

# Dados em Painel e Diff-in-Diff

Raphael Corbi

Universidade de São Paulo

Abril 2014

## Dados em Painel e Diff-in-Diff

- ▶ identificação causal = controlar por *confounding factors*
- ▶ *caso confounding factors não observados, podemos usar IV*
- ▶ *mas instrumentos bons são difíceis de achar... qual a alternativa?*
  
- ▶ *explorar dimensão do tempo/coorte para controlar por características não observadas, porém FIXAS no tempo*
- ▶ **hipótese:** *tendências observadas do controle/tratamento iguais*

## Efeitos Fixos Individuais

- ▶ **sindicalização e salários**: existe um efeito?

- ▶ podemos estimar uma regressão do tipo:

$$Y_{it} = \alpha_i + \lambda_t + \rho D_{it} + X'_{it}\beta + \epsilon_{it}$$

- ▶ quais as hipóteses para identificação?
- ▶ H1: sindicalização é aleatório condicional a habilidade e observadas
- ▶ H2: termo não observado (habilidade) é constante
- ▶ H3: forma funcional (modelo) é linear
- ▶ H4: efeito da sindicalização é aditivo e constante
  
- ▶ Nota: hipóteses mais restritas que IV (LOUSA)

## Efeitos Fixos Individuais

- ▶ modelo de efeitos fixos individuais:

$$Y_{it} = \alpha_i + \lambda_t + \rho D_{it} + X_{it}'\beta + \epsilon_{it}$$

- ▶  $\alpha_i$  = parâmetro a ser estimado (efeito fixo), com  $i \gg 0$
- ▶  $\lambda_t$  = parâmetro a ser estimado (efeito tempo)
- ▶ FE é equivalente a OLS com dados transformados:

$$\bar{Y}_i = \alpha_i + \bar{\lambda} + \rho \bar{D}_t + \bar{X}_i'\beta + \bar{\epsilon}_i$$

$$Y_{it} - \bar{Y}_i = \alpha_i + \lambda_t - \bar{\lambda} + \rho(D_{it} - \bar{D}_t) + (X_{it} - \bar{X}_i)'\beta + (\epsilon_{it} - \bar{\epsilon}_i)$$

## Efeitos Fixos Individuais

- ▶ alternativamente podemos usar primeiras-diferenças:

$$\delta Y_{it} = \Delta \lambda_t + \rho \Delta D_{it} + \Delta X'_{it} \beta + \Delta \epsilon_{it}$$

- ▶ onde  $\Delta$  é a diferença de um ano para o outro
- ▶ caso  $t = 2$ , desvios da média e primeiras-diferenças são equivalentes
- ▶ caso  $t > 2$ , primeiras-diferenças são menos eficientes pois os erros apresentam auto-correlação serial

## Freeman (1984) - estimativas de quatro bases de dados

Table 5.1.1: Estimated effects of union status on log wages

Survey	Cross section estimate	Fixed effects estimate
May CPS, 1974-75	0.19	0.09
National Longitudinal Survey of Young Men, 1970-78	0.28	0.19
Michigan PSID, 1970-79	0.23	0.14
QES, 1973-77	0.14	0.16

## Efeitos Fixos Individuais

- ▶  $\beta^{FE} < \beta^{OLS}$  = viés de seleção positivo?
- ▶  $\beta^{FE}$  particularmente suscetível a problemas de erro de medida
- ▶ viés de atenuação
  
- ▶ variáveis do tipo "sindicalizado" = 1 são persistentes
- ▶ erros de medida devido a misreporting/miscoding
- ▶ variação temporal pode ser quase toda ruído!

## Differences-in-Differences

- ▶  $\beta^{FE}$  requer dados em painel
- ▶ e se regressor de interesse variar só para um grupo (mais agregado)?
- ▶ ex: salário-mínimo estadual
- ▶ VVO será relevante somente no nível estado-ano
- ▶ **group-level-FE = identificação por DID**

## Differences-in-Differences

- ▶ pioneiro: John Snow (1855)
- ▶ causa da epidemia de cólera em Londres
- ▶ água contaminada ou "bad air"?
  
- ▶ Southwark and Vauxhall Company vs *Lambeth Company*
- ▶ *Lambeth mudou a captação de água para região livre de esgoto*
- ▶ *taxa de mortalidade em Lambeth caíram relativamente muito*

## Differences-in-Differences

- ▶ Card e Krueger (1994)
- ▶ efeito do salário mínimo
- ▶ mercado competitivo:  $\uparrow$  salários e  $\downarrow$  no emprego
- ▶ prejudica exatamente os beneficiários da lei
  
- ▶ Abril 1992: New Jersey aumenta SM de \$4.25 para \$5.05
- ▶ dados de restaurantes fast-food NJ e PENN
- ▶ PENN manteve SM de \$4.25 entre Fev e Nov

## Differences-in-Differences

- ▶ DID é uma versão do FE que usa dados agregados
- ▶ seja  $Y_{1ist}$  emprego potencial
  - ▶ caso salário mínimo seja alto ( $Y_{0ist}$  se baixo)
  - ▶ restaurante  $i$ , estado  $s$ , período  $t$
  - ▶ como sempre, somente observamos  $Y_1$  ou  $Y_0$
- ▶ Abril 1992: New Jersey aumenta SM de \$4.25 para \$5.05
- ▶ H1: estrutura aditiva para resultados potenciais
- ▶ H2:  $E[Y_{1ist} - Y_{0ist}] = \delta$  constante
- ▶ LOUSA

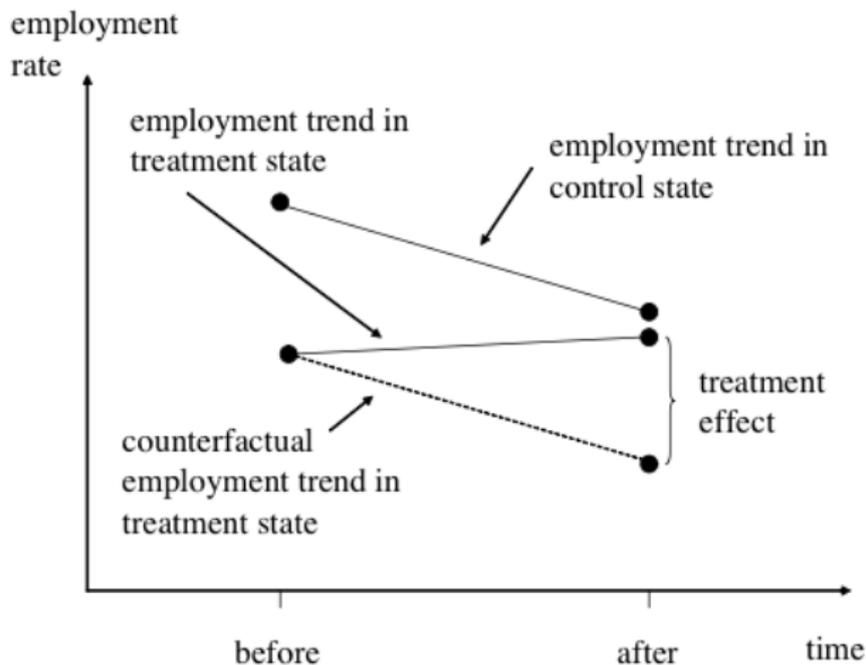
## Card e Krueger (1994) - emprego antes e depois

Variable	PA (i)	NJ (ii)	Difference, NJ-PA (iii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (0.51)	-2.89 (1.44)
2. FTE employment after, all available observations	21.17 (0.94)	21.03 (0.52)	-0.14 (1.07)
3. Change in mean FTE employment	-2.16 (1.25)	0.59 (0.54)	2.76 (1.36)

## Differences-in-Differences

- ▶ efeito é positivo!
- ▶ qual a hipótese de identificação?
  - ▶ na falta do tratamento, tendência de emprego deve ser a mesma em ambos os estados
  - ▶ estados podem ser diferentes, mas de maneira constante

## Card e Krueger (1994) - emprego antes e depois



## Differences-in-Differences

- ▶ a estimativa é tão boa quanto a hipótese!
- ▶ como testamos se a hipótese é boa? DADOS!
  
- ▶ Card e Krueger (2000) revisitam com mais dados históricos
- ▶ *payroll data de 1991 ate 1997*
- ▶ *idéia: quanto mais parecidas as tendências, melhor o contrafactual*

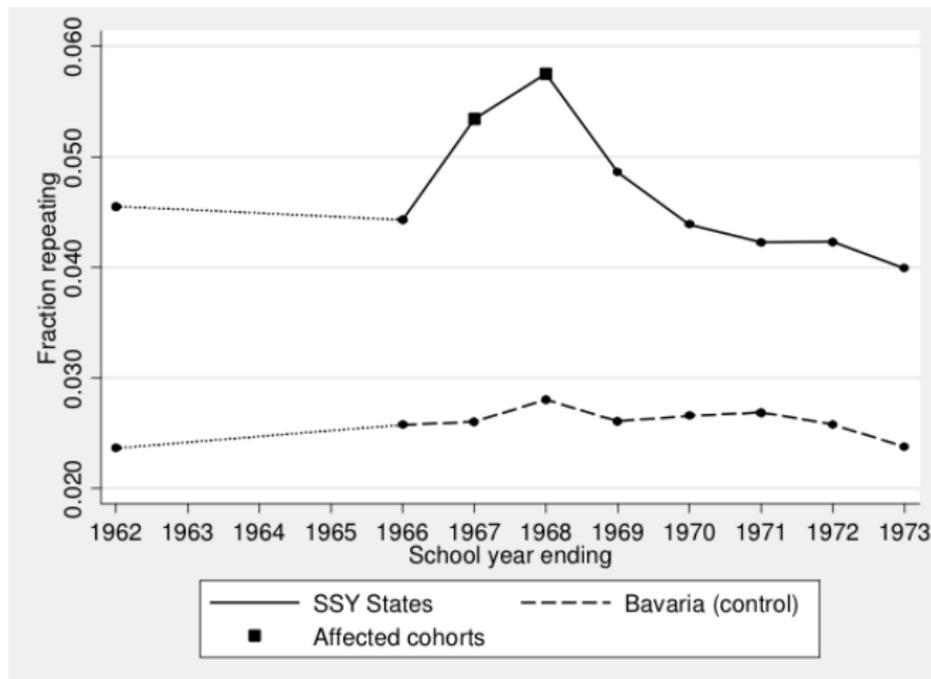
## Card e Krueger (2000) - tendência NJ e PENN



## Differences-in-Differences

- ▶ Pischke (2007)
- ▶ duração ano escolar sobre performance
- ▶ até 1960's, início escolar se dava na primavera em várias regiões da Alemanha (exceto Baviera)
- ▶ em 1966-67, mudança para outono
- ▶ transição: 2 coortes de alunos acabam tendo 2 anos mais curtos
- ▶ currículo foi mantido
- ▶ repetência, matrícula no ensino médio, salário e emprego futuro

## Pischke (2007) - ano escolar mais curto



## Differences-in-Differences

- ▶ Como estimar DID usando regressões?

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta(NJ_s d_t) + \epsilon_{ist}$$

- ▶ onde  $NJ_s d_t = D_{st}$  que indica tratamento
- ▶ esse modelo é saturado
- ▶ qual o link entre os parâmetros do modelo e da função da média condicional (LOUSA)?

## Differences-in-Differences

- ▶ vantagens do DID
  - ▶ forma conveniente de estimar efeito tratamento e SD
  - ▶ fácil de adicionar mais estados e períodos no tempo
- ▶ podemos também analisar tratamento de intensidade variável:
  - ▶ Card (1992): SM em todos os estados
  - ▶ alguns tem SM próximo ao nacional, alguns mais altos
  - ▶ Em 1990 SM nacional sobe de \$3.25 para \$3.80
  - ▶  $FA_s$  = fração de jovens que ganham entre \$3.25 para \$3.80

$$Y_{ist} = \gamma_s + \lambda_t + \delta(FA_s d_t) + \epsilon_{ist}$$

## Card (1992) - emprego e salário mínimo nos EUA

Table 5.2.2: Regression-DD estimates of minimum wage effects on teens, 1989 to 1992

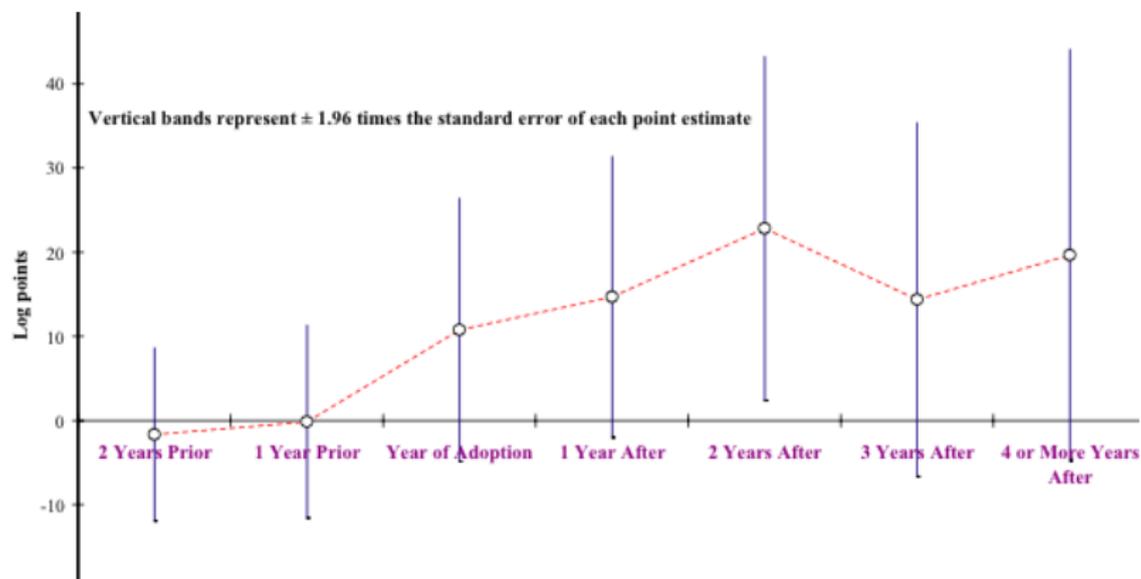
Explanatory Variable	Equations for Change in Mean Log Wage:		Equations for change in Teen Employment-Population Ratio:	
	(1)	(2)	(3)	(4)
1. Fraction of Affected Teens	0.15 (0.03)	.14 (0.04)	0.02 (0.03)	-.01 (0.03)
2. Change in Overall Emp./Pop. Ratio	–	0.46 (0.60)	–	1.24 (0.60)
3. R-squared	0.30	0.31	0.01	0.09

## Differences-in-Differences: causalidade de Granger

- ▶ no caso de muitos períodos, podemos fazer o teste de causalidade de Granger
- ▶ não é um teste formal de inferência causal, mas explora o timing dos acontecimentos
- ▶ teste se  $D_{st-\tau}$  (passado) preve  $Y_{ist}$ , mas  $D_{st+\tau}$  (futuro) não preve

$$Y_{ist} = \gamma_s + \lambda_t + \sum_{\tau} \delta_{-\tau} D_{s,t-\tau} + \sum_{\tau} \delta_{+\tau} D_{s,t+\tau} + X'_{ist} \beta + \epsilon_{ist}$$

## Autor (2002) - emprego temporário e proteção legal



## Besley e Burgess (2004) - proteção legal e produtividade

Table 5.2.3: Effect of labor regulation on the performance of firms in Indian states

	(1)	(2)	(3)	(4)
Labor regulation (lagged)	-0.186 (.0641)	-0.185 (.0507)	-0.104 (.039)	0.0002 (.02)
Log development expenditure per capita		0.240 (.1277)	0.184 (.1187)	0.241 (.1057)
Log installed electricity capacity per capita		0.089 (.0605)	0.082 (.0543)	0.023 (.0333)
Log state population		0.720 (.96)	0.310 (1.1923)	-1.419 (2.3262)
Congress majority			-0.0009 (.01)	0.020 (.0096)
Hard left majority			-0.050 (.0168)	-0.007 (.0091)
Janata majority			0.008 (.0235)	-0.020 (.0333)
Regional majority			0.006 (.0086)	0.026 (.0234)
State-specific trends	NO	NO	NO	YES
Adjusted R-squared	0.93	0.93	0.94	0.95

## efeitos fixos vs variável dependente defasada

- ▶ FE e DID assumem variáveis omitidas constantes no tempo
- ▶ lembrem-se do paper Dehejia e Wahba (1999):
  - ▶ participação em programas de treinamento subsidiado
- ▶ Ashenfelter e Card (1985) mostram que participantes geralmente observam grande queda de renda logo antes do programa
- ▶ logo, renda antes do tratamento varia com o tempo

## Efeitos Fixos vs Variável Dependente Defasada

- ▶ podemos incluir renda passada como controle:

$$Y_{it} = \alpha_i + \theta Y_{it-1} + \lambda_t + \delta D_{it} + X'_{it}\beta + \epsilon_{ist}$$

- ▶ mas misturar lagY e FE gera um grande problema
- ▶ tirando a diferença para eliminarmos o FE, temos:

$$\Delta Y_{it} = \theta \Delta Y_{it-1} + \Delta \lambda_t + \delta \Delta D_{it} + \Delta X'_{it}\beta + \Delta \epsilon_{it}$$

- ▶  $\Delta \epsilon_{it}$  é necessariamente correlacionado com  $\theta \Delta Y_{it-1}$
- ▶ isso é chamado de Nickell bias

## Efeitos Fixos vs Variável Dependente Defasada

- ▶ qual a solução para o Nickell bias?
- ▶ podemos usar  $Y_{it-2}$  como instrumentos de  $Y_{it-1}$ , mas isso somente seria válido caso resíduos não fossem auto-correlacionados
- ▶ Alternativa de Guryan (2004):
  - ▶ caso correto seja modelo com somente defasagem, efeito estimado de FE será grande demais
  - ▶ caso correto seja modelo com FE, efeito estimado com defasagem será pequeno demais
- ▶ logo, eles podem ser vistos como extremos de uma banda possível do efeito tratamento verdadeiro