar (2015); Mattos, Rocha e Arvate (2011); Meireles (2019); Nunes (2015); Prado (2001); Puttomatti (2013); Sakurai (2013); Soares e Melo (2016);- Alesina (1987); Hibbs Jr. (1977); Nordhaus (1975); Rogoff e Sibert (1988); Sakurai (2009); Sakurai e Gremaud (2007)- Batista (2015); Cingolani (2013); Cingolani, Thomsson e Crombrughe (2015); Hou (2007); Ingraham (2007); Ingraham, Joyce e Donahue (2003); Marengo, Strohschoen e Joner (2017); Veloso et al. (2011)

Nota: * Por exemplo, transferências voluntárias do estado ou da União para entidades privadas da assistência social e investimentos diretos dessas esferas em infraestrutura de transportes.

Fonte: elaboração própria.

Efeitos fixos: exemplo

O modelo empírico (Equação I)¹⁰ estrutura-se em torno de variações espaciais e temporais, podendo ser representado pelos mesmos quadrantes do Quadro 4. Tem o IFGF como variável dependente, e combinações município-ano (i - t) como unidade de análise.

$$y_{it} = \alpha + \beta C_{it,(t-1)} + a_i + b_t + \varepsilon_{it} \quad (I)$$

Fatores institucionais, incidentes sobre todos os municípios e invariantes na janela temporal considerada (2007-2016), não requerem operacionalização, justamente por não terem se alterado. Demais fatores estruturais, com diferenciação entre localidades, são captados via modelo para dados em painel, que considera o agrupamento das observações (tomadas ao longo do tempo) em municípios aos quais elas pertencem. O teste de Hausman indica que um modelo com efeitos fixos de município (a_i) é preferível em relação à especificação com interceptos aleatórios por município (p -valor = 0,0000). No modelo com efeitos fixos de município, um conjunto de *dummies* identificando cada município controla pelas diferenças entre essas unidades, desde que tais diferenças tenham permanecido constantes de 2007 a 2016, a exemplo do ano de criação e da condição de capital.

Condicionantes do desempenho fiscal que se modificam no tempo, porém atingindo igualmente todos os municípios, são operacionalizados por efeitos fixos de ano (b_t). Exemplos de fatores dessa natureza são a situação fiscal da União e do Estado de Minas Gerais; também, especificamente em 2016, as receitas municipais aumentaram perto de 4%, em média, como resultado da Lei de Repatriação (Lei Federal nº 13.254/2016, sobre regularização de recursos de origem lícita), já que IR e multas apuradas na repatriação foram distribuídos aos municípios seguindo o esquema de repartição do FPM (SISTEMA FIRJAN, 2017a).

IDB WORKING PAPER SERIES N° IDB-WP-679

Do Longer School Days Improve Student Achievement? Evidence from Colombia

Diana Hincapié

Inter-American Development Bank
Education Division

March 2016

Efeitos fixos: exemplo

Efeitos fixos: exemplo

Resumo: Este artigo analisa o impacto de jornadas escolares mais longas sobre o desempenho dos alunos na Colômbia. Para identificar esse impacto, o estudo explora variações plausivelmente exógenas na duração da jornada escolar dentro das próprias escolas. Utilizando as notas obtidas pelos alunos em um teste aplicado ao 5º e 9º anos nos anos de 2002, 2005 e 2009, juntamente com dados administrativos escolares, esta pesquisa emprega modelos com efeitos fixos por escola para estimar variações na média das notas entre diferentes coortes de escolas para escolas que passaram de meio turno para turno integral, ou vice-versa. Os resultados indicam que os alunos expostos ao turno integral apresentam notas aproximadamente um décimo de desvio padrão superiores às daqueles que frequentaram apenas meio turno. O impacto do turno integral é maior em matemática do que em linguagem, e é mais acentuado no 9º ano do que no 5º ano. Os efeitos são maiores nas escolas mais pobres e naquelas localizadas em áreas rurais. Os resultados sugerem que a ampliação da jornada escolar pode ser uma política eficaz para melhorar o desempenho dos alunos, especialmente entre os estudantes de baixa renda na Colômbia e em outros países em desenvolvimento.

Hincapie (2016; tradução nossa)
<http://dx.doi.org/10.18235/0000268>

Efeitos fixos (FE) vs. Efeitos aleatórios (RE)

- Voltemos à equação inicial:

$$\text{Crime}_{it} = B_0 + B_1 \text{Desemprego}_{it} + a_i + u_{it}$$

RE são tema para um curso inteiro!

- RE assume que $\text{cov}(\text{Desemprego}_{it}, a_i) = 0$, e que, portanto, a_i pode ser considerado parte do erro aleatório:

$$\text{Crime}_{it} = B_0 + B_{1\text{RE}} \text{Desemprego}_{it} + \varepsilon_{it},$$

em que $\varepsilon_{it} = a_i + u_{it}$

Este é um modelo de regressão com interceptos aleatórios.

- RE não exige dados longitudinais e aceita variáveis explicativas invariantes no tempo. Ainda, RE consome muito menos graus de liberdade que FE.
- Todavia, RE não controla por a_i .
- A grande questão é se faz sentido assumir $\text{cov}(\text{Desemprego}_{it}, a_i) = 0$. Se já controlamos por tudo que poderia causar viés, ou se temos um desenho experimental, RE pode ser adequado.

Vide apêndice sobre [viés de variável omitida](#).

Efeitos fixos vs. aleatórios:

trecho da tese de Felipe Nunes (2015)

[...] the data also contain multiple observations for each municipality, which introduces within-municipality correlation. The correlations between spending in adjacent years varies from 0.01 to 0.99 from year to year and across countries, but they are always present to some extent. Ignoring that observations are correlated over time may lead to: exaggerated goodness of fit measures, estimated standard errors that tend to be smaller than the 'true' standard errors, and possible parameter inconsistency. The literature suggests many approaches for dealing with this. One approach is to assume that each unit has a fixed but unknown intercept that can be estimated by adding a dummy variable for each unit — the 'fixed effects' (FE) model (Angrist and Pischke, 2009). However, the fixed effects remove any of the average unit-to-unit variation from the analysis, and simply ask whether intra-unit changes in y are associated with intra-unit changes in x (Beck and Katz, 2011). Moreover, using fixed effects also makes it difficult to discern the impact of variables that change only slowly (Greene, 2003).

Another approach is to assume that each unit has random effects that are part of a distribution family, and that these random variables could be further modeled properly. This second approach — also known as the 'mixed model' or 'random effects model' — adds a hierarchical structure to the random effects and models them accordingly. I use a mixed model with municipal random intercepts and year random slopes that not only deals with the problems mentioned above, but that also allows each municipality to have its own over-time expenditure trajectory.

A major complaint lodged against random effect models relates to the restrictive assumption that independent variables be orthogonal with the random effects term. Since a variable varies both within and between municipalities, many argue that this is an unrealistic assumption, because unobserved heterogeneity will almost always be correlated with the independent variables.

(Nunes, 2015, p. 82-83).

Escolha entre FE e RE: teste de Hausman

Remarks and examples

stata.com

`hausman` is a general implementation of Hausman's (1978) specification test, which compares an estimator $\hat{\theta}_1$ that is known to be consistent with an estimator $\hat{\theta}_2$ that is efficient under the assumption being tested. The null hypothesis is that the estimator $\hat{\theta}_2$ is indeed an efficient (and consistent) estimator of the true parameters. If this is the case, there should be no systematic difference between the two estimators. If there exists a systematic difference in the estimates, you have reason to doubt the assumptions on which the efficient estimator is based.

```
. hausman fixed ., sigmamore
```

	Coefficients		(b-B) Difference	sqrt(diag(V_b-V_B)) Std. err.
	(b) fixed	(B) .		
age	-.005485	-.0069749	.0014899	.0004803
msp	.0033427	.0046594	-.0013167	.0020596
t1l_exp	.0383604	.0429635	-.0046031	.0007181

b = Consistent under H0 and Ha; obtained from `xtreg`.

B = Inconsistent under Ha, efficient under H0; obtained from `xtreg`.

Test of H0: Difference in coefficients not systematic

chi2(3) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 260.40


Prob > chi2 = 0.0000

Under the current specification, our initial hypothesis that the individual-level effects are adequately modeled by a random-effects model is resoundingly rejected. This result is based on the rest of our model specification, and random effects might be appropriate for some alternate model of wages.

Estes são excertos da explicação do teste de Hausman segundo sua documentação no Stata:

<https://www.stata.com/manuals/rhausman.pdf>

Pooled OLS vs. FD vs. FE vs. RE



Ben Lambert
 97,6 mil inscritos

INÍCIO
 VÍDEOS
 PLAYLISTS
 COMUNIDADE
 CANAIS
 SOBRE

Fixed Effects Estimator

$$H_{it} = \beta_0 + \beta_1 \text{Chave}_{it} + \beta_2 \text{Unemp}_{it} + \epsilon_{it}$$

$$\bar{H}_i = \frac{1}{T} \sum_{t=1}^T H_{it}$$


7:58

Fixed Effects estimators: an introduction

Ben Lambert • 145 mil visualizações • há 7 anos

This video explains the motivation, and mechanics behind Fixed Effects estimators in panel econometrics. Check out <http://oxbridge-tutor.co.uk/undergraduate-econometrics-course> for course...

FE/FD vs Pooled OLS



7:02

Fixed Effects, First Differences and Pooled OLS - intuition

Ben Lambert • 101 mil visualizações • há 7 anos

This video provides intuition as to why Fixed Effects, First Differences and Pooled OLS panel estimators can yield significantly different results. Check out <http://oxbridge-tutor.co.uk/undergradu...>

Random Effects vs Fixed Effects

$$y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 \bar{x}_{it} + \epsilon_{it}$$

$$\text{Cov}(x_{it}, \epsilon_{it}) = 0$$

$$\text{Cov}(\bar{x}_i, \epsilon_{it}) = 0$$

5:25

Random Effects vs Fixed Effects estimators

Ben Lambert • 87 mil visualizações • há 7 anos

This video provides a comparison between Random Effects and Fixed Effects estimators. Check out <http://oxbridge-tutor.co.uk/undergraduate-econometrics-course> for course materials, and information...

First Differences vs Fixed Effects

$$T=2: FD = FE$$

$$T \geq 3: FD \neq FE$$

Fixed Effects and First Differences comparison - part 1

Fixed Effects and First Differences comparison - part 1

Ben Lambert • 54 mil visualizações • há 7 anos

This video explains some of the differences between Fixed Effects and First Differences estimators.

Fixed Effects, First Differences and Pooled OLS – intuition
<https://youtu.be/1SchyQ77VFg>

Fixed Effects estimators: an introduction
<https://youtu.be/sFvV9b1cGFc>

Fixed Effects and First Differences comparison - part 1
<https://youtu.be/G7WqK2o474Y>

Fixed Effects and First Differences comparison - part 2
<https://youtu.be/zZnk-HpGizI>

Fixed Effects and First Differences comparison - part 3
<https://youtu.be/SLvdik53JMQ>

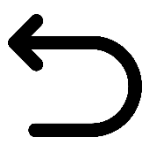
Random Effects Estimator - an introduction
<https://youtu.be/bQampZBzU9Q>

Random Effects vs Fixed Effects estimators
<https://youtu.be/weylmScboA>

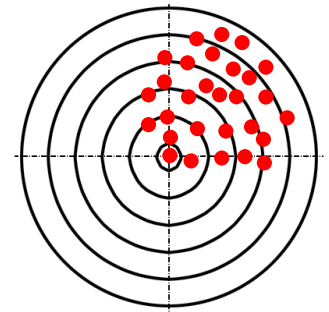
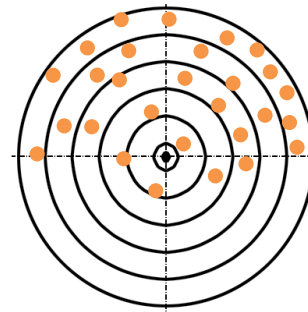
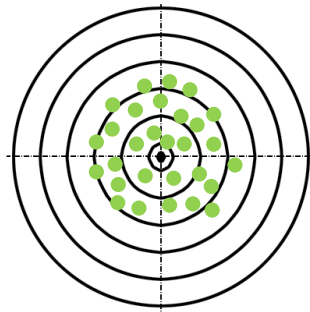
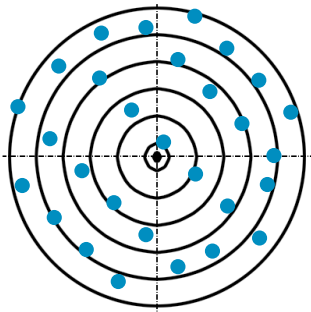
Efeitos fixos: limitações

- Abordagem **não elimina *confounders*** que **variaram no período de tempo analisado e afetaram os grupos de maneira diferente.**
- **Remove toda a variação entre grupos** (pois esta é captada pelos efeitos fixos de entidade e ou tempo); calcula o **efeito de X considerando apenas as variações de X dentro de cada agrupamento.**
- **Diferentemente de efeitos aleatórios** (que têm suas próprias limitações), efeitos fixos **exigem dados longitudinais e não são capazes de calcular coeficientes para:**
 - Covariáveis de interesse se estas não variarem no tempo (no caso de efeitos fixos de entidade) ou
 - Covariáveis de interesse se estas variarem no tempo, mas não entre entidades (no caso de efeitos fixos de tempo). Dito de outro modo, efeitos fixos de tempo absorverão todas as variáveis que variam no tempo, mas não variam entre entidades.

apêndice: recordatório de viés de variável omitida



Acurácia, precisão e a distribuição teórica de β hats



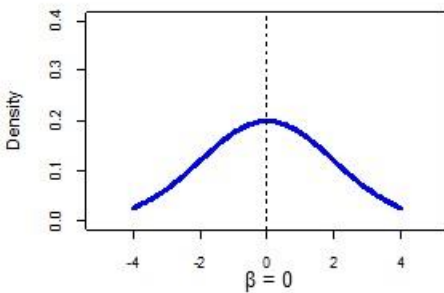
Acurácia
(ausência
de viés)



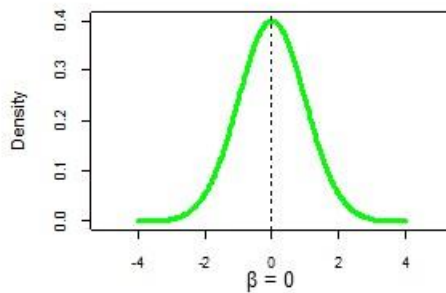
Precisão
(baixa
dispersão)



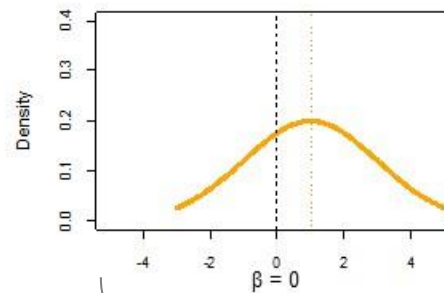
PDF Beta-hat



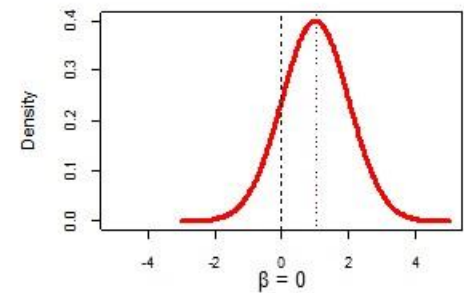
PDF Beta-hat



PDF Beta-hat



PDF Beta-hat



Estes cenários mostram viés positivo (centro da distribuição de β hats está à direita do verdadeiro β). Viés negativo é igualmente prejudicial à acurácia do modelo

Endogeneidade e viés de variável omitida

- O modelo deveria ser:

$$y_i = \beta_0 + \beta_1 t_i + \beta_2 m_i + \varepsilon_i$$

- Onde Y = Rendimento, T = Exposição ao tratamento (participação voluntária em programa de qualificação profissional), M = Motivação do trabalhador
- Todavia, o modelo estimado foi:

$$y_i = \beta_0 + \beta_1 t_i + \varepsilon_i$$

- A exposição ao tratamento é provavelmente determinada por fatores que também causam Y (e.g., M), mas que estão omitidos na regressão; no modelo que omite M, T é considerada uma **variável endógena**, porque ela é correlacionada com o termo de erro
- Modelo não consegue separar efeito de T do efeito de M; o estimador de β_1 produz uma combinação desses efeitos
- Como consequência, o estimador do efeito de T (i.e., o estimador de β_1) está carregando o efeito de T mas também de M; é como se $\hat{\beta}_1$ “absorvesse” parte do efeito de M. Assim, a $E(\hat{\beta}_1)$ se afasta do verdadeiro β_1 , o que caracteriza o viés

Efeito sistemático das variáveis omitidas é captado pelos coeficientes estimados

Numa conversa mais estruturada....

- Na **equação populacional “completa”**, ε carrega apenas o efeito não sistemático de variáveis omitidas; tudo que afeta sistematicamente o Y aparece como variável explicativa
- Na **equação populacional “simplificada”**, aquela que realmente estimamos, nem todos os fatores formadores de Y estão presentes como variáveis explicativas; ε carrega o efeito médio das variáveis omitidas
 - **Premissa:** Variáveis omitidas correlacionam-se com Y, porém não com X, i.e., **$\text{corr}(\varepsilon, X) = 0$**
 - **Premissa:** **$E(\varepsilon) = 0$** , portanto ε não precisa ser “estimado”, e não aparece na equação ajustada
- **Para onde vai, então, o efeito sistemático médio das variáveis omitidas?**

Variáveis omitidas **não correlacionadas com X** têm seu efeito médio atribuído a **$\beta_0\text{hat}$** , afetando o nível de \hat{Y} (intercepto estimado), mas não o efeito de X sobre Y (inclinação da reta ajustada, ou seja, efeito estimado de X em Y)



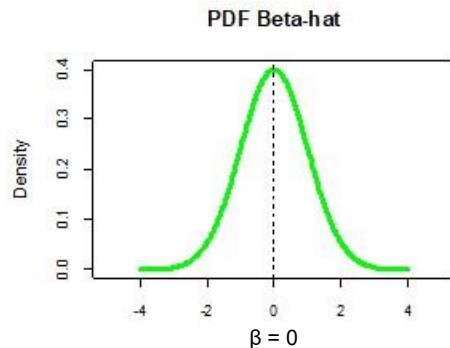
Variáveis omitidas **correlacionadas com X violam premissa de MQO**, têm seu efeito parcialmente absorvido pelo estimador do **βhat** de inclinação respectivo, **e o enviesam**

- O modelo não conseguirá distinguir o efeito de X sobre Y do efeito de Z sobre Y, e **atribuirá erroneamente a X parte do efeito de Z**
- Dada essa **afinidade (correlação)** entre X e Z, é como se o βhat “**atraísse**” parte do efeito médio das variáveis omitidas (que, na ausência de correlação entre X e Z, teria sido totalmente incorporado pelo estimador de $\beta_0\text{hat}$)

Exemplo de viés: Distribuição de $\hat{\beta}$ de inclinação não está centrada no verdadeiro β

Ausência de viés = acurácia (i.e., estimador não tenderá a super ou subestimar β)

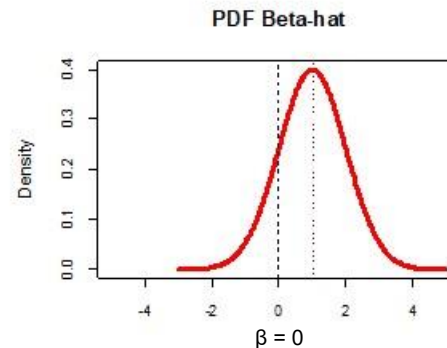
- $E(\hat{\beta}) = \beta$
- Distribuição de $\hat{\beta}$ está centrada no verdadeiro valor de β



- Good news: Na maioria dos casos, estimador não enviesado produzirá “bom” $\hat{\beta}$
- Bad news: Estimador não enviesado pode produzir $\hat{\beta}$ bem distante de β

Viés = “inacurácia” (i.e., estimador tenderá a super ou subestimar β)

- $E(\hat{\beta}) \neq \beta$
- Distribuição de $\hat{\beta}$ **não** está centrada no verdadeiro valor de β



- Modelo não consegue separar efeito de X do efeito de Z e produz $\hat{\beta}$ que é uma combinação desses efeitos. Exemplo:

$$Violent\ crime_t = \beta_0 + \beta_1 Ice\ cream\ sales_t + \epsilon_t$$

Uma regressão simples de crimes violentos e venda de sorvetes provavelmente captará uma associação entre essas variáveis; todavia, essa associação é espúria

Prevendo a direção do viés

- O problema é:

$$E(\hat{\beta}_1) \neq \beta_1$$

- Qual a direção do viés?

- Positivo (“para cima”): $E(\hat{\beta}_1) > \beta_1$

- Negativo (“para baixo”): $E(\hat{\beta}_1) < \beta_1$

- Como prever a direção do viés? Através de uma multiplicação de sinais:

$$\text{Sinal do viés} = \text{Sinal do } \beta_{om} * \text{Sinal de } f(\text{correlação}_{in, om})$$

- β_{om} é o efeito (não observável) da variável omitida sobre Y
- $f(\text{correlação}_{in, om})$ é uma função da correlação entre a variável incluída e a omitida (T e M, no exemplo)

Endogeneidade e viés de variável omitida

- O modelo deveria ser:

$$y_i = \beta_0 + \beta_1 t_i + \beta_2 m_i + \varepsilon_i$$

- Onde Y = Rendimento, T = Exposição ao tratamento (participação voluntária em programa de qualificação profissional), M = Motivação do trabalhador
- Todavia, o modelo estimado foi:

$$y_i = \beta_0 + \beta_1 t_i + \varepsilon_i$$

- **Como prever a direção do viés? Através de uma multiplicação de sinais:**

$$\text{Sinal do viés} = \text{Sinal do } \beta_{om} * \text{Sinal de } f(\text{correlação}_{in, om})$$

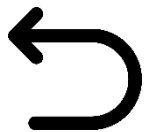
- **B_{om}** é o efeito (não observável) da variável omitida (M) sobre Y
- **f(correlação_{in, om})** é uma função da correlação entre a variável incluída (T, para cujo coeficiente estamos analisando a direção do viés) e a variável omitida (M)

Sinal do viés = positivo * positivo
Sinal do viés = positivo

Estes sinais nos são desconhecidos;
consideramos nossas hipóteses sobre eles.

Conclusão: diante da omissão de M, estimador de β_1 tende a inflacionar o efeito do tratamento

[fim do apêndice]



Efeitos fixos e aleatórios

Avaliação de Políticas Públicas B

16, 18 e 30 de junho de 2025

Leitura básica:

FACURE, Matheus. **Causal inference for the brave and true**. 2022. [Capítulo 14: Panel data and fixed effects.] Disponível em: <https://matheusfacure.github.io/python-causality-handbook/landing-page.html>

Leitura complementar:

HUNTINGTON-KLEIN, Nick. **The effect: an introduction to research design and causality**. 2023. [Capítulo 16: Fixed effects.] Disponível em: <https://theeffectbook.net/>

STOCK, James H.; WATSON, Mark W. **Introduction to Econometrics**. 4. ed. New York: Pearson, 2019. [Capítulo 10: Regression with panel data.] Edições anteriores também contém este capítulo.

TOWNSEND, Zac et al. The choice between fixed and random effects. In: SCOTT, Marc A.; SIMONOFF, Jeffrey S.; MARX, Brian D. (org.). **The SAGE handbook of multilevel modeling**. Londres: SAGE, 2013, p. 73-88. Disponível em: http://agfda.userweb.mwn.de/ALD_2016/downloads/Uebung2/Townsend_et_al_2013.pdf