Australian National University

Impact Evaluation Workshop

IRSA Annual Conference, Solo, 2018

Difference in Difference

Dr. Sarah Dong

Fellow, Indonesia Project

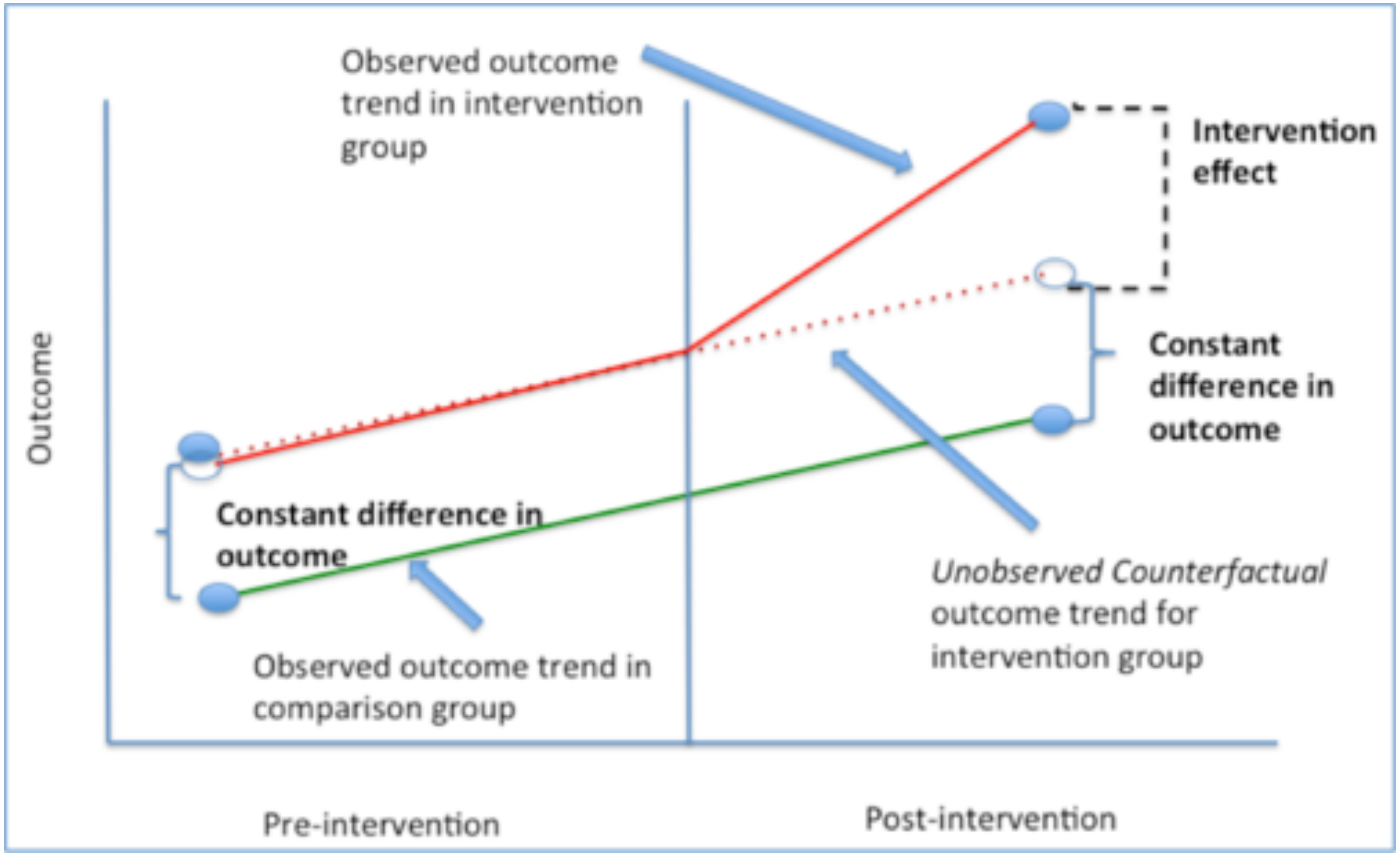
Australian National University

sarah.dong@anu.edu.au

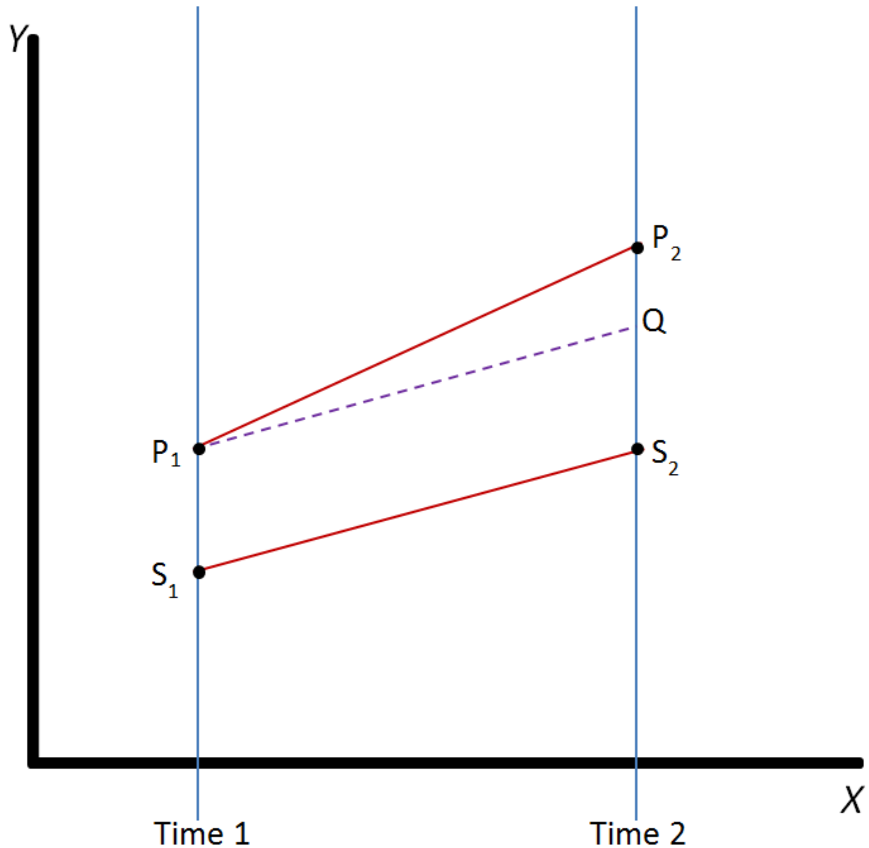
Intuition

- Assume we observe two groups of people for a long time: group A and group B
- They exhibit similar trend in some outcome overtime
- Then a policy is allocated to group A but not to group B
- It is likely that group A will still have the same trend as group B if the policy wasn't introduced
- Therefore the trend in group B is a potential counterfactual to evaluate the effect of the policy

Intuition illustrated using graph



Difference in difference



Treatment effect =

$(P_2 - P_1) - (S_2 - S_1)$

=

$(P_2 – S_2) - (P_1 - S_1)$

=

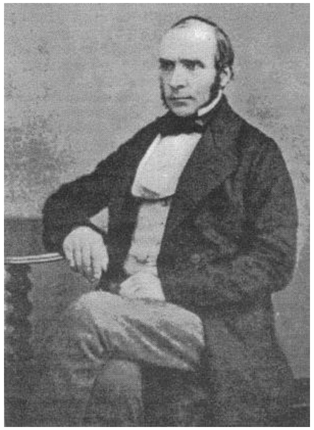
QP_2

Key assumption

- Without the intervention, both groups would have the same trend
  - Fundamentally not testable
  - Usually assessed by looking at the trend pre-program
  - In fact more convincing if have long pre-program data to assess parallel trend assumption

First study that uses DID

- First application DID: John Snow (1855)
  - Cholera London epidemic mid-nineteenth century
  - Prevailing theory: “bad air”
- Snow’s analysis
  - Hypothesis: contaminated drinking water
  - Compared death rates from cholera in districts served by two water companies
  - In 1849 both companies obtained water from the dirty Thames
  - In 1852, one of them moved water works upriver to an area free of sewage
  - Death rates fell sharply in districts served by this water company



DID implementation

Panel:

$$y_{it} = \alpha + \gamma Treat * I(t > t_{treat})_{it} + \beta X_{it} + \lambda_t + \mu_i + \epsilon_{it}$$

where γ is the treatment effect

Cross-section equivalent:

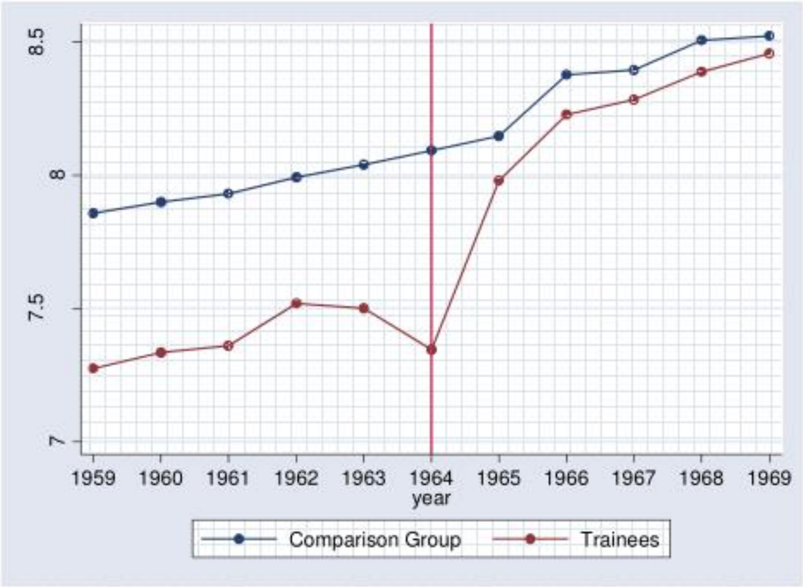
$$y_{\mathrm{idt}} = \alpha + \gamma Treat * I(t > t_{treat})_{dt} + \beta X_{\mathrm{idt}} + \lambda_t + \mu_{\mathrm{d}} + \epsilon_{it}$$

Two group equivalent:

$$y_{\mathrm{it}} = \alpha + \gamma Treat * I(t > t_{treat})_{dt} + \beta X_{it} + \lambda_t + \mu \mathrm{Treat} + \epsilon_{it}$$

Ashenfelter's dip

Earnings for period 1959-69



DID as an evaluation tool

- Widely used
- Doesn’t require stringent assumptions
- Although much more convincing if have pre-program long term trends

Case study 1: Privatization of water

Sebastian Galiani, Paul Gertler, and Ernesto Schargrodsky, "Water for Life: The Impact of the Privatization of Water Services on Child Mortality," Journal of Political Economy 113, no. 1 (February 2005): 83-120.

Background: Privatization of water services in Argentina

- Local governments responsible for water services
  - Supply of clean water
  - Treatment and removal of sewage
- In 1990's 28% of municipalities privatized water services
  - Covering 60% of the population

- Research question:
  - ➢Did privatization of water services in Argentina lead to improved health outcomes?

Privatization of water services in Argentina

TABLE 1
CHANGE IN OWNERSHIP OF WATER SYSTEMS, 1990–99

Ownership	Number of Municipalities	Percentage
Always public	196	39.7
Always private not-for-profit cooperative	143	28.9
Transferred from public to private for-profit	137	27.7
Always private for-profit	1	.2
No service or missing information	17	3.4
Total	494	100.0

NOTE.—In municipalities in which more than one company provides water services, we defined the ownership status of the municipality as the ownership of the company supplying the largest fraction of the population. Source: SPIDES, ENOHSA.

Privatization of water services in Argentina

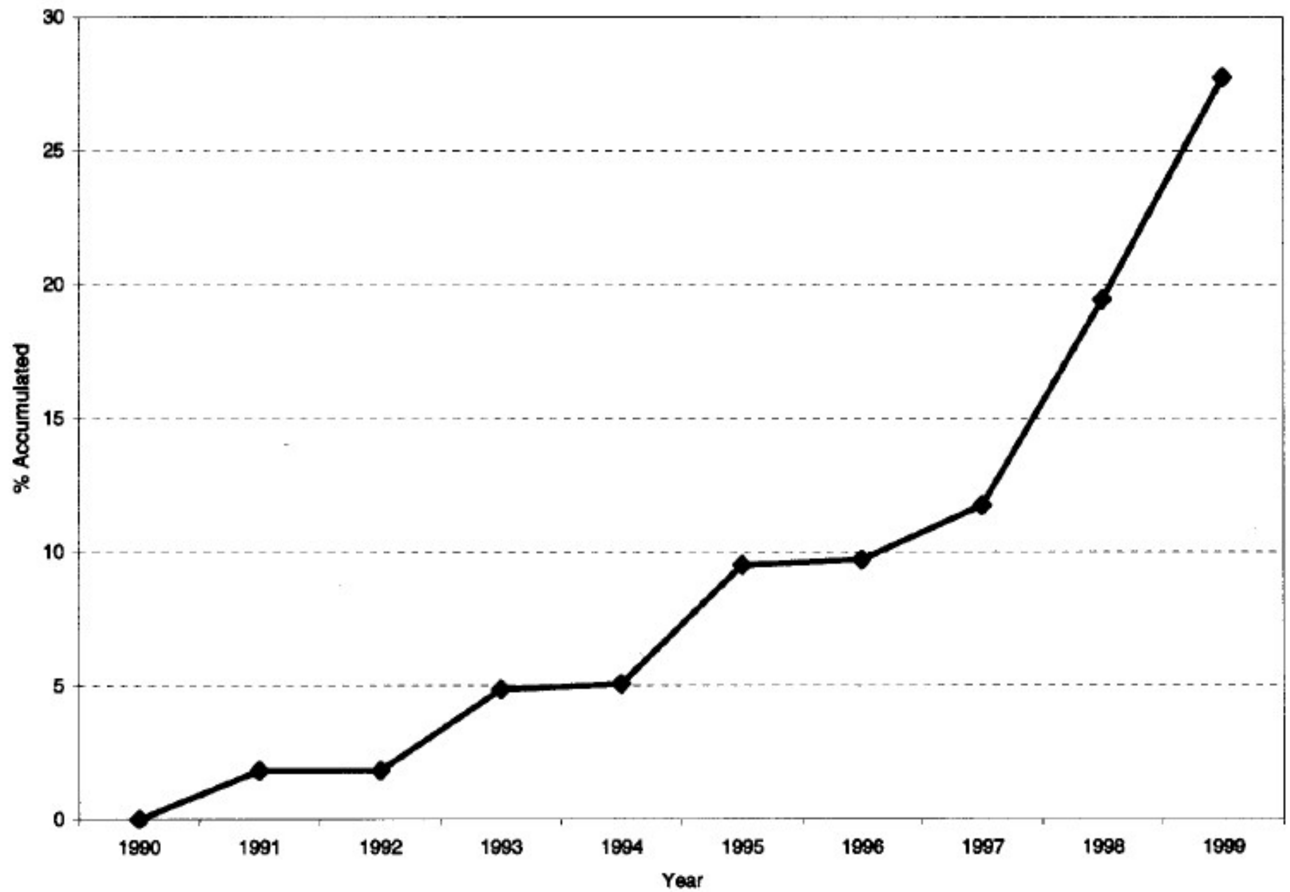


FIG. 2.—Percentage of municipalities with privatized water systems

Evaluation problem

- Privatization of water services is not random
  - ➢It is a choice by local governments!
- Possible motivation for privatization
  - – Poor municipalities with low tax base or underdeveloped infrastructure may be more prone to privatize water services
  - – Privatization driven by economic shocks and recession
  - – Privatization may coincide with other policy reforms

Evaluation design

- Difference-in-difference analysis
- Compare changes in child mortality rates
  - Over time
  - Between municipalities that privatized and those that did not

$$yit = \alpha dIit + \beta Xit + \lambda t + \mu i + \epsilon it$$

- Combine with PSM kernel matching: common support
- Assess scope for bias from time varying factors
  - Assess probability of being privatized
  - Assess whether pre-intervention time trends are similar for treatments and controls

Probability of privatization

TABLE 2
DISCRETE-TIME HAZARD ESTIMATE OF THE PROBABILITY OF BEING PRIVATIZED

	Mean (Standard Deviation) (1)	Model 1 (2)	Model 2 (3)
Time-varying covariates:			
Federal government operates services (=1)	.018	15.975***	16.035***
	(.134)	(2.719)	(2.727)
Local government by Radical party (=1)	.139	−3.198***	−3.204***
	(.346)	(1.067)	(1.067)
Local government by Peronist party (=1)	.719	−.042	−.054
	(.449)	(.401)	(.402)
Δ log GDP per capita$_{t-1}$	.047	4.294	4.259
	(.135)	(3.567)	(3.561)
Δ unemployment rate$_{t-1}$	.006	−6.692	−6.805
	(.029)	(5.696)	(5.711)
Δ income inequality$_{t-1}$	.005	.483	.139
	(.014)	(7.483)	(7.503)
Δ child mortality rate$_{t-1}$	−.266		.034
	(2.994)		(.043)

Australian National University

Fixed pretreatment characteristics as of 1991:			
GDP per capita	60.601	−.022***	−.022***
	(30.388)	(.007)	(.008)
Unemployment rate	.045	12.871**	12.790**
	(.023)	(5.384)	(5.383)
Income inequality	.452	−3.591	−3.469
	(.021)	(5.820)	(5.805)
Child mortality rate	6.208		−.009
	(3.683)		(.036)
Population is 5,000–25,000 (=1)	.419	.227	.225
	(.493)	(.471)	(.480)
Population is 25,000–50,000 (=1)	.202	.106	.110
	(.402)	(.535)	(.540)
Population is 50,000–100,000 (=1)	.114	−.261	−.256
	(.318)	(.605)	(.610)
Population is 100,000–250,000 (=1)	.079	.663	.668
	(.269)	(.612)	(.615)
Population is more than 250,000 (=1)	.066	1.159*	1.151*
	(.249)	(.631)	(.640)
Proportion of families with UBN	.246	−13.660**	−13.328**
	(.151)	(6.067)	(6.226)
Proportion of families living in over-crowded housing	.097	13.560*	13.444*
	(.059)	(7.150)	(7.200)
Proportion of families living in poor housing	.060	6.980**	6.987**
	(.049)	(3.472)	(3.451)
Proportion of families living below subsistence	.036	5.221	4.917
	(.022)	(7.418)	(7.449)
Proportion of houses with no toilet	.095	10.143**	9.798**
	(.117)	(4.429)	(4.563)
No sewerage connection (=1)	.280	−.182	−.171
	(.449)	(.323)	(.328)
Proportion of household heads with more than high school education	.025	−27.242**	−27.182**
	(.012)	(10.971)	(11.003)

TABLE 2
(*Continued*)

	Mean (Standard Deviation) (1)	Model 1 (2)	Model 2 (3)
Mean household head's age between 45 and 52 (=1)	.653	.279	.288
	(.476)	(.343)	(.343)
Mean household head's age above 52 (=1)	.144	.506	.513
	(.351)	(.456)	(.456)
Duration dependence[a]		yes	yes
Observations		2,281	2,281

NOTE.—Standard errors are in parentheses.
* Statistically different from zero at the .1 level.
** Statistically different from zero at the .05 level.
*** Statistically different from zero at the .01 level.
[a] A fifth-order polynomial in time controls for duration dependence. Each coefficient in the polynomial is statistically different from zero at the .1 level.

Trends in child mortality rates

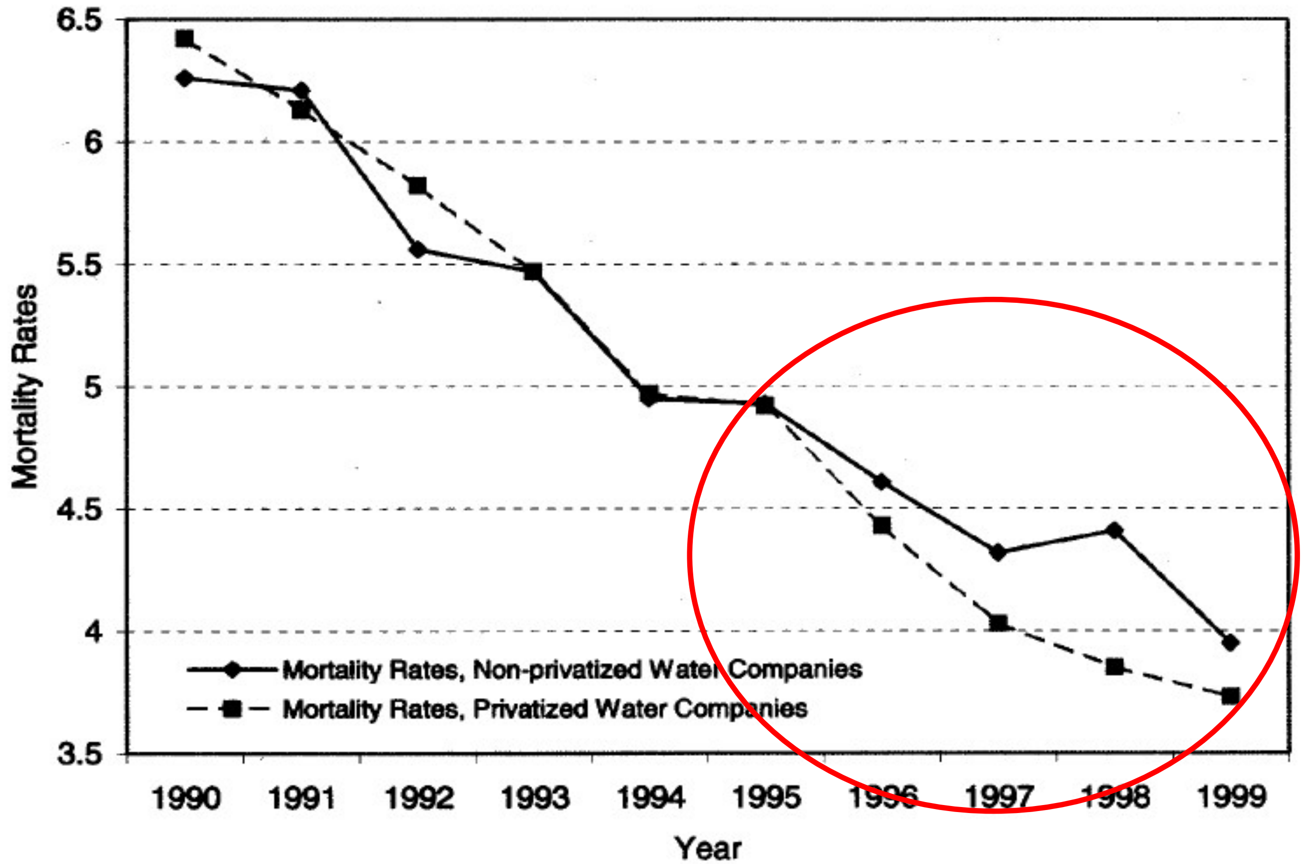


FIG. 1.—Evolution of mortality rates for municipalities with privatized vs. nonprivatized water services

Impact on child mortality

TABLE 3
IMPACT OF PRIVATIZATION OF WATER SERVICES ON CHILD MORTALITY

	FULL SAMPLE			USING OBSERVATIONS ON COMMON SUPPORT			KERNEL MATCHING ON COMMON SUPPORT[a]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Private water services (=1)	−.334	−.320	−.283	−.540	−.541	−.525	−.604
	(.169)**	(.170)*	(.170)*	(.177)***	(.178)***	(.178)***	(.168)***
	[.157]**	[.163]**	[.162]*	[.191]***	[.198]***	[.195]***	
	{.195}*	{.203}	{.194}	{.261}**	{.274}**	{.266}**	
%Δ in mortality rate	−5.3	−5.1	−4.5	−8.6	−8.6	−8.4	−9.7
Other covariates:							
Real GDP per capita		.007	.009		.005	.006	
		(.005)	(.006)		(.006)	(.006)	
		[.006]	[.006]		[.007]	[.007]	
		{.007}	{.007}		{.007}	{.008}	
Unemployment rate		−.555	−.636		−.778	−.836	
		(1.757)	(1.758)		(1.797)	(1.802)	
		[2.161]	[2.166]		[2.249]	[2.263]	
		{2.862}	{2.846}		{2.635}	{2.635}	

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income inequality		5.171	5.085		2.932	3.052	
		(2.868)*	(2.880)*		(2.907)	(2.926)	
		[3.468]	[3.445]		[3.314]	[3.289]	
		{3.696}	{3.691}		{3.833}	{3.838}	
Public spending per capita		−.028	−.035		−.068	−.070	
		(.038)	(.038)		(.039)*	(.039)*	
		[.055]	[.055]		[.059]	[.059]	
		{.054}	{.055}		{.049}	{.050}	
Local government by Radical party (=1)			.482			.166	
			(.267)*			(.284)	
			[.281]*			[.301]	
			{.288}*			{.365}	
Local government by Peronist party (=1)			−.202			−.168	
			(.191)			(.193)	
			[.202]			[.230]	
			{.254}			{.309}	
R^2	.1227	.1256	.1272	.1390	.1415	.1420	
Observations	4,732	4,597	4,597	3,970	3,870	3,870	3,970

NOTE.—Each column reports the estimated coefficients of a separate regression model in which the dependent variable is the child mortality rate, whose mean was 6.25 per thousand in 1990. Standard errors are in parentheses. Standard errors clustered at the municipality level are in brackets. Standard errors clustered at the province-year level are in braces. All the regressions include year and municipality fixed effects. The sample includes the municipalities with always-public, privatized, and nonprofit cooperative water companies (see table 1).

[a] Standard errors for the kernel matching estimate are bootstrapped standard errors using 100 replications.

* Statistically different from zero at the .1 level of significance.

** Statistically different from zero at the .05 level of significance.

*** Statistically different from zero at the .01 level of significance.

Impact by socioeconomic status

TABLE 5
IMPACT OF PRIVATIZATION ON CHILD MORTALITY BY POVERTY LEVEL

	1990 Mean Mortality Rate	Estimated Impact Coefficients	%Δ in Mortality Rate
Nonpoor municipalities	5.07	.114 (.233) [.165] {.159}	...
Poor municipalities	6.97	−1.004 (.279)*** [.297]*** {.278}***	−14.4
Extremely poor municipalities	9.11	−2.415 (.544)*** [1.051]** {.605}***	−26.5

Impact on access to water: non-parametric DID

TABLE 8
DIFFERENCE-IN-DIFFERENCES ESTIMATES OF THE IMPACT OF PRIVATIZATION ON THE PROPORTION OF HOUSEHOLDS CONNECTED TO THE WATER NETWORK, 1991–97

	All Municipalities	Excluding Buenos Aires
	Municipalities That Were Not Privatized before 1997	
Proportion of households connected in 1991 (p_{91}^{public})	.866	.866
Proportion of households connected in 1997 (p_{97}^{public})	.898	.898
Difference 1997 − 1991 ($p_{97}^{public} - p_{91}^{public}$)	.032	.032
	Municipalities That Were Privatized before 1997	
Proportion of households connected in 1991 ($p_{91}^{private}$)	.730	.640
Proportion of households connected in 1997 ($p_{97}^{private}$)	.780	.714
Difference 1997 − 1991 ($p_{97}^{private} - p_{91}^{private}$)	.050	.074
Difference-in-differences ($p_{97}^{private} - p_{91}^{private}$) − ($p_{97}^{public} - p_{91}^{public}$)	.018	.042
Z-test for difference-in-differences estimate[a]	2.83***	5.78***

Case study 2: Salience and taxation

Raj Chetty, Adam Looney, and Kory Kroft (2009) Salience and Taxation: Theory and Evidence, American Economic Review 2009, 99:4, 1145-1177

Evaluation design: RCT+DDD

- Experiment (Intervention)
  - Posting the tax-inclusive price in addition to before-tax price
  - Outcome: quantify sold
- Counterfactuals:
  - In treatment store, have control categories for which no intervention is conducted
    - Assuming the change in demand will be the same between the categories had intervention not occurred
  - In addition have control stores where no intervention is conducted
    - Assuming the difference in change in demand between the categories would be the same across similar stores had the intervention not occurred

EXHIBIT 1. TAX-INCLUSIVE PRICE TAGS

TABLE 3— EFFECT OF POSTING TAX-INCLUSIVE PRICES: *DDD* ANALYSIS OF MEAN QUANTITY SOLD

Period	Control categories	Treated categories	Difference
Panel A. Treatment store			
Baseline (2005:1–2006:6)	26.48	25.17	−1.31
	(0.22)	(0.37)	(0.43)
	[5,510]	[754]	[6,264]
Experiment (2006:8–2006:10)	27.32	23.87	−3.45
	(0.87)	(1.02)	(0.64)
	[285]	[39]	[324]
Difference over time	0.84	−1.30	$DD_{TS} = -2.14$
	(0.75)	(0.92)	(0.68)
	[5,795]	[793]	[6,588]
Panel B. Control stores			
Baseline (2005:1–2006:6)	30.57	27.94	-2.63
	(0.24)	(0.30)	(0.32)
	[11,020]	[1,508]	[12,528]
Experiment (2006:8–2006:10)	30.76	28.19	−2.57
	(0.72)	(1.06)	(1.09)
	[570]	[78]	[648]
Difference over time	0.19	0.25	$DD_{CS} = 0.06$
	(0.64)	(0.92)	(0.95)
	[11,590]	[1,586]	[13,176]
DDD Estimate			−2.20
			(0.59)
			[19,764]

Parametric estimation

$$Y = \alpha + \beta_1 TT + \beta_2 TS + \beta_3 TC + \gamma_1 TT * TC + \gamma_2 TT * TS + \gamma_3 TS * TC + \delta TT * TC * TS + \xi X + \epsilon$$

Where, TT , TS and TC indicates treatment, wrt time, store & category

TABLE 4—EFFECT OF POSTING TAX-INCLUSIVE PRICES: REGRESSION ESTIMATES

Dependent variable	Quantity per category (1)	Revenue per category ($) (2)	Log quantity per category (3)	Quantity per category (4)	Quantity (treat. categories only) (5)
Treatment	−2.20	−13.12	−0.101	−2.27	−1.55
	(0.60)	(4.89)	(0.03)	(0.60)	(0.35)
Average price	−3.15	−3.24		−3.04	−15.06
	(0.26)	(1.74)		(0.25)	(3.55)
Average price squared	0.05	0.06		0.05	1.24
	(0.00)	(0.03)		(0.00)	(0.34)
Log average price			−1.59		
			(0.11)		
Before treatment				−0.21	
				(1.07)	
After treatment				0.20	
				(0.78)	
Category, store, week FEs	x	x	x	x	x
Sample size	19,764	19,764	18,827	21,060	2,379

Stata exercise: tax salience

- Use DID.dta to estimate the effect of increasing salience on outcome variables
  - Weekly quantify sold per category
  - Weekly revenue per category

Steps for exercise

1. Test for differences in outcomes for the treatment and control categories before and after the experiment
   - Separately for both the treatment and control stores
2. Estimate the impact of the experiment on the outcome variables
   - Using DD method and using only treatment stores
3. Expand the analysis by controlling for other factors
4. Estimate the impact of the experiment on the outcome variables
   - Using DDD method with both treatment and control stores
   - Expand the analysis by controlling for other factors

Conclusion: which technique to choose?

- RCT if having a good pilot/theory to test and have large sample and big intervention
- RDD usually preferred over other non-RCT methods as its assumption is least demanding
- Next is DID with long-run pre-program data
- IV is very hard to find: unless have a really good IV, think twice
- Matching is usually last resort, should match using a large set of chars. More convincing if combined with DID

Last exercise: remember the assumptions for causal identification

- RCT?
- Matching?
- RDD?
- IV?
- DID?