内生性问题：处理方法与进展

连玉君

中山大学 岭南学院

电邮：arlionn@163.com

2017-09-17

提纲

- 公司金融中的内生性问题：如此之多！
- 内生性问题的来源
  - 遗漏变量（模型设定偏误）
  - 衡量偏误(变量的衡量)
  - 联立方程组（双向因果）
- 内生性问题的处理方法
  - IV-GMM
  - 面板数据模型(Panel Data)
  - Heckman 选择模型、Treatment effect 模型
  - 倍分法 (DID)、倾向得分匹配分析 (PSM)
  - 准自然实验：断点回归设计 (RDD)
  - 合成控制法 (SCM)
  - 结构方程模型 (SEM)

投稿时，我怕被问及“内生性问题”；

审稿时，我乐于问及

内生性问题：如此之多！

- 一些值得考虑的问题
  - 相关关系 ⇔ 因果关系?
  - 自然实验
- 一些潜伏着内生问题的研究主题
  - 教育水平对收入的影响(遗漏变量，ability，fat)
  - 经营绩效与社会责任（因果关系不明朗）
  - 投资-现金流敏感性（衡量偏误）
  - 在职培训与工资水平 (self-selection)
  - 建立政治关联有助于改善公司业绩吗？ (self-selection)

何谓内生性？

- 内生性：在回归分析中，干扰项和解释变量相关

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

- 回顾：确保估计量具有一致性的条件
  - 随机抽样 $(y, x_1, x_1, \cdots, x_k)$
  - 满秩 $rank(X'X) = k$
  - 外生 $Cov(X, \varepsilon) = 0$ *or* $E[\varepsilon | x_1, x_1, \cdots, x_k] = 0$
- 内生性的后果
  - 统计角度而言：OLS (MLE) 估计结果有偏 (不是我们想要的结果)
  - 实践角度而言：经验结果存在多种可能的解释 (并非“因果”推断)

    审稿人可以提出多种可能导致你的实证结果的解释

多数人的处理方法：

摆 Pose !

内生性问题的可能来源

- 互为因果
  - 资本结构、投资行为、现金持有、Tobin's Q
- 遗漏变量 \ 数据缺失
  - 理论分析和前期文献中提到的重要变量
  - 自我选择偏误
- 衡量偏误
  - Fazzari et al. (1988, JEL): 投资−现金流敏感性

$$Invest_{it}=\alpha_i+\beta_1 Q_{it}+\beta_2 CashFlow_{it}+\varepsilon_{it}$$

Refs：Fazzari et al. (1988) |**JEL**|，Kaplan and Zingales (1997) |**QJE**|，Fazzari et al. (2000) |**QJE**|，Kaplan and Zingales (2000) |**QJE**|，Erickson and Whited (2000) |**JPE**|，Alti (2003) |**JF**|

遗漏变量\数据缺失

Omitted Variable or missing value bias: 简介

$$True: \quad y = \alpha + \beta_1 x_1 + \boxed{\beta_2 x_2 + u_1}$$

$$Estimate: \quad y = \alpha + \beta_1 x_1 + u_2 \qquad (?)$$

$$if \ Corr(x_2, x_1) \neq 0, \ then \ Corr(u_2, x_1) \neq 0 \quad \text{Endog!}$$

- 评论：
  - 多数情况下，遗漏变量是我们的 |无奈之举|
  - 更多的情况下，我们都表现为 |过度自信| 或 |掩耳盗铃|
- 解决方法：
  - 尽量使用“丰满”一点的模型(要熟悉相关理论和文献)
  - IV or GMM (如何找？)

遗漏变量

Omitted Variable bias: 一个例子

- 教育的回报率

$$Income_i = \alpha + \beta_1 Controls_i + \beta_2 Education_i + \varepsilon_i$$

  - Q1: 是否存在内生性问题？
  - A1: Maybe yes, maybe not.
    - C1: 遗漏变量 (能力-马云、肥胖-相扑)
    - C2: 是否读大学？ (0/1)
  - Q2: 怎么办？
  - A2: IV—— 父母教育水平; 住所离大学的距离; 入学年份;
  - A2: RDD, 高考录取分为600分，可以对比分析599分和601分两组的差异
  - A3: PSM, 找到那些与上大学的学生特征相似的落榜生作为对照组

衡量偏误

Measurement Error (ME): 简介

$$True\ model: \quad y = \alpha + \beta x^{*} + u$$

$$Empirical\ model: \quad y = \alpha + \beta x + \upsilon$$

$$x = x^{*} + \varepsilon, \quad E[\varepsilon] = 0$$

$$\begin{aligned} y &= \alpha + \beta x^{*} + u \\ &= \alpha + \beta(x - \varepsilon) + u \\ &= \alpha + \beta x + \underline{(u - \beta\varepsilon)} \\ &= \alpha + \beta x + \upsilon \end{aligned}$$

$$Cov(x, \varepsilon) \neq 0 \ ?$$

- Stata commands: eivreg | sem | logitem | simex | cme | Ewreg | XTEWreg

衡量偏误

Measurement Error (ME): 一场争论

- 融资约束假说与投资–现金流敏感性
  - Fazzari et al. (1988) |**JEL**|， Kaplan and Zingales (1997) |**QJE**|，
  - Fazzari et al. (2000) |**QJE**|， Kaplan and Zingales (2000) |**QJE**|，
  - Erickson and Whited (2000) |**JPE**|， Alti (2003) |**JF**|，
  - Erickson and Whited (2012) |**RFS**|

$$\left(\frac{I_{it}}{K_{it-1}}\right) = \beta_0 + \beta_1 Q_{it} + \beta_2 \left(\frac{CF_{it}}{K_{it-1}}\right) + \varepsilon_{it}$$

- T. Whited 的处理方法:
  - Higher Order Moments GMM (HGMM) | Signs Estimator (SigE)
  - Erickson and Whited(2012) |**RFS**| *Average q* v.s. *Marginal q*
    - 对比了 HGMM, Dynamic Panel Data, IV
    - 提出了 Minimum Distance Technique (Stata codes)
    - Stata commands: | Ewreg | XTEWreg |

内生性问题的处理方法

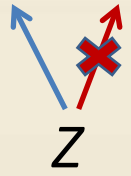
- 研究设计和模型设定：从根源上理清内生性问题
- 工具变量法与GMM估计(IV-GMM)
- 面板数据模型 (Panel Data Models)
- Heckman 选择模型、Treatment effect 模型
- 倍分法 (DID)
- 倾向得分匹配分析 (PSM)
- 断点回归设计 (RDD)
- 合成控制法 (SCM)
- 结构方程模型（SEM）

模型设定

- 理论依据
- 前期文献中普遍使用的模型设定
- 控制变量的选取
- 关键指标的界定和衡量方法(自控能力、文化、父母健康、公司业绩)
- 数据类型(线性回归、离散选择、计数模型、面板)
- 离群值的处理
- 结构变化
- 排他性解释(均值回复与动态权衡、11合一的事件研究)
- 稳健性检验(结论的适用范围、结果的敏感性)
- 安慰剂检验(运气，心理作用等)

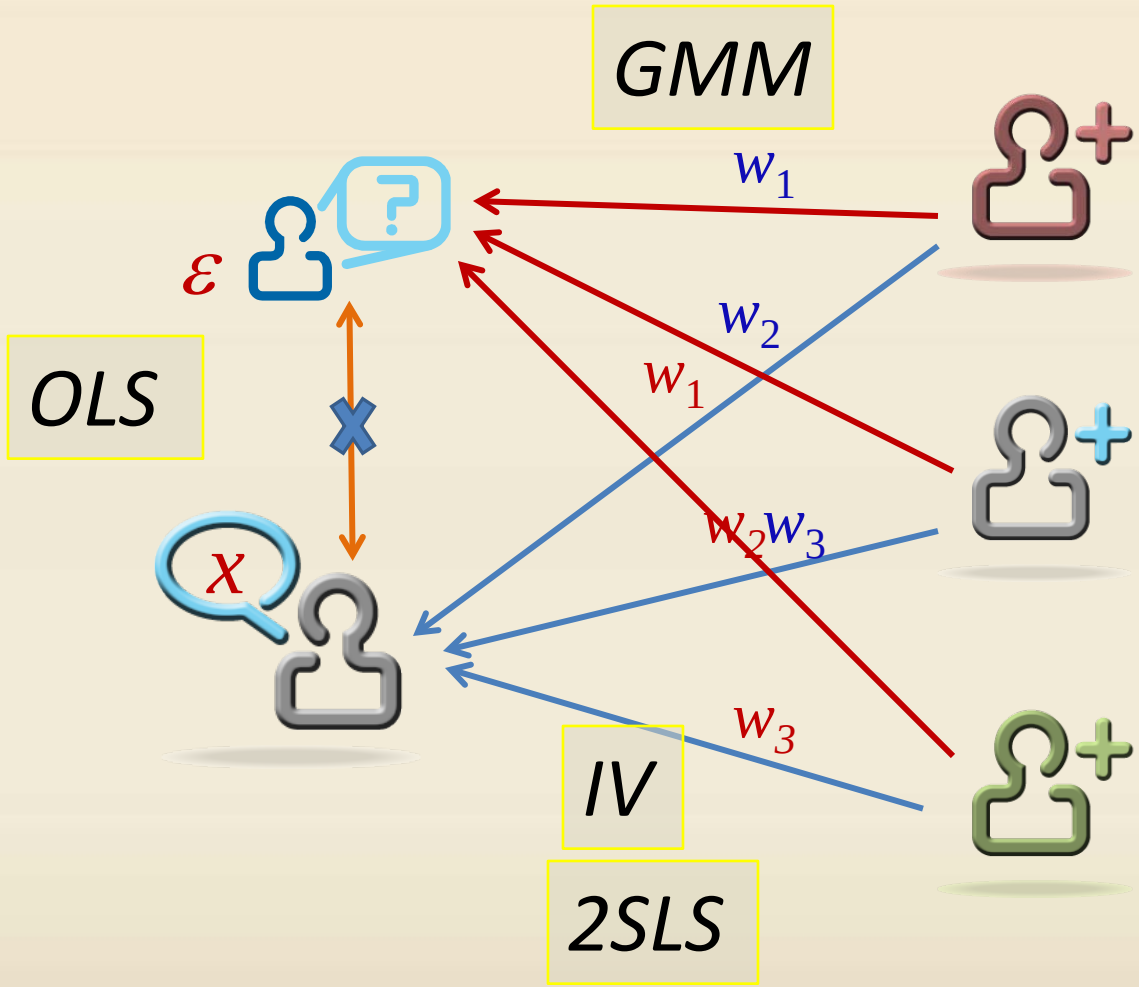
IV-GMM 估计

$$y = a + X\beta + \varepsilon$$



- IV：假设 Corr(Z, ε) = 0，一夫一妻
- 2SLS：假设 Corr(Z, ε) = 0，一夫多妻
  - 第一阶段的回归只是在分配 $Z1$, $Z2$ …… 的与 X 之间关系的权重
- GMM
  - E[$Z1'\varepsilon$] = 0,
  - E[$Z2'\varepsilon$] = 0,
  - ……

- Stata commands: ivregress | ivreg2 | gmm



IV-2SLS 估计

- IV

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$$

$$\begin{aligned} \mathbf{Z}'\mathbf{u} &= 0 \\ \mathbf{Z}'(\mathbf{y} - \mathbf{X}\beta) &= 0 \end{aligned}$$

$$\begin{aligned} \mathbf{Z}'\mathbf{y} - \mathbf{Z}'\mathbf{X}\widehat{\beta}_{\mathrm{IV}} &= 0 \\ \widehat{\beta}_{\mathrm{IV}} &= (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y} \end{aligned}$$

- 2SLS
  - Stage1: reg X on Z, get X_hat

$$\widehat{\mathbf{X}} = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}$$

  - Stata2: reg Y on X_hat, get

$$\begin{aligned} \widehat{\beta}_{2\mathrm{SLS}} &= \left(\widehat{\mathbf{X}}'\mathbf{X}\right)^{-1}\widehat{\mathbf{X}}'\mathbf{y} \\ &= \left\{\mathbf{X}'\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{X}\right\}^{-1}\left\{\mathbf{X}'\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{y}\right\} \\ &= \left(\mathbf{X}'\mathbf{P}_Z\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{P}_Z\mathbf{y} \end{aligned}$$

  - This is wrong! (SE is biased)
  - 正确设定: ivregress 2sls y x1 x2 (x3 x4 = z1 z2 z3)

GMM 估计

Lars Peter Hansen

- Moment Condition (MC, 矩条件)

$$g_i(\beta) = \mathbf{Z}_i' u_i = \mathbf{Z}_i'(y_i - \mathbf{x}_i\beta)$$

- 样本矩条件(SMC)

$$\bar{g}(\beta) = \frac{1}{N}\sum_{i=1}^{N} g_i(\beta) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{z}_i'(y_i - \mathbf{x}_i\beta) = \frac{1}{N}\mathbf{Z}'\mathbf{u}$$

- 目标函数

$$J\left(\widehat{\beta}_{\mathrm{GMM}}\right) = N\bar{g}\left(\widehat{\beta}_{\mathrm{GMM}}\right)' \mathbf{W}\bar{g}\left(\widehat{\beta}_{\mathrm{GMM}}\right)$$

固定效应模型
Fixed Effects Model (FE)

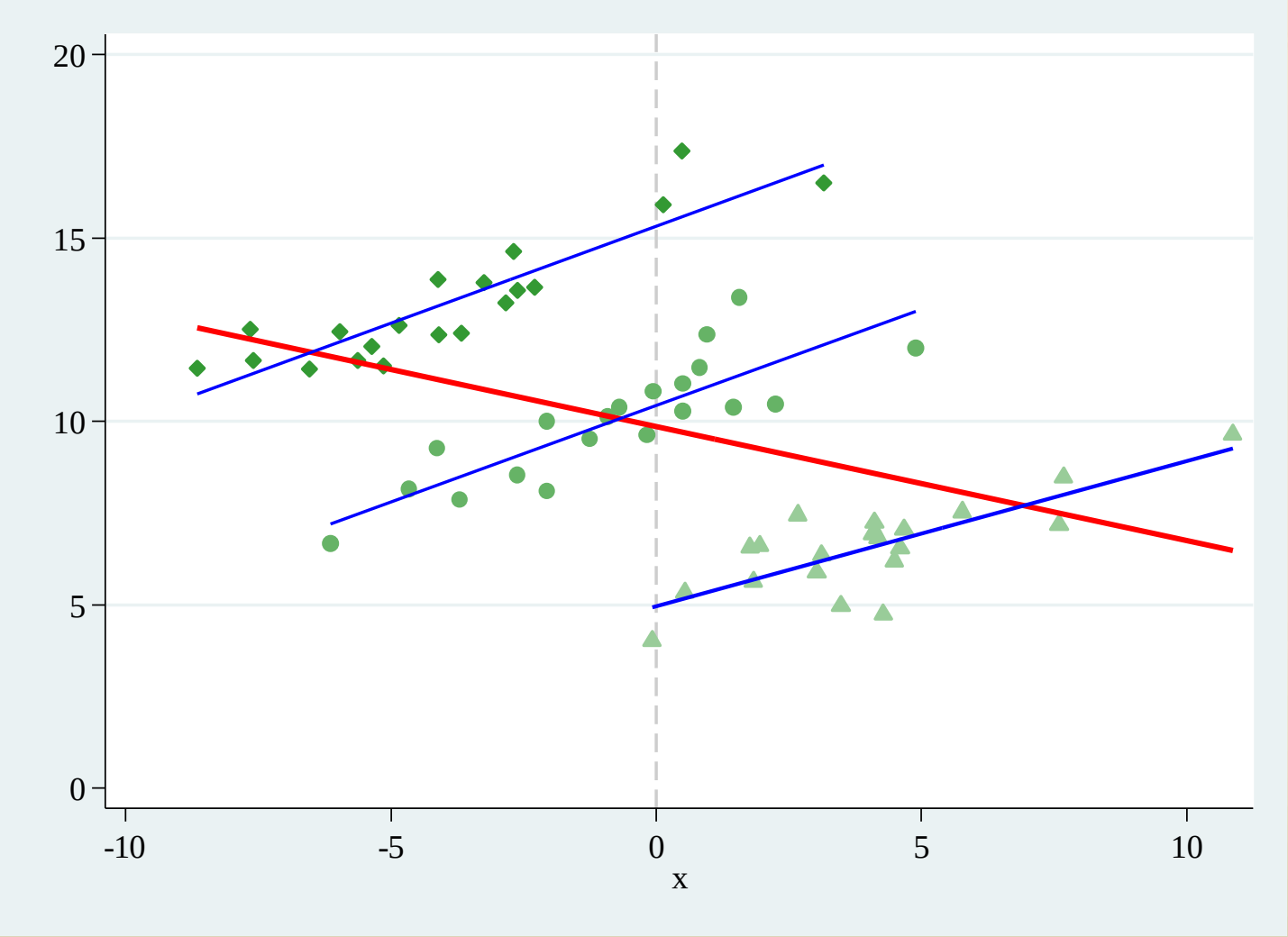
- 模型设定

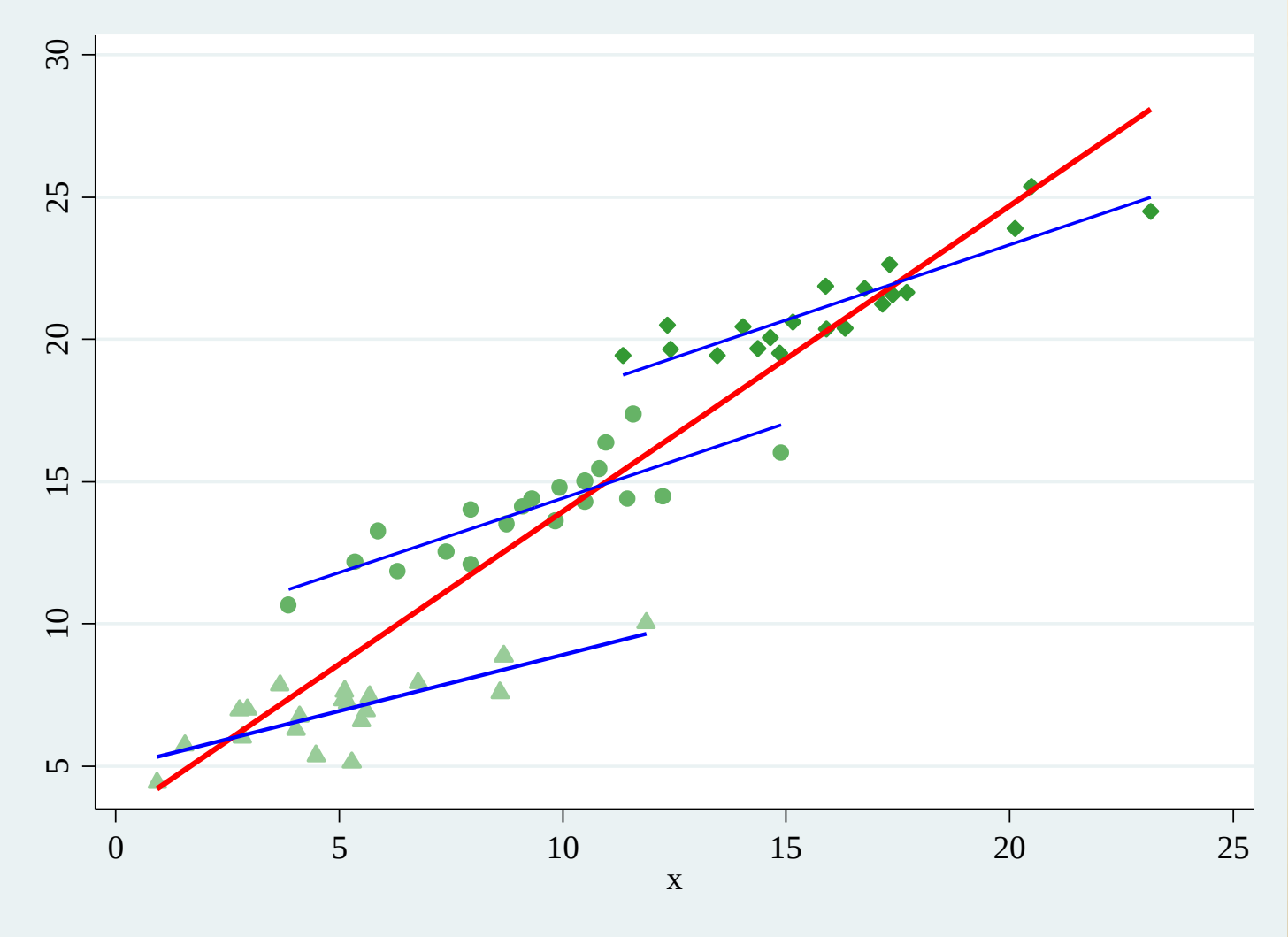
$$FE:\ y_{it} = \alpha_0 + \phi_i + X'_{it}\beta + \varepsilon_{it} \qquad POLS:\ y_{it} = \alpha_0 + X'_{it}\beta + u_{it}$$

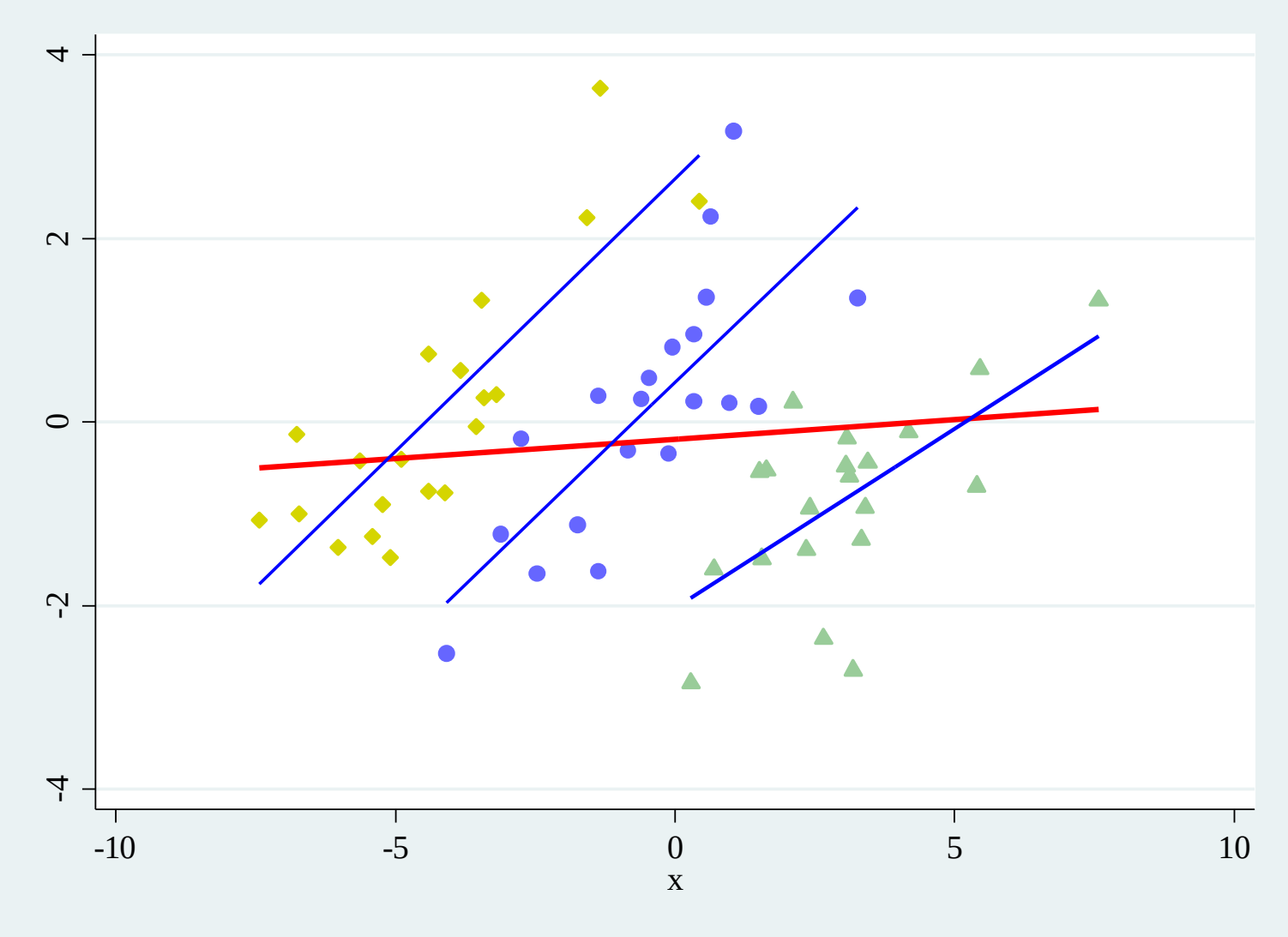
$$u_{it} \overset{?}{=} \phi_i + \varepsilon_{it}$$

  - ϕ_i：肥胖，CEO 特征，公司文化等
  - 所有不随时间变化的因素，包括：
    - 可观测的：性别、种族、地区
    - 不可观测的：性格、能力、文化
  - 因此，在FE模型中直接加入性别、种族、地区等虚拟变量

- Stata commands: xtreg, fe | xi: regress i.id | areg







固定效应模型
Fixed Effects Model (FE)

- 模型设定

$$FE:\ y_{it} = \alpha_0 + \phi_i + X_{it}^{'}\beta + \varepsilon_{it}$$

- FE估计的基本思想
  - 一阶差分变换:
  - 组内去心变换:

$$\Delta y_{it} = \Delta X_{it}^{'}\beta + \Delta\varepsilon_{it}$$

$$(y_{it} - \overline{y}_{it}) = (X_{it}^{'} - \overline{X}_{it}^{'})\beta + (\varepsilon_{it} - \overline{\varepsilon}_{it})$$

$$where,\ \overline{y}_{it} = \frac{1}{T_i}\sum_{t=1}^{T_i} y_{it}$$

固定效应模型
Fixed Effects Model (FE)

- 应用
  - Flannery and Rangan (2006) |JFE|，资本结构的动态调整
  - Lemmon et al. (2008) |JF|，资本结构的动态调整
  - Malmendier et al.(2011) |JF|，经理人特征（早期经历）与财务决策
  - Graham et al.(2012) |RFS|，经理人特征与高管薪酬
  - 叶德珠 等(2012) |经济研究|，国家文化与居民消费行为
  - Petersen(2009) |RFS|，面板模型中标准误的估计
  - Cameron and Miller (2015) |JHR|，聚类标准误

动态面板模型
Dynamic Panel Data Models

- 模型设定

$$y_{it} = \alpha_i + \rho y_{it-1} + X_{it}'\beta + \varepsilon_{it} \quad (1)$$ || 资本结构、投资行为、现金持有

$$y_{it-1} = \alpha_i + \rho y_{it-2} + X_{it-1}'\beta + \varepsilon_{it-1} \quad (2)$$ || 递归特征

$$\Delta y_{it} = \rho \Delta y_{it-1} + \Delta X_{it}'\beta + \Delta\varepsilon_{it} \quad (3)$$ || 一阶差分，可以去除个体效应

$\Delta y_{it-1} = y_{it-1} - y_{it-2}$ || OLS, FE 估计量都是有偏的，要采用 GMM

- $\Delta\varepsilon_{it} = \varepsilon_{it} - \varepsilon_{it-1}$ || IVs for Δy_{it-1}: ? y_{it-2} y_{it-3} y_{it-4} ……Δy_{it-2}

  $\Rightarrow Corr(\Delta y_{it-1}, \Delta\varepsilon_{it}) \neq 0$ || OLS, FE 估计量都是有偏的，要采用 GMM

- Stata commands: xtabond | xtdpdsys | xtdpd | xtlsdvc | xtregdhp | xtabond2

动态面板模型
Dynamic Panel Data Models

- 应用
  - Aghion et al.(2009) |JM|，汇率波动、金融发展与生产率（规范）
  - Brown et al.(2009) |JF|，金融创新与企业成长（规范）
  - Wintoki et al.(2012) |JFE|，非常细致地探讨了公司治理中的内生性问题，对各种动态面板估计方法进行了非常深入的对比分析（综合）
  - Flannery and Hankins(2013) |JCF|，综述：公司金融中的动态面板估计方法
  - Seo and Shin (2017) |JoE|，动态门槛面板

动态面板模型
Dynamic Panel Data Models：进展

- 长差分估计法(long-difference, LD)
  - Hahn et al.(2007) |JE|，适用于 *T* 较小，*y* 持续性较强的动态面板
  - Huang and Ritter(2009) |JFQA|，应用：资本结构调整速度估算
- Han-Phillips dynamic panel data model
  - Han and Phillips(2010) |ET|，Linear Dynamic Panel Data Regression 适用于*y* 持续性较强的动态面板，Panel Unit Root Test
- 分位数动态面板模型 (Quantile Dynamic Panel Data)
  - Galvao(2011) |ET|，Quantile regression for dynamic panel data
- 面板VAR模型 (Panel VAR models)
  - Holtz-Eakin et al.(1988) |E~trica|； Arellano and Bond(1991) |RES| ；
  - Love and Zicchino(2006) |QREF| Canova and Ciccarelli (2013, Survey)
  - Abrigo and Love (2016, Stata Journal)
- Stata commands: xtregdhp | gmm | pvar | pvar2 | xtvar

空间动态面板模型

Spatial Dynamic Panel Data Models

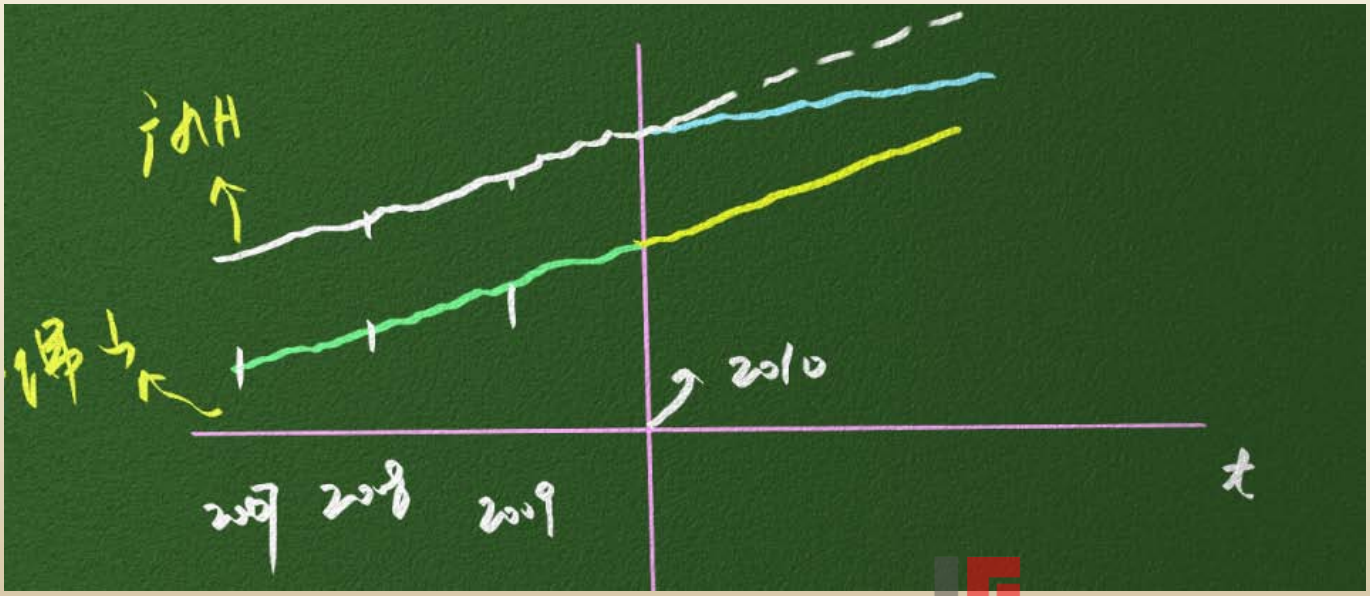
$$Y_{nt} = \lambda_0 W_n Y_{nt} + \gamma_0 Y_{n,t-1} + \rho_0 W_n Y_{n,t-1} + X_{nt}\beta_0 + \mathbf{c}_{n0} + \alpha_{t0} l_n + V_{nt}, \quad t = 1, 2, \ldots, T,$$

- **Lee, L.-f.**, J. Yu, 2010, A spatial dynamic panel data model with both time and individual fixed effects, *Econometric Theory, 26 (02), pp. 564-597.*
- Yu, J., R. de Jong, **L.-f. Lee**, 2012, Estimation for spatial dynamic panel data with fixed effects: The case of spatial cointegration, *Journal of Econometrics, 167 (1), pp. 16-37.*
- **Lee, L.-f.**, J. Yu, 2010, Some recent developments in spatial panel data models, *Regional Science and Urban Economics, 40 (5), pp. 255-271.* (综述)
- Yu, J., **L.-f. Lee**, 2012, Convergence: A spatial dynamic panel data approach, Global Journal of Economics, 1 (1), pp. forthcoming. (应用：经济收敛)
- **Lee, L.-f.**, J. Yu, 2011, Estimation of spatial panels, Now Publishers Inc. (Book)

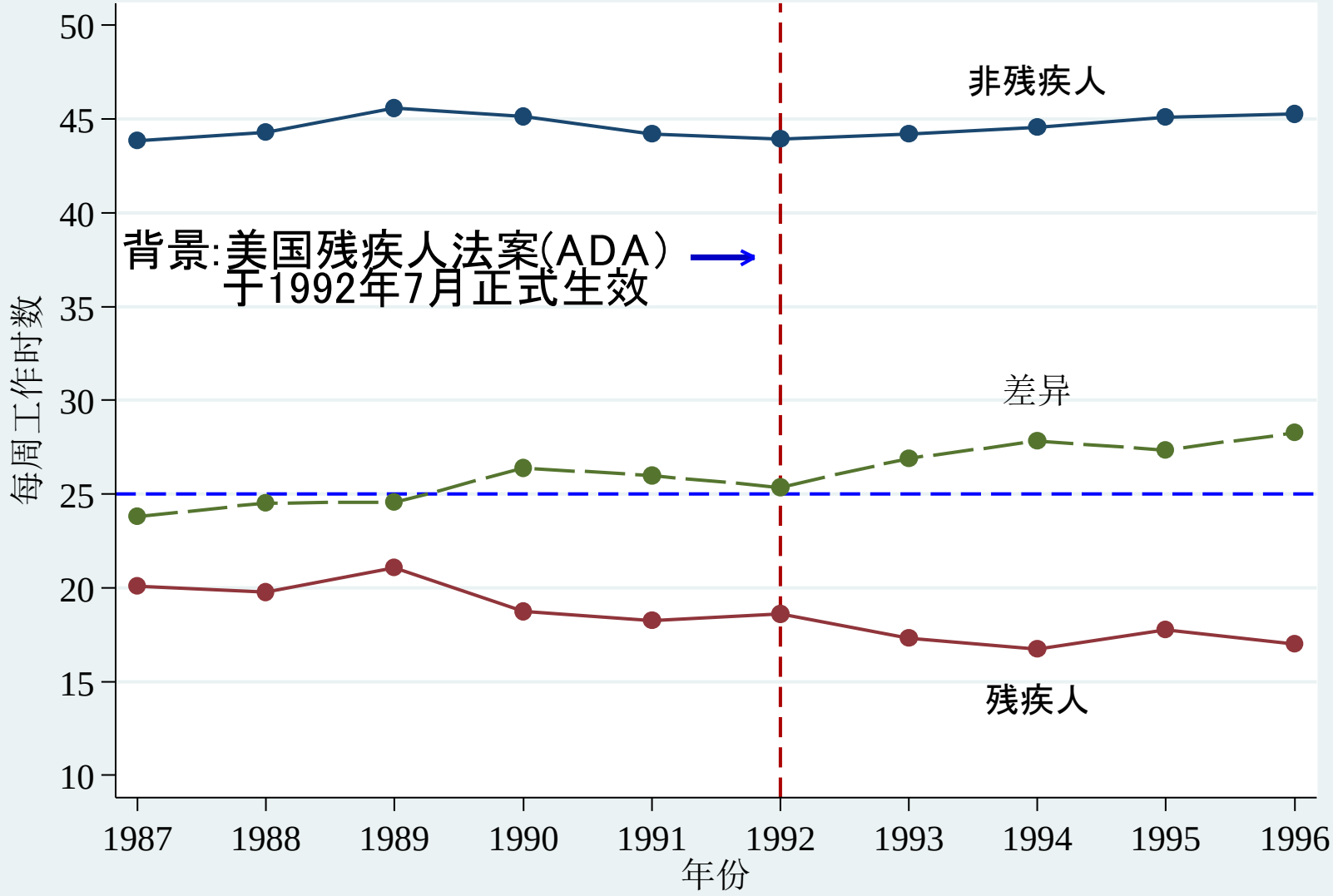
倍分法
Difference-In-Difference (DID)

- 房地产调控政策(限价)有效吗？

	2009	2011	Difference
广州(限价)	16,000	20,000	4,000 ↑
佛山(不限价)	12,000	17,000	5,000 ↑
Difference			-1,000



残疾人法案真的能保护残疾人吗？



Source: Acemoglu and Angrist (2001, JPE) FIG.2

倍分法
Difference-In-Difference (DID)

- 估计方法：

$$y_{it} = \beta_0 + \beta_1 Treat_i + \beta_2 Post_{it} + \gamma \cdot Treat_i \times Post_{it} + \beta_3 X_{it} + \varepsilon_{it}$$

- 假设条件：共同趋势

$$y_{it} = \beta_0 + \beta_1 Treat_i + \beta_2 \mathrm{Year}_t + \gamma \cdot Treat_i \times Year_t + \beta_3 X_{it} + \beta_4 X_{it} \times Year_t + \varepsilon_{it}$$

- Stata commands（检验共同趋势，估计DID效果）
  - global controls “z1 z2 z3 z4”
  - reg y Treat i.year i.Treat#i.year i.year##($controls)

- Stata commands: diff | ddid | regress

倍分法
Difference-In-Difference (DID)

- 关键问题
  - 配对样本的选择：二者随时间自然变化的部分应相同
  - PSM + DID
  - 面板数据：多次调控(Treat) ？　　help ddid

倍分法
Difference-In-Difference (DID)

- 应用
  - Cooper et al. (2005) |JF|，基金更名行为的影响
  - Villalonga (2004) |FM|，多元化经营，DID，Heckman
  - Chhaochharia and Grinstein (2009) |JF|，萨班斯法案与 CEO 薪酬
  - Frésard (2010) |JF|，产品市场竞争与现金持有
  - Black and Kim (2012) |JF|, 董事会结构与公司价值, DID, 2SLS, 3SLS
  - Tsoutsoura (2015) |JF|, 继承税对家族企业投资的影响

倾向得分匹配分析
Propensity Score Matching Method (PSM)

- 为何要配对？（自选择问题）

$$Y_i = \alpha + \gamma D_i + X_i^{'}\beta + \varepsilon_i$$

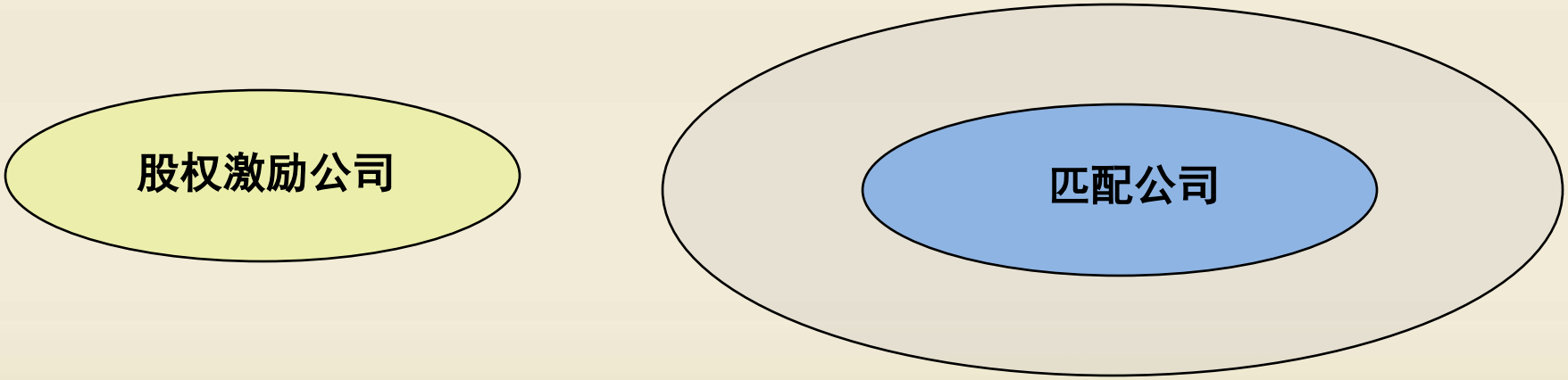
$$\gamma = E[Y_i \mid D_i = 1, X_i = x] - E[Y_i \mid D_i = 0, X_i = x]$$

{observable} {unobservable}

- 传统匹配方法：多维(规模、行业、盈利能力)
- PSM：Logit 模型，多维 → 一维 PS 值
- Stata commands: teffects psmatch | kmatch | psmatch2 | optmatch2 | ccmatch | cem

倾向得分匹配分析
Propensity Score Matching Method (PSM)

- 基本思路：



匹配指标： Propensity Score（PS 值）

Logit(Size, Industry, ROA, Leverage, Ownership, ….) → PS 值

降维：多维 → 一维

倾向得分匹配分析
Propensity Score Matching Method (PSM)

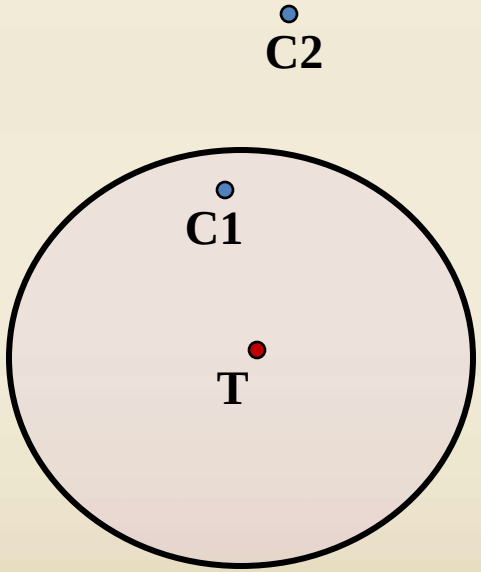
- 最近邻匹配

$$C(i) = \min_j \left\| p_i - p_j \right\|$$

- 半径匹配

$$C(i) = \left\{ p_j \,\middle|\, \left\| p_i - p_j \right\| < r \right\}$$

- 核匹配
  - 用所有 Control 组公司的加权平均虚构出一个配对公司



倾向得分匹配分析
Propensity Score Matching Method (PSM)

- 应用
  - Cooper et al. (2005) |JF|，基金更名行为的影响
  - Hellmann et al. (2008) |RFS|，银企关系
  - Campello et al. (2010) |JFE|，金融危机中 CFO 如何应对
  - Faulkender and Yang (2010) |JFE|，经理人薪酬激励
  - Michaely and Roberts (2012) |RFS|，私营企业的股利支付行为
  - Faccio, Marchica and Mura(2016) |JCF|，女性 CEO 与风险承担

自选择模型
Self-Selection Models

- 问题的根源：被解释变量(y)中经常包含缺漏值
  - Case I: 随机缺漏
  - Case II: 非随机缺漏(无法观察到)
  - 例如，y = 公司的研发支出；高管的在职消费；公司的游说支出
- 模型设定(Heckman selection model)
  - 回归方程 $y_j = \mathbf{x}_j\boldsymbol{\beta} + u_{1j}$
  - 选择方程: y is observed only if $\mathbf{z}_j\boldsymbol{\gamma} + u_{2j} > 0$

$$u_1 \sim N(0, \sigma)$$
$$u_2 \sim N(0, 1)$$
$$\text{corr}(u_1, u_2) = \rho$$

处理效应模型
Treatment Effect Models

- 模型设定：解释变量中包含一个内生的 0/1 变量

$$y_1 = x_1\alpha_1 + d\beta + u_1,$$

$$y_2^* = x_2\alpha_2 + u_2,$$

$$\begin{aligned} d &= 1 \quad \text{if } y_2^* \geqslant 0 \\ &= 0 \quad \text{if } y_2^* < 0, \end{aligned}$$

- Stata commands: etregress | heckman | ivprobit | cmp | itreatreg | mtreatreg | etpoisson | treatoprobit | etpoisson

处理效应模型
Treatment Effect Models

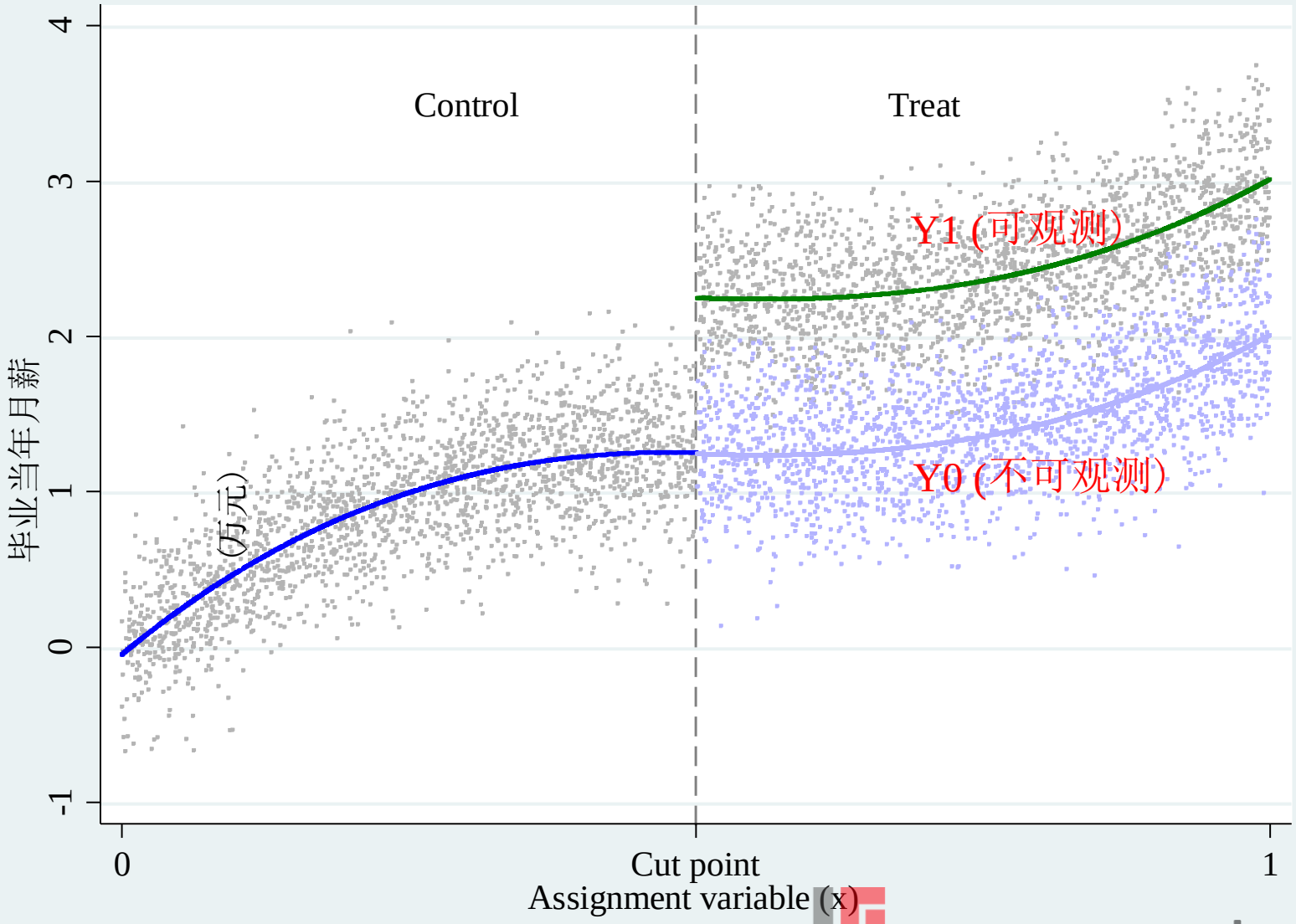
- 应用
  - Laeven and Levine (2007) |RFS|，多元化折价
  - Gompers et al. (2010) |RFS|，双重股权公司
  - Ayyagari et al. (2010) |RFS|，非正规融资，中国
  - Ross (2010) |RFS|，主导银行效应
  - Core and Guay (2001) |JFE|，股权激励
  - Lee and Masulis (2009) |JFE|，二次发行
  - Masulis and Mobbs (2011) |JF|， 独立董事市场
  - Huang, Lian and Li(2016) |CER|，子女外出对父母健康的影响，多元处理效应

断点回归设计
Regression Discontinuity Designs (RDD)

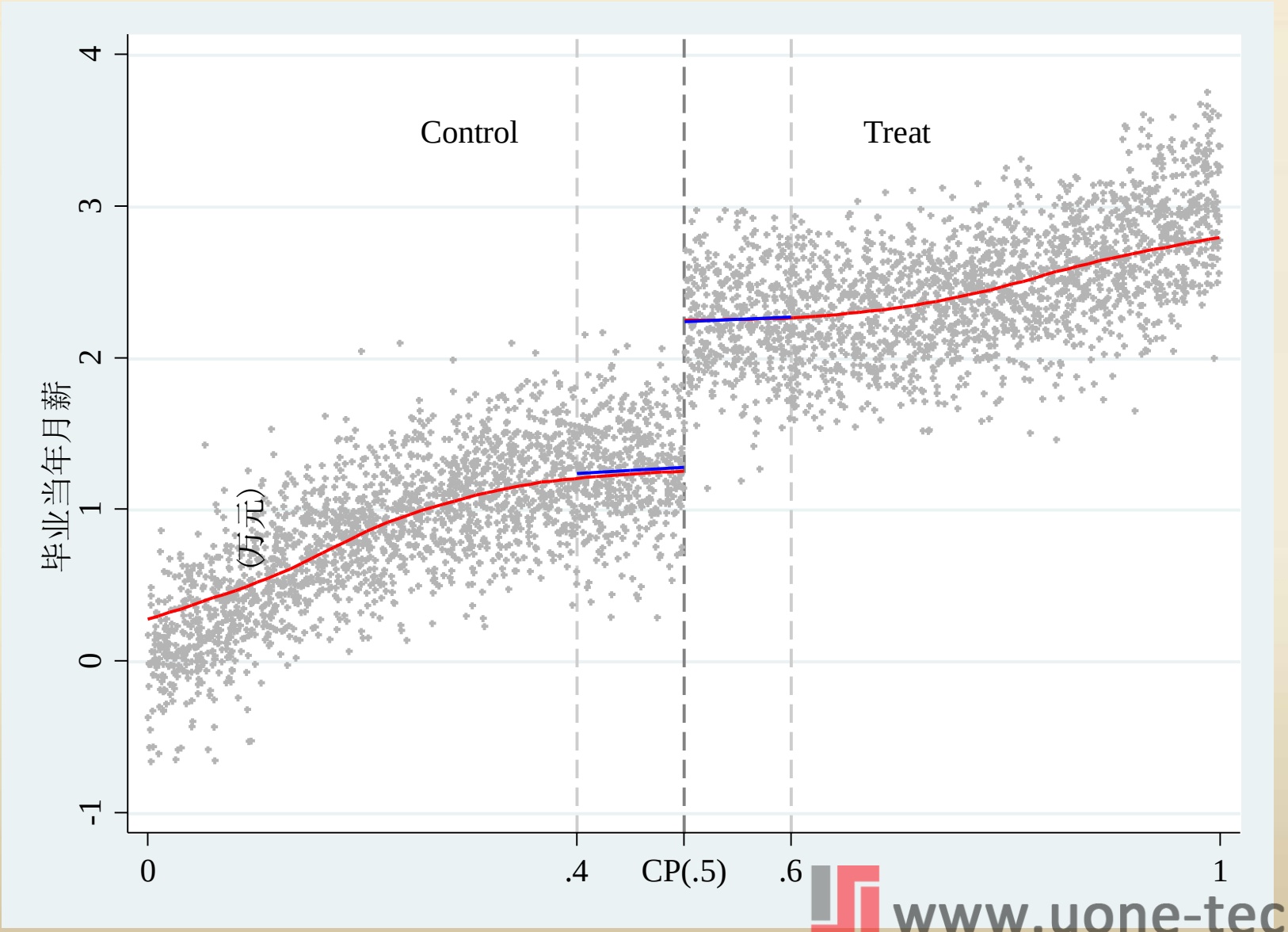
- RDD: 接近于自然实验的研究方法

- Stata commands: | rd | rdrobust | rdplot | rdcv | next

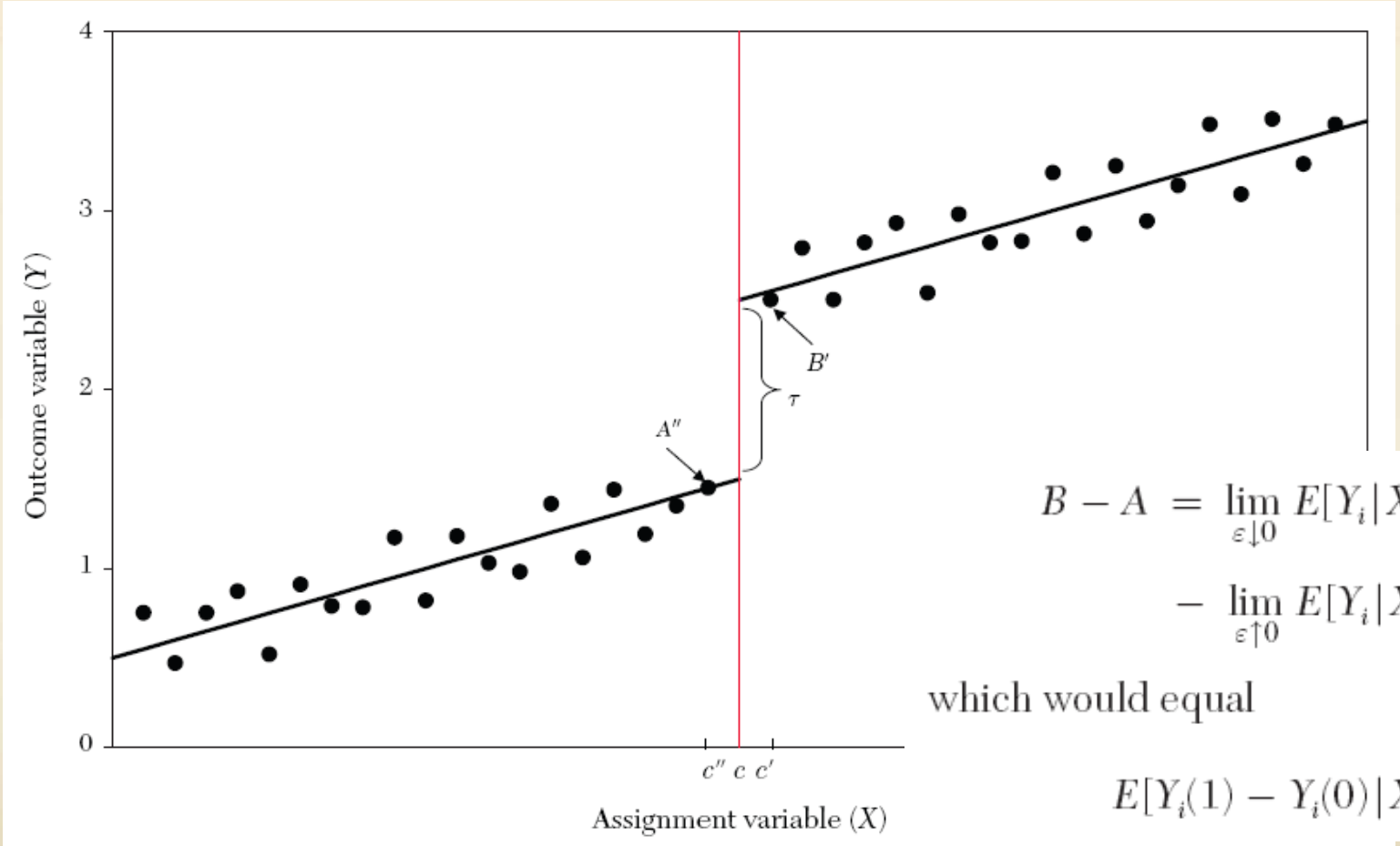
RDD 图示（反事实，Jump）



RDD 估计 (多项式回归, 局部回归)



RDD



$$B - A = \lim_{\varepsilon \downarrow 0} E[Y_i | X_i = c + \varepsilon] - \lim_{\varepsilon \uparrow 0} E[Y_i | X_i = c + \varepsilon],$$

which would equal

$$E[Y_i(1) - Y_i(0) | X = c].$$

Source: Lee and Lemieux (2010, |JEL|, Figure. 1)

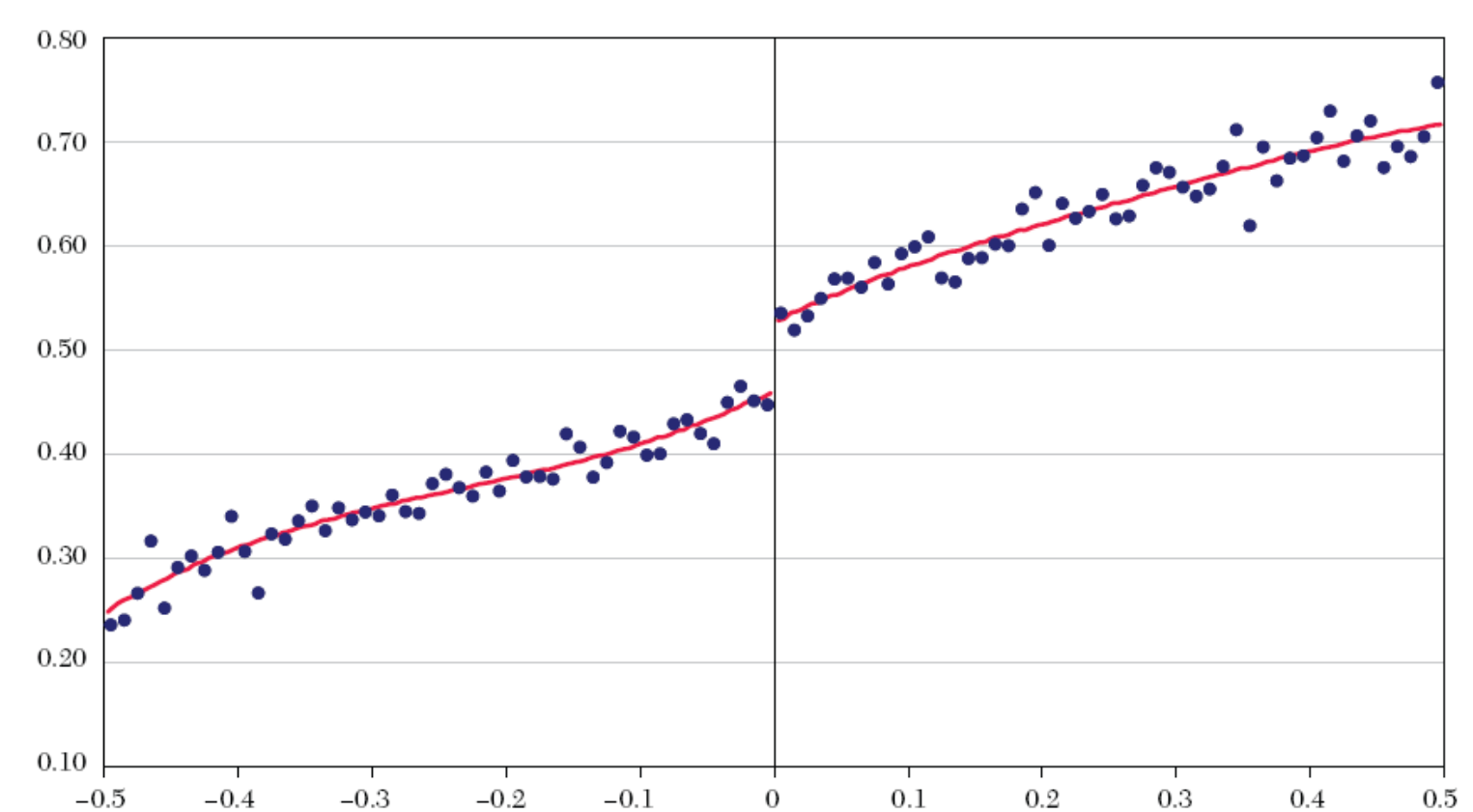


Figure 7. Share of Vote in Next Election, Bandwidth of 0.01 (100 bins)

Notes: 横轴为驱动变量 上一届选举中执政党与在野党票数比重之差

纵轴为结果变量 下一届选举中执政党获得的选票比重

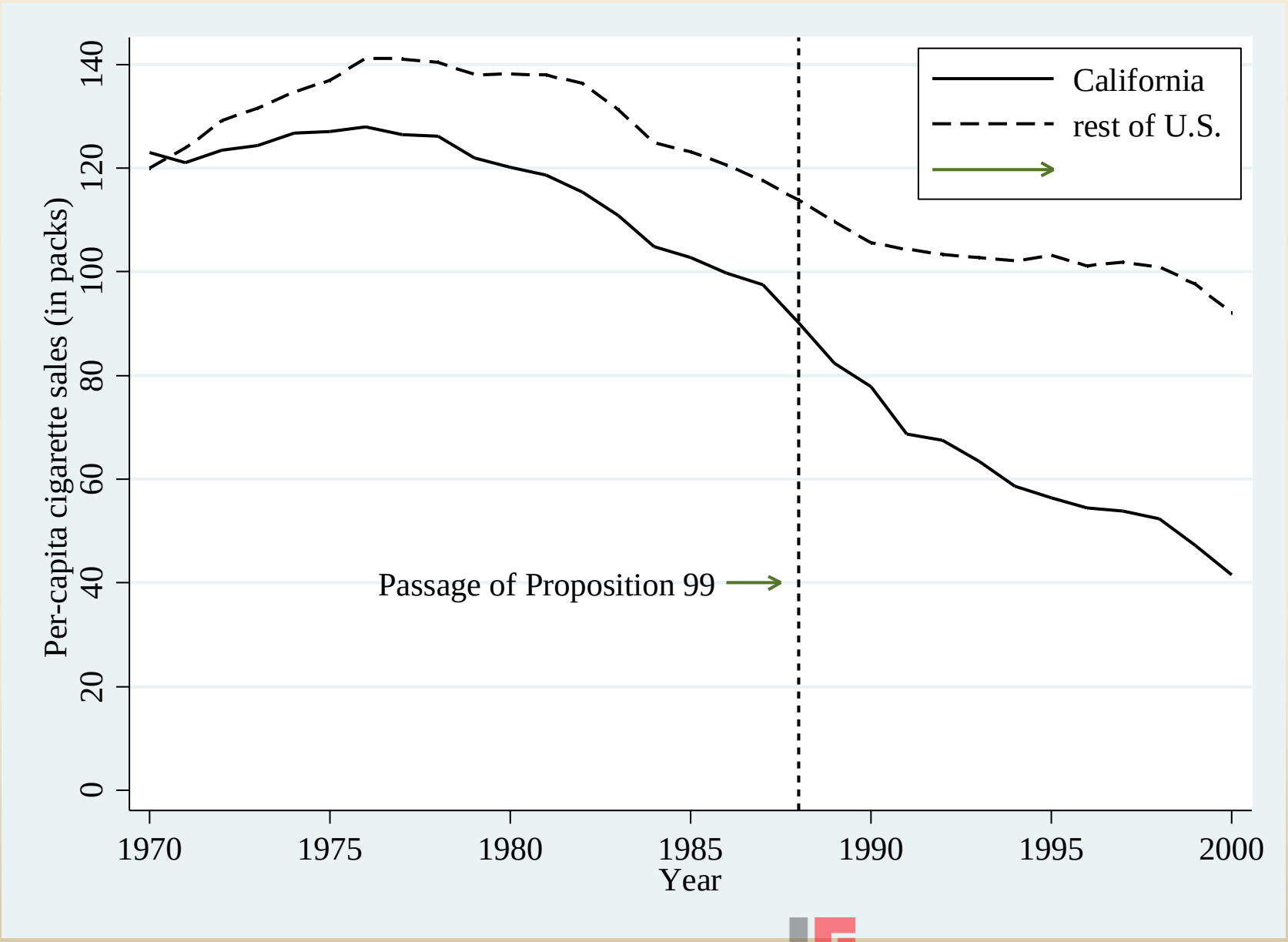
Source: Lee and Lemieux (2010, |JEL|, Figure. 7)

断点回归设计
Regression Discontinuity Designs (RDD)

- 应用
  - Chava and Roberts (2008) |JF|，债务契约与投资行为
  - Roberts and Sufi (2009) |JF|，控制权与资本结构
  - Iliev (2010) |JF|，萨班斯法案对融资成本、盈余管理和股价的影响
  - Garmaise and Natividad (2010) |RFS|，信息不对称与融资成本
  - Cuñat et al.(2010) |NBER|，公司治理与股东价值(股东年会投票数据)
  - Baker et al.(2011) |JFE|，参考价格与兼并收购行为
  - 刘生龙, 周绍杰, 胡鞍钢 (2016) |经济研究|, 义务教育法的政策效果

合成控制法
(Synthetic control methods, SMC)

- Abadie, A., A. Diamond, J. 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.
- Q: 如果政策对象只有一个国家或一个州或一个省，怎么办？
  - 比如，加州于 1989 年实施禁烟法案 (99号提案)
- A：以其它未未受政策冲击的州为样本池，人工制作一个“合成加州”，让它在 1989 年之前与“真实加州”无异 (Year<1989)
- 政策效果(Year>1989) = Y(真实加州) - Y(合成加州)



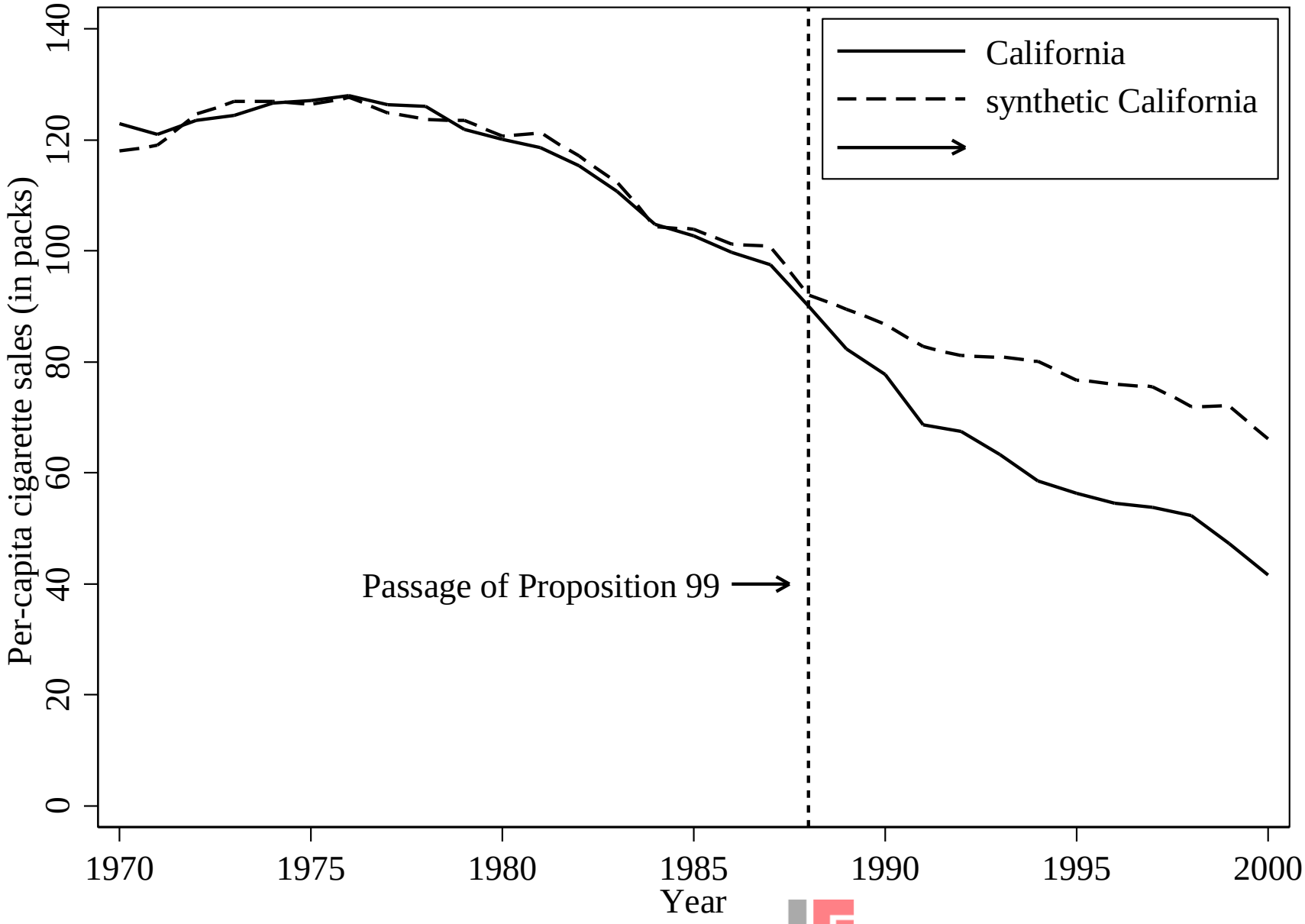
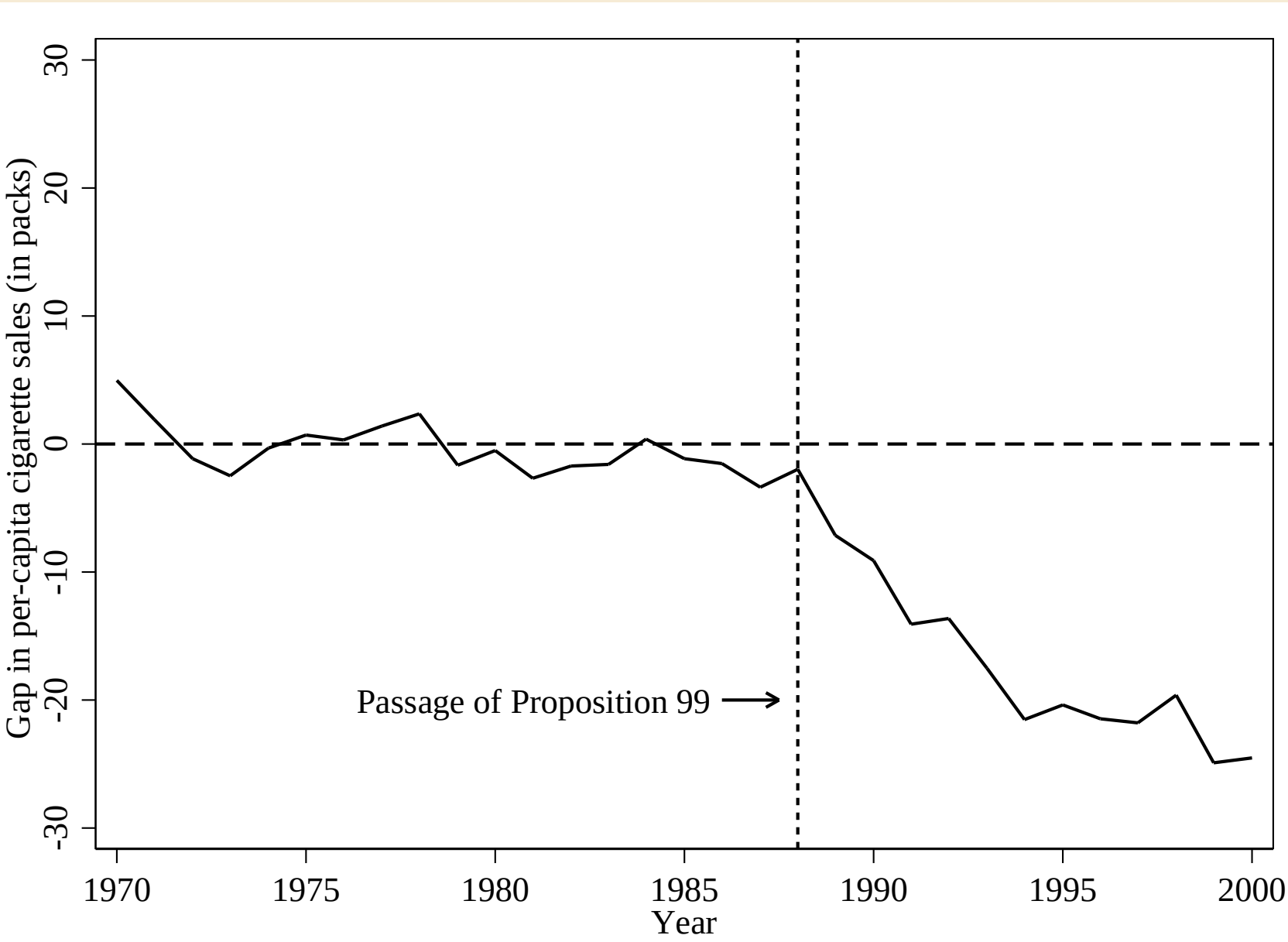
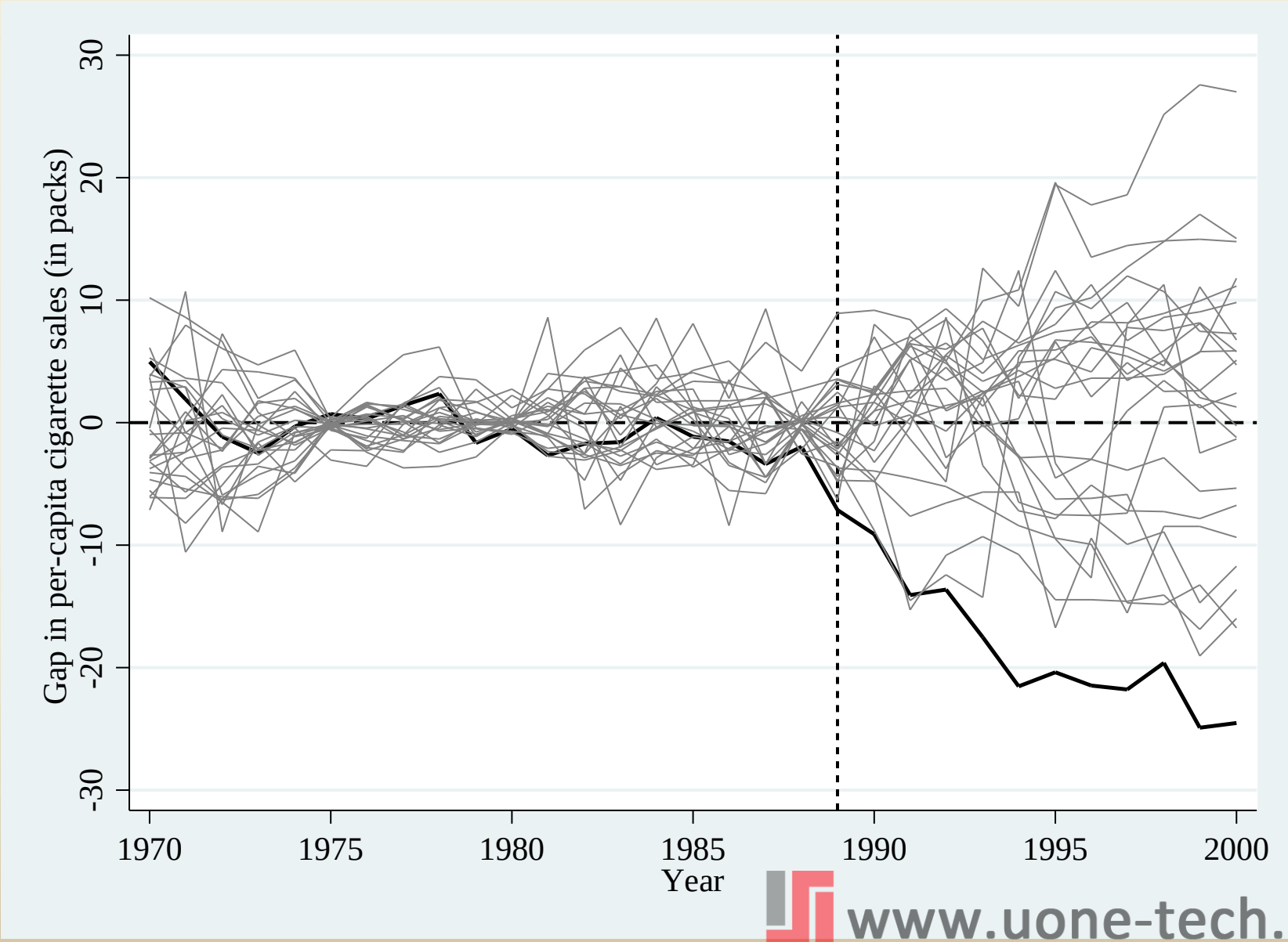


Table 2. State weights in the synthetic California

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



安慰剂检验



结 语

对于实证分析的建议

- 清晰界定你所研究的问题（重要的、有意义的）
- 数据总是有缺陷的，要通过巧妙的研究设计来保证统计推断的可靠性
  - e.g. Fazzari et al. (1988), 投资-现金流敏感性 → 融资约束假说
- 方法的实现不是问题，关键在于要选择合适的方法
- 研究设计：
  - 制度背景的深刻理解（很重要！）
  - 内生性问题的来源与后果（避免摆 Pose）
  - 采用何种方法能够恰当地进行统计推断（多种方法的配合使用）
  - 特殊的事件、特殊的数据：尽量接近于自然实验

让我们的**实证研究**

更接近于**自然实验**

……

附

- 内生性问题综述
  - Wintoki et al. (2008)； Coles et al. (2007)； Tucker (2011)； Lee (2005)
  - Roberts and Whited (2011)； Imbens and Wooldridge (2009)
  - Imbens and Lemieux(2008) JE, RDD
  - Lee and Lemieux(2010) JEL, RDD
- 相关模型和方法的Stata实现过程及范例
  - IV-GMM估计：Stata高级视频 B4_IV_GMM
  - 静态面板数据模型和动态面板数据模型：Stata高级视频 B7_Panel
  - 面板门槛模型：Stata学术论文视频 (说明书)Hansen_1999(附带Stata命令 xtthres)
  - 倾向得分匹配分析PSM：Stata学术论文视频 (说明书) Lian_2012_PSM

参考文献

- Aghion, P, Bacchetta P, Ranciere R, Rogoff K (2009). Exchange rate volatility and productivity growth: The role of financial development. Journal of Monetary Economics, 56 (4): 494-513.
- Alti, A (2003). How Sensitive Is Investment to Cash Flow When Financing Is Frictionless? Journal of Finance, 58 (2): 707-722.
- Arellano, M, Bond S (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. Review of Economic Studies, 58 (2): 277-297.
- Ayyagari, M, Demirgüç-Kunt A, Maksimovic V (2010). Formal versus informal finance: evidence from China. Review of Financial Studies, 23 (8): 3048-3097.
- Baker, M, Pan X, Wurgler J (2011). The effect of reference point prices on mergers and acquisitions. Journal of Financial Economics, 102 (1): 1-27.
- Black, B, Kim W (2012). The effect of board structure on firm value: A multiple identification strategies approach using Korean data. Journal of Financial Economics, 104 (1): 203-226.
- Black, B S, Jang H, Kim W (2006). Does corporate governance predict firms' market values? Evidence from Korea. Journal of Law, Economics, and Organization, 22 (2): 366-413.
- Brown, J R, Fazzari S M, Petersen B C (2009). Financing Innovation and Growth: Cash Flow, External Equity, and the 1990s R&D Boom. The Journal of Finance, 64 (1): 151-185.

参考文献

- Campello, M, Graham J R, Harvey C R (2010). The Real Effects of Financial Constraints: Evidence from a Financial Crisis. Journal of Financial Economics, 97 (3): 470-487.
- Cellini, S R, Ferreira F, Rothstein J (2010). The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design. The Quarterly Journal of Economics, 125 (1): 215-261.
- Chava, S, Roberts M R (2008). How Does Financing Impact Investment? The Role of Debt Covenants. Journal of Finance, 63 (5): 2085-2121.
- Chhaochharia, V, Grinstein Y (2009). CEO Compensation and Board Structure. Journal of Finance, 64 (1): 231-261.
- Coles, J L, Lemmon M L, Felix Meschke J (2012). Structural models and endogeneity in corporate finance: The link between managerial ownership and corporate performance. Journal of Financial Economics, 103 (1): 149-168.
- Coles, J L, Lemmon M L, Meschke F (2007). Structural Models and Endogeneity in Corporate Finance: the Link Between Managerial Ownership and Corporate Performance. EFA 2003 Glasgow, Forthcoming. Available at SSRN: http://ssrn.com/abstract=423510
- Cooper, M J, Gulen H, Rau P R (2005). Changing Names with Style: Mutual Fund Name Changes and Their Effects on Fund Flows. The Journal of Finance, 60 (6): 2825-2858.

参考文献

- Core, J E, Guay W R (2001). Stock option plans for non-executive employees. Journal of Financial Economics, 61 (2): 253-287.
- Cuñat, V, Gine M, Guadalupe M (2010). The Vote is Cast: The Effect of Corporate Governance on Shareholder Value. NBER Working Papers,
- Erickson, T, Whited T (2000). Measurement error and the relationship between investment and q. Journal of Political Economy, 108 (5): 1027-1057.
- Erickson, T, Whited T M (2012). Treating measurement error in Tobin's q. Review of Financial Studies, 25 (4): 1286-1329.
- Faulkender, M, Yang J (2010). Inside the black box: The role and composition of compensation peer groups. Journal of Financial Economics, 96 (2): 257-270.
- Fazzari, S, Hubbard R, Petersen B (2000). Investment-Cash Flow Sensitivities are Useful: A Comment on Kaplan and Zingales. Quarterly Journal of Economics, 115 (2): 695-705.
- Fazzari, S, Hubbard R, Petersen B, Blinder A,Poterba J (1988). Financing Constraints and Corporate Investment. Brookings Papers on Economic Activity, 1988 (1): 141-206.
- Flannery, M J, Hankins K W (2013). Estimating dynamic panel models in corporate finance. Journal of Corporate Finance, 19: 1-19.
- Flannery, M J, Rangan K P (2006). Partial adjustment toward target capital structures. Journal of Financial Economics, 79 (3): 469-506.

参考文献

- Frésard, L (2010). Financial Strength and Product Market Performance: The Real Effects of Corporate Cash Holdings. Journal of Finance, 65 (3): 1097-1122.
- Fracassi, C (2009). Corporate Finance Policies and Social Networks. working Paper, http://mif.bus.tu.ac.th/conference/Paper7.pdf,
- Galvao, A F (2011). Quantile regression for dynamic panel data with fixed effects. Journal of Econometrics, 164 (1): 142-157.
- Garmaise, M J, Natividad G (2010). Information, the Cost of Credit, and Operational Efficiency: An Empirical Study of Microfinance. Review of Financial Studies, 23 (6): 2560-2590.
- Giroud, X, Mueller H M (2010). Does corporate governance matter in competitive industries? Journal of Financial Economics, 95 (3): 312-331.
- Gompers, P A, Ishii J, Metrick A (2010). Extreme Governance: An Analysis of Dual-Class Firms in the United States. Review of Financial Studies, 23 (3): 1051-1088.
- Graham, J R, Li S, Qiu J P (2012). Managerial Attributes and Executive Compensation. Review of Financial Studies, 25 (1): 144-186.

参考文献

- Han, C, Phillips P C B (2010). Gmm Estimation for Dynamic Panels with Fixed Effects and Strong Instruments at Unity. Econometric Theory, 26 (1): 119-151.
- Hellmann, T, Lindsey L, Puri M (2008). Building relationships early: Banks in venture cap
- ital. Review of Financial Studies, 21 (2): 513-541.
- Holtz-Eakin, D, Newey W, Rosen H (1988). Estimating Vector Autoregressions with Panel Data. Econometrica, 56 (6): 1371-1395.
- Huang, R, Ritter J R (2009). Testing Theories of Capital Structure and Estimating the Speed of Adjustment. Journal of Financial and Quantitative Analysis, 44 (2): 237-271.
- Iliev, P (2010). The Effect of SOX Section 404: Costs, Earnings Quality, and Stock Prices. The Journal of Finance, 65 (3): 1163-1196.
- Imbens, G, Lemieux T (2008). Regression discontinuity designs: A guide to practice. Journal of Econometrics, 142 (2): 615-635.
- Imbens, G, Wooldridge J (2009). Recent developments in the econometrics of program evaluation. Journal of Economic Literature, 47 (1): 5-86.
- Kaplan, S, Zingales L (1997). Do investment-cash flow sensitivities provide useful measures of financing constraints? Quarterly Journal of Economics, 112 (1): 169-215.

参考文献

- Kaplan, S, Zingales L (2000). Investment-Cash Flow Sensitivities are not Valid Measures of Financing Constraints. Quarterly Journal of Economics, 115 (2): 707-712.
- Kuersteiner, G M, Prucha I R (2013). Limit Theory for Panel Data Models with Cross Sectional Dependence and Sequential Exogeneity. Journal of Econometrics,
- Laeven, L, Levine R (2007). Is there a diversification discount in financial conglomerates? Journal of Financial Economics, 85 (2): 331-367.
- Lee, D, Lemieux T (2010). Regression Discontinuity Designs in Economics. Journal of Economic Literature, 48: 281-355.
- Lee, D S (2008). Randomized experiments from non-random selection in US House elections. Journal of Econometrics, 142 (2): 675-697.
- Lee, G, Masulis R W (2009). Seasoned equity offerings: Quality of accounting information and expected flotation costs. Journal of Financial Economics, 92 (3): 443-469.
- Lee, L-f, Yu J (2010a). Some recent developments in spatial panel data models. Regional Science and Urban Economics, 40 (5): 255-271.

参考文献

- Lee, L-f, Yu J (2010b). A spatial dynamic panel data model with both time and individual fixed effects. Econometric Theory, 26 (02): 564-597.
- Lee, L-f, Yu J (2011). Estimation of spatial panels. Now Publishers Inc
- Lee, M (2005). Micro-econometrics for policy, program, and treatment effects. Oxford University Press, USA
- Lemmon, M L, Roberts M R, Zender J F (2008). Back to the Beginning: Persistence and the Cross-Section of Corporate Capital Structure. Journal of Finance, 63 (4): 1575-1608.
- Li, K, Prabhala N. 2007, Self-Selection Models in Corporate Finance, in E Eckbo ed, *Handbook of Corporate Finance: Empirical Corporate Finance* (Elsevier, North Holland) 37-86.
- Love, I, Zicchino L (2006). Financial development and dynamic investment behavior: Evidence from panel VAR. Quarterly Review of Economics and Finance, 46 (2): 190-210.
- Malmendier, U, Tate G, Yan J (2011). Overconfidence and Early-Life Experiences: The Effect of Managerial Traits on Corporate Financial Policies. Journal of Finance, 66 (5): 1687-1733.

参考文献

- Masulis, R W, Mobbs S (2011). Are All Inside Directors the Same? Evidence from the External Directorship Market. The Journal of Finance, 66 (3): 823-872.
- Michaely, R, Roberts M R (2012). Corporate dividend policies: Lessons from private firms. Review of Financial Studies, 25 (3): 711-746.
- Petersen, M A (2009). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. Review of Financial Studies, 22 (1): 435-480.
- Rastad, M (2011). Capital Structure Pre-Balancing: Evidence from Convertible Bonds. Working Paper,
- Roberts, M, Sufi A (2009). Control Rights and Capital Structure: An Empirical Investigation. Journal of Finance, 64 (4): 1657-1695.
- Roberts, M R, Whited T M. 2011, Endogeneity in Empirical Corporate Finance, in G Constantinides, R Stulz,M Harris eds, Handbook of the Economics of Finance (Elsevier, Amsterdam).
- Ross, D G (2010). The "Dominant Bank Effect:" How High Lender Reputation Affects the Information Content and Terms of Bank Loans. Review of Financial Studies, 23 (7): 2730-2756.
- Tucker, J (2011). Selection bias and econometric remedies in accounting and finance research. SSRN working paper,

参考文献

- Villalonga, B (2004). Does Diversification Cause the "Diversification Discount"? Financial Management, 33 (2): 5-27.
- Wald, J K, Long M S (2007). The effect of state laws on capital structure. Journal of Financial Economics, 83 (2): 297-319.
- Wintoki, M B, Linck J S, Netter J M (2008). Endogeneity and the Dynamics of Corporate Governance. SSRN working paper, Available at SSRN: http://ssrn.com/abstract=970986,
- Wintoki, M B, Linck J S, Netter J M (2012). Endogeneity and the dynamics of internal corporate governance. Journal of Financial Economics, 105 (3): 581-606.
- Yu, J, de Jong R, Lee L-f (2012). Estimation for spatial dynamic panel data with fixed effects: the case of spatial cointegration. Journal of Econometrics, 167 (1): 16-37.
- Yu, J, Lee L-f (2012). Convergence: a spatial dynamic panel data approach. Global Journal of Economics, 1 (1): forthcoming.
- 叶德珠, 连玉君, 黄有光, 李东辉 (2012). 消费文化、认知偏差与消费行为偏差. 经济研究, (2): 80-92.