高级计量经济学

理论经济学博士课程

Lecture 15: DID and Synthetic Control

黄嘉平

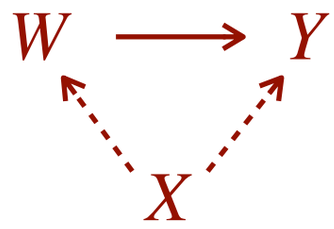
工学博士 经济学博士
深圳大学中国经济特区研究中心 讲师

Office 粤海校区汇文楼1510
Email huangjp@szu.edu.cn
Website https://huangjp.com

不可观测协变量的干扰

Confounding by Unobservables

当协变量 X 不可观测时，因果模型的 DAG 表达如下：



X 不可观测

由于 X 不可观测，我们无法对其进行控制，因此后门路径 $W \dashleftarrow X \dashrightarrow Y$ 无法被阻断，造成对平均处理效应的估计偏差。

针对这种情况，学术界提出了多种估计方法，分别建立在不同的附加条件下，主要包括：

- 双重差分（difference in differences）：用于面板数据或重复横截面数据
- 合成控制（synthetic control）：将 DID 中的控制组整合
- 工具变量（instrumental variables）：当存在不服从处理分配的情况时，估计服从者的 ATE
- 断点回归设计（regression discontinuity design）：用于处理变量是阈值函数的情况

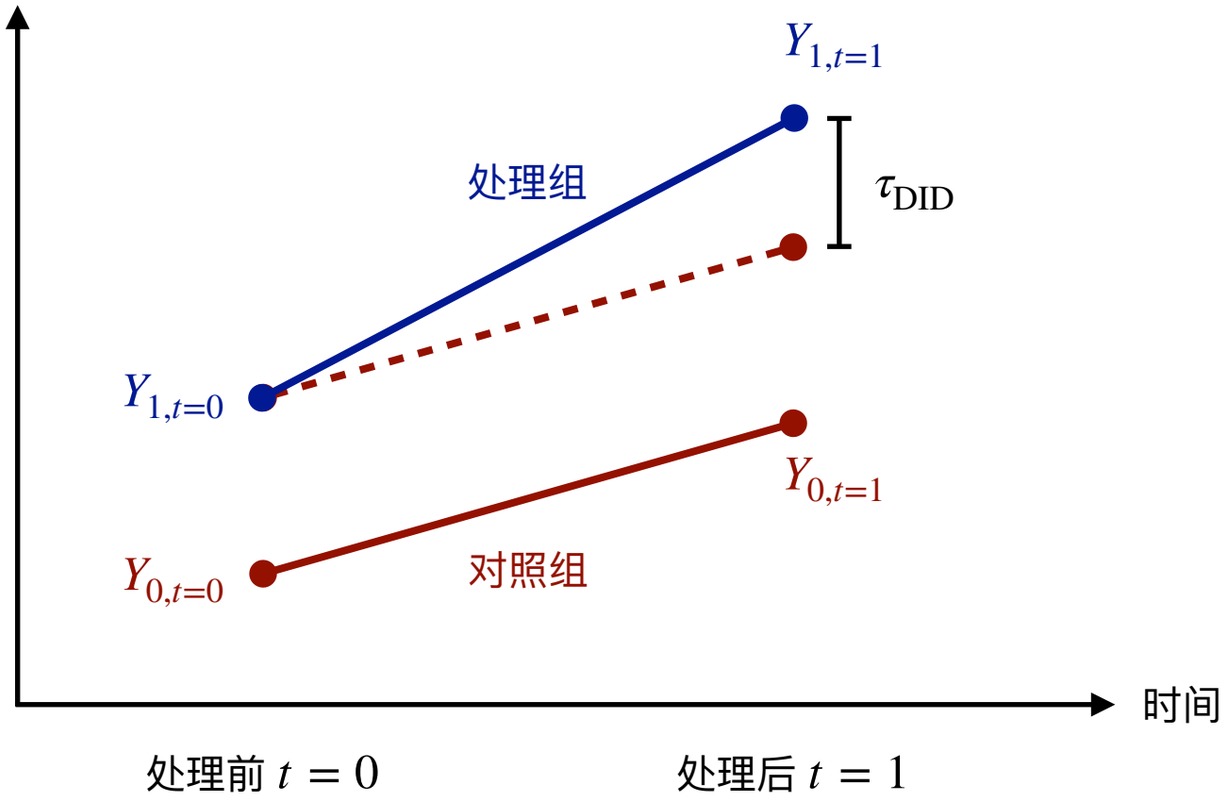
双重差分

双重差分原理

The Principle of DID

当我们可以获得处理前和处理后的观测结果时，我们可以把观测对象分为四组，即

	处理前	处理后
处理组	$Y_{1,t=0}$	$Y_{1,t=1}$
对照组	$Y_{0,t=0}$	$Y_{0,t=1}$



我们允许处理组和对照组间存在异质性，但假设处理前后的变化趋势相同。

此时，对照组的趋势可以用来估计处理组趋势的反事实结果，而两个趋势之差就是平均处理效应的估计量。

面板数据下的 DID 估计

DID with Panel Data

令 Y_{it}^{obs} 为个体 i 在时间 t 的观测结果，W_{it} 为处理状态。Y_{it}^{obs} 关于 W_{it} 的固定效应模型为

$$Y_{it}^{\text{obs}} = \tau_{it} W_{it} + \mu_i + \delta_t + \varepsilon_{it}$$

如果用潜在结果表达此模型，则有

$$\begin{aligned} &Y_{0it} = \mu_i + \delta_t + \varepsilon_{it} \\ &Y_{1it} = \tau_{it} + \mu_i + \delta_t + \varepsilon_{it} \end{aligned} \quad \Rightarrow \quad \tau_{it} = Y_{1it} - Y_{0it}$$

下面我们假设最简单的二期模型，即 $t \in \{0,1\}$。$t = 0$ 代表处理前（pre-treatment），因此 $W_{i0} = 0$；$t = 1$ 代表处理后（post-treatment），此时有一部分个体接受处理。μ_i 是不随时间变化的协变量，可以与 W_{it} 相关。

此模型中的关键假设是：

- 共同趋势假设（common trend assumption）：时间趋势不随个体变化，体现在 δ_t 不依赖 i 。
- 误差项独立于处理的分配：$E[\varepsilon_{it} \mid W_{it}] = E[\varepsilon_{it}] \quad \Rightarrow \quad E[\varepsilon_{i1} - \varepsilon_{i0} \mid W_{it}] = E[\varepsilon_{i1} - \varepsilon_{i0}]$。

从 $Y_{0it} = \mu_i + \delta_t + \varepsilon_{it}$ 和上面的假设可以得出

$$E[Y_{0i1} - Y_{0i0} \mid W_{it} = 1] = E[Y_{0i1} - Y_{0i0} \mid W_{it} = 0] = \delta_1 - \delta_0$$

即无论处理状态如何，未接受处理时的潜在结果的平均时间趋势是共通的。

面板数据下的 DID 估计

DID with Panel Data

对固定效应模型 $Y_{it}^{\text{obs}} = \tau_{it} W_{it} + \mu_i + \delta_t + \varepsilon_{it}$ 取条件期望，可得

$$
\begin{aligned}
E[Y_{i1}^{\text{obs}} \mid W_{i1} = 1] &= E[\tau_{i1} \mid W_{i1} = 1] + E[\mu_i \mid W_{i1} = 1] + \delta_1 + E[\varepsilon_{i1}] \\
E[Y_{i0}^{\text{obs}} \mid W_{i1} = 1] &= \qquad\qquad + E[\mu_i \mid W_{i1} = 1] + \delta_0 + E[\varepsilon_{i0}] \\
E[Y_{i1}^{\text{obs}} \mid W_{i1} = 0] &= \qquad\qquad + E[\mu_i \mid W_{i1} = 0] + \delta_1 + E[\varepsilon_{i1}] \\
E[Y_{i0}^{\text{obs}} \mid W_{i1} = 0] &= \qquad\qquad + E[\mu_i \mid W_{i1} = 0] + \delta_0 + E[\varepsilon_{i0}]
\end{aligned}
$$

由此可推出

$$
\begin{aligned}
\tau_{\text{ATET}} &= E[Y_{1it} - Y_{0it} \mid W_{i1} = 1] = E[\tau_{it} \mid W_{i1} = 1] = E[\tau_{i1} \mid W_{i1} = 1] \\
&= E[Y_{i1}^{\text{obs}} - Y_{i0}^{\text{obs}} \mid W_{i1} = 1] - E[Y_{i1}^{\text{obs}} - Y_{i0}^{\text{obs}} \mid W_{i1} = 0] = \tau_{\text{DID}}
\end{aligned}
$$

因此，我们可以获得 τ_{ATET} 的 DID 估计量

$$
\hat{\tau}_{\text{ATET}} = \hat{\tau}_{\text{DID}} = \frac{1}{N_1} \sum_{i:W_{i1}=1} (y_{i1}^{\text{obs}} - y_{i0}^{\text{obs}}) - \frac{1}{N_0} \sum_{i:W_{i1}=0} (y_{i1}^{\text{obs}} - y_{i0}^{\text{obs}})
$$

也可以用差分回归 $Y_{i1}^{\text{obs}} - Y_{i0}^{\text{obs}} = \alpha + \tau W_{i1} + u_i$ 中系数 τ 的 OLS 估计量估计，即 $\hat{\tau}_{\text{ATET}} = \hat{\tau}_{\text{OLS}}$。

即使协变量 μ_i 不可观测，面板数据也可以帮助我们估计 ATET。

重复横截面数据与 DID 估计

Repeated (Pooled) Cross-Section Data and DID

DID 是均值的比较，因此并不要求对同一个体进行持续观察，而只需要观察不同时间点上的处理组和对照组中的均值。也就是说，我们可以用重复横截面数据构建 DID 估计量。

考虑回归模型

$$Y_i^{\text{obs}} = \alpha + \beta T_i + \gamma G_i + \delta(T_i \times G_i) + \varepsilon_i$$

其中 T_i 是时间虚拟变量（$t = 1$ 时取值为 1），G_i 组别虚拟变量（i 在处理组时取值为 1）。只有在 T_i 和 G_i 的交互项取值为 1 时，i 才接受处理。也就是说 $W_i = T_i \times G_i$。

假设 $E[\varepsilon_i \mid G_i, T_i] = E[\varepsilon_i]$，可得

$$\begin{aligned}\delta = &\left(E[Y_i^{\text{obs}} \mid G_i = 1, T_i = 1] - E[Y_i^{\text{obs}} \mid G_i = 1, T_i = 0]\right) \\ &- \left(E[Y_i^{\text{obs}} \mid G_i = 0, T_i = 1] - E[Y_i^{\text{obs}} \mid G_i = 0, T_i = 0]\right) \\ = &\ \tau_{\text{DID}}\end{aligned}$$

因此，$\hat{\tau}_{\text{DID}} = (\bar{y}_{11}^{\text{obs}} - \bar{y}_{10}^{\text{obs}}) - (\bar{y}_{01}^{\text{obs}} - \bar{y}_{00}^{\text{obs}})$，$\bar{y}_{mt}^{\text{obs}} = \frac{1}{N_{gt}} \sum_{i:G_i=g,T_i=t} y_i^{\text{obs}}$。

用回归方法也可以求得 DID 估计量，即 $\hat{\tau}_{\text{DID}} = \hat{\delta}_{\text{OLS}}$。

多个分组与多个时间点

Multiple Groups and Multiple Time-Periods

我们可以把模型拓展到有 $G+1$ 个分组，即 $G_i \in \{0,\ldots,G\}$，和 $T+1$ 个时间点，即 $T_i \in \{0,\ldots,T\}$。

多个分组并不意味着有多个处理水平，而是意味着不同组别中的个体可能在不同的时间点接受处理。例如某一政策在不同地区的实施时间不同等情况。

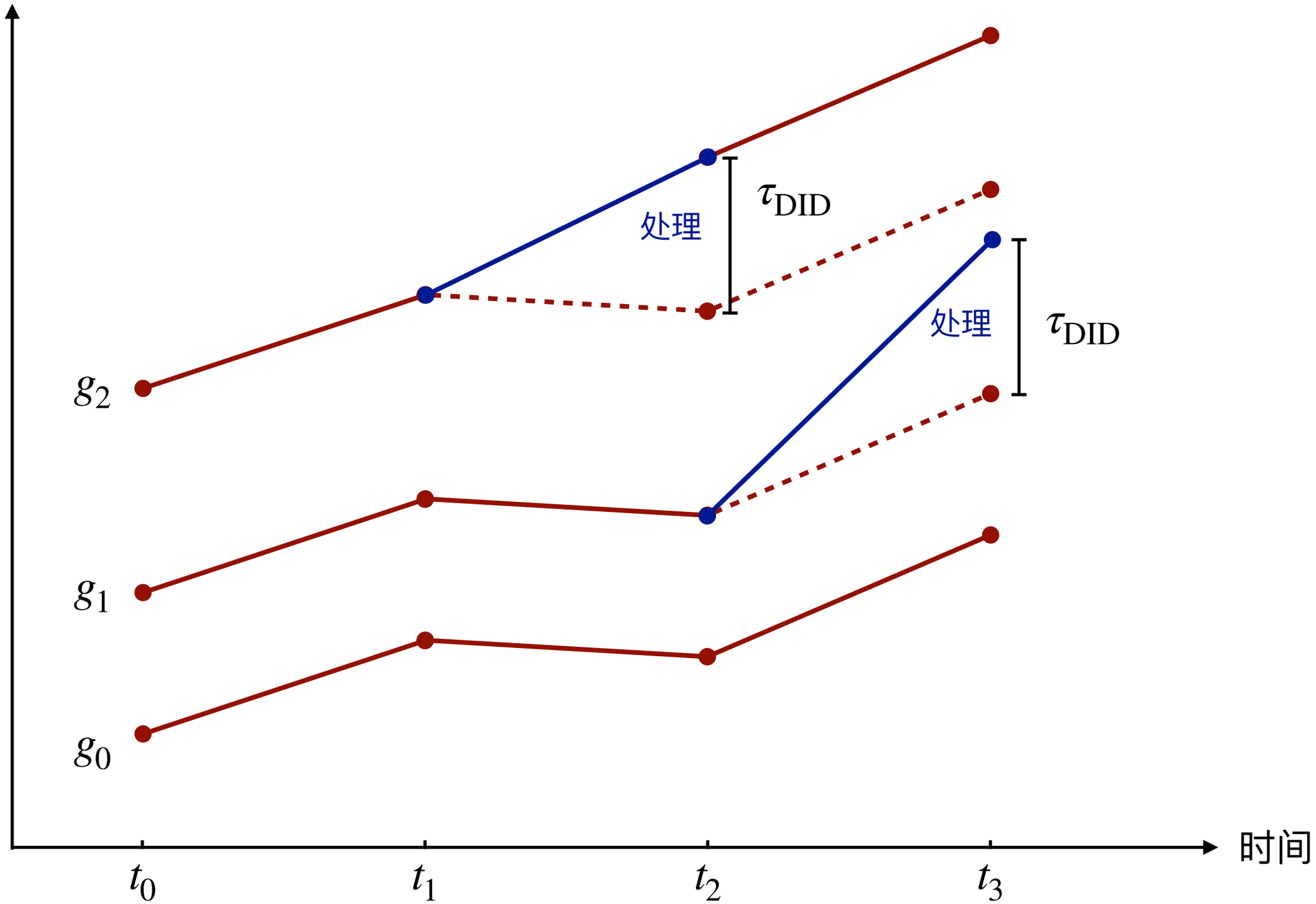
令 $1_X = \begin{cases} 1 & \text{if } X \text{ is true} \\ 0 & \text{if } X \text{ is false} \end{cases}$ 为条件 X 成立状态的指示函数。则前面的模型可以拓展为

$$Y_i^{\text{obs}} = \alpha + \sum_{t=1}^{T} \beta_t 1_{T_i=t} + \sum_{g=1}^{G} \gamma_g 1_{G_i=g} + \delta W_i + \varepsilon_i$$

处理变量 W_i 取 1 代表 i 在对应的分组与时间点的组合中接受处理。

这里依然假设存在共同趋势，即模型中的 β_t 不依赖 i 的分组。因此依然有 $\tau_{\text{DID}} = \delta$ 以及 $\hat{\tau}_{\text{DID}} = \hat{\delta}_{\text{OLS}}$。

与前面模型不同的是，此时我们可以检验共同趋势假设。如果分组 g_1 和 g_2 中的个体在时间点 t_1 和 t_2 都没有接受处理，则在共同趋势假设下，它们均值的双重差分应该为零。


τ_{DID}
处理
τ_{DID}
处理
g_2
g_1
g_0
t_0
t_1
t_2
t_3
时间

双重差分的实证研究

Card (1990)

The Impact of the Mariel Boatlift on the Miami Labor Market.
***Industrial and Labor Relations Review*, 43:2, 245-257.**

- 核心问题：低技能移民对国内低端劳动市场的影响

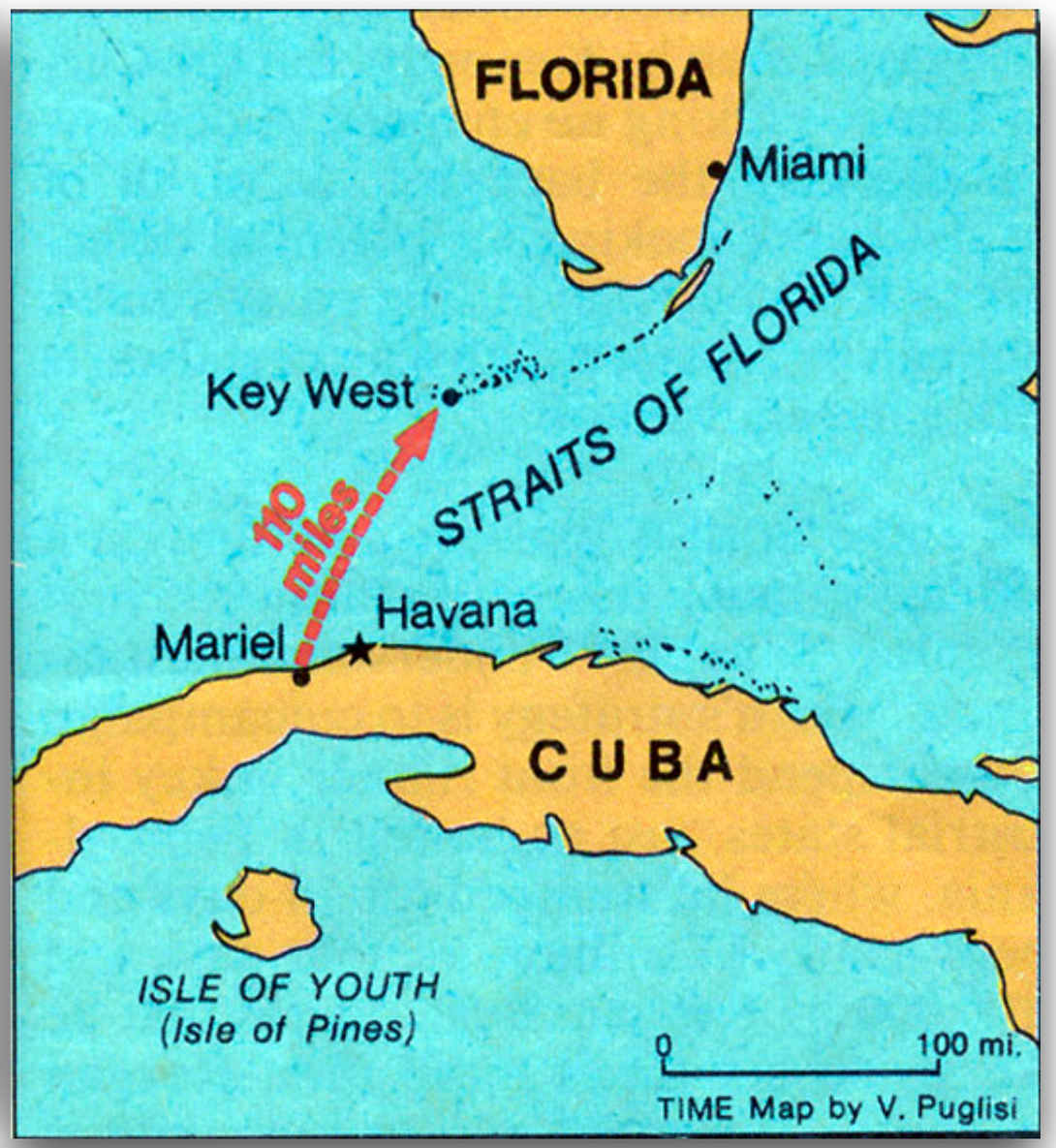
  经济学理论预测，低技能移民的大量流入会冲击当地的低端劳动市场，是原有的低技能劳动者收入降低或失业。

  然而现实中，移民和本地人会根据就业市场的需求选择居住地。如果流入地的就业市场饱和，则移民会更倾向去其他地方找工作。
  → 基于移民人口密度和工资间相关性的研究无法正确估计移民给劳动市场带来的冲击

- The 1980 Mariel Boatlift

  1980年4月20日，当时的古巴国家元首卡斯特罗宣布了一项政策，允许希望向国外移民的古巴人从哈瓦那附近的 Mariel 港自由离开古巴。卡苏特罗在1980年5月1日的一次讲演中说：
  “Those who have no revolutionary genes, those who have no revolutionary blood...we do not want them, we do not need them” (https://www.history.com/news/mariel-boatlift-castro-carter-cold-war)

  这一政策造成了当年5月至9月间，125000名古巴移民乘坐小船到达美国福罗里达州的迈阿密市。最终，50%的古巴移民留在了迈阿密，相当于之前在迈阿密工作的古巴人总数的20%，迈阿密总劳动力的7%。



Cooler
Highs in the low 80s. Lows in the low 60s on the coast and the mid 60s inland. Winds west-northwest 15 knots. Seas 3 feet. Bay, choppy.
Complete weather, 2A

The Miami News
A Cox Newspaper
Tuesday Afternoon, April 29, 1980
42 Pages

Tuesday
Metropolitan
★ Edition
15¢

New Cuban exodus due here

Graham declares state of emergency for Dade, Monroe

HEATHER DEWAR and ANA VECIANA
Miami News Reporters

pect the quota of 3,500 Cuban refugees set by the State Department will soon be shattered.

But Metro Mayor Steve Clark warned yesterday that four out of five of the refugees arriving here have no local relatives and need federal help.

the sand?" said Vincent Grimm, a city of Miami assistant administrator. "The city's already at the breaking point."

Card (1990)

The Impact of the Mariel Boatlift on the Miami Labor Market.
***Industrial and Labor Relations Review*, 43:2, 245-257.**

- 1980年的 Mariel Boatlift 可以看做是自然实验。它导致了低技能外来移民向迈 Miami 的大量流入。

- 作者选取的反映劳动力市场状况的变量 Y 是工资和失业率。

- 作者选取了1979年至1985年的数据作为处理前后的对比。Miami 作为处理组，而对照组中包含Atlanta、Los Angeles、Houston、Tampa-St. Petersburg。

  选取这四座城市的理由是：他们都有和 Miami 相似的种族构成（非裔和拉丁裔占比大）、以及相似的经济发展程度。

- 文章并没有发现新增的古巴移民给 Miami 当地的低技能劳动者（非裔、早年的古巴移民）的就业带来显著冲击。作者认为这和 Miami 的特殊性有很大关系。

- 本文虽不是第一个在劳动经济学中使用自然实验和 DID 方法的研究，但其影响力很大。

Table 3. Logarithms of Real Hourly Earnings of Workers Age 16–61 in Miami and Four Comparison Cities, 1979–85.

Group	*1979*	*1980*	*1981*	*1982*	*1983*	*1984*	*1985*
Miami:							
Whites	1.85	1.83	1.85	1.82	1.82	1.82	1.82
	(.03)	(.03)	(.03)	(.03)	(.03)	(.03)	(.05)
Blacks	1.59	1.55	1.61	1.48	1.48	1.57	1.60
	(.03)	(.02)	(.03)	(.03)	(.03)	(.03)	(.04)
Cubans	1.58	1.54	1.51	1.49	1.49	1.53	1.49
	(.02)	(.02)	(.02)	(.02)	(.02)	(.03)	(.04)
Hispanics	1.52	1.54	1.54	1.53	1.48	1.59	1.54
	(.04)	(.04)	(.05)	(.05)	(.04)	(.04)	(.06)
Comparison Cities:							
Whites	1.93	1.90	1.91	1.91	1.90	1.91	1.92
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Blacks	1.74	1.70	1.72	1.71	1.69	1.67	1.65
	(.01)	(.02)	(.02)	(.01)	(.02)	(.02)	(.03)
Hispanics	1.65	1.63	1.61	1.61	1.58	1.60	1.58
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.02)

Note: Entries represent means of log hourly earnings (deflated by the Consumer Price Index—1980 = 100) for workers age 16–61 in Miami and four comparison cities: Atlanta, Houston, Los Angeles, and Tampa–St. Petersburg. See note to Table 1 for definitions of groups.

Source: Based on samples of employed workers in the outgoing rotation groups of the Current Population Survey in 1979–85. Due to a change in SMSA coding procedures in 1985, the 1985 sample is based on individuals in outgoing rotation groups for January–June of 1985 only.

Table 4. Unemployment Rates of Individuals Age 16–61 in Miami and Four Comparison Cities, 1979–85.
(Standard Errors in Parentheses)

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

Note: Entries represent means of unemployment indicator variable for individuals age 16–61 in Miami and four comparison cities: Atlanta, Houston, Los Angeles, and Tampa–St. Petersburg. Samples are based on individuals in the labor force. See notes to Table 3 for definitions of groups and data sources.

Table 6. Comparison of Wages, Unemployment Rates, and Employment Rates for Blacks in Miami and Comparison Cities.
(Standard Errors in Parentheses)

	All Backs				*Low-Education Blacks*			
	Difference in Log Wages, Miami − Comparison		*Difference in Emp./Unemp., Miami − Comparison*		*Difference in Log Wages, Miami − Comparison*		*Difference in Emp./Unemp., Miami − Comparison*	
Year	*Actual*	*Adjusted*	*Emp. − Pop. Rate*	*Unemp. Rate*	*Actual*	*Adjusted*	*Emp. − Pop. Rate*	*Unemp. Rate*
1979	−.15	−.12	.00	−2.0	−.13	−.15	.03	−.8
	(.03)	(.03)	(.03)	(1.9)	(.05)	(.05)	(.04)	(3.8)
1980	−.16	−.12	.05	−7.1	−.07	−.07	.03	−8.2
	(.03)	(.03)	(.03)	(1.6)	(.05)	(.05)	(.04)	(3.5)
1981	−.11	−.10	.02	−3.0	−.05	−.11	.04	−7.7
	(.03)	(.03)	(.03)	(2.0)	(.05)	(.05)	(.04)	(4.2)
1982	−.24	−.20	−.06	3.3	−.17	−.20	−.04	.6
	(.03)	(.03)	(.03)	(2.4)	(.05)	(.05)	(.04)	(4.7)
1983	−.21	−.15	−.02	.1	−.13	−.11	.04	−3.3
	(.03)	(.03)	(.03)	(2.7)	(.06)	(.05)	(.04)	(4.7)
1984	−.10	−.05	−.04	2.1	−.04	−.03	.05	.1
	(.03)	(.03)	(.03)	(2.4)	(.06)	(.05)	(.04)	(4.7)
1985	−.05	−.01	−.06	−5.5	.18	.09	.00	−4.7
	(.04)	(.04)	(.04)	(2.6)	(.07)	(.07)	(.06)	(5.6)

Notes: Low-education blacks are those with less than 12 years of completed education. Adjusted differences in log wages between blacks in Miami and comparison cities are obtained from a linear regression model that includes education, potential experience, and other control variables; see text. Wages are deflated by the Consumer Price Index (1980 = 100). "Emp.-Pop. Rate" refers to the employment:population ratio. "Unemp. Rate" refers to the unemployment rate among those in the labor force.

下图取自
Angrist & Krueger (1990). Empirical Strategies in Labor Economics. In Ashenfelter & Card, *Handbook of Labor Economics*, Volume 3.

1994年夏，又有30000人从古巴入境美国，导致美国国内反移民情绪高涨。克林顿政府因此决定遣返企图入境的古巴移民，并结束了之前无条件接受古巴移民的政策。

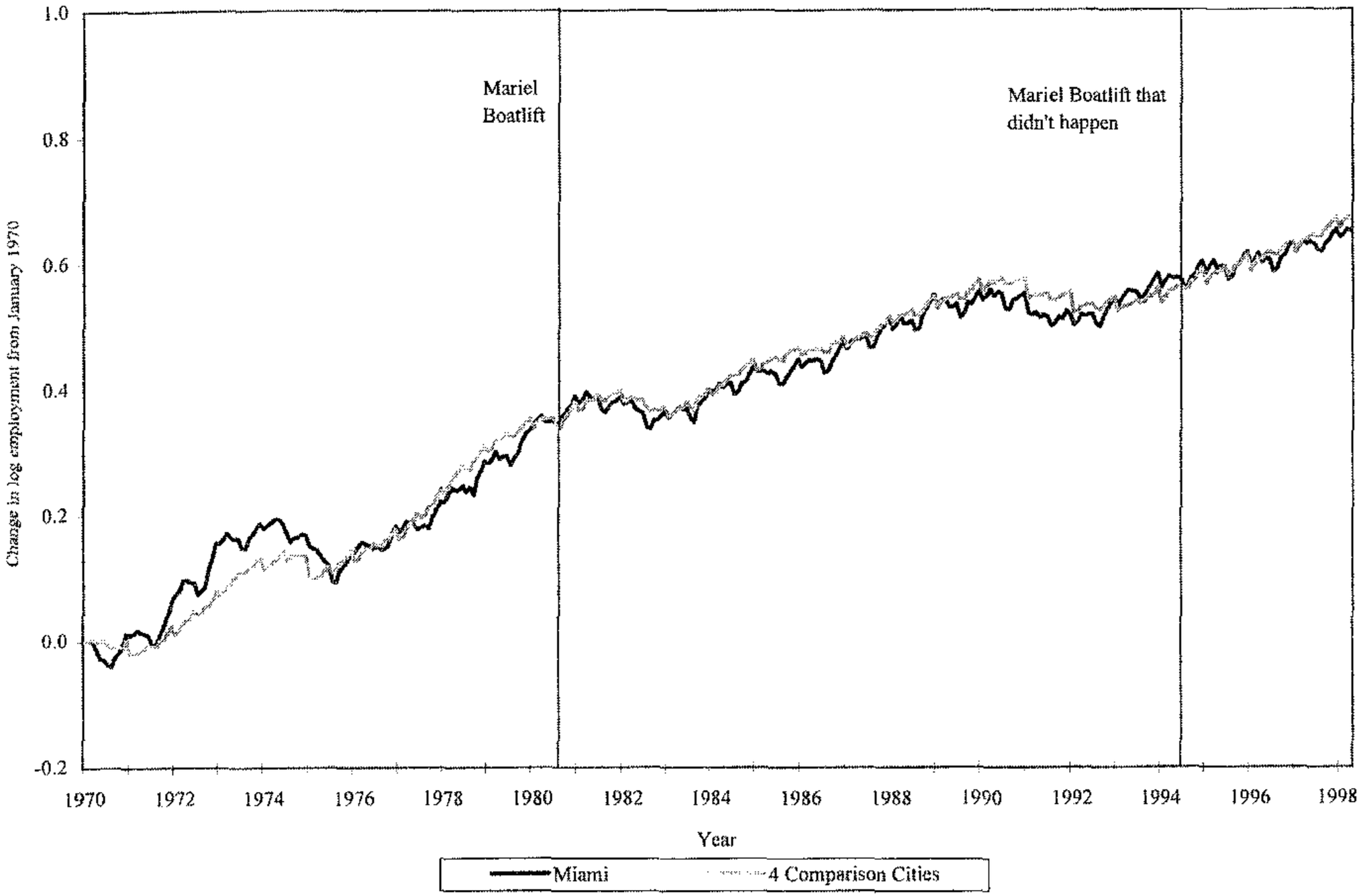


Fig. 1. Changes in employment in Miami and comparison cities. *Source*: authors' calculations from BLS State and Area Employment, Hours, and Earnings Establishment Survey.

下表取自
Angrist & Krueger (1990). Empirical Strategies in Labor Economics.
In Ashenfelter & Card, *Handbook of Labor Economics*, Volume 3.

Table 4
Differences-in-differences estimates of the effect of immigration on unemployment[a]

	Group	Year		
		1979 (1)	1981 (2)	1981–1979 (3)
	Whites			
(1)	Miami	5.1 (1.1)	3.9 (0.9)	−1.2 (1.4)
(2)	Comparison cities	4.4 (0.3)	4.3 (0.3)	−0.1 (0.4)
(3)	Miami-Comparison Difference	0.7 (1.1)	−0.4 (0.95)	−1.1 (1.5)
	Blacks			
(4)	Miami	8.3 (1.7)	9.6 (1.8)	1.3 (2.5)
(5)	Comparison cities	10.3 (0.8)	12.6 (0.9)	2.3 (1.2)
(6)	Miami-Comparison Difference	−2.0 (1.9)	−3.0 (2.0)	−1.0 (2.8)

[a] Notes: Adapted from Card (1990, Tables 3 and 6). Standard errors are shown in parentheses.

Angrist & Krueger (1999) 的验证

Empirical Strategies in Labor Economics.
In Ashenfelter & Card, *Handbook of Labor Economics*, Volume 3.

1994 年的第二次移民潮实际上并没有发生，但是 Angrist & Krueger (1999) 假设它是另一次自然实验，并进行了 DID 研究。这可以看作是针对共同趋势假设的检验。

Table 7
Unemployment rates of individuals age 16–61 in Miami and four comparison cities, 1988–1996[a]

	1988	1989	1990	1991	1992	1993	1994	1995	1996
Miami									
Whites	2.8	3.6	3.3	5.7	4.2	4.9	6.2	3.9	4.4
	(0.8)	(0.9)	(0.9)	(1.2)	(1.1)	(1.3)	(1.4)	(1.4)	(1.2)
Blacks	10.0	11.8	11.9	8.8	10.1	10.1	15.1	13.7	11.1
	(1.7)	(1.8)	(1.9)	(1.9)	(2.0)	(2.1)	(2.4)	(2.8)	(2.4)
Hispanics	5.5	7.6	7.2	9.1	10.3	8.5	9.4	8.4	8.9
	(1.4)	(1.5)	(1.4)	(1.6)	(1.7)	(1.6)	(1.8)	(1.8)	(1.6)
Comparison cities									
Whites	4.2	3.5	3.8	4.9	5.1	5.4	5.0	4.1	4.1
	(0.3)	(0.2)	(0.2)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)
Blacks	11.3	8.4	9.6	9.6	13.6	11.5	10.9	8.8	9.3
	(0.9)	(0.8)	(0.8)	(0.9)	(1.0)	(0.9)	(0.9)	(0.8)	(0.8)
Hispanics	7.2	7.5	5.8	9.1	10.9	11.3	11.0	10.0	9.4
	(0.7)	(0.6)	(0.4)	(0.5)	(0.6)	(0.6)	(0.6)	(0.7)	(0.6)

[a] Note: Standard errors are in parentheses. The four comparison cities (Atlanta, Houston, Los Angeles, and Tampa-St. Petersburg), are the same comparison cities used by Card (1990). The reported unemployment rates are from the authors' tabulations of CPS Outgoing Rotation Groups.

针对非裔的 DID 估计值是 6.3 ($t = 1.70$)，说明共同趋势假设可能不成立。

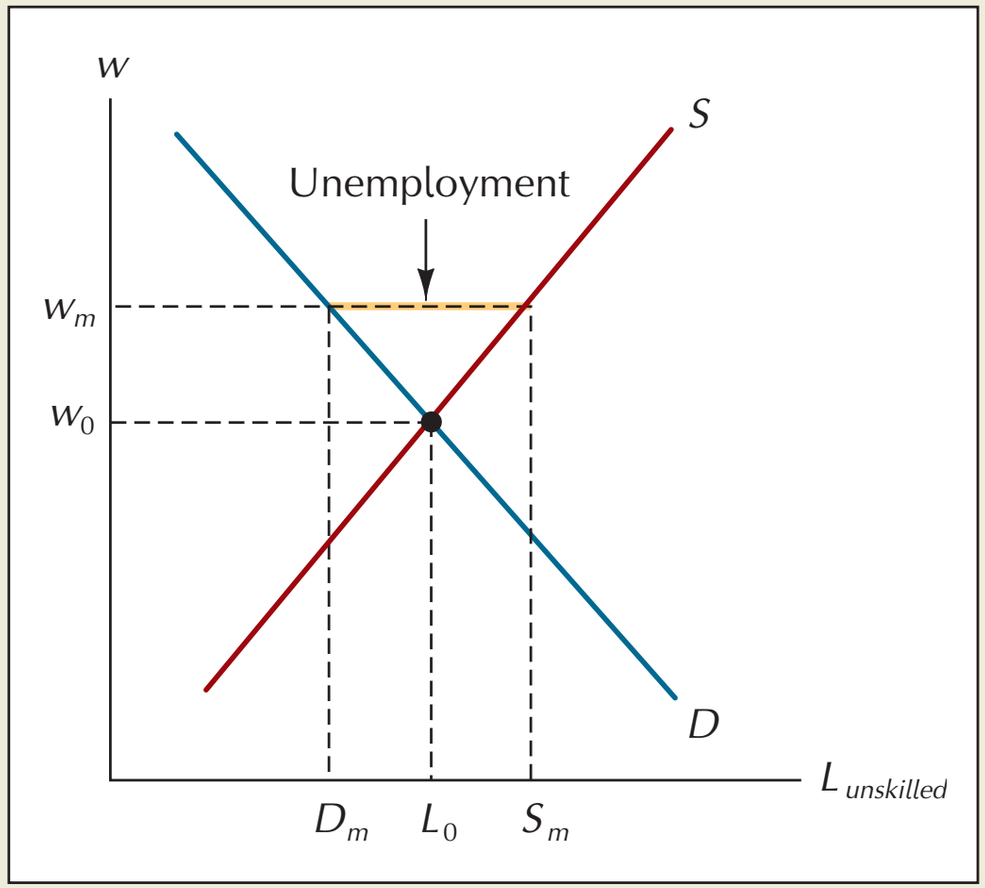
Card & Krueger (1994)

Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania
***American Economic Review*, 84:4, 772-793.**

- 核心问题：提高最低工资标准对低端劳动市场的影响

  根据传统经济学理论，在完全竞争市场中，提高最低工资标准会增加失业率

FIGURE 14.13
A Statutory Minimum Wage
The effect of the minimum wage is to reduce employment of unskilled labor from L_0 to D_m, while increasing supply from L_0 to S_m. The resulting difference, $S_m - D_m$, is the unemployment attributable to the minimum wage.



本图出自 Frank, R. Microeconomics and Behavior, 9th Edition, McGraw-Hill Education.

Card & Krueger (1994)

Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania
***American Economic Review*, 84:4, 772-793.**

- 很多实证研究并未发现提高最低工资标准对就业带来明显的负面影响。
- 本文中，作者选择美国东北部的 Pennsylvania 州和 New Jersey 州进行对比。

1990年代初的美国联邦最低工资标准变化
1990.4.1: $3.35/h → $3.80/h
1991.4.1: $3.80/h → $4.25/h

同时期，NJ 制定了自己的最低工资标准变化
1992.4.1: $4.25/h → $5.05/h

- 作者选取了 NJ 和 PA 东部的快餐店，
以电话采访的方式收集数据。

快餐店包括：
Burger King, KFC, Wendy's, Roy Rogers

采访时间：1992年2-3月，1992年11-12月

TABLE 1—SAMPLE DESIGN AND RESPONSE RATES

	All	Stores in: NJ	PA
Wave 1, February 15 – March 4, 1992:			
Number of stores in sample frame:[a]	473	364	109
Number of refusals:	63	33	30
Number interviewed:	410	331	79
Response rate (percentage):	86.7	90.9	72.5
Wave 2, November 5 – December 31, 1992:			
Number of stores in sample frame:	410	331	79
Number closed:	6	5	1
Number under rennovation:	2	2	0
Number temporarily closed:[b]	2	2	0
Number of refusals:	1	1	0
Number interviewed:[c]	399	321	78

[a]Stores with working phone numbers only; 29 stores in original sample frame had disconnected phone numbers.

[b]Includes one store closed because of highway construction and one store closed because of a fire.

[c]Includes 371 phone interviews and 28 personal interviews of stores that refused an initial request for a phone interview.

下表取自
Angrist & Pischke (2009). *Mostly Harmless Econometrics*.

Table 5.2.1
Average employment in fast food restaurants before and after the New Jersey minimum wage increase

Variable	PA (i)	NJ (ii)	Difference, NJ – PA (iii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (.51)	−2.89 (1.44)
2. FTE employment after, all available observations	21.17 (.94)	21.03 (.52)	−.14 (1.07)
3. Change in mean FTE employment	−2.16 (1.25)	.59 (.54)	2.76 (1.36)

Notes: Adapted from Card and Krueger (1994), table 3. The table reports average full-time-equivalent (FTE) employment at restaurants in Pennsylvania and New Jersey before and after a minimum wage increase in New Jersey. The sample consists of all restaurants with data on employment. Employment at six closed restaurants is set to zero. Employment at four temporarily closed restaurants is treated as missing. Standard errors are reported in parentheses.

FTE 翻译为全时人力工时，指企业雇佣的所有员工的总工作量相当于雇佣多少全职员工可以完成的总工作量。

DID 估计值为正，并没有观测到就业率的降低

Card & Krueger (2000) 的验证

学术界对 Card & Krueger (1994) 的质疑包括：样本不具有代表性，无法验证共同趋势假设是否成立等。

作者在本文中采用了 Bureau of Labor Statistics 关于就业的调查数据，将时间区间扩大至 1991:Q4-1997:Q3，并增加了 PA 州的样本量。

1996年10月，美国联邦最低工资标准从 $4.25/h 提升至 $4.75/h。这次提升影响了 PA 州，但没有影响 NJ 州。

这次变化可以看作第二次自然实验。

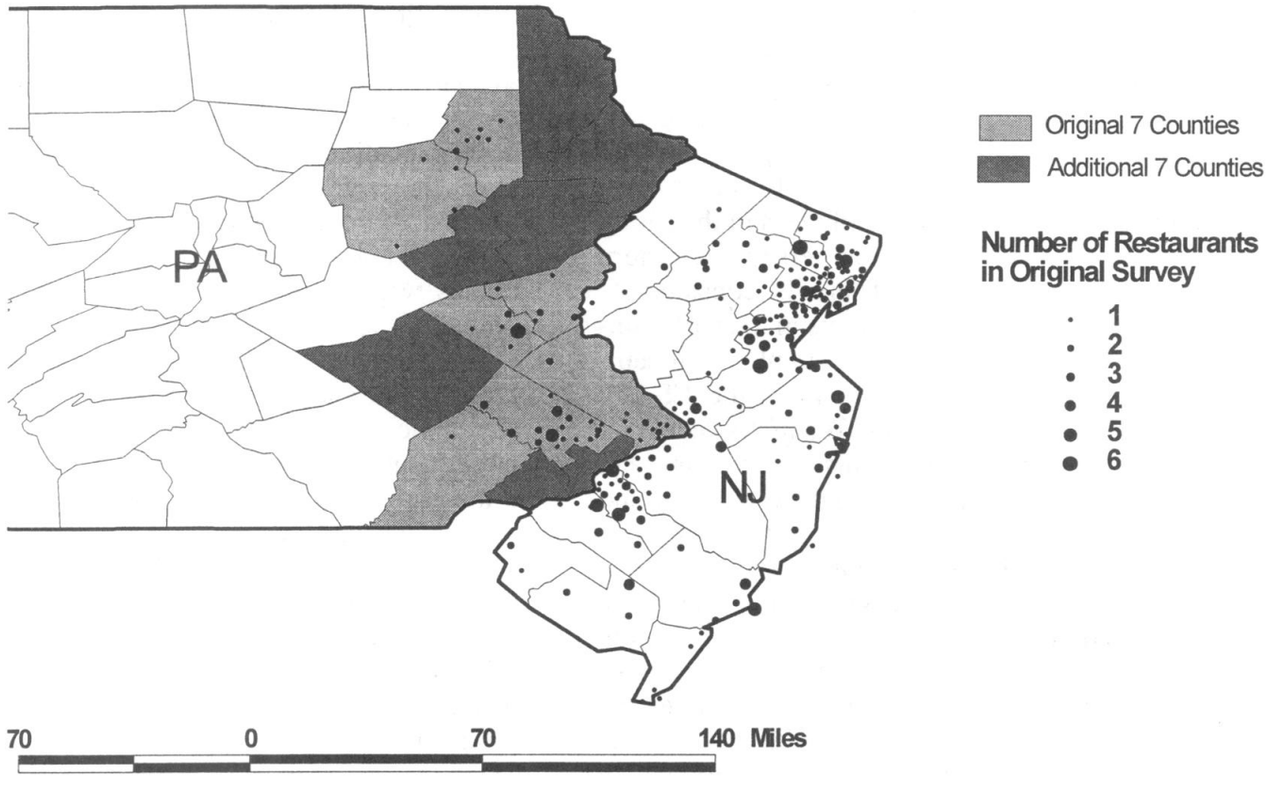


FIGURE 1. AREAS OF NEW JERSEY AND PENNSYLVANIA COVERED BY ORIGINAL SURVEY AND BLS DATA

TABLE 1—DESCRIPTIVE STATISTICS FOR FAST-FOOD RESTAURANTS DRAWN FROM BLS ES-202 DATA AND CARD-KRUEGER SURVEY

	Means with standard deviations in parentheses:								
	New Jersey			7 Pennsylvania counties			14 Pennsylvania counties		
	Before	After	Change	Before	After	Change	Before	After	Change
A. *BLS ES-202 Data*									
February–March 1992 to November–December 1992	37.2 (19.9)	37.6 (21.0)	0.41 (9.82)	42.5 (23.2)	42.4 (23.5)	−0.12 (10.94)	44.8 (53.7)	44.3 (59.9)	−0.53 (12.32)
February 1992 to November 1992	37.2 (19.9)	37.8 (20.9)	0.57 (10.12)	42.7 (23.8)	42.2 (23.2)	−0.54 (12.82)	44.9 (53.6)	44.4 (60.4)	−0.58 (13.83)
March 1992 to March 1993	37.2 (20.1)	34.8 (20.0)	−2.48 (13.99)	42.3 (22.8)	37.5 (18.6)	−4.80 (22.74)	44.7 (54.0)	40.7 (54.5)	−4.0 (18.1)
B. *Card-Krueger Survey Data*									
February 1992 to November 1992	29.8 (12.5)	30.0 (13.0)	0.19 (9.82)	33.1 (14.7)	30.9 (10.6)	−2.23 (11.98)	NA	NA	NA

Notes: Sample sizes for the first two rows are 437 for New Jersey, 127 for Pennsylvania 7 counties, and 250 for Pennsylvania 14 counties; sample sizes for third row are 436, 127, and 250, respectively; sample sizes for the last row are 309 for New Jersey and 75 for Pennsylvania. The 7 Pennsylvania counties used in the middle columns are the same counties used in Card and Krueger (1994); these 7 counties are a subset of the 14 counties in the last three columns (see text). The unit of observation for the BLS data is the "reporting unit," which in some cases includes multiple establishments. The unit of observation in the Card-Krueger data is the individual restaurant.

DID 估计值为正，基本确认了 Card & Krueger (1994) 的结论

下图取自
Angrist & Pischke (2009). *Mostly Harmless Econometrics*.

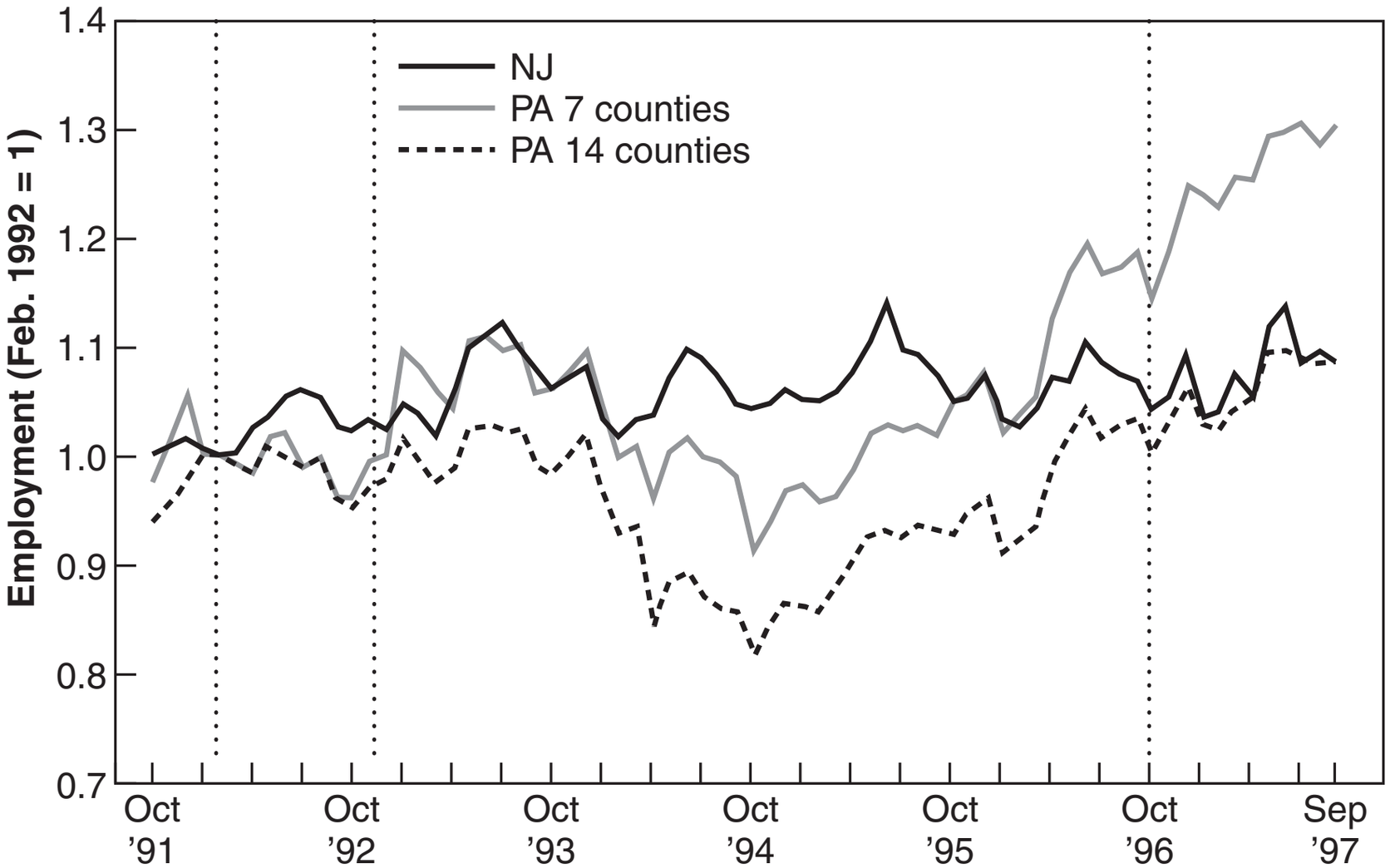


Figure 5.2.2 Employment in New Jersey and Pennsylvania fast food restaurants, October 1991 to September 1997 (from Card and Krueger 2000). Vertical lines indicate dates of the original Card and Krueger (1994) survey and the October 1996 federal minimum wage increase.

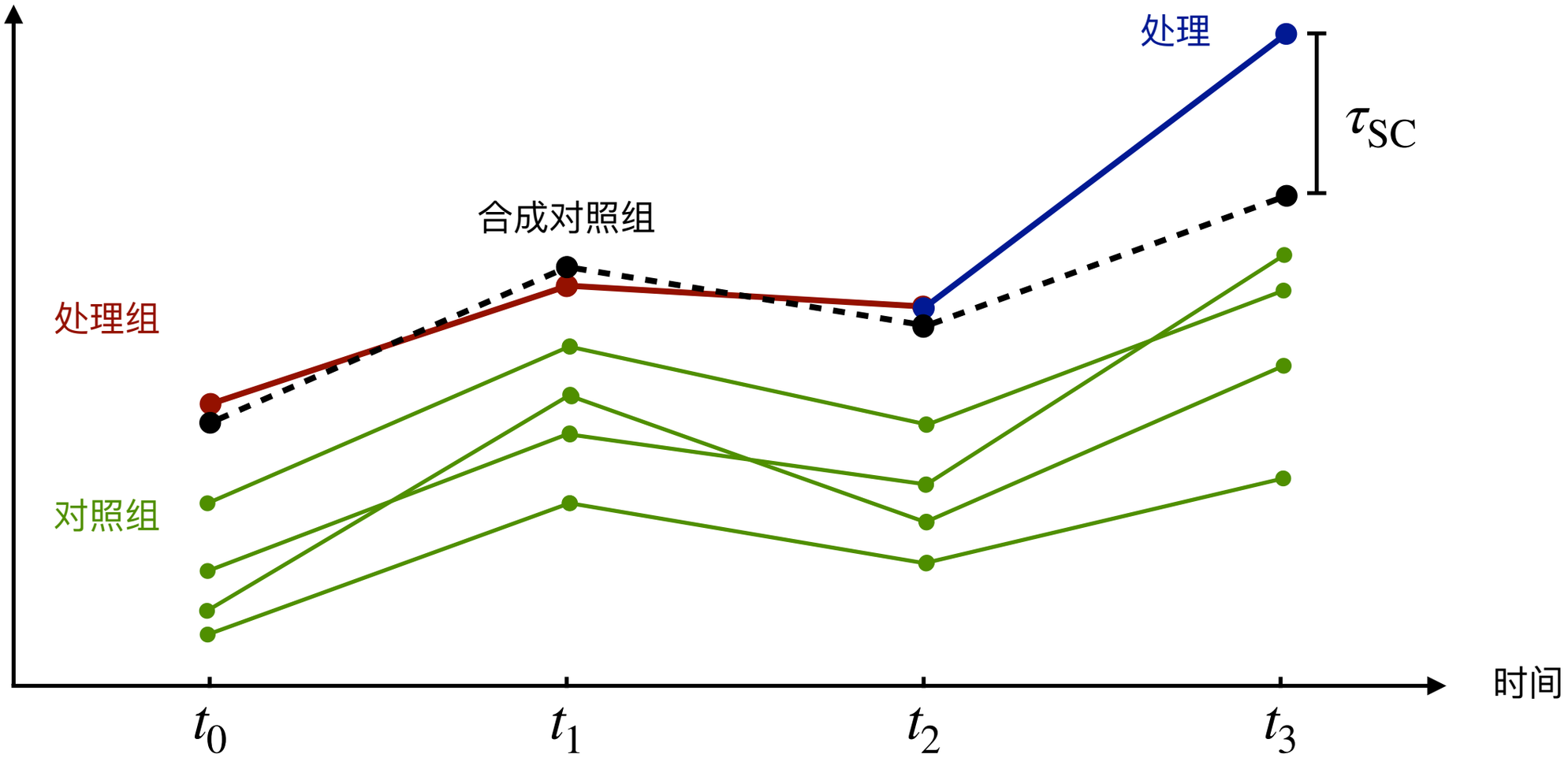
1993:Q3-1996:Q3之间的数据显示：PA 和 NJ 的就业率走势存在明显不同，并不支持共同趋势假设。

合成控制

合成控制原理

The Principle of Synthetic Control

当存在多个对照组和多个处理前时间点时，我们可以用这些信息构造一个合成对照组（synthetic control group），并生成处理后阶段的虚拟潜在对照结果（potential control）。我们用这个虚拟潜在对照结果和处理结果（可观察）之差作为处理效应的估计量。



合成控制法

Synthetic Control Methods

假设有 $G+1$ 个分组和 $T+1$ 个时间点，其中 $\{0,\ldots,G-1\}$ 为控制组（或 donor pool），G 为处理组。所有分组在时间点 $\{0,\ldots,T-1\}$ 时都不接受处理，只有处理组 G 在时间点 T 时才接受处理。

在时间点 T，我们可以观测到

$$Y^{\text{obs}}_{i:G_i=G,T_i=T} = Y_{1i:G_i=G,T_i=T}, \qquad Y^{\text{obs}}_{i:G_i\neq G,T_i=T} = Y_{0i:G_i\neq G,T_i=T}$$

这里考虑用对照组观测值的加权平均作为处理组的潜在对照结果，即

$$\hat{E}[Y_{0i} \mid G_i = G, T_i = T] = \sum_{g=0}^{G-1} \lambda_g \bar{Y}_{gT}$$

此时，

$$\begin{aligned}\tau_{\text{ATET}} &= E[Y_{1i} \mid G_i = G, T_i = T] - \hat{E}[Y_{0i} \mid G_i = G, T_i = T] \\ &= \bar{Y}_{GT} - \sum_{g=0}^{G-1} \lambda_g \bar{Y}_{gT}\end{aligned}$$

剩下的问题就是如何确定权重 $\lambda_0, \ldots, \lambda_{G-1}$。

如何确定权重

How to Determine Weights

在处理前的时间段 $\{0,\ldots,T-1\}$ 中，处理组和所有对照组的可观测结果都是 $Y^{\text{obs}}_{i:T_i<T} = Y_{0i:T_i<T}$。

我们可以选择使处理组观测结果和对照组观测结果的加权平均最相似的权重，即 $\boldsymbol{\lambda} = [\lambda_0, \ldots, \lambda_{G-1}]^\top$ 满足

$$\boldsymbol{\lambda} = \underset{\boldsymbol{\ell}:\boldsymbol{\ell}\geq\mathbf{0},\boldsymbol{\ell}^\top\boldsymbol{\iota}=1}{\arg\min} \left\| \begin{array}{c} \bar{Y}_{G0} - \sum_{g=0}^{G-1} \ell_g \bar{Y}_{g0} \\ \bar{Y}_{G1} - \sum_{g=0}^{G-1} \ell_g \bar{Y}_{g1} \\ \vdots \\ \bar{Y}_{G,T-1} - \sum_{g=0}^{G-1} \ell_g \bar{Y}_{g,T-1} \end{array} \right\|$$

$\|\cdot\|$ 为向量间的距离函数。也可以把可观测协变量加入上面的目标函数中。

当权重 $\lambda_i = 1/G$ 时，SC 估计量等于 DID 估计量。

Abadie, Diamond, & Hainmueller (2010)

Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105:490, 493-505.

- 核心问题：烟草税对香烟消费的影响

  1988年11月，美国加州通过了99号提按（Proposition 99，正式名称为 California Tobacco Tax and Health Protection Act of 1988），旨在通过提高烟草税（每盒烟增加 $0.25 的消费税）达到降低香烟消费的目的。这是美国现代史上第一个大规模控烟政策。该法案生效于1989年1月。

  99号提按不仅造成了广泛的社会影响，也带来了明显的控烟效果。截止到1999年，加州的青少年吸烟率降至全美最低，人均香烟消费量减少一半以上。

- 鉴于加州的成功，其他州也陆续通过了控烟法案，例如：

  马萨诸塞州：1993年，增加 $0.25/pack
  亚利桑那州：1994年，增加 $0.50/pack
  俄勒冈州：1996年，增加 $0.30/pack

Abadie, Diamond, & Hainmueller (2010)

Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105:490, 493-505.

- 数据：1970-2000 年间的年度州级面板数据

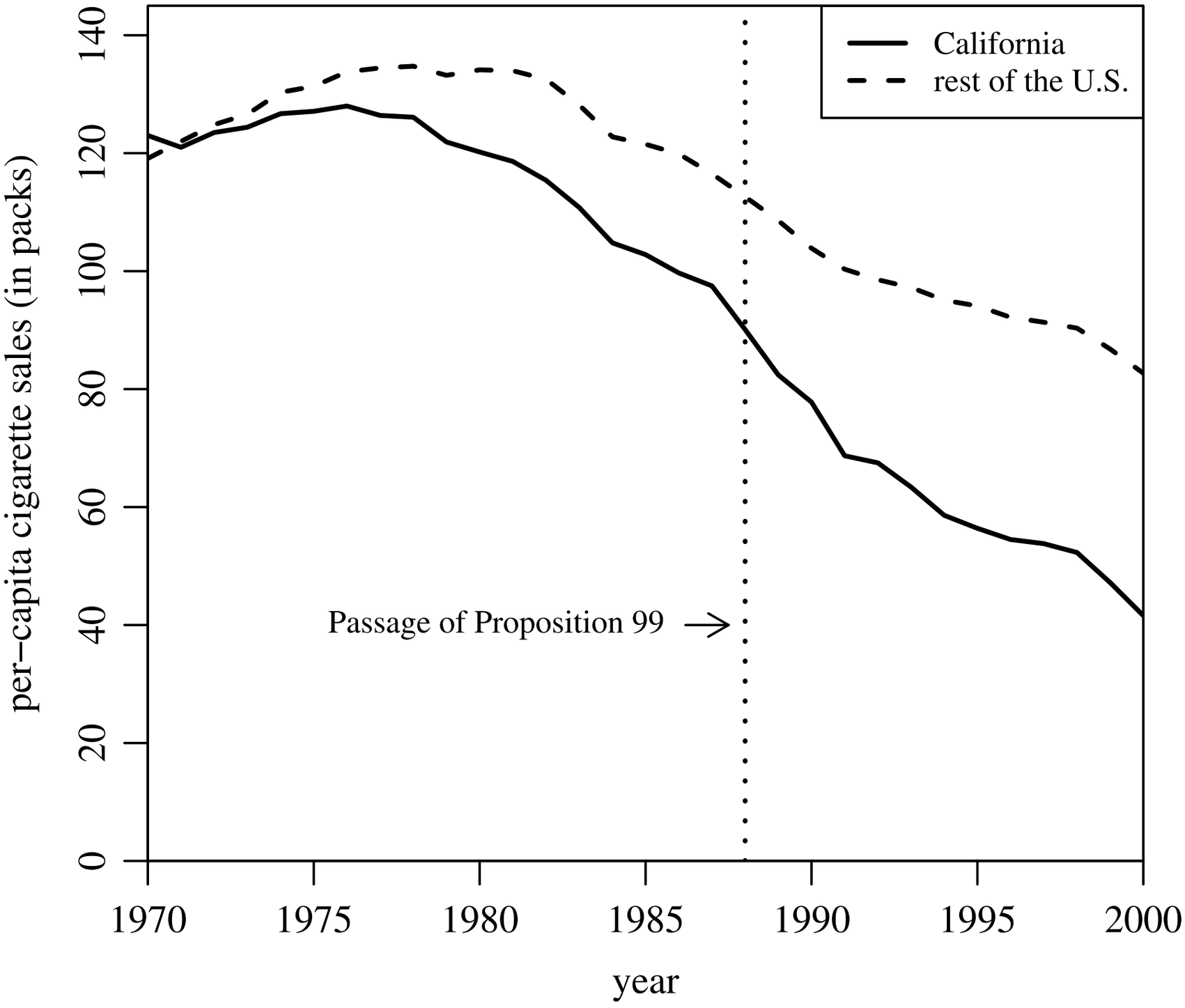
  州级香烟消费数据开始于1970年，而2000年后更多的州启动了控烟政策。

- 加州为处理组，控制组包括其他州，但除去了在1989-2000年间通过大规模控烟法案或大幅提升烟草税（超过 $0.5/pack）的州，也不包括华盛顿哥伦比亚特区：

  控烟法案：马萨诸塞州、亚利桑那州、俄勒冈州、佛罗里达州
  烟草税：阿拉斯加州、夏威夷州、马里兰州、密歇根州、新泽西州、纽约州、华盛顿州

  最终控制组中包含 38 个州

- 结果变量：州级年度人均香烟消费量（数据中体现为销售量、单位为盒）
- 协变量：香烟零售价格、人均收入、15-24岁人口比例、人均啤酒消费量等



从人均香烟销售量变化可以看出，加州与除加州外的美国之间不存在共同趋势，因此 DID 估计不合适

Figure 1. Trends in per-capita cigarette sales: California vs. the rest of the United States.

Table 1. Cigarette sales predictor means

Variables	California		Average of 38 control states
	Real	Synthetic	
Ln(GDP per capita)	10.08	9.86	9.86
Percent aged 15–24	17.40	17.40	17.29
Retail price	89.42	89.41	87.27
Beer consumption per capita	24.28	24.20	23.75
Cigarette sales per capita 1988	90.10	91.62	114.20
Cigarette sales per capita 1980	120.20	120.43	136.58
Cigarette sales per capita 1975	127.10	126.99	132.81

NOTE: All variables except lagged cigarette sales are averaged for the 1980–1988 period (beer consumption is averaged 1984–1988). GDP per capita is measured in 1997 dollars, retail prices are measured in cents, beer consumption is measured in gallons, and cigarette sales are measured in packs.

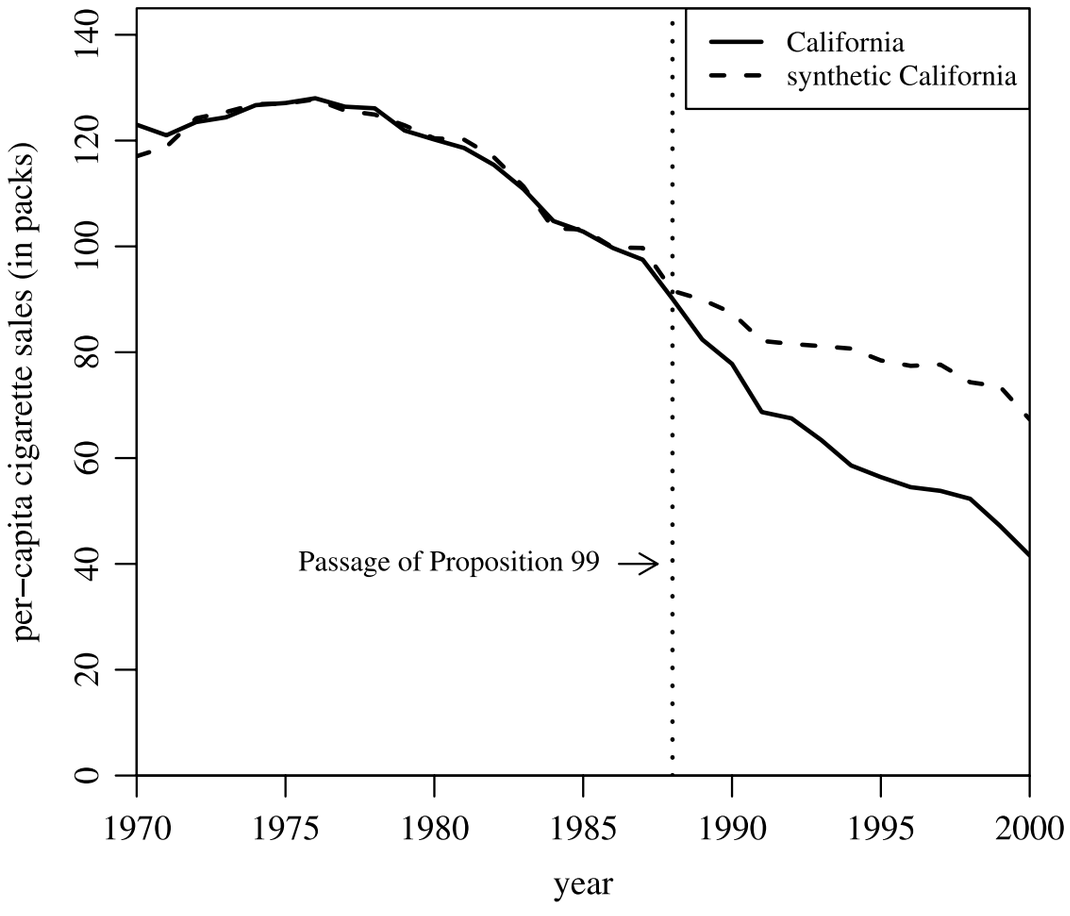
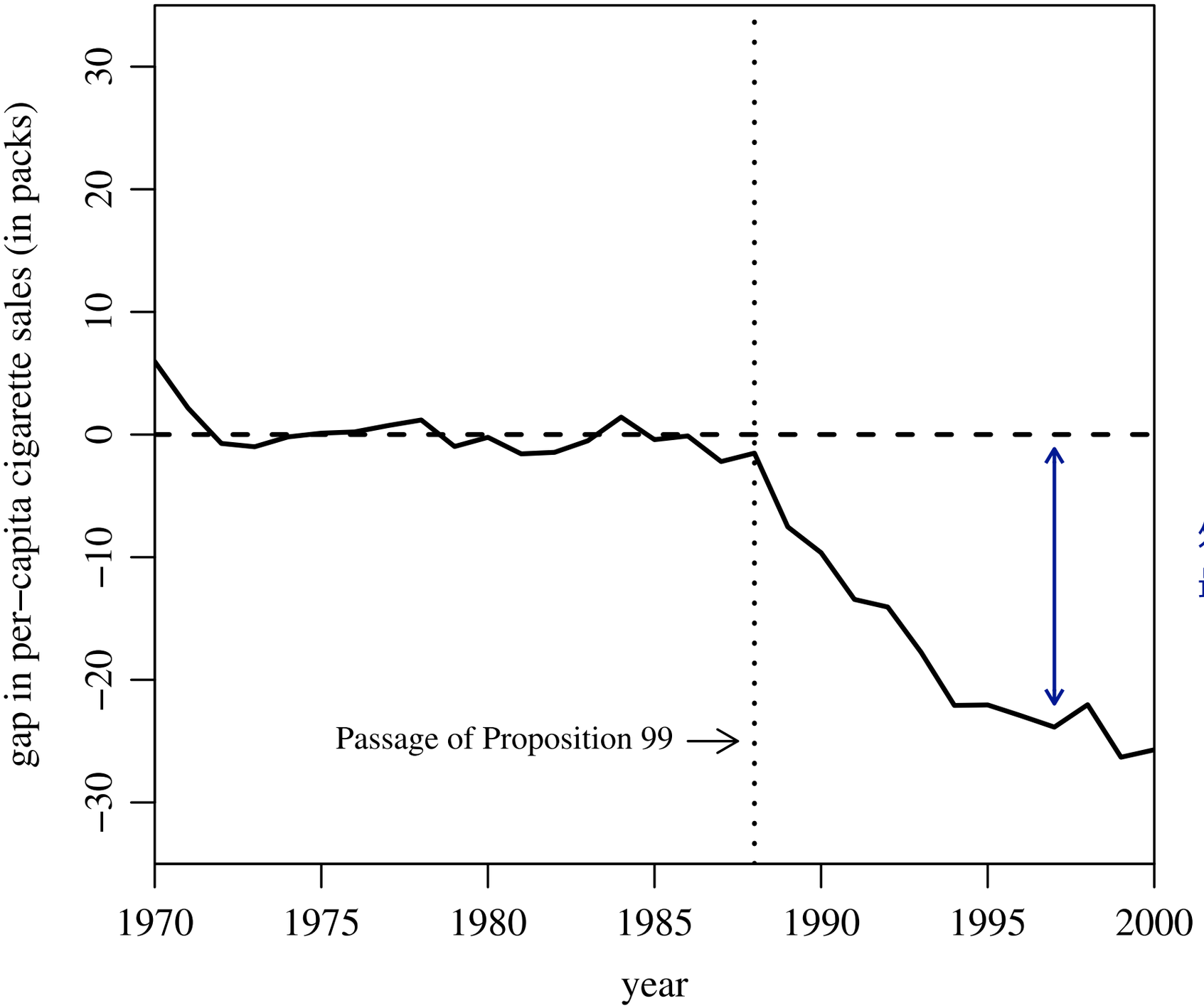


Figure 2. Trends in per-capita cigarette sales: California vs. synthetic California.

Table 2. State weights in the synthetic California

State	Weight	State	Weight
Alabama	0	Montana	0.199
Alaska	–	Nebraska	0
Arizona	–	Nevada	0.234
Arkansas	0	New Hampshire	0
Colorado	0.164	New Jersey	–
Connecticut	0.069	New Mexico	0
Delaware	0	New York	–
District of Columbia	–	North Carolina	0
Florida	–	North Dakota	0
Georgia	0	Ohio	0
Hawaii	–	Oklahoma	0
Idaho	0	Oregon	–
Illinois	0	Pennsylvania	0
Indiana	0	Rhode Island	0
Iowa	0	South Carolina	0
Kansas	0	South Dakota	0
Kentucky	0	Tennessee	0
Louisiana	0	Texas	0
Maine	0	Utah	0.334
Maryland	–	Vermont	0
Massachusetts	–	Virginia	0
Michigan	–	Washington	–
Minnesota	0	West Virginia	0
Mississippi	0	Wisconsin	0
Missouri	0	Wyoming	0



处理效应的 SC 估计值逐年增加，
最终达到人均减少 25 盒左右。

Figure 3. Per-capita cigarette sales gap between California and synthetic California.

结果是显著的吗？安慰剂检验

Is the Estimate Significant? A Placebo Test

本文中的安慰剂检验是指用对照组中的其他州替代加州作为处理组，然后进行同样的 SC 估计。如果估计结果和加州类似，说明对加州的估计结果不显著，反之则显著。

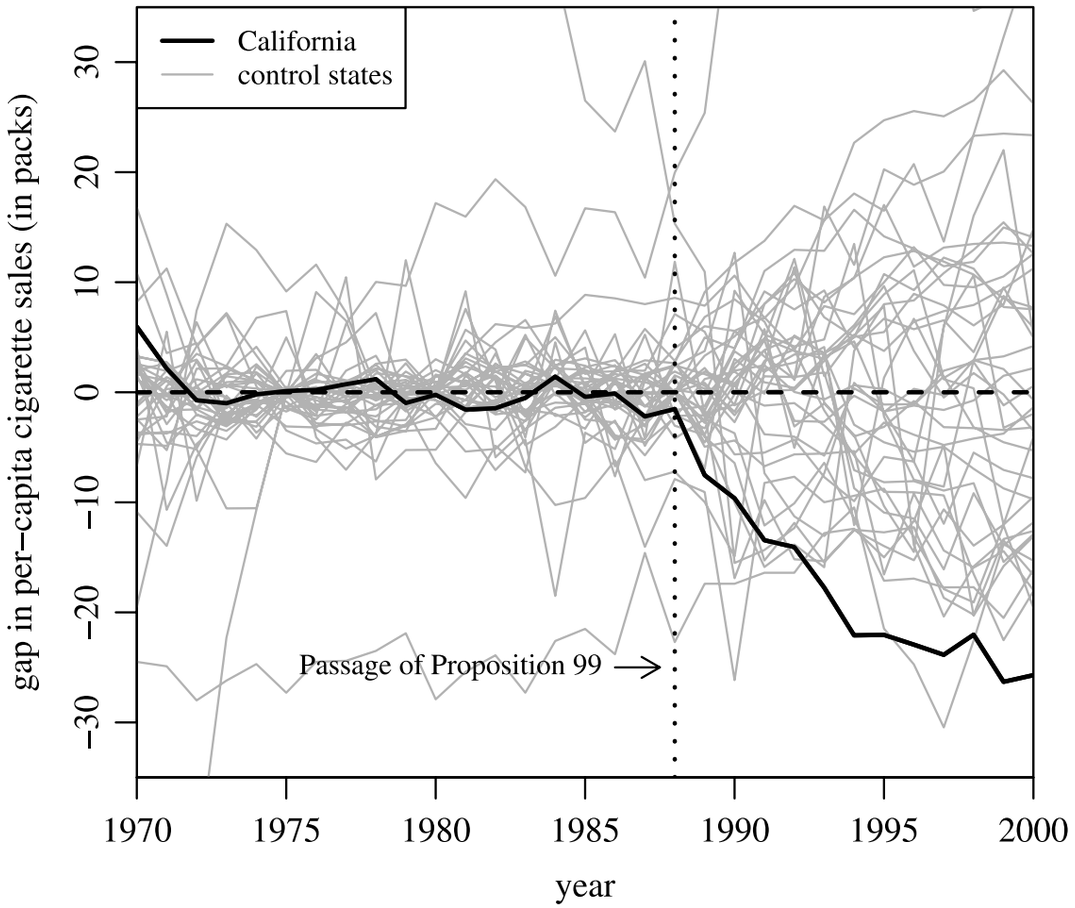


Figure 4. Per-capita cigarette sales gaps in California and placebo gaps in all 38 control states.

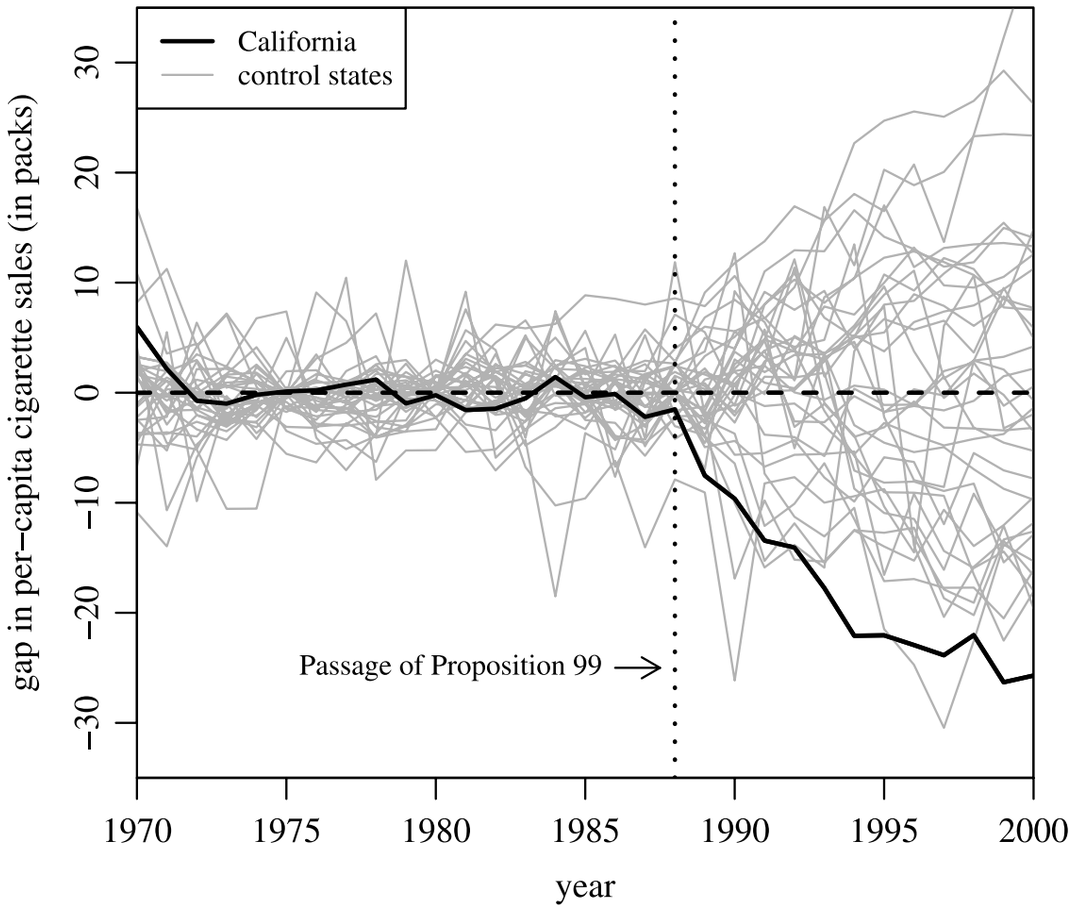


Figure 5. Per-capita cigarette sales gaps in California and placebo gaps in 34 control states (discards states with pre-Proposition 99 MSPE twenty times higher than California's).

课外阅读

- LaLonde, R. J. (1986).
  Evaluating the Econometric Evaluations of Training Programs with Experimental Data.
  American Economic Review, 76:4, 604-620.
  http://www.jstor.org/stable/1806062

- Dehejia, R. H. and Wahba, S. (1999).
  Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs.
  Journal of the American Statistical Association, 94:448, 1053-1062.
  https://www.tandfonline.com/doi/abs/10.1080/01621459.1999.10473858

- Gruber, J. (1994).
  The Incidence of Mandated Maternity Benefits.
  American Economic Review, 84:3, 622-641.
  http://www.jstor.org/stable/2118071