

# Differences in differences

Valutazione di politiche

Simona Comi  
Management e Design dei servizi  
AA 2021/2022

# Differences in differences

- Vediamo un esempio di DID, preso dal capitolo 5 del libro Mastering metrics
- Nel 1920 "Caldwell and Company" regnava sulla più grande catena di banche nel SUD degli USA e molte altre attività economiche. Erano nota come l'impero dei Caldwell.
- Nel novembre del 1930, la corruzione e una gestione scorretta ne causarono il fallimento
- Entro 6 mesi, 29 banche del Mississippi chiusero, perchè i correntisti furono presi dal panico (corsa allo sportello).

# La caduta dell'impero Caldwell

- Questo tipo di collasso finanziario necessariamente causa il declino delle attività economiche?
- I policy makers che fronteggiano una corsa allo sportello hanno due possibilità: chiudere il rubinetto o aprirlo maggiormente facilitando la circolazione del denaro.
  - L'aumento del circolante permette alle banche che hanno problemi di affrontare le domande di prelievo, evitando che si diffonda il panico e riducendo la necessità di ulteriore supporto in futuro
  - Ma ciò funziona solo se le crisi sono solo dovute a una carenza di fiducia.....altrimenti si rischia di sovvenzionare le banche peggiori....

# La caduta dell'impero di Caldwell

- Il sistema bancario federale americano è organizzato in 12 distretti, ognuno gestito da una Federal Reserve Bank. Al tempo della crisi del '29, i governatori di queste dodici banche godevano di una elevata autonomia.
- La FED di Atlanta (6° distretto) preferiva prestare soldi alle banche in difficoltà, mentre quella di St Louis (8° distretto) seguiva la politica della Real Bills Doctrine, e restringeva il credito alle banche in difficoltà.
- Il confine tra questi due distretti attraversa a metà verticalmente lo stato del Mississippi.
- Abbiamo un esempio di politiche diverse nello stesso stato!!!!

# Quasi-experimental evidence

- Richardson and Troost " Monetary Intervention Mitigated Banking Panics during The Great Recession: Quasi-Experimental Evidence from a Federal Reserve District Border, 1929-1933." JPE, 2009.
- In questo articolo gli autori usano l'ottavo distretto come gruppo di controllo per il sesto (gruppo dei trattati). La politica che valutano è l'aumentare i prestiti alle banche in difficoltà.
- Un primo outcome è il numero di banche (ancora) attive: 132 nell'ottavo, 121 nel sesto nel 1931.
- Ma attenzione, nel 1930 prima della crisi Caldwell, le banche erano 135 nel sesto e 165 nell'ottavo distretto.
- Possiamo usare DID o DD per tenere conto delle differenze pre-esistenti nei due distretti.

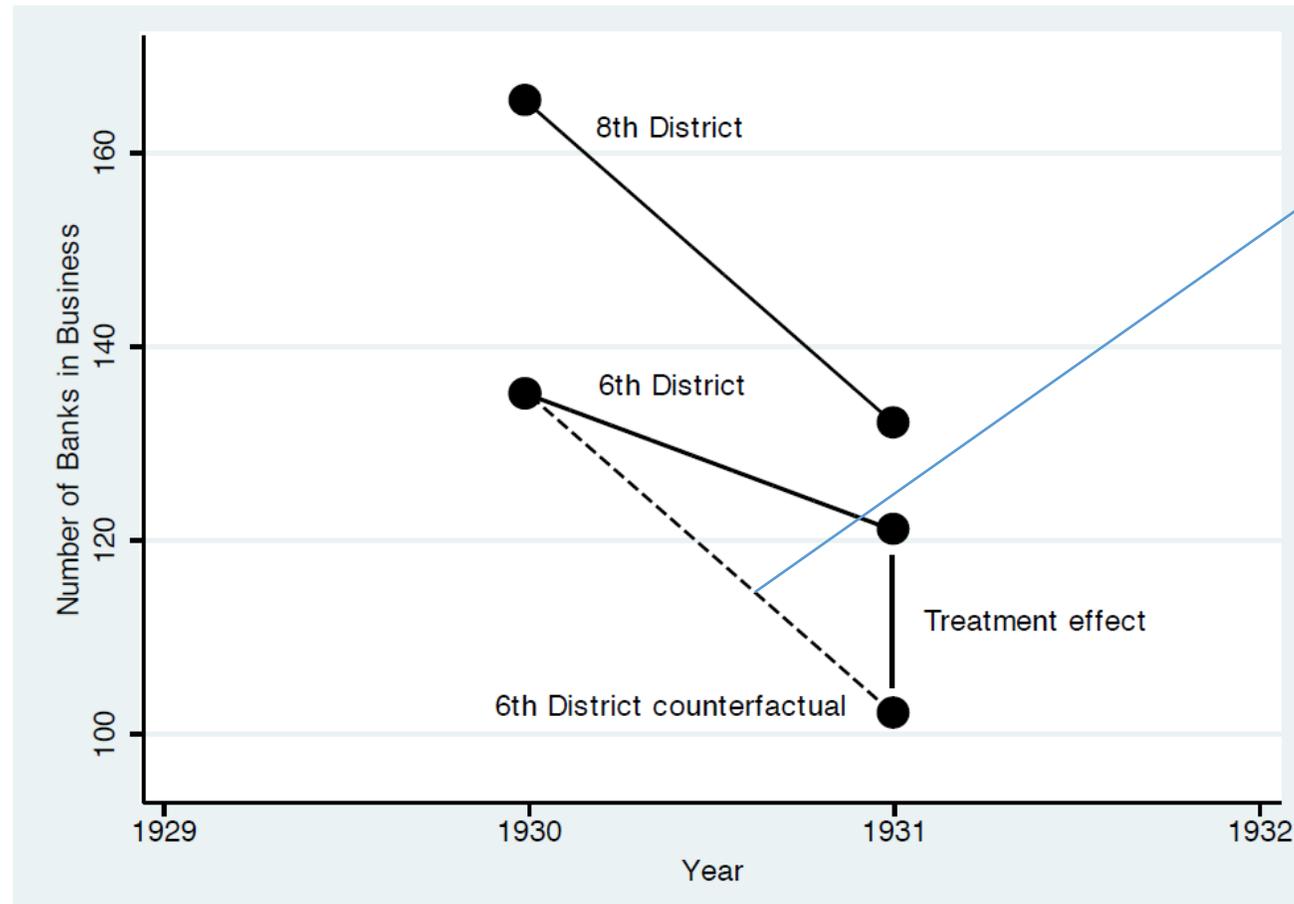
# Mondi paralleli

- Chiamiamo  $Y_{dt}$  il numero di banche aperte nell'anno  $t$  (1930 o 1931) nel distretto  $d$  (6 o 8)
- Possiamo calcolare l'effetto della politica come:

$$\begin{aligned}\delta_{DD} &= (Y_{6,1931} - Y_{6,1930}) - (Y_{8,1931} - Y_{8,1930}) \\ &= (121 - 135) - (132 - 165) \\ &= -14 - (-33) = 19.\end{aligned}$$

- Stiamo confrontando la variazione nel numero di banche tra i due distretti, non il loro livello.
- La logica del DD è illustrata dalla seguente figura

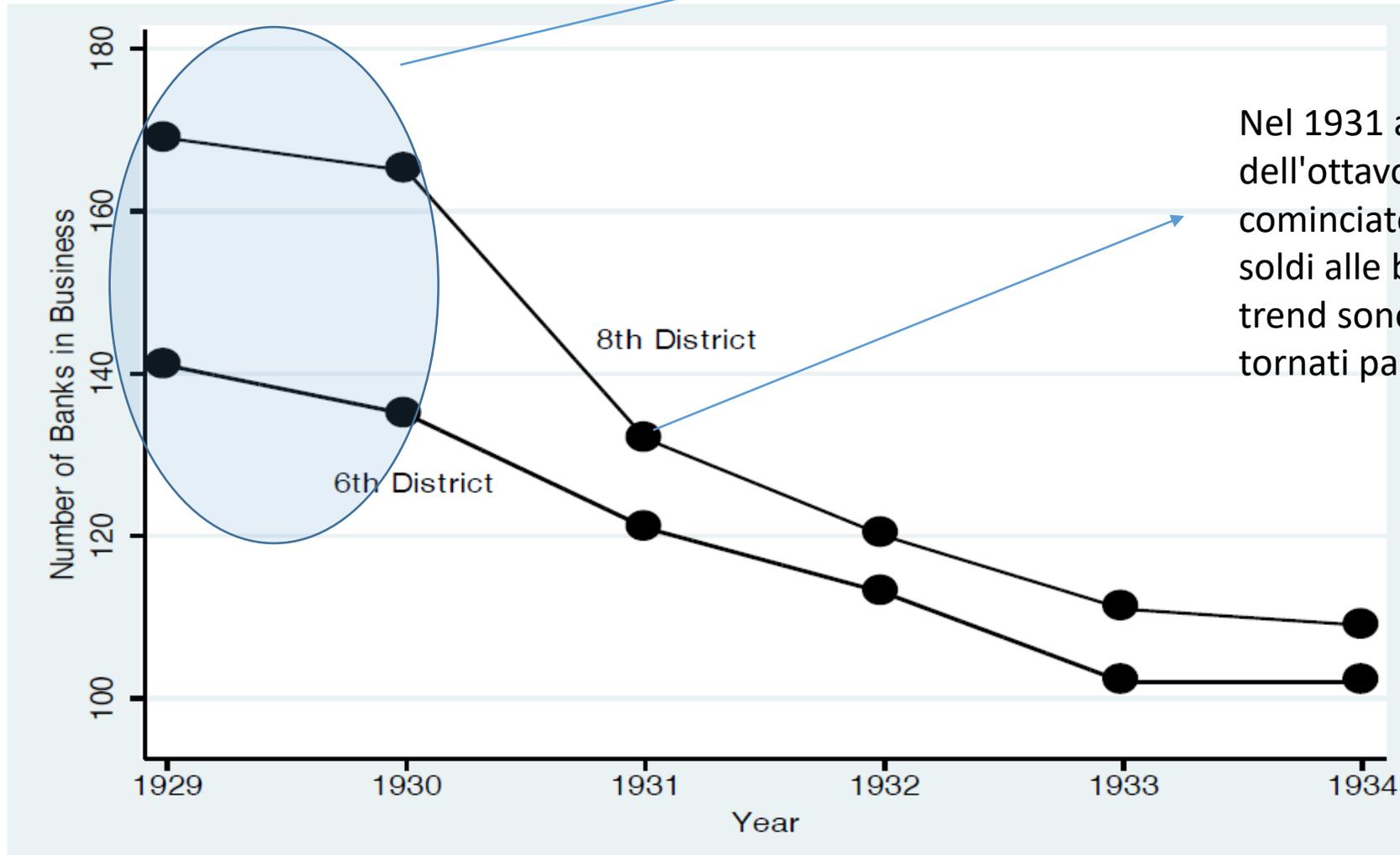
# DD in chiave grafica



Questa linea ci dice quello che sarebbe successo nel sesto distretto senza l'intervento della banca federale.

# Una visione più lunga

Common trend prima del trattamento: le due rette sono parallele



Nel 1931 anche la Fed dell'ottavo distretto ha cominciato a prestare soldi alle banche in crisi: i trend sono di nuovo tornati paralleli!!!

# DD con la regressione – già in differenze

- Il modo più semplice per stimare il differences in differences è lavorare con le variazioni ( $\Delta$ ).
- Definiamo un dummy che assuma valore 1 per chi è stato trattato (dummy treat)
- Calcoliamo  $\Delta_{6,1931-1930} = Y_{6,1931} - Y_{6,1930}$  e  $\Delta_{8,1931-1930} = Y_{8,1931} - Y_{8,1930}$ . E applichiamo queste differenze ai diversi anni. Possiamo indicare le differenze più in generale con  $\Delta_d$
- In regressione, possiamo stimare:

$$\Delta Y_d = \beta_0 + \beta_1 X_d + u_d$$

- $\beta_1$  è lo stimatore DD

# DD con la regressione non in differenze

- Per stimare DD con dati in «livello», abbiamo bisogno di tre ingredienti :
  - Una dummy per il distretto trattato, che chiamiamo  $TREAT_d$ , questa è sempre uguale a 1 per il distretto trattato (non varia nel tempo), controlla per le caratteristiche medie del distretto trattato (fisse nel tempo) (**Dummy trattamento**)
  - Una dummy per il periodo post trattamento ( $POST_t$ , che invece varia nel tempo, e prende valore 1 dal 1931 in avanti (0 prima del 1931) (**Dummy post**)
  - Una interazione  $TREAT_d \times POST_t$  ottenuta moltiplicando queste due dummies tra di loro: il coefficiente di questa variabile è la stima DD. Questa variabile indica il sesto distretto dopo la politica (**Interazione**).
- Possiamo quindi stimare:

$$Y_{d,t} = \alpha + \beta TREAT_d + \gamma POST_t + \delta_{rDD}(TREAT_d * POST_t) + e_{d,t}$$

# La stima

- Con le 12 osservazioni che abbiamo (2 distretti x 6 anni) si ottiene:

$$Y_{dt} = 167 - \underset{(8.8)}{29} \text{TREAT}_d + \underset{(7.6)}{49} \text{POST}_t + \underset{(10.7)}{20.5} (\text{TREAT}_d \times \text{POST}_t) + e_{dt}$$

- Infine, andiamo a vedere l'effetto su altri due outcome che misurano meglio la situazione economica nei due distretti, il numero di venditori all'ingrosso (wholesale) e il totale delle vendite all'ingrosso

TABLE 5.1

## Wholesale firm failures and sales in 1929 and 1933

	1929	1933	Difference (1933–1929)
Panel A. Number of wholesale firms			
Sixth Federal Reserve District (Atlanta)	783	641	–142
Eighth Federal Reserve District (St. Louis)	930	607	–323
Difference (Sixth–Eighth)	–147	34	181
Panel B. Net wholesale sales (\$ million)			
Sixth District Federal Reserve (Atlanta)	141	60	–81
Eighth District Federal Reserve (St. Louis)	245	83	–162
Difference (Sixth–Eighth)	–104	–23	81

# Le due ipotesi fondamentali perché il DID sia valido

- 1. Nulla deve essere avvenuto oltre al trattamento nell'anno (o al momento) del trattamento. Il DID infatti consiste in una differenza prima e dopo un certo istante nel tempo. Coglie quindi l'effetto di tutto quello che è avvenuto contemporaneamente al trattamento.
- 2. Trattati e controlli devono avere un trend parallelo prima del trattamento.

# 1. Solo il trattamento....

- Capiamo l'importanza di questa ipotesi con un altro esempio classico, noto come "The Mariel Boatlift experiment"
- Il 20 aprile del 1980 Fidel Castro ha dichiarato che i cubani erano liberi di migrare verso gli Stati Uniti dal porto di Mariel.
- Da maggio a settembre di quell'anno(1980), 125.000 Cubani sono arrivati a Miami su una flottiglia di navi.

# The mariel boatlift

- Questa variazione improvvisa ed enorme nella popolazione residente a Miami costituisce un esperimento naturale, che ci permette di stimare l'effetto di uno shock esogeno positivo nell'immigrazione sulle condizioni del mercato del lavoro locale. La domanda di ricerca è: che effetto ha un flusso migratorio sulle condizioni (occupazione e salari) del mercato del lavoro locale?
- Questo evento causò:
  - L'aumento del 7% della popolazione residente a Miami tra maggio a settembre
  - L'aumento del 20% della popolazione di cubani a Miami
  - Ci si aspetta un effetto sui tassi di occupazione e sui salari anche di altri gruppi etnici.

# The mariel boatlift

- Utilizzando i dati sulla popolazione e sulle condizioni del mercato del lavoro di Miami prima e dopo il mariel boatlift, e confrontandoli con quelli di altre città simili a Miami è possibile misurare l'effetto di questo shock con il DID.
- Le altre città rappresentano il controfattuale di Miami e ci dicono quello che sarebbe successo a Miami in assenza dell'intensa migrazione.

# Miami population by ethnicity before Mariel Boatlift

Table 1. Characteristics of 16–61-Year-Olds in Miami, 1979.

Characteristic	Whites	Blacks	Cubans	Hispanics	All
<i>Characteristics of Population Age 16–61</i>					
1. Estimated Number (1000's)	319.3	244.1	252.4	102.9	928.4
2. Mean Education	12.8	11.4	11.0	11.6	11.8
3. Percent in Labor Force	75.6	68.3	77.2	68.8	73.1
<i>Characteristics of Those in Labor Force</i>					
4. Estimated Number (1000's)	241.3	166.6	194.7	70.8	678.2
5. Mean Education	13.1	11.8	11.3	11.9	12.1
6. Percent Age 16–24	21.1	24.1	22.0	26.0	22.8
<i>Occupation Distribution (Percent of Employed)</i>					
7. Professional and Technical	19.1	10.9	9.5	10.1	13.2
8. Managers	15.7	2.8	8.6	8.1	9.4
9. Sales	6.2	4.4	7.8	7.6	6.5
10. Clerical	21.9	21.0	19.1	20.9	20.9
11. Craftsmen	13.3	9.4	15.1	12.7	12.8
12. Operatives	4.4	8.4	19.4	16.7	11.1
13. Transportation Operatives	2.6	8.1	5.4	5.9	5.2
14. Laborers	5.1	10.5	4.7	4.0	6.3
15. Farm Workers	1.1	0.1	0.4	0.8	0.6
16. Less-Skilled Service Workers	5.0	13.3	6.1	10.2	8.0
17. More-Skilled Service Workers	5.7	10.9	4.0	3.0	6.2

CENSUS 1980: **35.5** of residents in MIAMI were foreign born, compared to **22.4** in Los Angeles and **6.1** nationwide. **Rows 1-17**: crude indication of the degree of labour market competition among the four groups. Cubans and Hispanics have very similar occupation distributions, while Blacks are more concentrated in in service related occupations and are under-represented in managerial occupations.

# Descriptive stats of Cuban Population in march 1985

## Mariels have:

- lower education
- Younger
- Male

as compared to other Cubans.

Characteristic	Mariel	All other
	Immigrants	Cubans
Educational Attainment (Percent of Population in Each Category):		
No High School	56.5	25.4
Some High School	9.1	13.3
Completed High School	9.5	33.4
Some College	6.8	12.0
Completed College	18.1	15.8
Percent Male	55.6	50.7
Percent Under 30 in 1980	38.7	29.6
Mean Age in 1980 (Years)	34.9	38.0
Percent in Miami in 1985	53.9	52.4
Percent Worked in 1984	60.6	73.4
Mean Log Hourly Earnings	1.37	1.71
Occupation Distribution (Percent Employed in Each Category):		
Professional/Managers	19.3	21.0
Technical	0.0	1.5
Sales	4.5	11.2
Clerical	2.5	13.5
Craftsmen	9.5	19.9
Operatives	19.1	13.8
Transportation Ops.	3.8	4.3
Laborers	10.8	3.3
Farm Workers	0.0	1.8
Less-Skilled Service	26.0	7.4
More-Skilled Service	4.6	2.3
Sample Size	50	528
Weighted Count	42,300	476,900

## 2. Unadjusted wage

gap: Mariels VS  
Cubans = - 34%

Attributable to lower  
education and  
younger age.

## 3. Adjusted wage

gap (education,  
pot exp, gender)  
: Mariels VS  
Cubans = - 18%

## Unemployment rates by ethnic group and city

<i>Group</i>	<i>1979</i>	<i>1980</i>	<i>1981</i>	<i>1982</i>	<i>1983</i>	<i>1984</i>	<i>1985</i>
<i>Miami:</i>							
Whites	5.1 (1.1)	2.5 (0.8)	3.9 (0.9)	5.2 (1.1)	6.7 (1.1)	3.6 (0.9)	4.9 (1.4)
Blacks	8.3 (1.7)	5.6 (1.3)	9.6 (1.8)	16.0 (2.3)	18.4 (2.5)	14.2 (2.3)	7.8 (2.3)
Cubans	5.3 (1.2)	7.2 (1.3)	10.1 (1.5)	10.8 (1.5)	13.1 (1.6)	7.7 (1.4)	5.5 (1.7)
Hispanics	6.5 (2.3)	7.7 (2.2)	11.8 (3.0)	9.1 (2.5)	7.5 (2.1)	12.1 (2.4)	3.7 (1.9)
<i>Comparison Cities:</i>							
Whites	4.4 (0.3)	4.4 (0.3)	4.3 (0.3)	6.8 (0.3)	6.9 (0.3)	5.4 (0.3)	4.9 (0.4)
Blacks	10.3 (0.8)	12.6 (0.9)	12.6 (0.9)	12.7 (0.9)	18.4 (1.1)	12.1 (0.9)	13.3 (1.3)
Hispanics	6.3 (0.6)	8.7 (0.6)	8.3 (0.6)	12.1 (0.7)	11.8 (0.7)	9.8 (0.6)	9.3 (0.8)

A sizable increase in Cuban unemployment in Miami following the Mariel immigration.

Assuming unemployment rates of earlier Cuban immigrants were unaffected by Mariel Boatlift, this effect is consistent with unemployment rates of around **20% among the Mariels themselves**. This suggests that the increase in Cuban unemployment rates could easily be explained as a result of the addition of the Mariel refugees to the Cuban population.

# La strategia di identificazione

- Supponiamo di avere:
- due città
  - Miami (M) esposta alla migrazione
  - Los Angeles (L) non esposta alla migrazione
- Due periodi t:
  - $T_1=79$  prima della ondata migratoria
  - $T_2=81$  dopo l'ondata migratoria
- Ipotesi identificatrice:
- Le città hanno avuto trend di immigrazioni simili prima dell'ondata cubana.

- I trend in termini di salari e disoccupazione di Miami tra il 1980 e il 1985 furono simili a quelli di città come Atlanta, Houston e Los Angeles, che non hanno sperimentato il flusso di migranti da Mariel. Le usiamo come gruppo di controllo.
- Nel 1979, prima del Mariel flow, il tasso di disoccupazione della popolazione afro-americana di Miami era del 8,3% ed è salito al 9,6% dopo il Mariel flow.
- Prima di concludere che i Marielitos furono responsabili di **questo aumento del'1,3% nella disoccupazione degli afroamericani**, dobbiamo vedere cosa successe nelle altre città.
- Scopriamo che nelle città di controllo il tasso di disoccupazione degli afro-americani stava **crescendo ancora più rapidamente, dal 10.3 al 12.6 %, (un aumento di 2.3 punti percentuali)**—probabilmente perchè le condizioni macroeconomiche stavano peggiorando in quel periodo anche in altre parti degli US.
- Sembra quindi che **il Mariel flow abbia rallentato l'aumento del tasso di disoccupazione** degli afro-americani.
- Il calcolo difference-in-differences (1.3–2.3) suggerisce che il Mariel flow abbia ridotto del 1.0 % il tasso di disoccupazione degli afro-americani.

# L'effetto dell'immigrazione sul mercato del lavoro di Miami

	<u>The Mariel Flow</u>	
	Before	After
Unemployment rate of blacks in Miami	8.3	9.6
Comparison cities	10.3	12.6
Difference-in-differences	<b>-1.0</b>	

- La conclusione del paper di Card sembra essere confermata da quanto osservato in corrispondenza di grandi flussi migratori in altri paesi.
- 900.000 persone di origine Europea ritornarono in Francia in un anno dopo che l'Algeria ottenne l'indipendenza dell'Algeria nel 1962, aumentando la forza lavoro francese del 2%. Non si registrarono effetti significativi sul tasso di disoccupazione.
- Allo stesso modo, quando il Portogallo perse le colonie africane del Mozambico e dell'Angola a metà degli anni '70, quasi 600000 persone tornarono in Portogallo aumentando la popolazione del 7%.
- I *retornados* non ebbero un grande effetto sull'economia portoghese.

# Should Diff in diff be trusted?

- Lo studio del Mariel è un eccellente esempio della metodologia difference-in-differences.
- Recenti ricerche però hanno sollevato alcune domande sulla solidità dell'evidenza prodotta da questa metodologia che sfrutta quelli che abbiamo chiamato natural experiments

# The Mariel boatlifts that did not happen

- Nel 1994, le condizioni politiche ed economiche a Cuba erano tali da far pensare che sarebbe partita una nuova ondata migratoria verso Miami: migliaia di cubani partirono sui gommoni.
- L'amministrazione Clinton per impedire ai rifugiati di raggiungere le coste della Florida ordinò alla Marina Militare di deviare i migranti verso la base di Guantanamo.
- Alla fine pochissimi rifugiati raggiunsero Miami.

- E' possibile replicare la stima vista prima intorno a questa ondata migratoria "mancata" confrontando Miami con le stesse città di controllo usate nello studio del Mariel boatlifts ("the Mariel boatlift that didn't happen)
- La stima mostra che questo "non evento" ebbe un effetto negativo significativo sul tasso di disoccupazione dei lavoratori di colore di Miami.

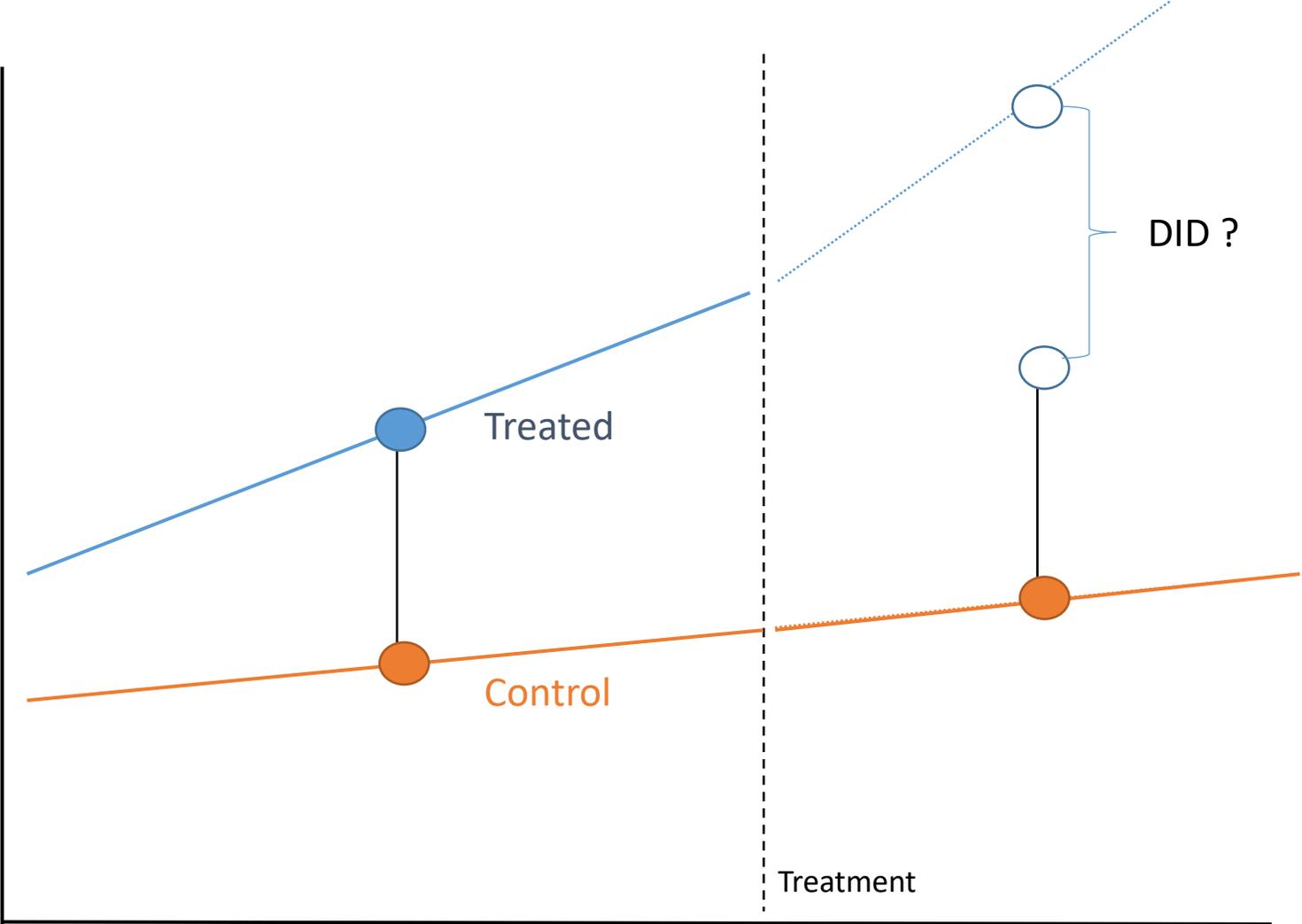
	<u>The Mariel Flow</u>		<u>The Mariel Flow That Did Not Happen</u>	
	Before	After	Before	After
Unemployment rate of blacks in				
Miami	8.3	9.6	10.1	13.7
Comparison cities	10.3	12.6	11.5	8.8
Difference-in-differences	<b>-1.0</b>		<b>+6.3</b>	

- Il tasso di disoccupazione dei neri a Miami salì dal 10.1 al 13.7 tra il 1993 e il 1995, mentre nelle città di controllo scese dal 11.5 al 8.8%.
- La stima difference-in-differences imputa quindi ad un flusso migratorio che non ci fu un aumento del tasso di disoccupazione della popolazione nera di Miami del 6.3%.
- Come possiamo interpretare questo risultato?
- Semplicemente, per come è costruito il difference-in-differences coglie l'effetto di ogni evento avvenuto nello stesso periodo dello shock.... Ed è impossibile distinguere i due effetti!!

## 2. Ipotesi del parallel trend

- E' una ipotesi fondamentale, che valida l'uso del DD. In sua assenza, il gruppo di controllo non è il controfattuale ideale per i trattati → imputo al trattamento una differenza sistematica tra trattati e controlli.
- Come possiamo testare la presenza di un trend parallelo tra trattati e controlli prima del trattamento?
- Finora ci siamo limitati all'analisi grafica

Outcome



Treated

Control

Treatment

DID ?

Time

CYCLING TO SCHOOL:  
INCREASING SECONDARY SCHOOL ENROLLMENT FOR GIRLS IN INDIA

Karthik Muralidharan  
Nishith Prakash

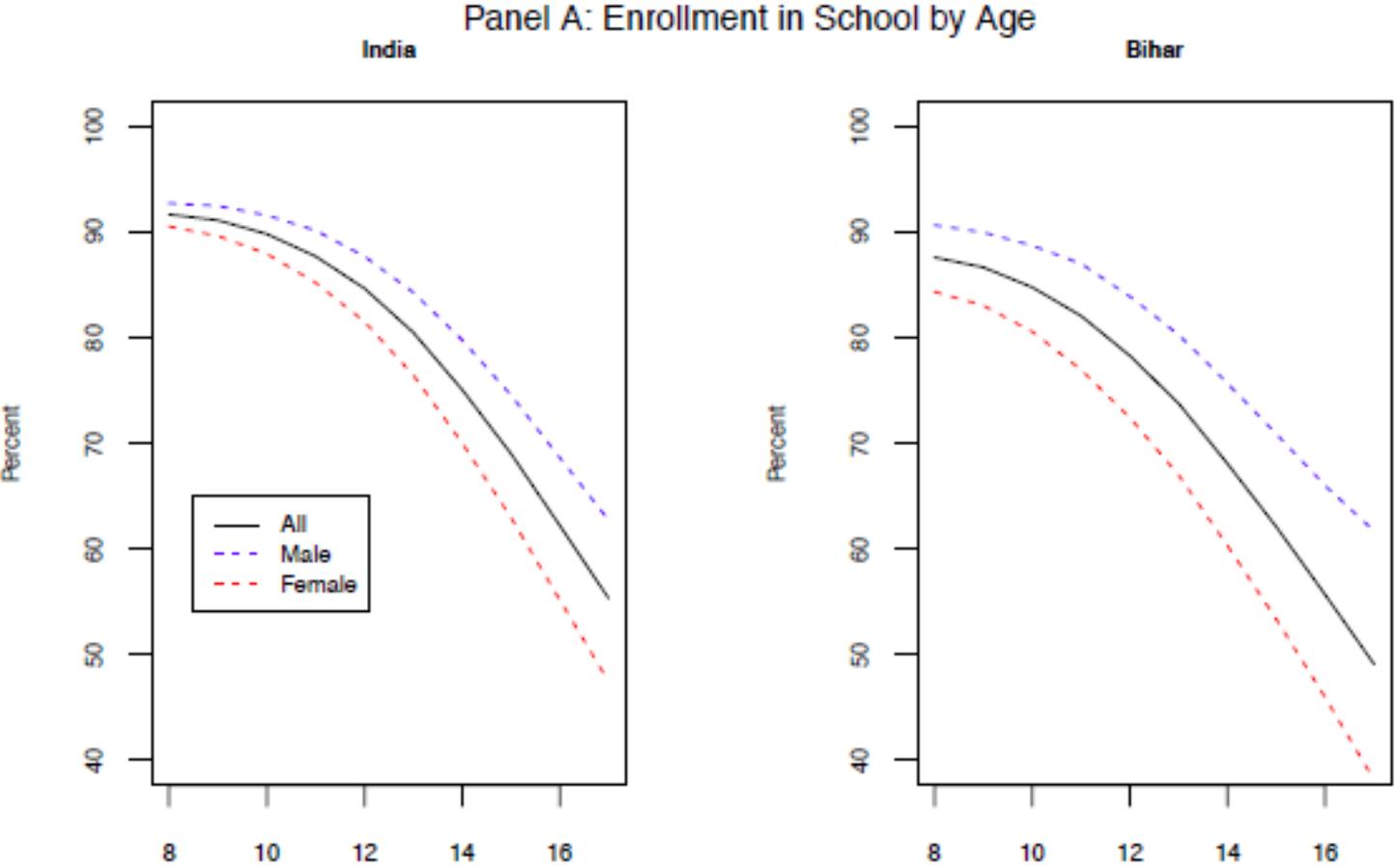
Working Paper 19305  
<http://www.nber.org/papers/w19305>



# Cycling to school

- Nel 2006 lo stato del Bihar ha introdotto una politica per aumentare i tassi di iscrizione alla scuola superiore per le ragazze
- La politica prevedeva un voucher per l'acquisto di una bicicletta.
- Le biciclette potevano essere acquistate dalle ragazze iscritte al 9 e 10 anno di scuola (1 e 2 superiore, 14 e 15 anni).

# Tassi di enrollment (iscrizione)



# Dati e strategia di stima

- Usano una survey condotta nel 2008 (circa 1,5 anni dopo la politica)
- Trattati dopo la politica: ragazze di 14 -15 anni
- Trattati prima: ragazze di 16-17 anni
- Gruppo di controllo prima e dopo la politica: i ragazzi 14-15 e ragazzi di 16-17 anni
- Outcome: iscritte o che hanno già completato grade 9 (prima superiore)

Questa è una strategia abbastanza comune nel diff-in-diff

# IL modello econometrico: DD

$$y_{ihv} = \beta_0 + \beta_1 \cdot F_{ihv} \cdot T_{ihv} + \beta_2 \cdot F_{ihv} + \beta_3 \cdot T_{ihv} + \varepsilon_{ihv}$$

- Dove  $Y_{ihv}$  è l'outcome che corrisponde al ragazzo  $i$ , nella famiglia  $h$  e nel villaggio  $v$ 
  - $F_{ihv}$  è una dummy female
  - $T_{ihv}$  è una dummy che indica la coorte trattata.
- Rispetto a quanto già visto, qui il DD va composto considerando:
  - TREAT= FEMALE
  - POST=COORTE TRATTATA
  - TREAT\* POST=INTERAZIONE TRA FEMALE E COORTE TRATTATA
- NB: più controlli sociodemografici a livello di famiglia e villaggio [perché? .....

# Risultati della stima DD

## Difference in Differences Estimate for the Impact of Cycle Program on Girl's Enrollment (Comparing Changes in Enrollment for Girls and Boys in Bihar alone)

	(1)	(2)	(3)	(4)
<b>Treat *Female dummy</b>	<b>0.123***</b> <b>(0.0149)</b>	<b>0.114***</b> <b>(0.0144)</b>	<b>0.0908***</b> <b>(0.0135)</b>	<b>0.0904***</b> <b>(0.0134)</b>
Treat	-0.192*** (0.0108)	-0.184*** (0.0106)	-0.167*** (0.00992)	-0.166*** (0.00992)
Female dummy	-0.186*** (0.0117)	-0.178*** (0.0112)	-0.168*** (0.0103)	-0.167*** (0.0103)
Constant	0.475*** (0.00980)	0.823*** (0.0831)	0.487*** (0.0622)	0.502*** (0.0673)
Demographic Controls	NO	YES	YES	YES
Household Asset & Literacy Controls	NO	NO	YES	YES
Village-Level Controls	NO	NO	NO	YES
Observations	18,453	18,453	18,353	18,331
R-squared	0.038	0.106	0.225	0.227

# TESTIAMO IL PARALLEL TREND

- La validità del DD dipende dall'esistenza di un trend parallelo tra trattati e controlli prima del trattamento.
- Questo significa escludere l'esistenza di un trend nel tasso di enrolment diverso per genere
- Usano dati amministrativi per il periodo 2002-2006 in cui stimano la stessa specificazione del DD, solo che sostituiscono alla dummy trattamento una variabile TREND (chiamata YEAR che prende valore 1 nel 2002, 2 nel 2003, 3 nel 2004 etc)
- Se l'interazione con il genere risulta non significativa, allora esiste un trend parallelo, se è significativa, allora non esiste un trend parallelo.

### Table 3: Testing the Parallel Trend Assumption

Dependent variable: Log (9th Grade Enrollment)

PANEL A: Testing Parallel Trends for the Difference-in-Difference (DD)

Female Dummy×Year	0.0518*** (0.00)
Female Dummy	-0.870*** (0.06)
Year (time trend)	0.0852*** (0.01)
Constant	4.235*** (0.05)
Observations	20,266
R-squared	0.167

Il coefficiente del trend interagito con female (trattate) è statisticamente significativo → Negli anni prima del trattamento le ragazze stavano facendo catching up con i ragazzi nei tassi di enrollment. **E' violata l'ipotesi di trend paralleli.**

# Quindi? DDD

- Come vediamo il DD non è valido, perché non esiste un parallel trend prima del trattamento tra maschi e femmine.
- Che fare?
- **Triple differences in differences!**
- Esiste una differenza tra maschi e femmine che probabilmente ritroviamo anche negli stati vicini che non hanno però la politica
- Introduciamo nelle stime anche almeno 1 di questi stati: il Jharkhand. Tra l'altro era uno stato unico con il Bihar fino al 2001 e sono stati divisi amministrativamente solo in seguito.



# IL DDD

- L'idea di base del DDD è quella di confrontare (=calcolare la differenza) il DD calcolato per il Bihar con quello calcolato esattamente allo stesso modo per lo stato di Jharkhand
- Quindi servono un gruppo di trattati (femmine) e controlli (maschi) nel paese 1, osservati prima e dopo il trattamento (qui osservano due coorti diverse) e lo stesso schema nel paese 2, dove il trattamento non è avvenuto.

# Stima del ddd

$$y_{ihv} = \beta_0 + \beta_1 \cdot F_{ihv} \cdot T_{ihv} \cdot BH + \beta_2 \cdot F_{ihv} \cdot BH + \beta_3 \cdot T_{ihv} \cdot BH + \beta_4 \cdot F_{ihv} \cdot T_{ihv} + \beta_5 \cdot F_{ihv} + \beta_6 \cdot T_{ihv} + \beta_7 \cdot BH_{ihv} + \varepsilon_{ihv} \quad (2)$$

- Dove  $Y_{ihv}$  è l'outcome che corrisponde al ragazzo  $i$ , nella famiglia  $h$  e nel villaggio  $v$ 
  - $F_{ihv}$  è una dummy female (TREAT)
  - $T_{ihv}$  è una dummy che indica la coorte trattata. (POST)
  - $BH$  è la dummy per il Bihar (PAESE TRATTATO)
- A differenza di prima, ora la stima DDD ha l'interazione di tre variabili, e la specificazione comprende tutte le possibili doppie interazioni e le tre variabili dicotomiche (TREAT, POST, PAESE TRATTATO)

# Stime DDD

## Triple Difference Estimate for the Impact of Cycle Program on Girl's Enrollment (Comparing the Double Difference between Bihar and Jharkhand)

	(1)	(2)	(3)	(4)
<b>Treat*Female dummy*Bihar dummy</b>	<b>0.103***</b> <b>(0.0302)</b>	<b>0.0912***</b> <b>(0.0294)</b>	<b>0.0525**</b> <b>(0.0252)</b>	<b>0.0523**</b> <b>(0.0253)</b>
Treat*Female dummy	0.0195 (0.0263)	0.0235 (0.0256)	0.0380* (0.0214)	0.0381* (0.0215)
Treat*Bihar dummy	-0.0437** (0.0179)	-0.0418** (0.0177)	-0.0290* (0.0160)	-0.0281* (0.0161)
Female dummy*Bihar dummy	-0.0942*** (0.0233)	-0.0905*** (0.0226)	-0.0686*** (0.0200)	-0.0673*** (0.0201)
Treat	-0.148*** (0.0143)	-0.143*** (0.0142)	-0.138*** (0.0127)	-0.138*** (0.0127)
Female dummy	-0.0915*** (0.0202)	-0.0880*** (0.0196)	-0.0986*** (0.0172)	-0.0994*** (0.0172)
Bihar dummy	0.0115 (0.0163)	-0.0437*** (0.0165)	-0.0247* (0.0146)	-0.0378** (0.0148)
Constant	0.464*** (0.0130)	0.771*** (0.0240)	0.503*** (0.0240)	0.463*** (0.0393)
Demographic Controls	NO	YES	YES	YES
Household Asset & Literacy Controls	NO	NO	YES	YES
Village-Level Controls	NO	NO	NO	YES
Observations	30,295	30,295	30,147	30,112
R-squared	0.035	0.088	0.208	0.210

# Parallel trend per il DDD

- I due DD hanno un trend parallelo?
- Possiamo stimarlo sostituendo alla dummy treatment un trend e usando i dati di entrambi gli stati.

### PANEL B: Testing Parallel Trends for the Triple Difference (DDD)

Female Dummy×Year×Bihar dummy	-0.0100 (0.01)
Female Dummy×Year	0.0618*** (0.01)
Female Dummy×Bihar dummy	0.175 (0.11)
Bihar dummy×Year	0.0290** (0.01)
Female dummy	-1.045*** (0.09)
Year (time trend)	0.0562*** (0.01)
Bihar dummy	-0.123 (0.12)
Constant	4.358*** (0.11)
Observations	22,279
R-squared	0.171