



THE WORLD BANK



技术专题 III 倍差法和面板数据

Christel Vermeersch
中国北京, 2009

内容框架

- 何时使用倍差法 (Diff-in-diff 或 DD)?
- 评价策略: 3种倍差法的评价方法
- 实例:
 - 印度尼西亚教育服务的扩展
 - 生活用水 (阿根廷)

何时使用倍差法？

- 我们不能总是随机抽取受益者.....
 - 例如：评价以往项目的影响
- 通常，我们需要确认
 - 哪组受到政策变化的影响（“干预”）
 - 哪组不受影响（“对照”）
- 我们试图找到能够确定政策影响的“自然实验”
 - 例如：政策中某一个不可预见的变化
 - 例如：一项只影响16岁而非15岁人群的政策
- 对照组的质量决定了评价的质量

倍差法的3种评价方法

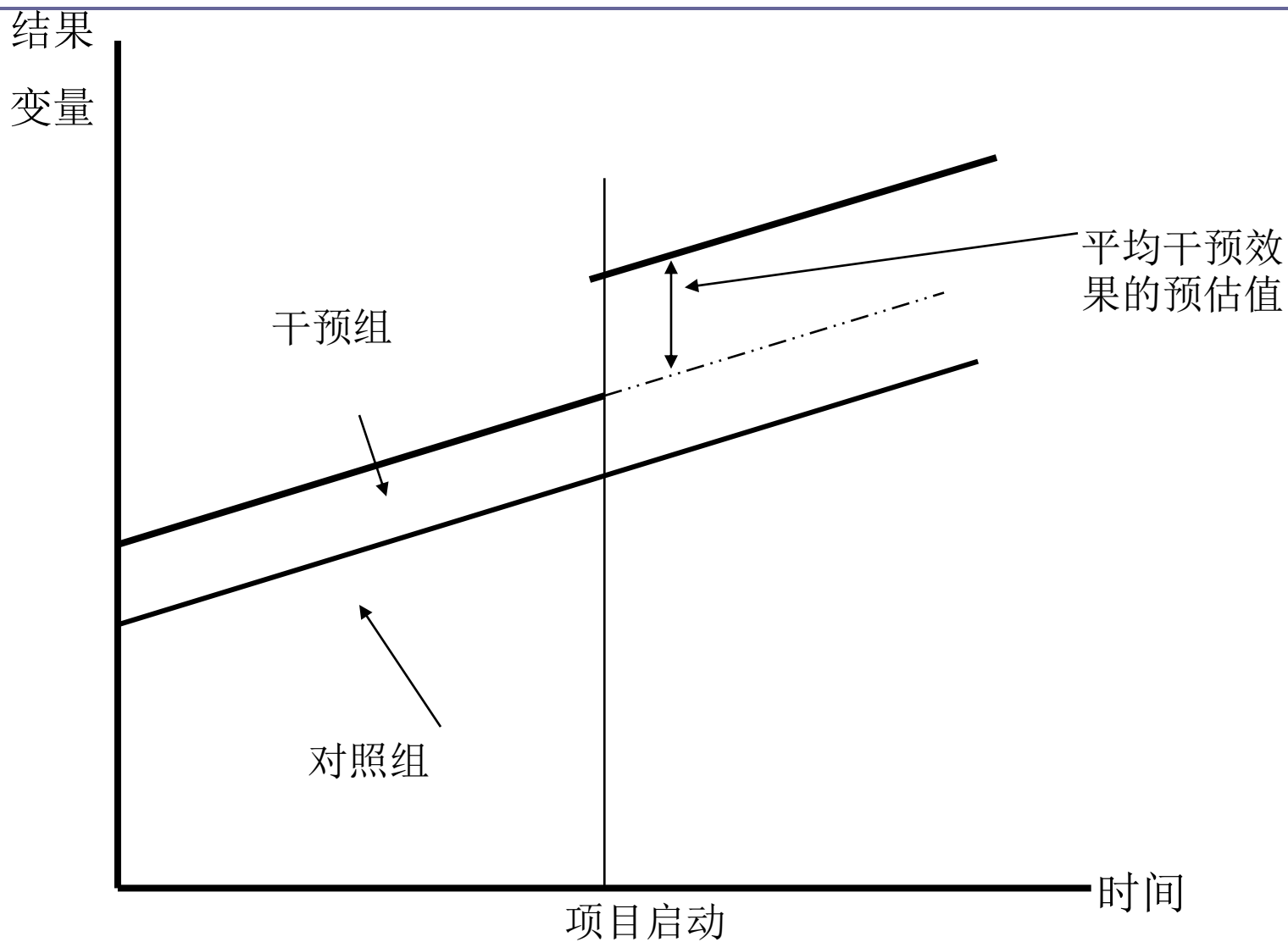
- 框格法
- 绘图法
- 回归法

框格法

	受政策变化 影响组 (干预)	不受政策变化 影响组 (对照)
项目启动 之后	$Y_1(u_i) D_i = 1$	$Y_1(u_i) D_i = 0$
项目启动 之前	$Y_0(u_i) D_i = 1$	$Y_0(u_i) D_i = 0$
	$(\bar{Y}_1 D = 1)$ $-(\bar{Y}_0 D = 1)$	$(\bar{Y}_1 D = 0)$ $-(\bar{Y}_0 D = 0)$

$$DD = ((\bar{Y}_1 | D = 1) - (\bar{Y}_0 | D = 1)) - ((\bar{Y}_1 | D = 0) - (\bar{Y}_0 | D = 0))$$

绘图法



回归法 (2个时间周期)

$$Y_{it} = \alpha + \beta.1(t = 1) + \gamma.1(D_i = 1) + \delta.(t = 1).1(D_i = 1) + \varepsilon_{it}$$

⇓

$$E(Y_{i1} | D_i = 1) = ???$$

$$E(Y_{i0} | D_i = 1) = ???$$

$$E(Y_{i1} | D_i = 0) = ???$$

$$E(Y_{i0} | D_i = 0) = ???$$

⇓

$$\begin{aligned} DD &= (E(Y_{i1} | D_i = 1) - E(Y_{i0} | D_i = 1)) - (E(Y_{i1} | D_i = 0) - E(Y_{i0} | D_i = 0)) \\ &= ??? \end{aligned}$$

回归法 (2 个时间周期)

$$Y_{it} = \alpha + \beta.1(t=1) + \gamma.1(D_i = 1) + \delta.1(t=1).1(D_i = 1) + \varepsilon_{it}$$

⇓

$$E(Y_{i1} | D_i = 1) = \alpha + \beta.1 + \gamma.1 + \delta.1.1 + E(\varepsilon_{i1} | D_i = 1) = \alpha + \beta + \gamma + \eta$$

$$E(Y_{i0} | D_i = 1) = \alpha + \beta.0 + \gamma.1 + \delta.0.1 + E(\varepsilon_{i0} | D_i = 1) = \alpha + \gamma$$

$$E(Y_{i1} | D_i = 0) = \alpha + \beta.1 + \gamma.0 + \delta.1.0 + E(\varepsilon_{i1} | D_i = 0) = \alpha + \beta$$

$$E(Y_{i0} | D_i = 0) = \alpha + \beta.0 + \gamma.0 + \delta.0.0 + E(\varepsilon_{i0} | D_i = 0) = \alpha$$

⇓

$$DD = (E(Y_{i1} | D_i = 1) - E(Y_{i0} | D_i = 1)) - (E(Y_{i1} | D_i = 0) - E(Y_{i0} | D_i = 0))$$

$$= (\beta + \delta) - \beta$$

$$= \delta$$

当有2个以上的时间周期/组

我们对时间和分组采用固定效应回归分析：

$$Y_{it} = \alpha + \sum_{\tau=1}^T \beta_{\tau} \cdot 1(t = \tau) + \sum_{l=1}^I \gamma_l \cdot 1(i = l) + \delta \cdot D_{it} + \varepsilon_{it}$$

D_{it} 为第 t 时间段，第 i 组的干预 D 的强度

倍差法的确认

- 干预效果的确定基于组间跨期的变异程度。
- 即: 经过一段时间后, 结果变量 Y 的变化, 尤其是针对干预组
- 即: 干预开始时仅干预组(而非对照组)的结果变量发生变化

注意...

- 倍差法/ 固定效应控制：
 - 固定的组别效应
 - 如： 拥有土地的农民, 无地农民
 - 在一特定时间点，对所有的组有共同效应，或“共同趋势”
 - 如： 所有的农民，不论其是否拥有土地，都受到了2006年干旱的影响
- 只有当政策变化对结果变量产生即时效应时才有效。
 - 如果政策变化的影响延迟，我们则需要使用滞后干预变量。

注意...

- 倍差法是将干预发生时干预组和对照组趋势的差异归因于该干预。
- 若有其他因素影响两组之间的趋势差异，评价结果便会产生偏倚！

倍差法的质量控制 ...

- 实行“安慰剂”倍差法,即利用一个“假的”干预组
 - 延伸至前几年(如: 前年, 去年).
 - 或用一个干预组, 已知该人群不受影响
 - 若倍差法的估计值不等于0, 趋势并不相似, 那么我们最初的倍差法很可能有偏倚。
- 使用不同的对照组
 - 两次倍差法的估计值应相同
- 使用结果变量 Y ~ 已知该变量不受干预的影响
 - 使用相同的对照组和干预年份
 - 若倍差法估计值不为0, 就有问题

倍差法的常见问题

- 干预实施前结果的差异是参与的基础
 - 项目前评估(“Ashenfelter dip”)
- 功能相依
- 当应答的大小以非线性的方式依存于干预的大小时，我们比较高强度干预组与低强度干预组
- 当时间单位内/组内的观察呈相关性。



THE WORLD BANK



实例 #1: 学校教育和劳动力市场 印度尼西亚学校建设的结果: 一项独特的政策试验实证

Esther Duflo (MIT)
美国经济评论 2001年9月

研究问题

- 学校基础设施→教学成绩？
- 教学成绩 → 薪酬水平？
- 学校教育的经济收益是什么？

项目描述

- 1973-1978: 印度尼西亚政府建造了61000所学校（相当于每所学校有500名5至14岁的儿童）
- 1973-1978年间入学率由69%增至85%
- 每个地区内学校的数量取决于1972年，即该项目启动之前本地区校外儿童的数量。

确定干预效果

针对每一个特定的个体，其项目强度有2个变异来源：

- 就地区而言：每个地区学校数量的变异
- 就年龄而言：
 - 1972年时12岁以上的儿童未从该项目中受益。
 - 1972年时，年龄越小的儿童受益越多 – 因为她在新学校就学的时间更长

数据来源

- 1995 年人口普查: 个体水平的数据包括:
 - 出生日期
 - 1995年工资水平
 - 1995年教育水平
- 样本中每个人的出生地中, 建设项目的的作用强度
- 样本: 1950 - 1972 年间出生的男性

最初的效果评价...

第一步: 让我们简化问题, 评价项目的效果。

- 简单评价项目的强度: 高或低
- 简化进行儿童分组:
 - “年幼组” 受益的儿童队列
 - “年长组” 未受益的儿童队列

结果变量的平均值

建筑项目的
强度

		高	低	
1974年 时的 年龄	2-6岁 (年幼组)	8.49	9.76	
	12-17岁 (年长组)	8.02	9.4	
差异		0.47	0.36	0.12 DD (0.089)

结果变量的平均值

		建筑项目的强度		差异
		高	低	
1974年 时的 年龄	2-6岁 (年幼组)	8.49	9.76	-1.27
	12-17岁 (年长组)	8.02	9.4	-1.39
				0.12 DD (0.089)

倍差法的安慰剂组 (比较798页, 表 3, B组)

想法: 选择2个未受益的组, 计算DD值, 并检验估计效应值是否为0。若不是0, 就有问题....

		建筑项目的强度		
		高	低	
1974年 时的 年龄	12-17岁	8.02	9.40	
	18-24岁	7.70	9.12	
差异		0.32	0.28	0.034 DD (0.098)

第二步: 应用回归法估算

$$S_{ijk} = c + \alpha_j + \beta_k + \gamma \cdot (P_j \cdot T_i) + \delta \cdot (C_j \cdot T_i) + \varepsilon_{ijk}$$

with

S_{ijk} = 在k队列中j地区的i人的教育水平

P_j = 若在该地出生的人受到高强度项目影响, 则为1。

T_i = 若此人属于年幼组, 则为1

C_j = j地区的哑变量

ε_{ijk} = i, j, k 人的误差项

第三步：使用补充信息

我们在每个地区使用项目作用强度：

$$S_{ijk} = c + \alpha_j + \beta_k + \gamma \cdot (P_j \cdot T_i) + \delta \cdot (C_j \cdot T_i) + \varepsilon_{ijk}$$

其中： P_j = j 地区建筑活动的强度

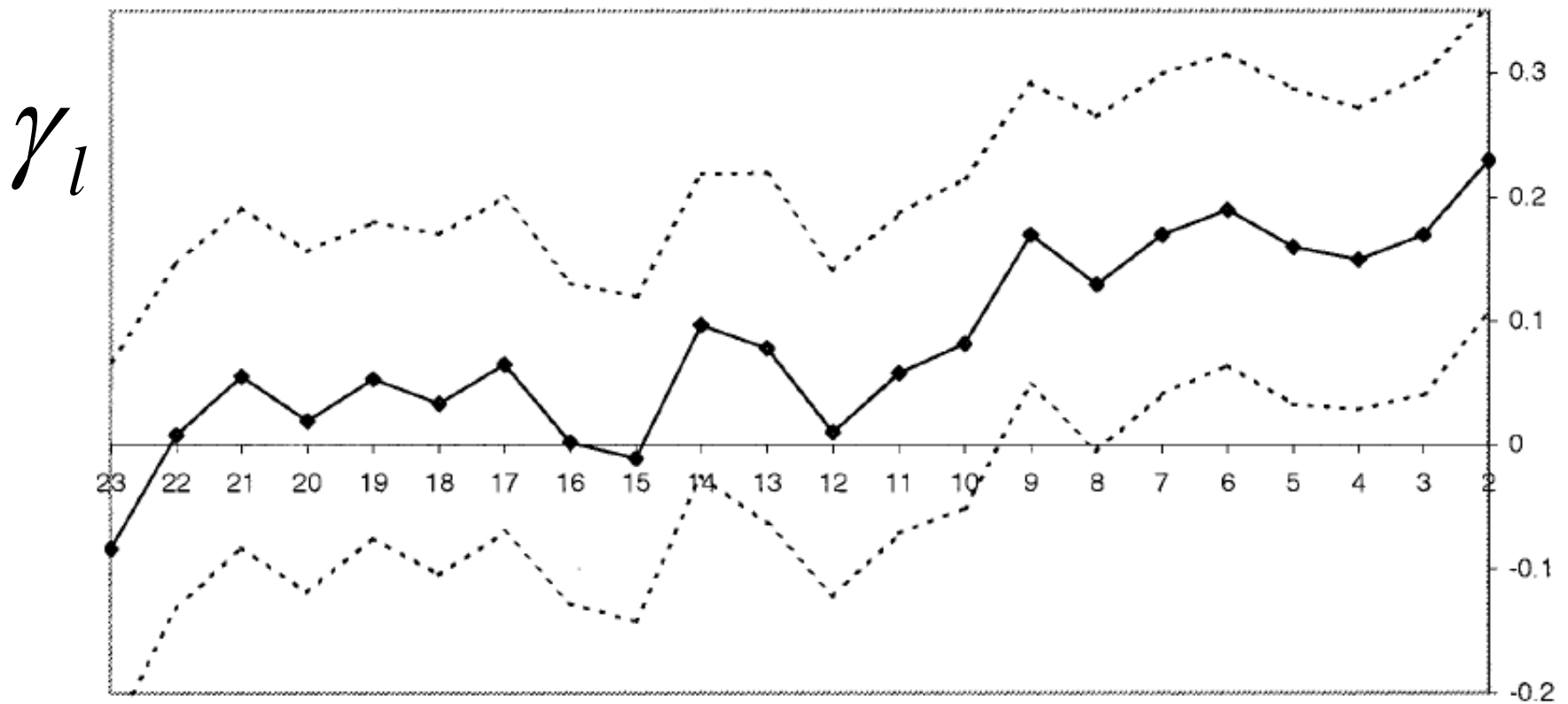
C_j = 地区特征向量

我们分别估算每一队列受到项目的影响：

$$S_{ijk} = c + \alpha_j + \beta_k + \sum_{l=2}^{23} \gamma_l \cdot (P_j \cdot d_i) + \sum_{l=2}^{23} \delta_l C_j T_i + \varepsilon_{ijk}$$

其中： d_i = i 队列中的哑变量

每个队列的作用效果



1974年时的年龄

Y = 因变量 = 工资

	记录（工资）		
	出生地区的项目水平		
	高	低	差异
	(4)	(5)	(6)
A组：实验组			
在1974年年龄为2-6岁	6.61 (0.0078)	6.73 (0.0064)	-0.12 (0.010)
在1974年年龄为12-17岁	6.87 (0.0085)	7.02 (0.0069)	-0.15 (0.011)
差异	-0.26 (0.011)	-0.29 (0.0096)	0.026 (0.015)
B组：对照组			
在1974年年龄为12-17岁	6.87 (0.0085)	7.02 (0.0069)	-0.15 (0.011)
在1974年年龄为18-24岁	6.92 (0.0097)	7.08 (0.0076)	-0.16 (0.012)
差异	0.056 (0.013)	0.063 (0.010)	0.0070 (0.016)

结论

- 结果: 对每千名学生兴建的每所学校
 - 平均教学成绩增加0.12-0.19 年
 - 平均薪酬水平提高 2.6 – 5.4 %
- 确保倍差法估算值的准确性:
 - “安慰剂”组的估算作用为“0”
 - 使用多种可替代标准
 - 核对每个年龄队列的影响都有意义



THE WORLD BANK

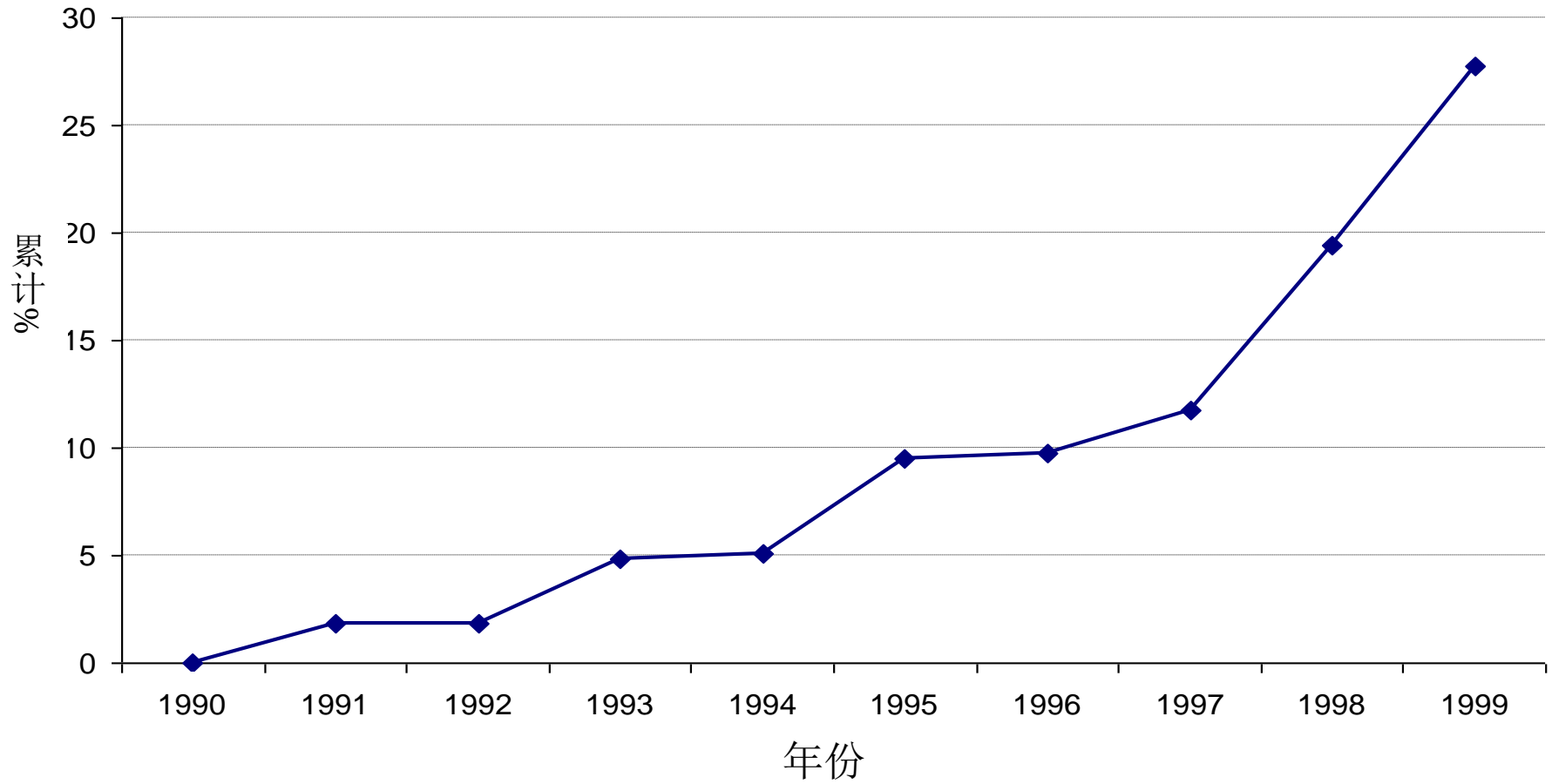


实例 #2:
生活用水: 供水私有化对儿童死亡率的影响
Sebastián Galiani, Universidad de San Andrés
Paul Gertler, UC Berkeley
Ernesto Schargrodsky, Universidad Torcuato Di Tella
JPE (2005)

1990-1999供水服务的变化

供水方法分类	市区数量	%
公有	196	39.7%
非营利性合作	143	28.9%
由公转私	138	27.9%
私有	1	0.2%
不详	16	3.2%
合计	494	100.0%

图1：拥有私有化供水系统的城市所占比例



使用“外部”因素确定谁来私有化...

- 城市的执政党
 - 联邦政府和庇隆主义省级政党: 允许私有化
 - 激进党派: 不允许私有化
- 哪个政党执政 和供水私有化并不取决于:
 - 城市收入水平, 失业率, 不公平程度
 - 近期婴儿死亡率的变化

回归法

$$y_{it} = \alpha dI_{it} + \beta x_{it} + \lambda_t + \mu_i + \varepsilon_{it}$$

注明:

y_{it} = t 年 i 市的婴儿死亡率

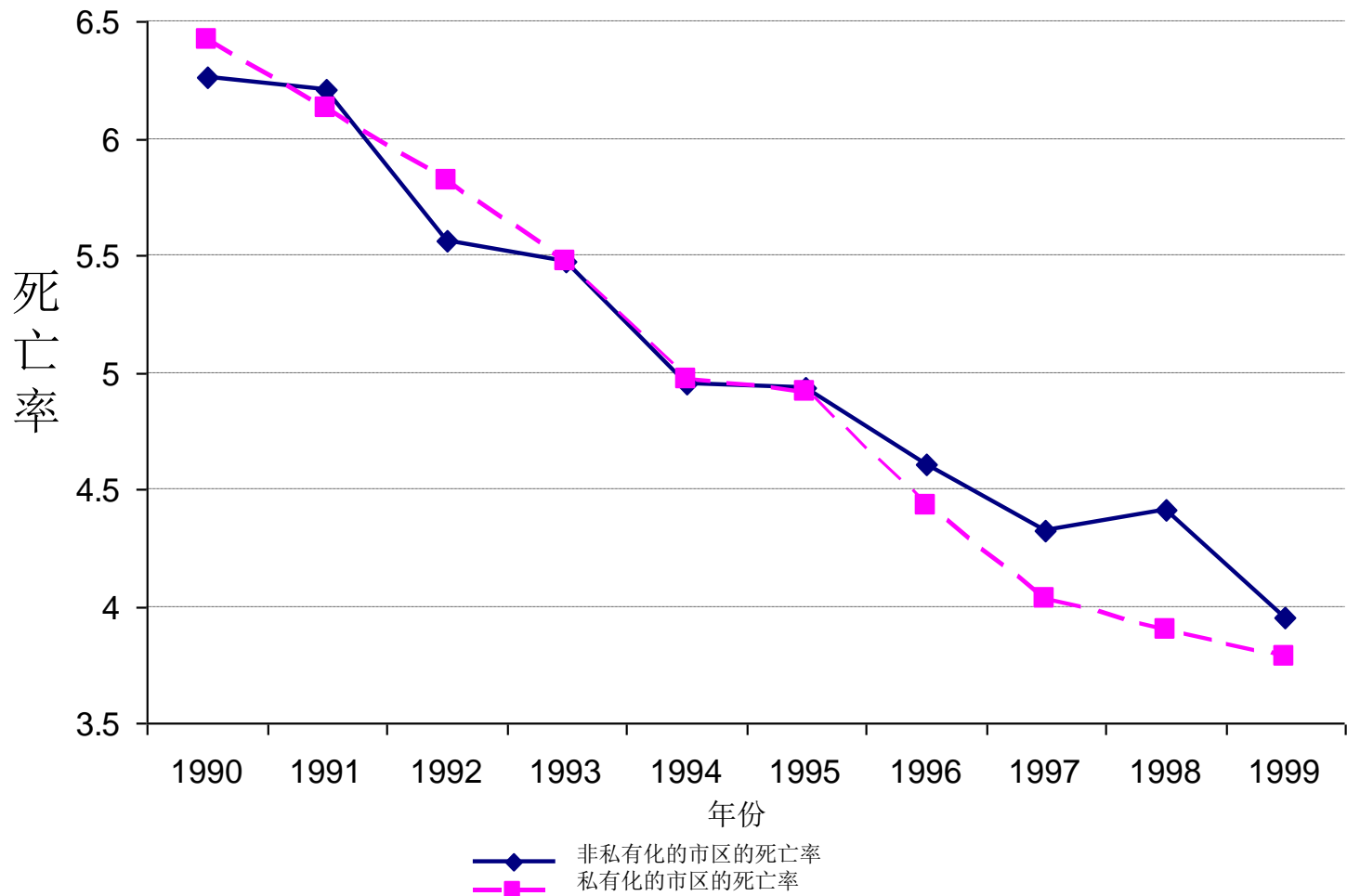
dI_{it} = 若 i 市在 t 年有私人供水者, 则哑变量赋值为1

x_{it} = 协变量向量

λ_t = 固定时间效应

μ_i = 固定城市效应

图4：私有化和非私有化的供水服务城市
死亡率演变的比较



倍差法结果: 私有化减少婴儿死亡率

		全样本			共同的支集			匹配
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
私立供水 (=1)	1)	-0.33 **	-0.32 *	-0.29 *	-0.54 ***	-0.54 ***	-0.53 ***	-0.60 ***
% 死亡率	y	-5.3 %	-5.1 %	-4.5 %	-8.6 %	-8.6 %	-8.4 %	-10.0 %
人均实际GDP	ta		0.01	0.01		0.01	0.01	
失业率	u		-0.56	-0.64		-0.78	-0.84	
不公平性 (Gini)			5.17 *	5.09 *		3.05	3.05	
人均公共费用	ap		-0.03	-0.04		-0.07 *	-0.07 *	
激进政党 (=1)	=1)			0.48 *			0.17	
庇隆主义党 (=1))			-0.20			-0.17	
市区固定效益F值		13.84***	11.92***	11.51***	10.39***	8.65***	8.32***	
年份固定效益F值		55.03***	19.88***	18.25***	52.25***	15.59***	12.98***	



倍差法的质量检查

1. 检查发现：两类城市在私有化之前的婴儿死亡率变化趋势相同
 - 你可以用相同的方程式计算干预前的作用效果，干预效果在那些年应该为0
 - 在私有化之前，我们不能否定在干预组和对照组之间相同趋势的无效假设
2. 检查发现，私有化只有在与水 and 卫生有关时才会影响死亡率。
 - 例如，私有化对心血管疾病和意外事故的死亡率并无影响。

私有化对各种原因所致的死亡的影响 倍差法的共同的支集

	1990年 平均死亡率	估算的 影响系数	私人企业 死亡率中差异所占的 %
感染性和寄生虫性疾病	.565	-.108 (.048)** [.055]* {.068}	-18.2
围产期死亡人数	2.316	-.266 (.105)** [.107]** {.123]**	-11.5
其他原因综合	2.565	-.082 (.114) [.101] {.109}	-3.2
其他原因分解:			
意外事故	.399	-.004 (.057)	...
先天性畸形	.711	-.022 (.056)	...
皮肤和软组织疾病	.000	.000 (.001)	...
血液和出血性疾病	.024	-.002 (.008)	...
神经系统紊乱	.163	.025 (.026)	...
心血管系统疾病	.236	.006 (.030)	...

私有化对贫困和特困城市的影响比非贫困城市大

城市	1990年每 100人平均 死亡率,	预期影响	死亡率 变化率%
非贫困	5.15	0.105	...
贫困	7.17	-0.767***	-10.7%
特贫	9.46	-2.214***	-23.4%

结论: 应用综合方法, 我们发现....

- 供水服务的私有化使婴儿死亡率降低5-7%。
- 死亡率降低的原因与:
 - 感染性疾病和寄生虫疾病所导致的死亡数减少有关
 - 与非水和非卫生因素所导致的死亡率改变无关
- 低收入城市的婴儿死亡率降幅最大

参考文献

- Duflo, E. (2001). “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence From an Unusual Policy Experiment,” *American Economic Review*, Sept 2001
- Sebastian Galiani, Paul Gertler and Ernesto Schargrotsky (2005): “Water for Life: The Impact of the Privatization of Water Services on Child Mortality,” *Journal of Political Economy*, Volume 113, pp. 83-120.
- Chay, Ken, McEwan, Patrick and Miguel Urquiola (2005): “The central role of noise in evaluating interventions that use test scores to rank schools,” *American Economic Review*, 95, pp. 1237-58.
- Gertler, Paul (2004): “Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA’s Control Randomized Experiment,” *American Economic Review*, 94, pp. 336-41.