Quasi-Experiments for Causal Evidence

Sebastian Galiani

Washington University in St. Louis

Identifying Causal Relations

- **Identifying Causal Relations in Social Sciences is not straightforward.**
  - **There are confounders.**
  - **Imply "*What if…*" questions.**

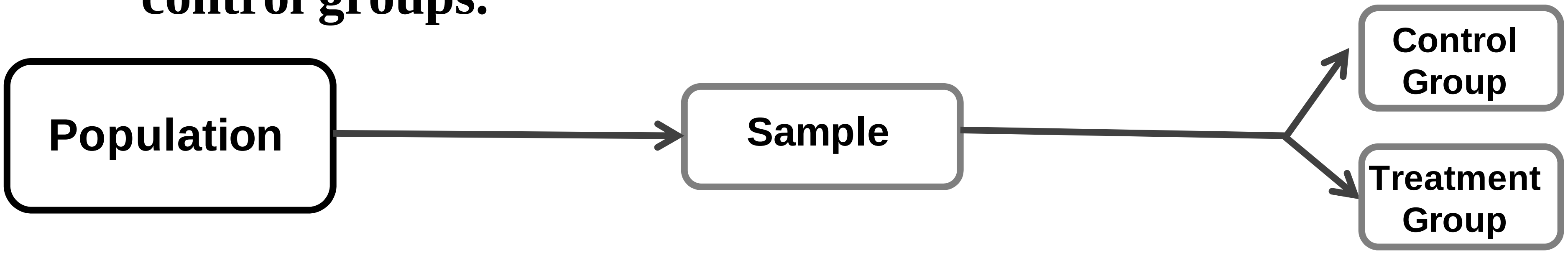
    …"what would have happened if another course of action was chosen…"

Dealing with Counterfactuals

Fundamental Problem of Causal Inference

Dealing with Counterfactuals

- Consider the evaluation of an intervention. How do we deal with Counterfactuals in practice?
- We select from a sample of the population: treatment and control groups.



•So, the question is: choosing any control and treatment group would allow us estimate the treatment effect?

•How should assignment into treatment status be, so as to find an estimator of the causal parameter of interest?

Dealing with Counterfactuals

- **The control group should play the role of the treated units in the absence of treatment.**
- **But… if selection into treatment is related to observables and unobservables individual characteristics, we cannot assume that the groups are on average similar!**

- **For example, self-selection into job-training program. Mean difference of participants and non-participants earnings estimates ATE?**

- **Probably those participating had on average lower earnings than those who do not participate.**

Randomization

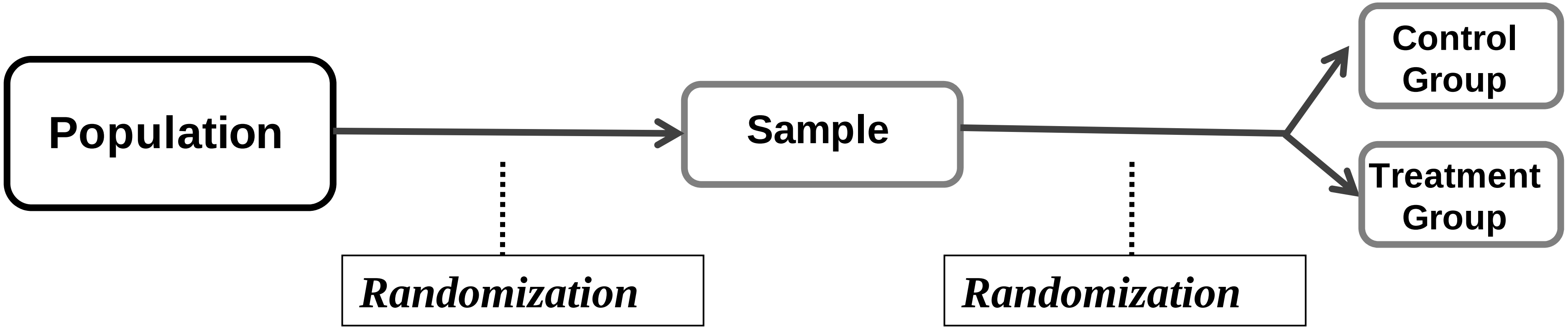
- **If the researcher assigns the subjects to the groups at random or by chance, the two groups will be on average balanced with respect to all observable and unobservable factors other than treatment.**

- **In principle, randomized trials ensure that outcomes in the control group really do capture the counterfactual for a treatment group.**

- **Random assignment is achieved by any procedure that assigns units to conditions based only on chance (toss of a coin, random numbers), in which each unit has the same nonzero probability of being assigned to a condition.**

Two-Stage Randomization

- **Ideally randomization must be performed in the two stages:**



1st Stage:

ensures that the results in the sample will represent the results in the population within a defined level of sampling error

External Validity

2nd Stage:

ensures that the observed effect on the dependent variable is due to some aspect of the treatment rather than other confounding factors

Internal Validity

Estimating ATE under Randomization

- **Under two-stage randomization we know that:**

$$[\overline{\mathrm{Y}}_1 \mid \mathrm{D}=1]=[\overline{\mathrm{Y}}_1 \mid \mathrm{D}=0] \text{ and } [\overline{\mathrm{Y}}_0 \mid \mathrm{D}=1]=[\overline{\mathrm{Y}}_0 \mid \mathrm{D}=0]$$

- **Remember that ATE is:**

$$ATE = \overline{\delta} = \overline{\mathrm{Y}}_1 - \overline{\mathrm{Y}}_0$$

- **So, the following estimator, consistently estimates ATE:**

$$\hat{\overline{\delta}} = [\hat{\overline{\mathrm{Y}}}_1 \mid \mathrm{D}=1] - [\hat{\overline{\mathrm{Y}}}_0 \mid \mathrm{D}=0]$$

Estimating ATE under general assignment rules

$$E\left[\hat{\bar{\delta}}\right] = \bar{\delta} + \underbrace{\left(\left[\bar{Y}_1 \mid D=1\right] - \left[\bar{Y}_0 \mid D=0\right]\right)}_{\text{Selection Bias: Baseline Difference}} + \underbrace{(1-\pi)\,(\bar{\delta}_{\{D=1\}} - \bar{\delta}_{\{D=0\}})}_{\text{Treatment Heterogeneity}} = \bar{\delta} + 0 + 0 = \bar{\delta}$$

Estimating ATE under Randomization

- **Furthermore, notice that under two-stage randomization ATE is equal to TOT:**

$$TOT = [\overline{Y}_1 | D = 1] - [\overline{Y}_0 | D = 1] = ATE$$

- **But if randomization takes place only in 2nd stage –on a selected sub-population-, the estimator:**

$$\hat{\overline{\delta}} = [\hat{\overline{Y}}_1 \mid D = 1] - [\hat{\overline{Y}}_0 \mid D = 0]$$

Only estimates TOT consistently.

Quasi-experiments

- I do not believe that reduced-form evaluation needs only to rely on experimental designs. Randomized evaluations are preferred (even with the caveats that has been extensively discussed in the literature).

- However, I am still sympathetic to quasi-experimentation. Here, the question is to explain plausible rival hypothesis. If possible, one should not restrict oneself to a single research design or approach when trying to assess the impact of a program.

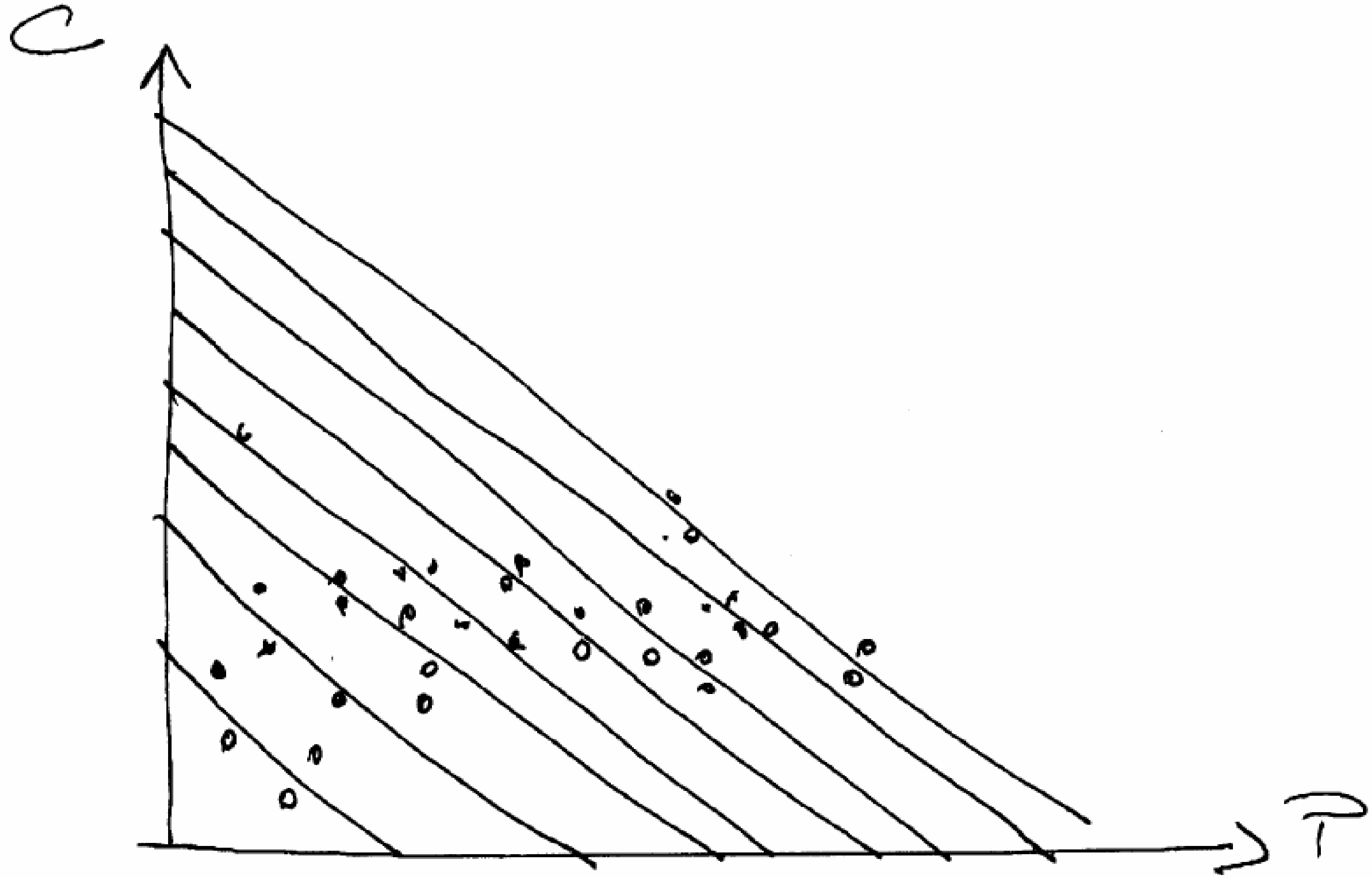
- Any and all methods that help to control or explain plausible rival hypothesis are useful.

Quasi-experiments

- When constructing a causal hypothesis one should envisage as many different consequences of its truth as possible (falsification tests (e.g., Duflo (2001), Angrist and Kruger (1992) and Berlinski, Galiani and Gertler (2006)).

- Also, rely on Multimethods when it is possible –Lavy (2002), Berlinski and Galiani (2007) and Berlinski, Galiani, McEwan and Shapiro (2007).

Police & Crime



Difference in Differences

- Suppose that the process that determines crime is:

$$C_{it} = \alpha_0 + \alpha_1 P_{it} + \lambda_t + \mu_i + \varepsilon_{it}$$

- Where $\beta_t = \alpha_0 + \lambda_t$ and $u_{it} = \mu_i + \varepsilon_{it}$ for t = 0, 1. $E(\varepsilon_{it}) = 0$ for all i and t.

- And where:

$$P_{it} = \begin{cases} P_{i0} = 0 \ \forall \ i \\ P_{i1} \begin{cases} = 0 \text{ for } i = 1,..,C \\ = 1 \text{ for } i = 1,...,T \end{cases} \end{cases}$$

Difference in Differences

- The change in crime for treated units is:

$$\Delta C_{it} = \alpha_1 + \Delta\lambda_{it} + \Delta\varepsilon_{it}$$

- While its mean is:

$$E(\Delta C_{it} \mid \Delta P_{it} = 1) = \alpha_1 + \Delta\lambda_{it} + E(\Delta\varepsilon_{it} \mid \Delta P_{it} = 1) \quad (1)$$

- We now assume that:

$$E(\Delta\varepsilon_{it} \mid \Delta P_{it} = 1) = 0$$

Difference in Differences

- Which also implies:

$$E(\Delta\varepsilon_{it} \mid \Delta P_{it} = 0) = 0$$

- Note that this average before-after comparison is not an unbiased estimate of TOT

$$E(\Delta C_{it} \mid \Delta P_{it} = 1) = \alpha_1 + \Delta\lambda_{it}$$

Difference in Differences

- The change in crime for untreated units is:

$$\Delta C_{it} = \Delta\lambda_{it} + \Delta\varepsilon_{it}$$

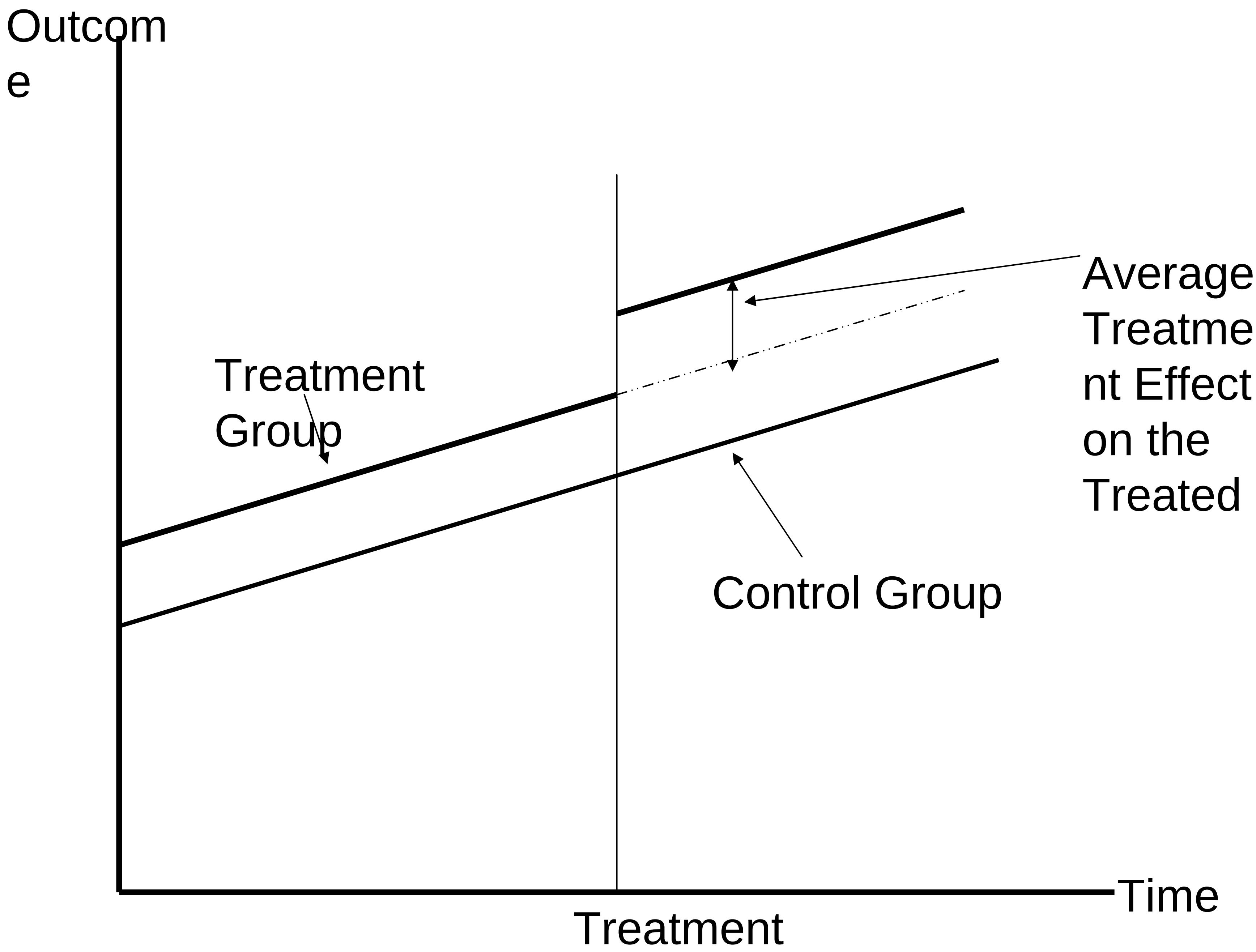
- While its mean is:

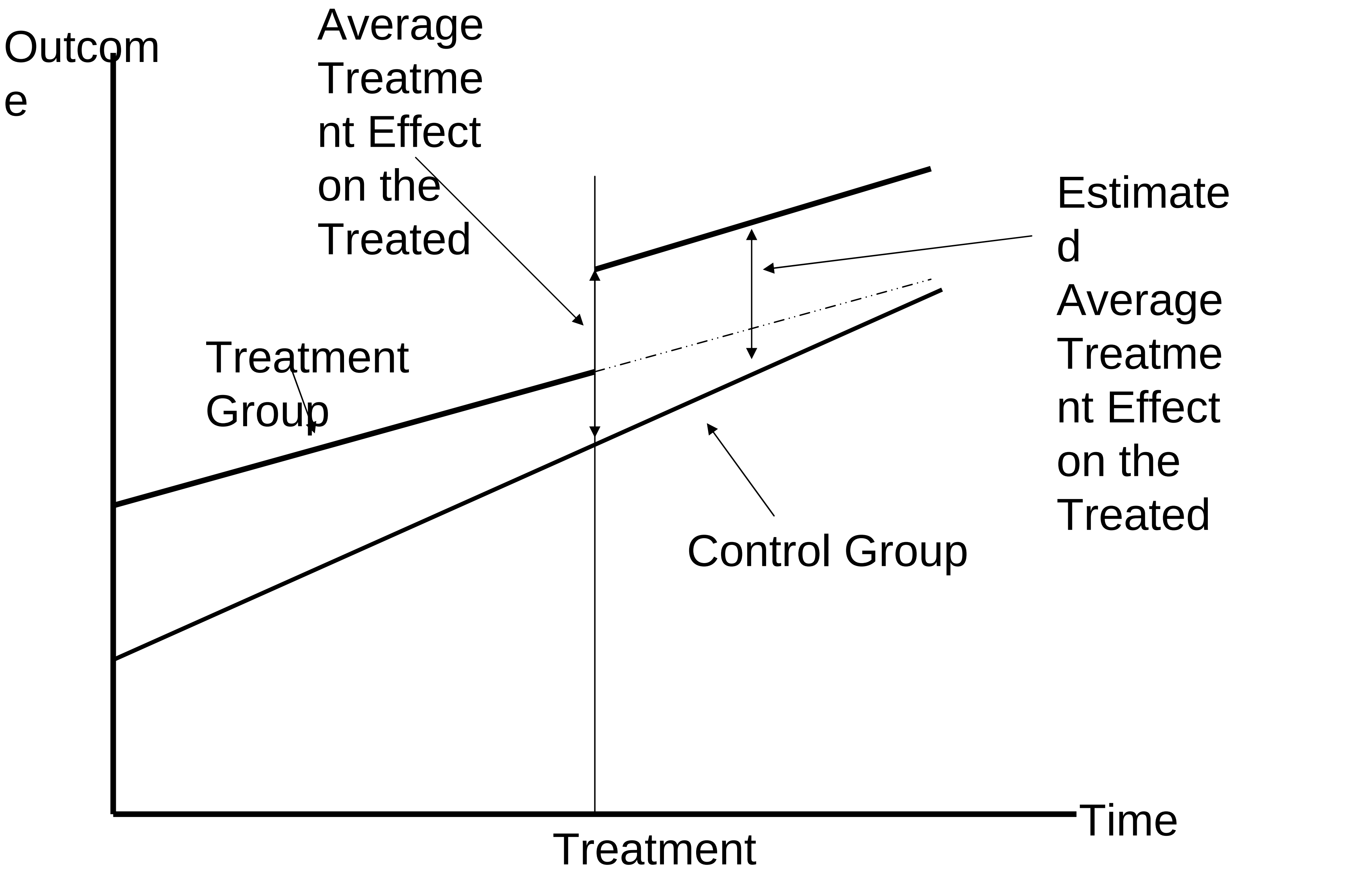
$$E(\Delta C_{it} \mid \Delta P_{it} = 1) = \Delta\lambda_{it} \quad (2)$$

- Then, the difference of the differences (1) – (2) is a consistent estimator of α_1.

Difference in Differences

$$E(\Delta C_{it} \mid \Delta P_{it} = 1) - E(\Delta C_{it} \mid \Delta P_{it} = 0) = \alpha_1$$


Outcome
Treatment Group
Average Treatment Effect on the Treated
Control Group
Time
Treatment


Outcome
Average Treatment Effect on the Treated
Estimated Average Treatment Effect on the Treated
Treatment Group
Control Group
Time
Treatment

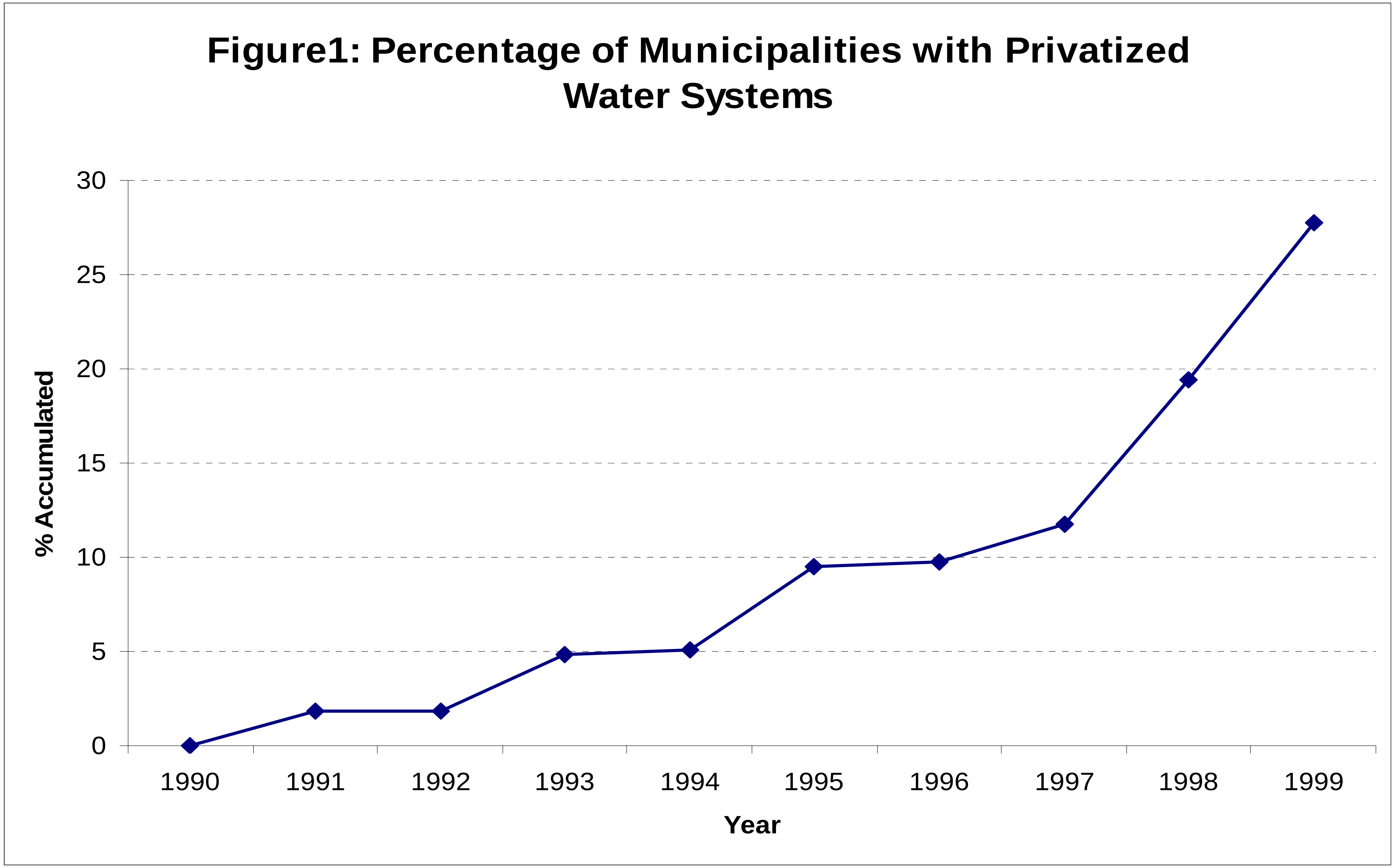


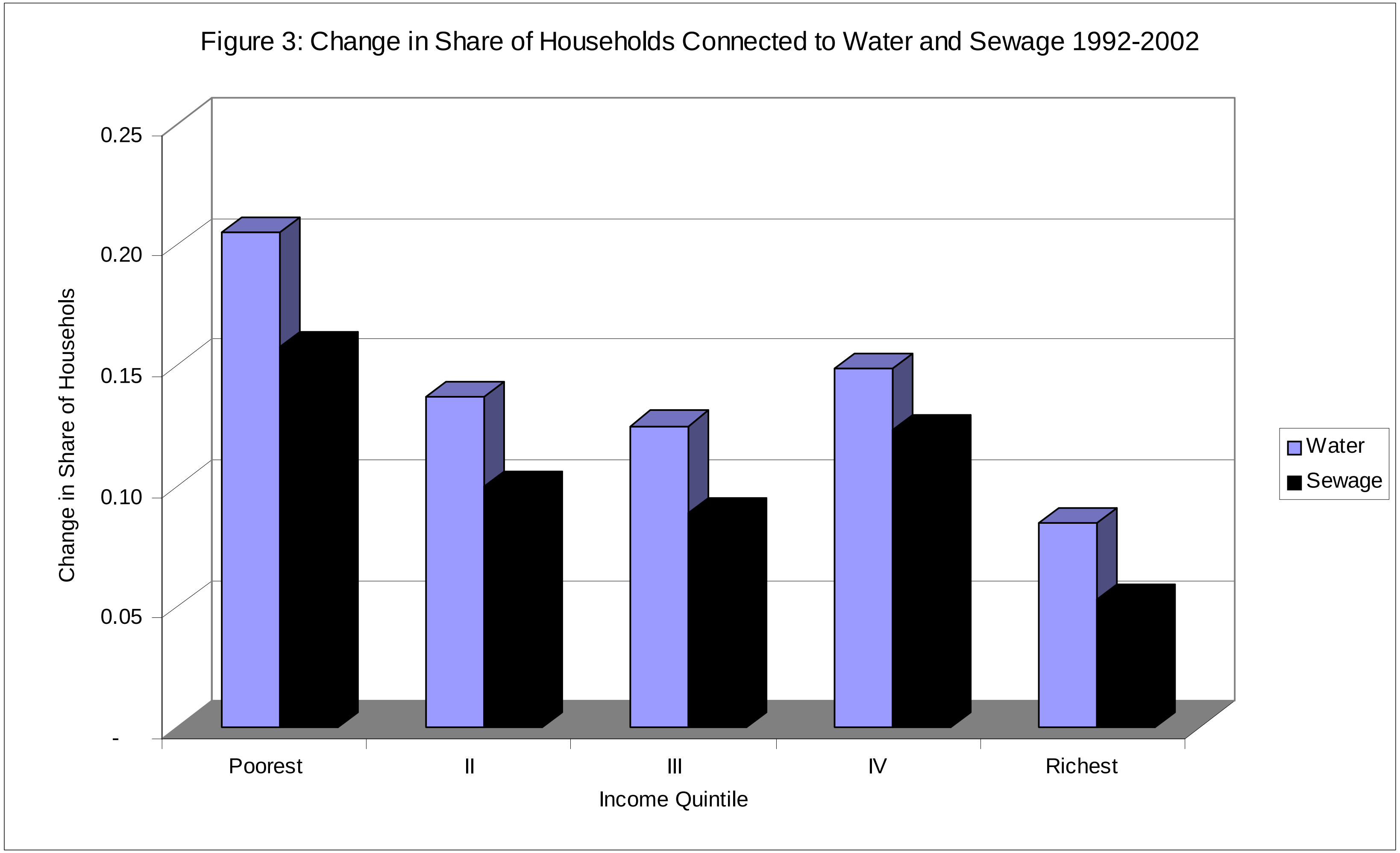
Figure1: Percentage of Municipalities with Privatized Water Systems
30
25
20
15
10
5
0
% Accumulated
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
Year

Use Discrete Time Hazard to estimate determinants of privatization…Find

- Political party controlling municipal government
  - Federal, Peronist and Provincial more likely to privatize
  - Radical less likely to privatize

- Does not depend on
  - Lagged changes in income, unemployment, inequality
  - Lagged changes in mortality rates

World Bank Survey 2000 shows access expanded most among poor 1993-2000



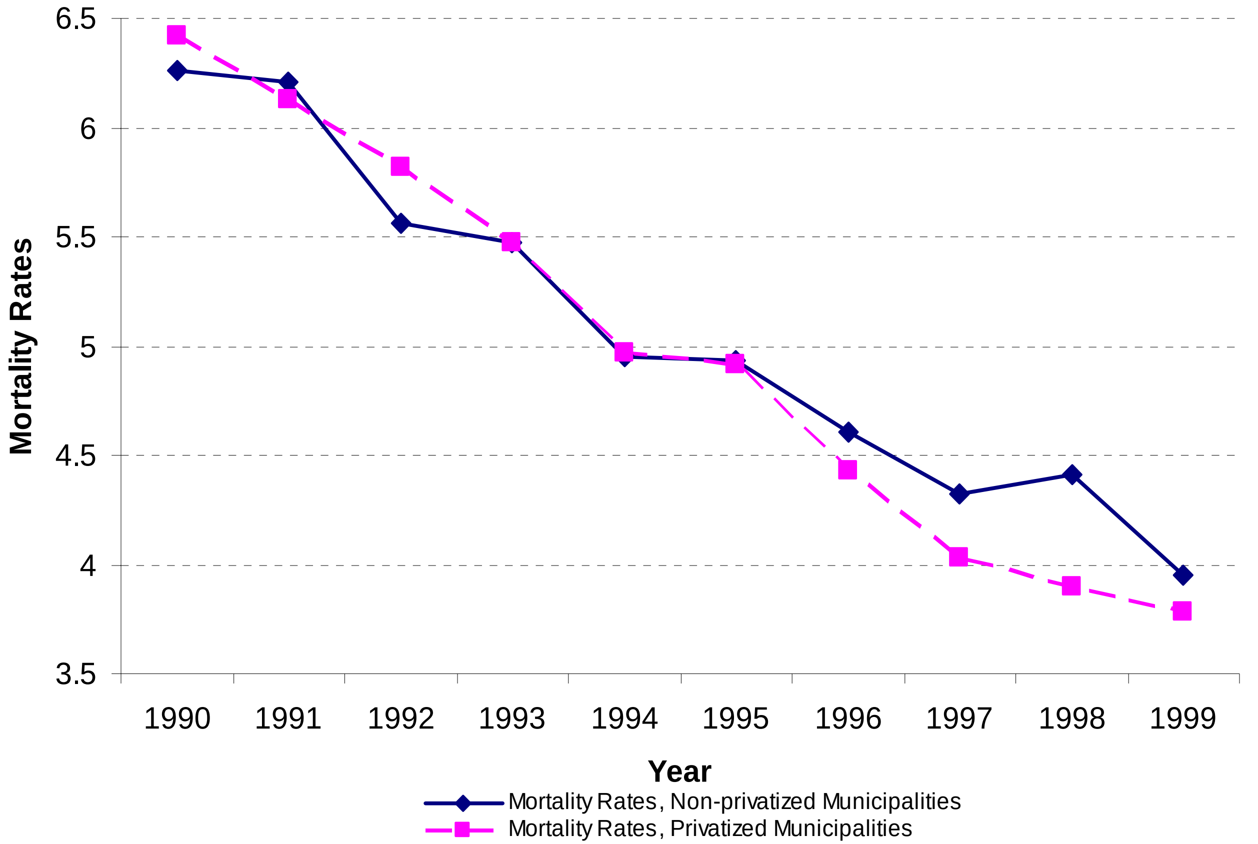
Actual D in D Implementation

- Difference in Difference in regression form

$$y_{it} = \alpha \, dI_{it} + \beta \, \boldsymbol{x}_{it} + \lambda_t + \mu_i + \varepsilon_{it}$$

- Test using only pre-intervention years that time trend in controls is same as time trend in treatments
  - Cannot reject the hypothesis of same trends between treatments and controls.

Figure 4: Evolution of Mortality Rates for Municipalities with Privatized vs. Non-Privatized Water Services



<u>Dif-in-Dif Estimates:</u> Privatization Significantly Reduced Child Mortality

	Full Sample			Common Support			Matched
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Private Water (=1)	- 0.33 **	- 0.32 *	- 0.29 *	- 0.54 ***	- 0.54 ***	- 0.53 ***	- 0.60 ***
% Δ in Mortality	*- 5.3 %*	*- 5.1 %*	*- 4.5 %*	*- 8.6 %*	*- 8.6 %*	*- 8.4 %*	*- 10.0 %*
Real GDP/Capita		0.01	0.01		0.01	0.01	
Unemployment Rate		- 0.56	-0.64		-0.78	-0.84	
Inequality (Gini)		5.17 *	5.09 *		3.05	3.05	
Public Spending/Cap		- 0.03	- 0.04		-0.07 *	- 0.07 *	
Radical Party (=1)			0.48 *			0.17	
Peronist Party (=1)			- 0.20			- 0.17	
F-Stat Municipal FE	13.84***	11.92***	11.51***	10.39***	8.65***	8.32***	
F-Stat for year FE	55.03***	19.88***	18.25***	52.25***	15.59***	12.98***	

Conclusions: Using a combination of methods find that …

- Privatization of water services is associated with a reduction in child mortality of 5 to 7 percent

- The reduction in mortality is from
  - a drop in deaths caused by infectious/parasitic diseases,
  - not from causes unrelated to water

- Most of the reduction in mortality occurred in low-income areas

Contrary to the concerns about negative health effects & worsening inequality

Our evidence suggests that…

1. The deterioration in Argentine water systems under public management was so large that privatization
   - generated profits, attracted investments, expanded service, and
   - reduced child mortality.
2. While private sector may provide sub-optimal services, it does a better job and the poor are benefiting from it

Effect on Mortality by Income Group

- Find higher impact in poor municipalities
  - Most of increase in access was in lower income groups
  - High-income groups already had a high rate of connection to the water network prior to privatization.
  - Even when service quality was unsatisfactory, high income sectors enjoyed better access to substitutes
    - e.g. pumped wells, septic tanks, or bottled water
- Again rules out many other explanations of results