Please do not distribute without permission.

应用微观计量经济学

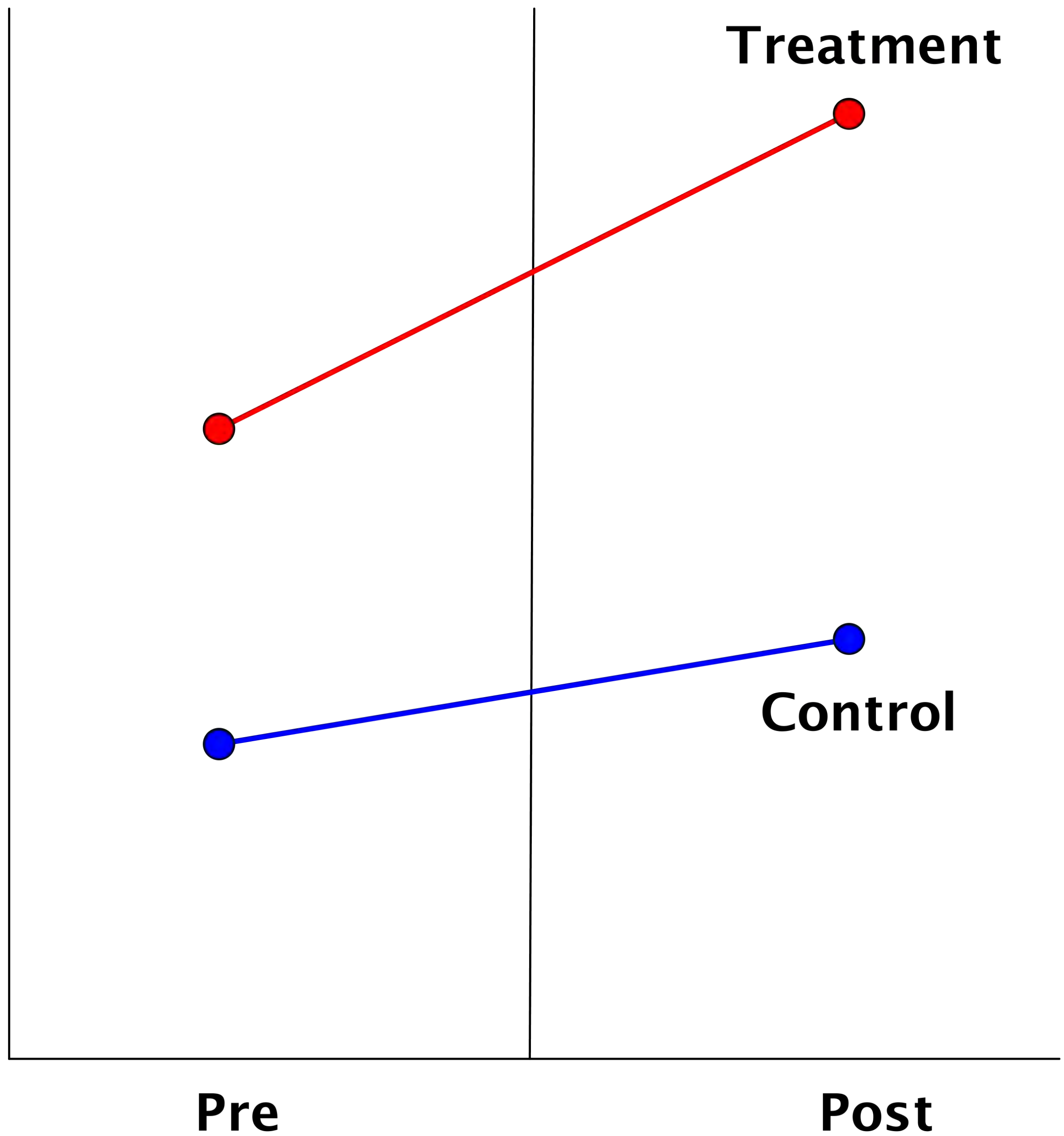
Applied Microeconometrics

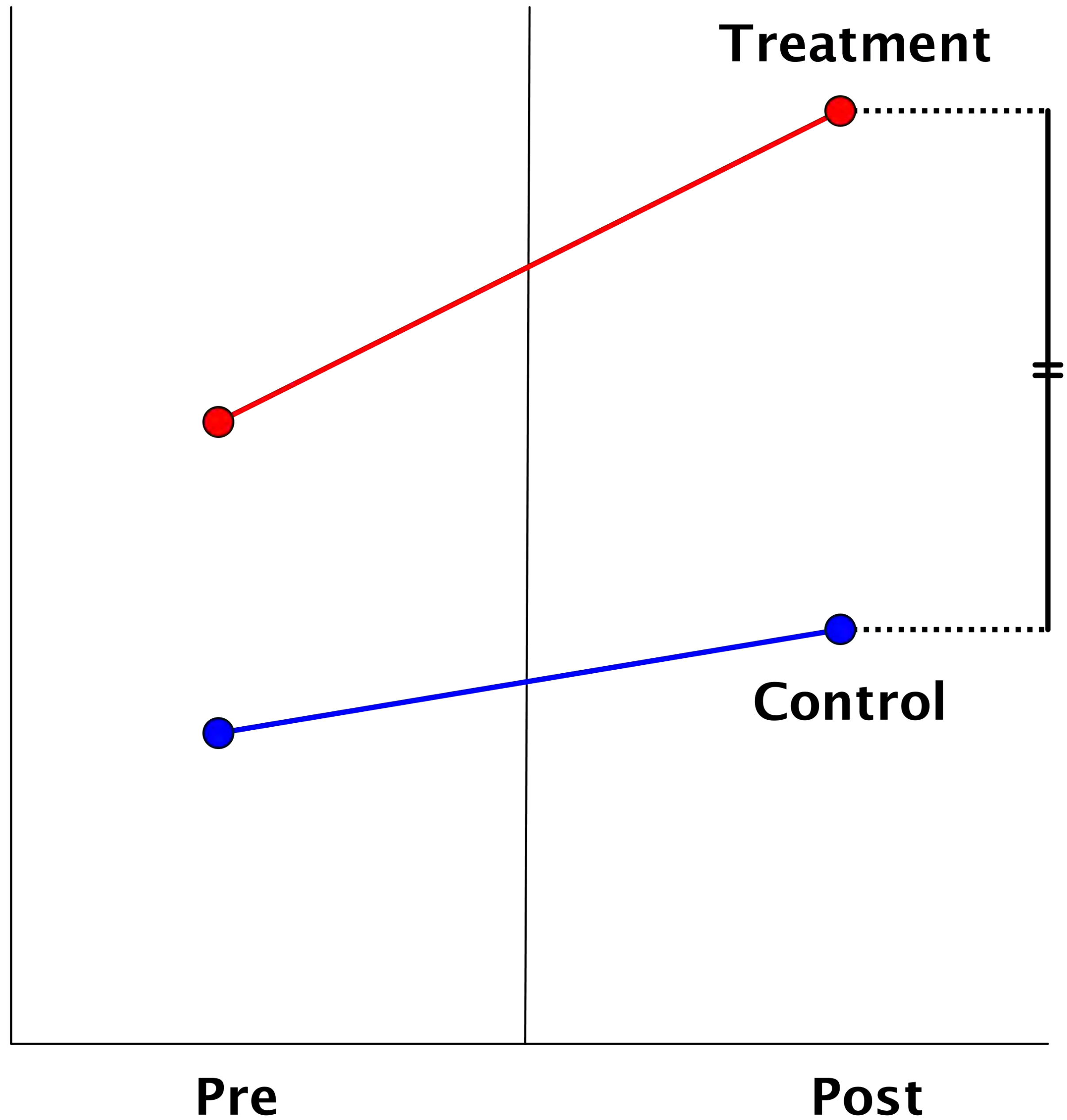
江　艇

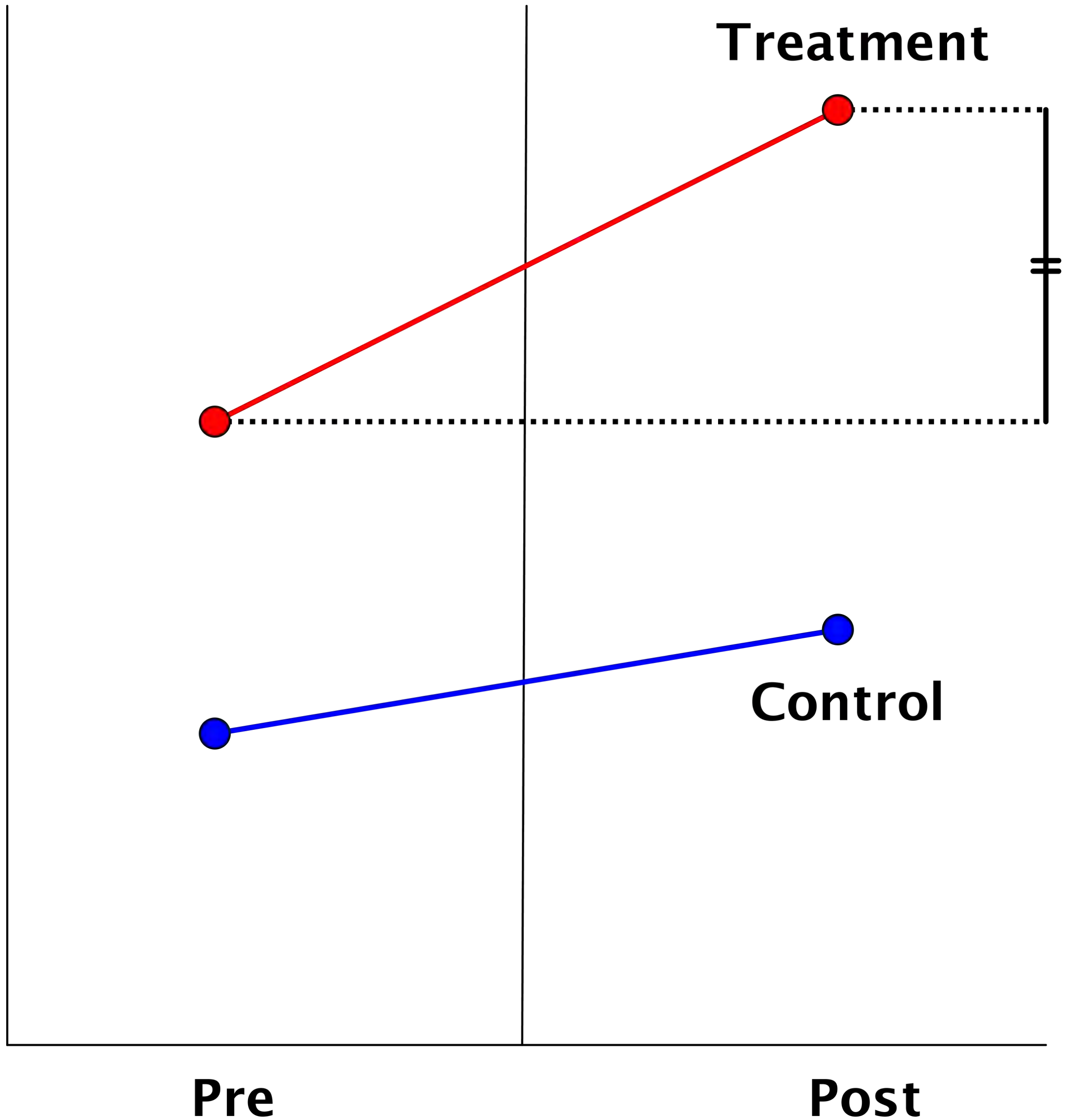
中国人民大学经济学院

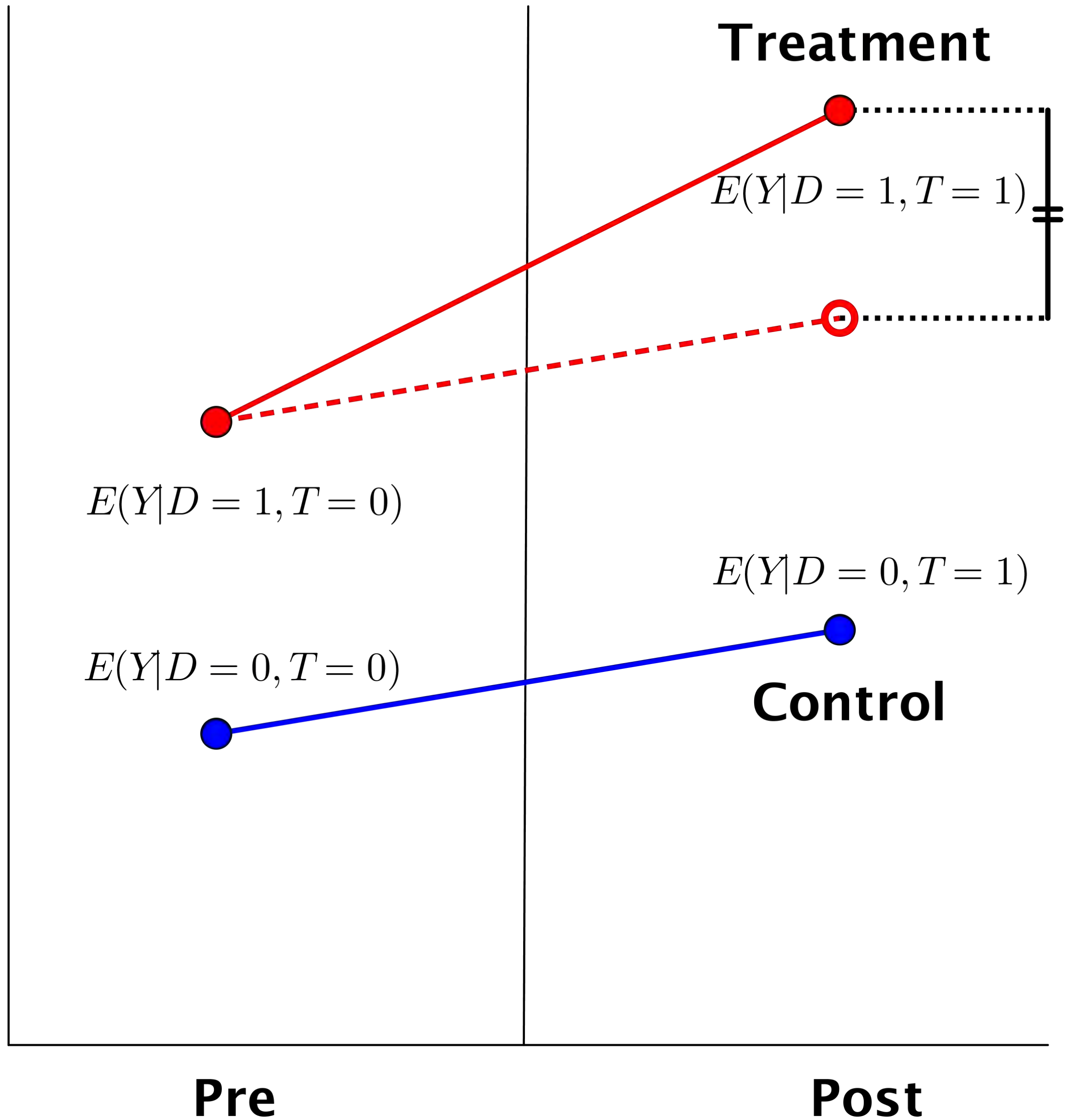
Last updated: April 22, 2019

Time:	周一晚 18:00-21:15
Venue:	明法 0502
Office:	明主 952
E-mail:	`ting.jiang@ruc.edu.cn`
Phone:	8250-0201
Grading:	期末闭卷考试


Treatment
Control
Pre
Post


Treatment
Control
Pre
Post


Treatment
Control
Pre
Post


Treatment
$E(Y|D=1,T=1)$
$E(Y|D=1,T=0)$
$E(Y|D=0,T=1)$
$E(Y|D=0,T=0)$
Control
Pre
Post

5.1 DID 的识别

$$
\begin{aligned}
\tau(X) =& E(Y|X, D=1, T=1) - E(Y|X, D=1, T=0) \\
& - \left[E(Y|X, D=0, T=1) - E(Y|X, D=0, T=0)\right] \\
=& E(Y^1|X, D=1, T=1) - E(Y^0|X, D=1, T=0) \\
& - \left[E(Y^0|X, D=0, T=1) - E(Y^0|X, D=0, T=0)\right] \\
=& \left[E(Y^1|X, D=1, T=1) - E(Y^0|X, D=1, T=1)\right] \\
& + \left[E(Y^0|X, D=1, T=1) - E(Y^0|X, D=1, T=0)\right] \\
& - \left[E(Y^0|X, D=0, T=1) - E(Y^0|X, D=0, T=0)\right]
\end{aligned}
$$

- 识别假设 III：

$$
\begin{aligned}
&E(Y^0|X, D=1, T=1) - E(Y^0|X, D=1, T=0)\\
=&E(Y^0|X, D=0, T=1) - E(Y^0|X, D=0, T=0)
\end{aligned}
$$

- 在此假设下，DID 估计量识别了处理组接受处理后的平均处理效应 (ATT at the post-treatment period).

$$
\tau_{DID} = E\Big(Y^1 - Y^0\Big|D=1, T=1\Big) = \int \tau(X) dF(X|D=1, T=1)
$$

- 识别假设 III 对分配机制的要求是什么？(DID 要求随机分组么？)
  识别假设 III 可以简写作
  $$E\big(\Delta Y^0|X, D=1\big) = E\big(\Delta Y^0|X, D=0\big)$$
  也就是说给定 X，ΔY^0 均值独立于 D，即关于 ΔY^0 是随机分组的！
- 识别假设 III 比识别假设 II 更弱么？
  回忆识别假设 II：
  $$E\big(Y^0|X, D=1\big) = E\big(Y^0|X, D=0\big)$$
  – 从数学上说，答案显然是否定的。
  – 识别假设 III 可以（部分）检验。但识别假设 II 也可以（找到某个 pseudo-outcome，比如 lagged outcome）。

5.2 DID 的估计

- 重复横截面数据

$$y_{it} = \tau D_i \times T_t + \mathbf{x}'_{it}\gamma + \rho D_i + \eta_t + \varepsilon_{it}$$

- 面板数据

$$y_{it} = \tau D_i \times T_t + \mathbf{x}'_{it}\gamma + u_i + \eta_t + \varepsilon_{it}$$

- 对于两期面板，等价于一阶差分

$$\Delta y_i = \tau D_i + \mathbf{x}'_i\gamma + \varepsilon_i$$

- 对于多期面板，可以采用更灵活的形式

$$y_{it} = \sum_{l=2}^{T} \tau_l (D_i \times T_t^l) + \mathbf{x}_{it}'\gamma + u_i + \eta_t + \varepsilon_{it}$$

$$T_t^l = \begin{cases} 1, & t = l \\ 0, & t \neq l \end{cases}$$

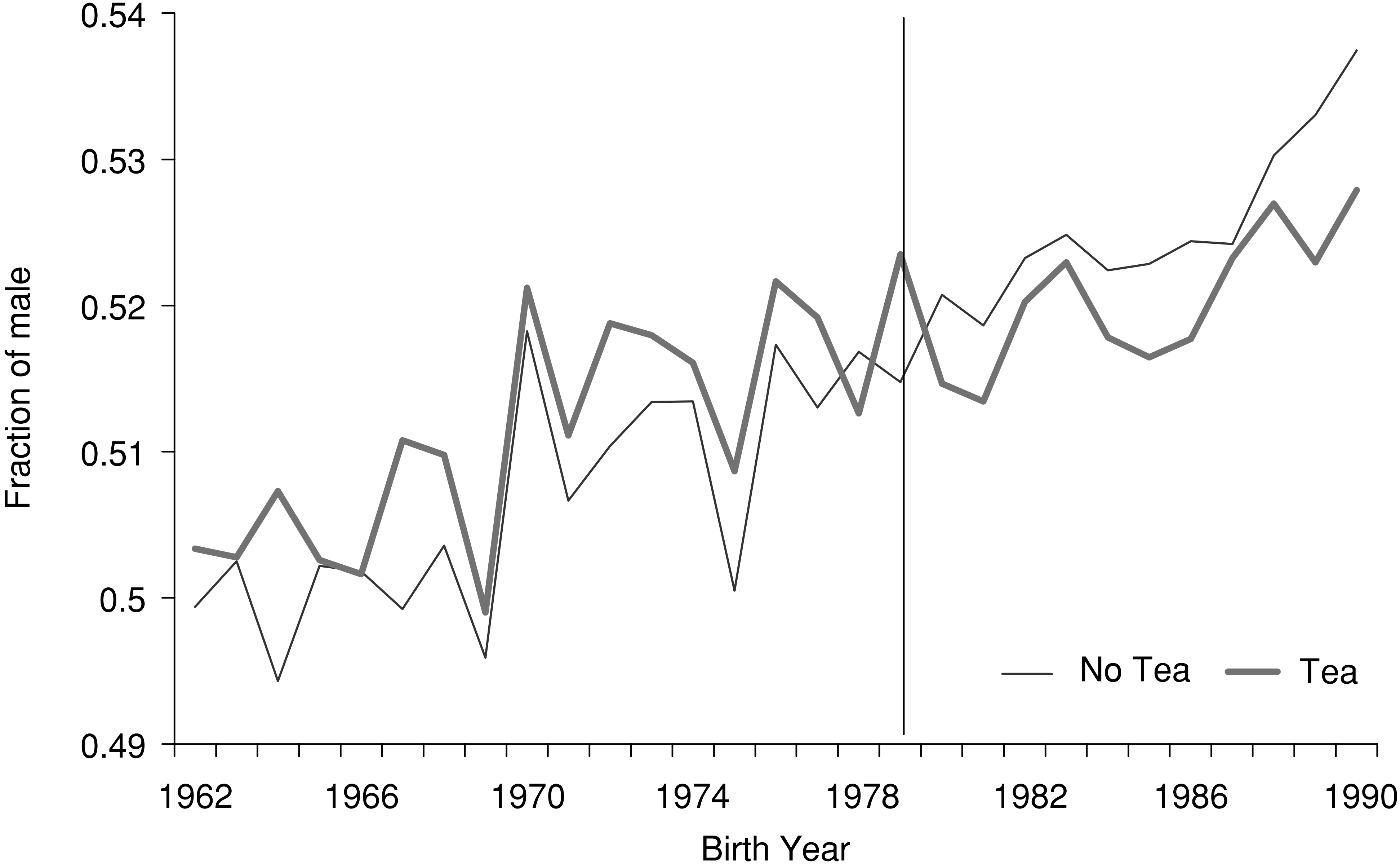
示例 1. 茶叶的价格与消失的女性 (Qian, 2008, *QJE*).

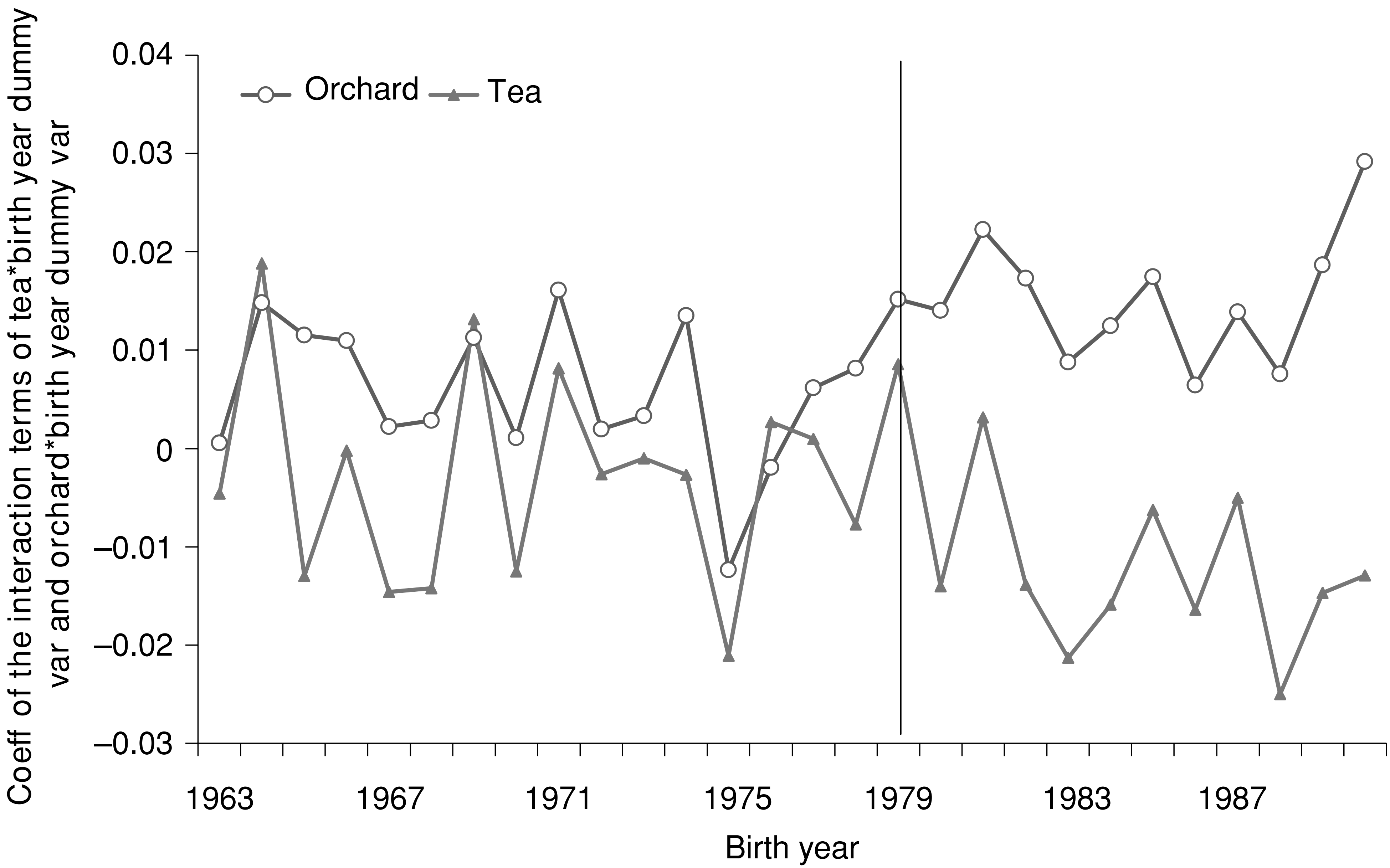
示例 2. 土豆与人口增长 (Nunn and Qian, 2011, *QJE*).

- 当处理组个体接受处理时间不一致时，直接构造表示接受处理的虚拟变量。

$$y_{it} = \tau D_{it} + \mathbf{x}_{it}'\gamma + u_i + \eta_t + \varepsilon_{it}$$

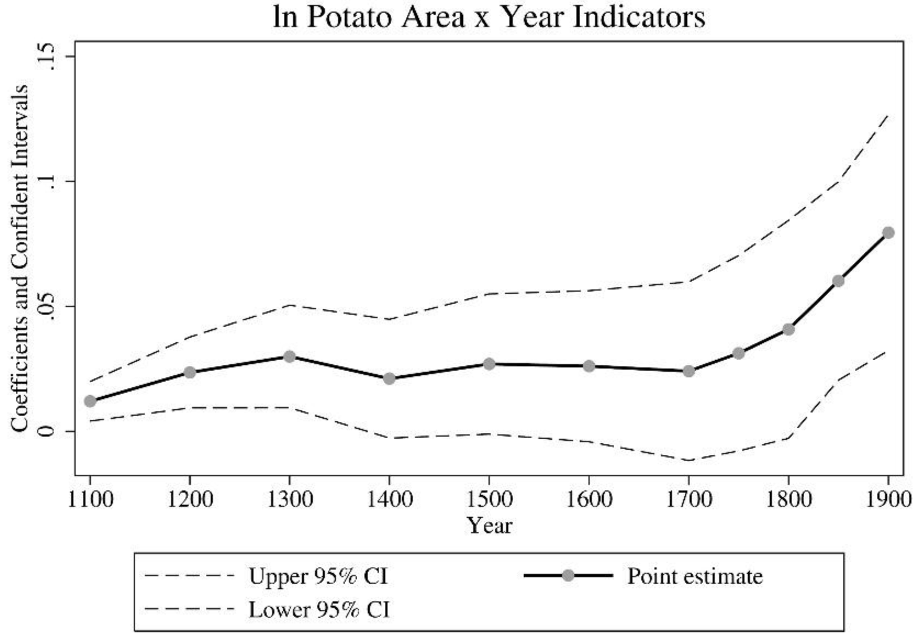
$$D_{it} = \begin{cases} 1, & \text{个体 } i \text{ 在第 } t \text{ 期接受处理} \\ 0, & \text{其它情形} \end{cases}$$


0.54
0.53
0.52
0.51
0.5
0.49
Fraction of male
1962
1966
1970
1974
1978
1982
1986
1990
Birth Year
No Tea
Tea

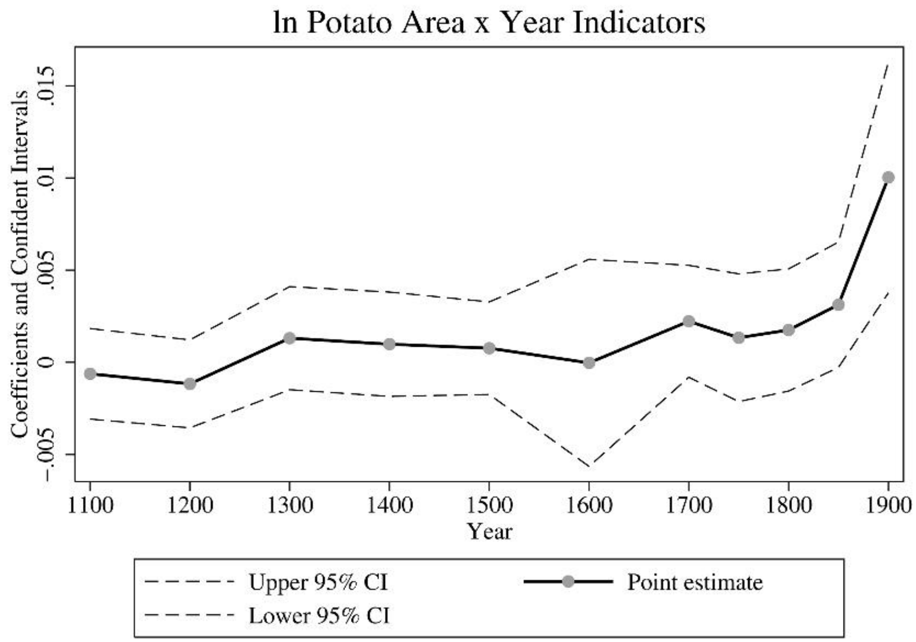

Orchard
Tea
Coeff of the interaction terms of tea*birth year dummy var and orchard*birth year dummy var
0.04
0.03
0.02
0.01
0
–0.01
–0.02
–0.03
1963
1967
1971
1975
1979
1983
1987
Birth year

Flexible Estimates: The Relationship Between Potato-Suitable Land Area and Population or City Population Share by Time Period

	Dependent Variable					
	ln total population			City population share		
	(1)	(2)	(3)	(4)	(5)	(6)
ln *Potato-Suitable Area* × 1100	0.013	0.011	0.012	–0.0018	–0.0013	–0.0006
	(0.003)	(0.003)	(0.004)	(0.0014)	(0.0009)	(0.0013)
ln *Potato-Suitable Area* × 1200	0.029	0.024	0.024	–0.0011	–0.0013	–0.0012
	(0.005)	(0.005)	(0.007)	(0.0009)	(0.0009)	(0.0012)
ln *Potato-Suitable Area* × 1300	0.039	0.031	0.030	0.0002	–0.0005	0.0014
	(0.007)	(0.007)	(0.010)	(0.0008)	(0.0011)	(0.0014)
ln *Potato-Suitable Area* × 1400	0.019	0.004	0.021	0.0008	0.0002	0.0010
	(0.008)	(0.008)	(0.012)	(0.0012)	(0.0015)	(0.0014)
ln *Potato-Suitable Area* × 1500	0.034	0.014	0.027	0.0003	–0.0002	0.0008
	(0.009)	(0.010)	(0.014)	(0.0009)	(0.0012)	(0.0013)
ln *Potato-Suitable Area* × 1600	0.041	0.021	0.026	0.0002	–0.0010	–0.0000
	(0.009)	(0.011)	(0.015)	(0.0014)	(0.0025)	(0.0029)
ln *Potato-Suitable Area* × 1700	0.043	0.018	0.024	0.0020	0.0017	0.0022
	(0.012)	(0.013)	(0.018)	(0.0010)	(0.0013)	(0.0015)
ln *Potato-Suitable Area* × *1750*	*0.055*	*0.030*	*0.031*	*0.0015*	*0.0011*	*0.0013*
	(0.012)	*(0.014)*	*(0.020)*	*(0.0009)*	*(0.0013)*	*(0.0018)*
ln *Potato-Suitable Area* × *1800*	*0.073*	*0.048*	*0.041*	*0.0020*	*0.0016*	*0.0018*
	(0.014)	*(0.015)*	*(0.022)*	*(0.0009)*	*(0.0013)*	*(0.0017)*
ln *Potato-Suitable Area* × *1850*	*0.095*	*0.069*	*0.060*	*0.0024*	*0.0022*	*0.0031*
	(0.015)	*(0.017)*	*(0.020)*	*(0.0011)*	*(0.0014)*	*(0.0017)*
ln *Potato-Suitable Area* × *1900*	*0.121*	*0.092*	*0.080*	*0.0118*	*0.0123*	*0.0100*
	(0.017)	*(0.021)*	*(0.024)*	*(0.0023)*	*(0.0024)*	*(0.0032)*
Baseline Controls (× Year fixed effects):						
ln *Old World Crops Area*	N	Y	Y	N	Y	Y
ln *Elevation*	N	N	Y	N	N	Y
ln *Ruggedness*	N	N	Y	N	N	Y
ln *Tropical Area*	N	N	Y	N	N	Y
Observations	1552	1552	1552	1552	1552	1552
R-squared	0.99	0.99	0.99	0.42	0.42	0.46
F Stat for Joint Signficance 1750–1900	17.88	13.60	4.20	8.02	8.82	4.89



(a) ln Total Population



(b) City Population Share

5.3 对 DID 的威胁

- $D \cdot T$ 反映的可能是其它处理效应。
- 可能不满足 SUTVA 假设。
- 不好的控制组。
  - 是否进入处理组取决于处理之前的结果差异。例如 "Ashenfelter's dip": selection on idiosyncratic temporary shocks (Ashenfelter, 1978, *REStat*).
  - 处理组和控制组越接近越好，检验协变量 X 的差异。
  - 检验平行趋势。平行趋势本质上是不可检验的——以前平行不见得今后就平行。如果政策干预不是随机的，而是由某因素 W 所决定，那么这个因素 W 在影响政策干预的同时，很有可能会影响趋势。如果 W 是可观测的，就要尽量控制起来，甚至控制它与时间虚拟变量的交互项。控制变量的变化在处理组和控制组间存在明显差异，可能意味着政策变化是内生的，某个经济条件在控制组不存在，这一不可观测的经济条件可能产生遗漏变量偏误。

- 检查结果对于个体特定的线性趋势的加入是否敏感。

$$y_{it} = \tau D_i \times T_t + \mathbf{x}'_{it}\gamma + u_i + \eta_t + u_i \times t + \varepsilon_{it}$$

- 如果处理组和控制组之间存在明显差异，对于不同的控制组选择和不同的方程设定，结果需保持稳健。
- 处理组和控制组在处理前后存在成分变化 (compositional changes)。随着处理的实施，分配机制发生变化，从而改变处理组的规模和构成。此时平行趋势假定尤其可疑。
  - ▷ 情形一：一些个体在样本区间内反复进入和退出处理组。
  - ▷ 情形二：处理实施是在（例如）地区层面，但处理针对的是个体。
  - ▷ 如前所述，控制个体特征（甚至是与时间虚拟变量的交互项）一定程度上能缓解这一问题。
  - ▷ 如果回归本身就在地区层面，可能要深入考察政策变化的决定因素，以此来选择恰当的随时间变化的地区控制变量，甚至考虑政策干预的工具变量。

Another problem of the empirical strategy is that if, at the time of the reforms, there is a change in the attitudes that drive sex preference in tea-planting counties, then the estimate of the effect of planting tea will capture both the relative female income effect and the effect of the attitude change. Or, if the increase in the value of tea changed the reason for women to pick tea, then the prereform cohort will not be an adequate control group. Although I cannot resolve the former problem, the latter is addressed by instrumenting for tea planting with time-invariant geographic data.[14]

$$\begin{aligned} \text{tea}_i \times \text{post}_c = {} & (\text{slope}_i \times \text{post}_c)\lambda + (\text{cashcrop} \times \text{post}_c)\varphi \\ & + \text{Han}_{ic}\zeta + \alpha + \psi_i + \text{post}_c\gamma + \varepsilon_{ic}. \end{aligned} \tag{4}$$

The second-stage regression is as follows:

$$\begin{aligned} \text{sex}_{ic} = {} & (\text{tea}_i \times \text{post}_c)\beta + (\text{cashcrop} \times \text{post}_c)\varphi \\ & + \text{Han}_{ic}\zeta + \alpha + \psi_i + \text{post}_c\gamma + \varepsilon_{ic}. \end{aligned} \tag{5}$$

5.4 敏感性分析

- 假的处理时点。

THE IMPACT OF THE POTATO WITH ALTERNATIVE CUT-OFFS

	Placebo Treatment Periods											
	1200–1500; Post = 1400, 1500		1300–1600; Post = 1500, 1600		1400–1700; Post = 1600, 1700		1500–1800; Post = 1700, 1800		1600–1900; Post = 1800, 1850, 1900		1600–1900; Post = 1750, 1800, 1850, 1900	
	ln pop.	City share	ln pop.	City share	ln pop.	City share	ln pop.	City share	ln pop.	City share	ln pop.	City share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ln *Potato*	–0.002	0.0008	0.001	–0.0008	0.001	0.0002	0.006	0.0014	0.033	0.0038	0.028	0.0030
Area × Post	(0.006)	(0.0012)	(0.005)	(0.0017)	(0.006)	(0.0016)	(0.007)	(0.0017)	(0.011)	(0.0015)	(0.009)	(0.0017)
Observations	516	516	518	518	520	520	650	650	780	780	780	780
R-squared	0.99	0.56	0.99	0.56	0.99	0.59	0.99	0.59	0.99	0.62	0.99	0.62

- 假的结果（明显不受处理影响的结果）。

内容提要　本文从 IPO 定价效率的视角探讨了地理位置和新型交通工具——高铁的修建对经济运行效率的影响。投资者与公司信息不对称造成了 IPO 折价，而企业的地理位置是反映信息不对称程度的重要方面。本文研究了 A 股上市公司是否位于“中心地带”对公司 IPO 折价率的影响，发现与三大中心城市距离更近的公司，IPO 折价率更低；经营不确定性程度更高的企业，地理位置对 IPO 折价率的影响更大。高铁带来的信息沟通便利弥补了地理距离对 IPO 定价的影响，降低发行价格相对于真实价值的扭曲，提高资本市场定价效率。通过构建“虚拟高铁站”进行安慰剂检验，并以 IPO 公司所在城市与中央政府政策规定的高铁规划线之间的距离构建工具变量，进一步证实了我们的结论。

尽管如此,在这一部分中,我们依然通过一个安慰剂检验(placebo test) 解决这一潜在的内生性问题。我们的研究子样本针对企业上市年份尚未有 50 公里以内高铁站的公司(即只保留哑变量 $Dum\_train = 0$ 的公司) 。针对这个子样本,我们构建一个“虚拟高铁站”哑变量($Placebo\_train$) ,如果上市后三年内,公司所在地 50 公里以内修建了高铁站,则该哑变量取值为 1,否则取值为 0。具体而言,$Placebo\_train$ 取值为 1 的企业,在 IPO 后的一段时间中被纳入高铁轨道布局,而在 IPO 当年,其附近事实上并未有高铁站,因而投资者并没有享有高铁的交通便利所带来的信息优势,我们可以推断: 如

- 不同的控制组。

示例 3. 最低工资与就业 (Card and Krueger, 1994, *AER*).

$$\text{(1a)} \qquad \Delta E_i = a + \mathbf{b}\mathbf{X}_i + c\,\mathrm{NJ}_i + \varepsilon_i$$

or

$$\text{(1b)} \qquad \Delta E_i = a' + \mathbf{b}'\mathbf{X}_i + c'\,\mathrm{GAP}_i + \varepsilon_i'$$

$\mathrm{GAP}_i = 0$ for stores in Pennsylvania

$= 0$ for stores in New Jersey with $W_{1i} \geq \$5.05$

$= (5.05 - W_{1i}) / W_{1i}$

for other stores in New Jersey.

Since GAP_i varies within New Jersey, it is possible to add both GAP_i and NJ_i to the employment model. The estimated coefficient of the New Jersey dummy then provides a test of the Pennsylvania control group. When we estimate these models, the coefficient of the New Jersey dummy is insignificant (with t ratios of 0.3–0.7), implying that inferences about the effect of the minimum wage are similar whether the comparison is made across states or across stores in New Jersey with higher and lower initial wages.

- 假的处理组。

A final specification check is presented in row 12 of Table 5. In this row we exclude stores in New Jersey and (incorrectly) define the GAP variable for Pennsylvania stores as the proportional increase in wages necessary to raise the wage to \$5.05 per hour. In principle the size of the wage gap for stores in Pennsylvania should have no systematic relation with employment growth. In practice, this is the case. There is no indication that the wage gap is spuriously related to employment growth.

5.5 DID 与匹配方法的结合

- 先通过匹配方法构造控制组，然后进行 DID 估计。
  - $1:1$ 匹配的样本，进行简单 OLS 回归。
  - $1:m$ 匹配的样本，进行加权的简单 OLS 回归。

    示例 4. 管制的成本 (Cicala, 2015, *AER*).
- 先差分，然后对差分结果进行匹配估计。

  示例 5. 排污权交易的环境后果 (Fowlie et al., 2012, *AER*).

Table 4—Average Treatment Effect using Nearest Neighbors Matching

	Levels	Logs	RECLAIM facilities	Controls
Panel A. Change in NO_x emissions between periods 1 and 4				
OLS	−32.58**	−0.30***	212	1,222
	(13.77)	(0.10)		
Nearest neighbor matching (base specification)	−20.59***	−0.25***	212	1,222
	(7.63)	(0.09)		
Nearest neighbor matching (alternative specification)	−18.12	−0.11	211	1,191
	(11.51)	(0.08)		
Nearest neighbor matching (restricted sample)	−14.16**	−0.20**	199	1,222
	(6.86)	(0.09)		
Panel B. Change in NO_x emissions between periods 2 and 3				
OLS	−6.84	−0.22***	255	1,577
	(6.65)	(0.04)		
Nearest neighbor matching (base specification)	−8.29**	−0.26***	255	1,577
	(3.85)	(0.06)		
Nearest neighbor matching (alternative specification)	−6.18	−0.16***	252	1,493
	(5.06)	(0.06)		
Nearest neighbor matching (unrestricted sample)	−6.37	−0.23***	268	1,577
	(4.57)	(0.06)		

Notes: We define periods as averages of positive emissions in two years: 1990 and 1993 (period 1); 1997–1998 (period 2); 2001–2002 (period 3); and 2004–2005 (period 4). All observations are from historic nonattainment counties. The OLS estimates control for average NO_x emissions during period 1 and four-digit SIC code indicator variables, with standard errors clustered by air basin. For all semiparametric matching, we match on the three closest neighbors with linear bias adjustment in levels and quadratic bias adjustment in logs. The baseline nearest neighbor matching model matches on historic emissions and exactly on four-digit SIC codes. In the alternative specification, industry-specific emissions quartile indicators are added to the exact matching variables; predetermined demographic characteristics (race and income) are added to the matching variables. Panel A's restricted sample omits 13 facilities removed from the program in 2001. Panel B's unrestricted sample includes these facilities. For the log specifications, emissions differences are defined as $\ln(\text{EmitX} + 1) - \ln(\text{Emit1} + 1)$, and all matching is on $\ln(\text{Emit1} + 1)$. Standard errors are reported in parentheses.