体育经济分析: 原理与应用

DiD

周正卿

18 October 2023

大纲

大纲

- Level 1
  - 一个例子

- Level 2
  - 基本概念

- Level 3
  - 具体实战

潜在结果框架帮助理解"真相"

因果识别方法

- 随着计量经济学"可信性革命"（credibility revolution）席卷经济学的各个领域，基于潜在因果模型的因果效应识别策略，如匹配法（matching）、工具变量法（instrumental variable）、双重差分法（difference-in-differences）和断点回归设计（regression discontinuity design）等，逐渐成为了经济学等社会科学领域实证研究的通行研究范式。

面板数据框架

面板数据框架

- 因果推断工具包中最重要的工具之一就是**面板数据估计量**（panel data estimator）
  - 该估计量为**纵贯数据**（longitudinal data） 专门设计
  - 不同时间点上对个体的重复观测
- 面板数据的优点:
  1. 更多的变化 (横截面数据和时间序列变化)
  2. 可以处理**特定类型数据带来的遗漏变量偏差**（不可观测的且不随时间推移而变化的因素）
- 示例:
  1. 个体（球员）i 在年份 t 的收入
  2. 公司（球队）i 在年份 t 的业绩
 3. 国家 i 在年份 t 的GDP
  4. 学校 i 在年份 t 的升学率

纵贯数据的有向无环图（Imai 和 Kim，2017）

- 结果变量 Y_i
  - 包含 3 期，分别是 Y_{i1}, Y_{i2}, Y_{i3}
- 协变量矩阵 D_i 随时间推移而变化 D_{i1}, D_{i2}, D_{i3}
- 随个体变化但不随时间变化，可被观测的 X_i
- 随个体变化但不随时间变化，但无法被观测到的 u_i
  - 无法观测
  - 随个体而变动
  - 给定个体 i，不随时间推移而变化

第 s 组别
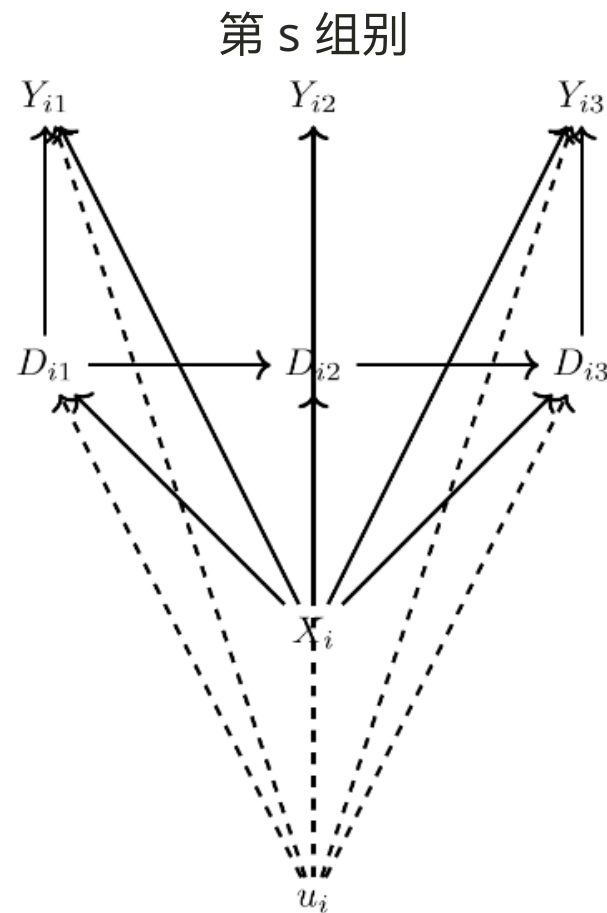
纵贯数据的有向无环图（Imai 和 Kim，2017）

- 第1层：Y_{i1} 与 Y_{i2} 和 Y_{i3} 相互不影响

- 第2层：D_{i1} 影响第1层 Y_{i1} 和同层 D_{i2}

- 第3层：X_i 影响第2层 $D_{i1} - D_{i3}$ 和第1层 $Y_{i1} - Y_{i3}$

- 第4层：不可观测的混杂因子 u_i 影响第3层 X_i 和 第2层 $D_{i1} - D_{i3}$ 以及第1层 $Y_{i1} - Y_{i3}$

- 处理变量 D 产生了**内生性**问题 → 直接估计存在偏误

- **注意几点**
  - **不存在**与第2层 D_{it} 相关的、不能被观测到且**随时间的推移而变化的其他混杂因素**

  - 第1层**过去结果** Y_{it-1} **不直接影响当前结果** Y_{it}

  - 1→2**过去结果** Y_{it-1} **不直接影响当前处理变量** D_{it}

  - 2→1**过去处理变量** D_{it-1} **不直接影响当前结果** Y_{it}

第 s 组别
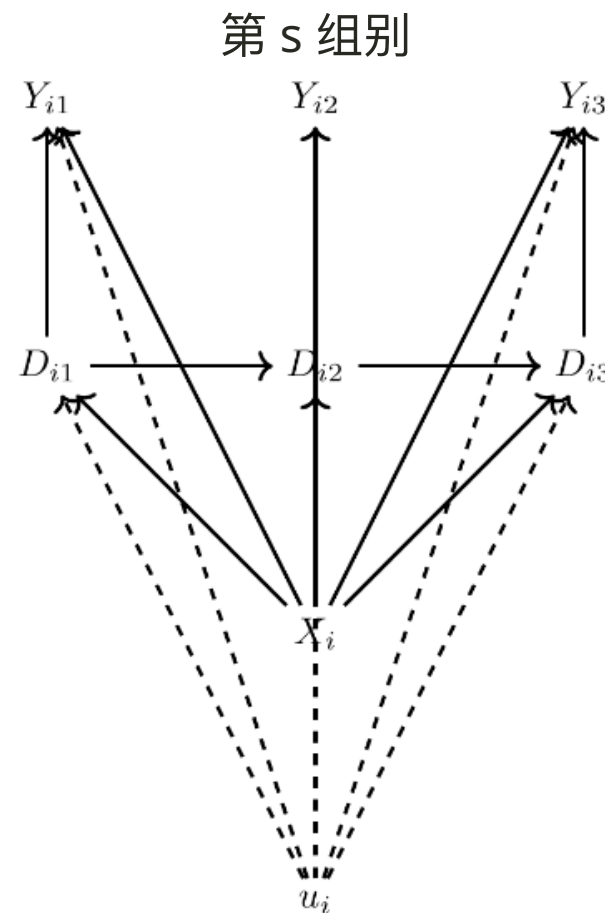
纵贯数据的有向无环图（Imai 和 Kim，2017）

- 基于数据的特殊结构和假设：可以使用**固定效应**（fixed effects）的特定面板方法来分离 D 对 Y 的因果效应

- 若固定效应中包含**年份固定效应**，该估计量通常称为**双向固定效应估计量**（Two-Way Fixed Effects Estimator，简称TWFE）

- 回归方程为：$Y_{it} = \alpha_0 + \delta D_{it} + \alpha_i + \alpha_t + X_{it} + \varepsilon_{it}$
  - $u_i = \alpha_i + \alpha_t$
  - u_i 称**未观测到的异质性**(unobserved heterogeneity)
  - ε_{it} 称**特异性误差**（idiosyncratic error）
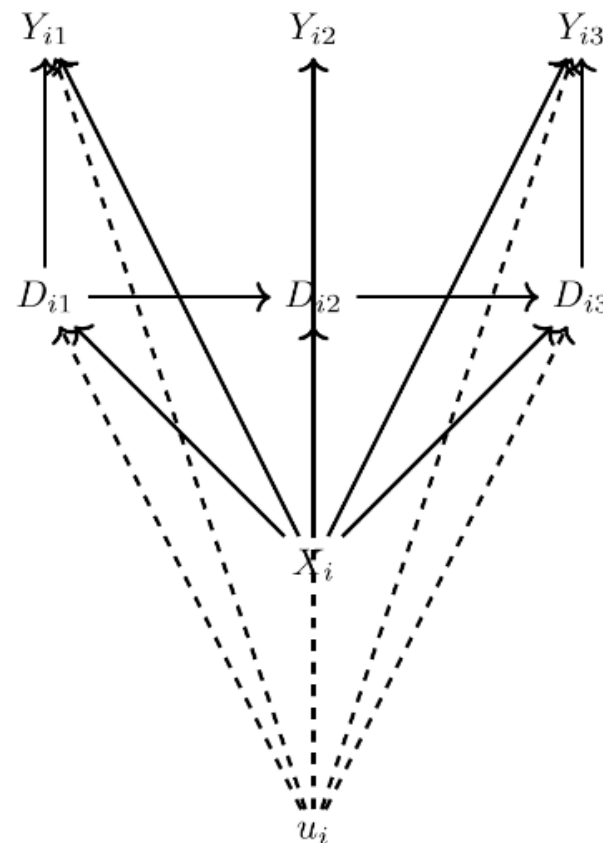
- 在 Y_{it} 对 D_{it} 回归时会发生什么?

估计方法 1：合并普通最小二乘（pooled OLS）

- 考虑**综合误差项** $\eta_{it} \equiv c_i + \varepsilon_{it}$
  - 这里的 c_i **视具体研究情境**而定。若以双向固定效应模型视角，则写作 $c_i = \alpha_i + \alpha_t$
- 则TWFE 可以改写为 $Y_{it} = \delta D_{it} + \eta_{it}; \quad t = 1, 2, \ldots, T$
- 如果

$$E\left[\eta_{it} \mid D_{i1}, D_{i2}, \ldots, D_{iT}\right] = E\left[\eta_{it} \mid D_{it}\right] = 0 \quad \text{对于所有的 } t = 1, 2, \ldots, T$$

  用 pooled OLS 估计该模型可以获得一致的 δ 估计值
  - 这相当假设**未观测到的异质性任何时期的** D_{it} **都不相关**，但前面的 DAG 并不支持该假设 → 存在**内生性**问题 → **遗漏变量偏差**

  - **序列相关问题**： u_i 在每个时期的 η_{it} 中，序列相关问题导致**异方差稳健标准误偏小** → **系数不显著**

估计方法 2：固定效应（组内估计量）[fixed effects（within estimator）]

- 假设没有年份固定效应，因此 TWFE 变为 $Y_{it} = \delta D_{it} + u_i + \varepsilon_{it}$

- 对该式进行**时间中心化**处理

$$Y_{it} = \delta D_{it} + u_i + \varepsilon_{it}$$
$$\bar{Y}_i = \delta \bar{D}_i + u_i + \bar{\varepsilon}_i$$
$$\left(Y_{it} - \bar{Y}_i\right) = \left(\delta D_{it} - \delta \bar{D}\right) + (u_i - u_i) + (\varepsilon_{it} - \bar{\varepsilon}_i)$$
$$\ddot{Y}_{it} = \delta \ddot{D}_{it} + \ddot{\varepsilon}_{it}$$

- 注意：
  - "组内"估计量="固定效应"估计量="双向固定效应"估计量（当含年份固定效应时）
  - 由于每个作者的遵循的命名传统不同，若文献中出现时，要该识别的方程具体形式
- 操作：
  - 对数据时间中心化处理后，进行 $\ddot{Y}_{it}$ 对 $\ddot{D}_{it}$ 的回归（需要矫正自由度）
  - 对原始方程添加个体虚拟变量来估计，**最小二乘虚拟变量 (LSDV) 估计量**
  - 在 STATA 固定效应模型中选择 fe 选项

识别假设1: 所有"右手变量"严格外生

- $E\left[\varepsilon_{it} \mid D_{i1}, D_{i2}, \ldots, D_{iT}, u_i\right] = 0$

- 比一般OLS中的假设更宽松，允许 D_{it} 与 u_i 任意相关
  - 含义：所有不可观测的影响因素中，那些不随时间推移而变化的固定效应全部都可以被 u_i 捕获，不存在遗漏。在这种情况下，回归过程中只限定与 ε_{it} 的相关性，不涉及与 u_i 的相关性

识别假设2: 无完全多重共线性与秩条件

- 考虑模型 $wage_{it} = \beta_0 + \beta_1 experience_{it} + \beta_2 male_i + \beta_3 white_i + u_i + \varepsilon_{it}$
  - 由于 $male_i$ 和 $white_i$，存在多重共线性问题。
  - 不能一致性估计系数 β_2 和 β_3，因为这些**时间恒定的不可观测因素固定效应 u_i 捕获**
- $\text{rank}\left(\sum_{t=1}^{T} E\left[\ddot{D}_{it}' \ddot{D}_{it}\right]\right) = K$
  - 估计系数 $\widehat{\beta_1} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$ 总是一个**缩放的**（scaled）协方差，而缩放要用到方差项
    → 进入回归的变量至少有部分的个体 i 是要随时间推移而变化的，并且不共线，才能确保 $\hat{\delta} \approx \delta$

推断与要点

- 该数据框架中的标准误必须以面板中的个体进行聚类，以允许同一个个体 i 的 ε_{it} 可以随时间相关

- 只要聚类的数量足够大，就可以产生有效推断
  - 经验法则：聚类数量要大于 30 个

  - 比如省级面板

- **要点 #1**: Fixed effects cannot solve reverse causality
  - 当存在反向因果关系时，DAG 就不再是前面提供的样子，会有 Y → D 的渠道出现

- **要点 #2**: Fixed effects cannot address time-variant unobserved heterogeneity
  - 变量中心化处理只是简单去除了一个未被观察到的随时间推移而变化的变量的均值

  - 当 $u_i \to u_{it}$ 时，中心化后的 $\ddot{u}_{it}$ 仍然与 $\ddot{D}_{it}$ 相关，所以后者仍然产生内生性

面板数据的总结

- 作用：

  - 固定效应（组内）能够消除所有与处理变量相关的**不随时间推移而变化的协变量**，无论该这些变量是否被观测到

  - 只要处理和结果随时间推移而变化，并且存在严格的外生性，那么固定效应（组内）估计量就能识别处理对结果的因果效应

- 限制：

  1. 不能处理**随时间推移而变化的未被观测到的异质性**
  2. 不能处理**强反向因果关系**

- 当存在上述2点限制时，固定效应（组内）不能估计一致因果效应，必须专项其他框架

DiD

介绍

- **差分法** (DiD) 是在非实验环境中估计因果效应的最受欢迎的策略之一。
  - 在NBER WPs中使用率超过20%

- 近几年，关于DiD的计量经济学研究激增

- 出现时间比随机实验还要早大约 85 年

- 区分为经典模型（一组个体同时受处理）和常见模型（一组个体在不同时间接受处理）

经典 2×2 DiD 模型

- Goodman-Bacon(2019)
    - 处理组和控制组各有2个时期：
    - 处理组有一个前期pre(k)和一个后期post(k)
    - 控制组有一个前期pre(U)和一个后期post(U)

经典 2×2 DiD 模型

$$
\hat{\delta}_{kU}^{2\times2}(1) = (\bar{y}_k^{\text{post}(k)} - \bar{y}_k^{\text{pre}(k)}) - (\bar{y}_U^{\text{post}(k)} - \bar{y}_U^{\text{pre}(k)})
$$

$$
(2) = (E[Y_k \mid \text{Post}] - E[Y_k \mid \text{Pre}]) - (E[Y_U \mid \text{Post}] - E[Y_U \mid \text{Pre}])
$$

$$
(3) = \underbrace{(E[Y_k^1 \mid \text{Post}] - E[Y_k^0 \mid \text{Pre}]) - (E[Y_U^0 \mid \text{Post}] - E[Y_U^0 \mid \text{Pre}].)}_{\text{潜在结果转换方程}}
$$

$$
+ \underbrace{E[Y_k^0 \mid \text{Post}] - E[Y_k^0 \mid \text{Post}]}_{\text{加上一个 "0"}}
$$

$$
(4) = \underbrace{E[Y_k^1 \mid \text{Post}] - \overbrace{E[Y_k^0 \mid \text{Post}]}^{\textcolor{red}{\text{反事实}}}}_{\text{ATT}}
$$

$$
+ \underbrace{[E[Y_k^0 \mid \text{Post}] - E[Y_k^0 \mid \text{Pre}]] - [E[Y_U^0 \mid \text{Post}] - E[Y_U^0 \mid \text{Pre}]]}_{2\times2 \text{ 情况下的非平行趋势偏差}}
$$

- 等号（1）：$\hat{\delta}_{kU}^{2\times2}$ 为 k 组估计的 ATT，右边第一项为处理组k 处理后减去处理前的差值
- 等号（2）：重新写成条件期望形式

经典 2×2 DiD 模型

$$\hat{\delta}_{kU}^{2\times2}(1) = (\bar{y}_k^{\text{post}(k)} - \bar{y}_k^{\text{pre}(k)}) - (\bar{y}_U^{\text{post}(k)} - \bar{y}_U^{\text{pre}(k)})$$

$$(2) = (E[Y_k \mid \text{Post}] - E[Y_k \mid \text{Pre}]) - (E[Y_U \mid \text{Post}] - E[Y_U \mid \text{Pre}])$$

$$(3) = \underbrace{(E[Y_k^1 \mid \text{Post}] - E[Y_k^0 \mid \text{Pre}]) - (E[Y_U^0 \mid \text{Post}] - E[Y_U^0 \mid \text{Pre}].)}_{\text{潜在结果转换方程}}$$

$$+ \underbrace{E[Y_k^0 \mid \text{Post}] - E[Y_k^0 \mid \text{Post}]}_{\text{加上一个 "0"}}$$

$$(4) = \underbrace{E[Y_k^1 \mid \text{Post}] - \overbrace{E[Y_k^0 \mid \text{Post}]}^{\text{反事实}}}_{\text{ATT}}$$

$$+ \underbrace{[E[Y_k^0 \mid \text{Post}] - E[Y_k^0 \mid \text{Pre}]] - [E[Y_U^0 \mid \text{Post}] - E[Y_U^0 \mid \text{Pre}]]}_{2\times2\ \text{情况下的非平行趋势偏差}}$$

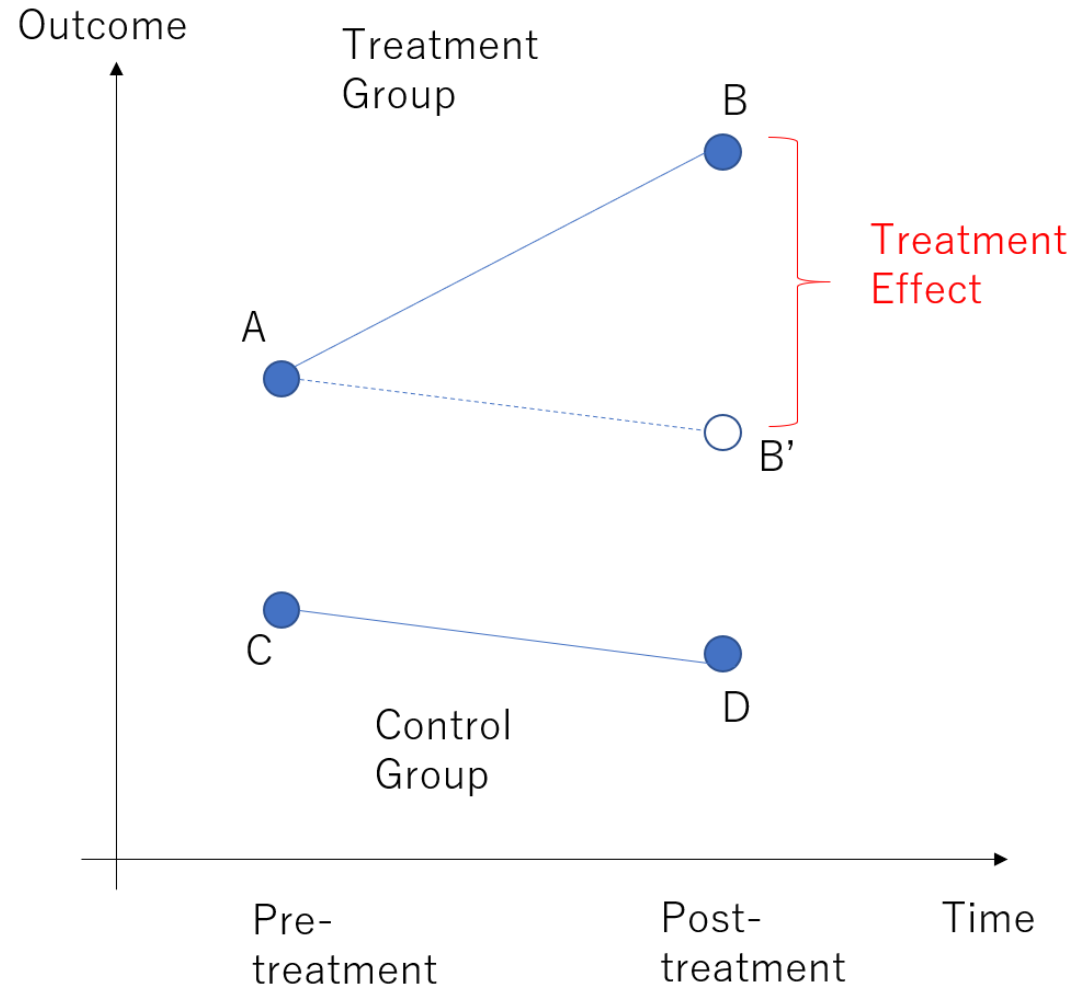- 等号（3）：潜在结果转换方程讲 Y 的历史取值转换成潜在结果形式

- 等号（4）：重新派丽，根据潜在结果的条件期望对 2×2 DiD进行分解 → 第二项包含两个差值，第一个差值（涉及处理组 k，是反事实），第二个差值（涉及控制组）

DID in Figure

Plan of the Lecture

- Framework

- Implementation in a regression framework

- Discussion on Parallel Trend Assumption

Reference

- Angrist and Pischke "Mostly Harmless Econometrics" Chapter 5

- Bertrand, Duflo, and Mullainathan (2004, QJE)
  - Discuss issues of calculating standard errors in the DID method.

- Ishise, Kitamura, Kudamatsu, Matsubayashi, and Murooka (2019)
  - Guidance on how to conduct policy evaluation with DID
  - Slide
  - Paper

Framework

Framework

- Two periods: $t = 1, 2$. Treatment implemented at $t = 2$.

- Y_{it}: observed outcome for person i in period t

- G_i: dummy for treatment group

- D_{it}: treatment status
  - $D_{it} = 1$ if $t = 2$ and $G_i = 1$

- **Potential outcomes**
  - $Y_{it}(1)$: outcome for i when she is treated
  - $Y_{it}(0)$: outcome for i when she is not treated

- Then

$$Y_{it} = D_{it} Y_{it}(1) + (1 - D_{it}) Y_{it}(0)$$

Identification

- Goal: ATT at $t = 2$

$$E[Y_{i2}(1) - Y_{i2}(0)|G_i = 1] = E[Y_{i2}(1)|G_i = 1] - E[Y_{i2}(0)|G_i = 1]$$

- What we observe

	Pre-period ($t = 1$)	Post ($t = 2$)
Treatment ($G_i = 1$)	$E[Y_{i1}(0) \mid G_i = 1]$	$E[Y_{i2}(1) \mid G_i = 1]$
Control ($G_i = 0$)	$E[Y_{i1}(0) \mid G_i = 0]$	$E[Y_{i2}(0) \mid G_i = 0]$

- Under what assumptions can we the ATT?

Simple Comparisons

- Within period comparison
  - If $E[Y_{i2}(0)|G_i = 1] = E[Y_{i2}(0)|G_i = 0]$, ATT is identified by

$$ATT = E[Y_{i2}(1)|G_i = 1] - E[Y_{i2}(0)|G_i = 0]$$

  - Selection bias?

- Before-after comparison

  - If $E[Y_{i2}(0)|G_i = 1] = E[Y_{i1}(0)|G_i = 1]$, ATT is identified by

$$ATT = E[Y_{i2}(1)|G_i = 1] - E[Y_{i1}(0)|G_i = 1]$$

  - Time trend?

- More resonable assumption?

Parallel Trend (並行トレンド) Assumption

- Assumption:

$$E[Y_{i2}(0) - Y_{i1}(0)|G_i = 0] = E[Y_{i2}(0) - Y_{i1}(0)|G_i = 1]$$

- Interpretation: Change in the outcome *without treatment* is the same across two groups.

- This puts an assumption on **counterfactual** trend in treatment group if it had not been treated.

Difference-in-differences

- Note that

$$\underbrace{E[Y_{i2}(1) - Y_{i2}(0)|G_i = 1]}_{ATT} = E[Y_{i2}(1)|G_i = 1] - E[Y_{i2}(0)|G_i = 1]$$

$$= E[Y_{i2}(1)|G_i = 1] - E[Y_{i1}(0)|G_i = 1]$$
$$- \underbrace{(E[Y_{i2}(0)|G_i = 1] - E[Y_{i1}(0)|G_i = 1])}_{=E[Y_{i2}(0)-Y_{i1}(0)|G_i=0] \ (pararell\ trend)}$$

- Thus,

$$ATT = E[Y_{i2}(1) - Y_{i1}(0)|G_i = 1] - E[Y_{i2}(0) - Y_{i1}(0)|G_i = 0]$$

- ATT is identified by **the difference of two differences**

Estimation

Estimation Approach 1: Plug-in Estimator

- Remember that the ATT is identified by

$$ATT = E[Y_{i2}(1) - Y_{i1}(0)|G_i = 1] - E[Y_{i2}(0) - Y_{i1}(0)|G_i = 0]$$

- Replace them with the sample average.

$$\hat{ATT} = \{\bar{y}(t = 2, G = 1) - \bar{y}(t = 1, G = 1)\}$$
$$- \{\bar{y}(t = 2, G = 0) - \bar{y}(t = 1, G = 0)\}$$

where $\bar{y}(t, G)$ is the sample average for group G in period t .

Example: Card and Kruger (1994, AER)

- Question: Effects of the minimum wage increase on employment.

- On April 1, 1992, New Jersey's minimum wage rose from 4.25 to 5.05 USD.

- Compare fast-food restaurants in New Jersey (treament group) and eastern Pennsylvania (control group) before and after the rise.

| Variable | Stores by state | | |
	PA (i)	NJ (ii)	Difference, NJ − PA (iii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (0.51)	− 2.89 (1.44)
2. FTE employment after, all available observations	21.17 (0.94)	21.03 (0.52)	− 0.14 (1.07)
3. Change in mean FTE employment	− 2.16 (1.25)	0.59 (0.54)	2.76 (1.36)

Figure 1: image

Estimation Approach 2: Linear Regression

- Run the following regression

$$y_{it} = \alpha_0 + \alpha_1 G_i + \alpha_2 T_t + \alpha_3 D_{it} + \beta X_{it} + \epsilon_{it}$$

  - G_i: dummy for treatment group
  - T_t :dummy for treatment period
  - $D_{it} = G_i \times T_t$. α_3 captures the ATT.

- Regression framework can incorporate covariates X_{it}, which is important to control for observed confounding factors.

Two-way Fixed Effects Estimator

- With panel data

$$y_{it} = \alpha D_{it} + \beta X_{it} + \epsilon_i + \epsilon_t + \epsilon_{it}$$

  - ϵ_i is individual FE

  - ϵ_t is time FE.

- Use the cluster-robust standard errors for inference! (Bertrand, Duflo, and Mullainathan 2004, QJE)

Event Study (イベントスタディ) Specification

- With multiple periods, we can estimate **the treatment effect in the post-treatment periods**

- The specification with k-periods lags and leads

$$y_{it} = \epsilon_i + \epsilon_t + \sum_{\tau=-k}^{k} \gamma^\tau D_{it}^\tau + \beta X_{it} + \epsilon_{it}$$

  - D_{it}^τ takes 1 if unit i is in treatment group and period t is τ period far from the treatment timing.
  - Normalize $\gamma^{-1} = 0$.

- Why include lags (i.e., $\tau = -k, \cdots, -1$)? See this later.

Discussion on Parallel Trend Assumption

Discussions on Parallel Trend

- Parallel trend assumption can be violated in various situations.

- If treatment status depends on **time-varying factors**, treatment and control group may have differential time trend in their outcome.
    - Note: DID can only deal with **time-invariant factors**.

- Example: Self-selection into treatment based on time-varying factor
    - People participate in worker training program because they expected a future decrease in earnings before they enter the program.

Diagnostics for Parallel Trends: Pre-treatment trends

- Check if the trends are parallel in the pre-treatment periods (プレトレンド)

- Requires data on multiple pre-treatment periods.

- Should do this if you have multiple pre-treatment periods.

- Note: this is only **diagnostics**. NOT a direct test of the assumption!

  - Remember: Parallel trend assumption is not testable.
  - You **should never say** "the key assumption for DID is satisfied if the pre-treatment trends are parallel.

# Other Diagnostics: Placebo Effects in Pre-Treatment

- Idea: If both groups are indeed similar, there should be **no treatment effect in the pre-treatment period.** (placebo outcome)

- We run the following event-study specification

$$y_{it} = \epsilon_i + \epsilon_t + \sum_{\tau=-k}^{k} \gamma^\tau D_{it}^\tau + \beta X_{it} + \epsilon_{it}$$

- The estimates of γ^τ should be close to zero up to the beginning of treatment

Example: Autor (2003) (from Angrist and Pischke MHE)

- Figure 5.2.4: Estimated impact of state courtsí adoption of an implied-contract exception to the employmentat- will doctrine on use of temporary workers (from Autor 2003). The dependent variable is the log of state temporary help employment in 1979 - 1995. Estimates are from a model that allows for eﬀects before, during, and after adoption.
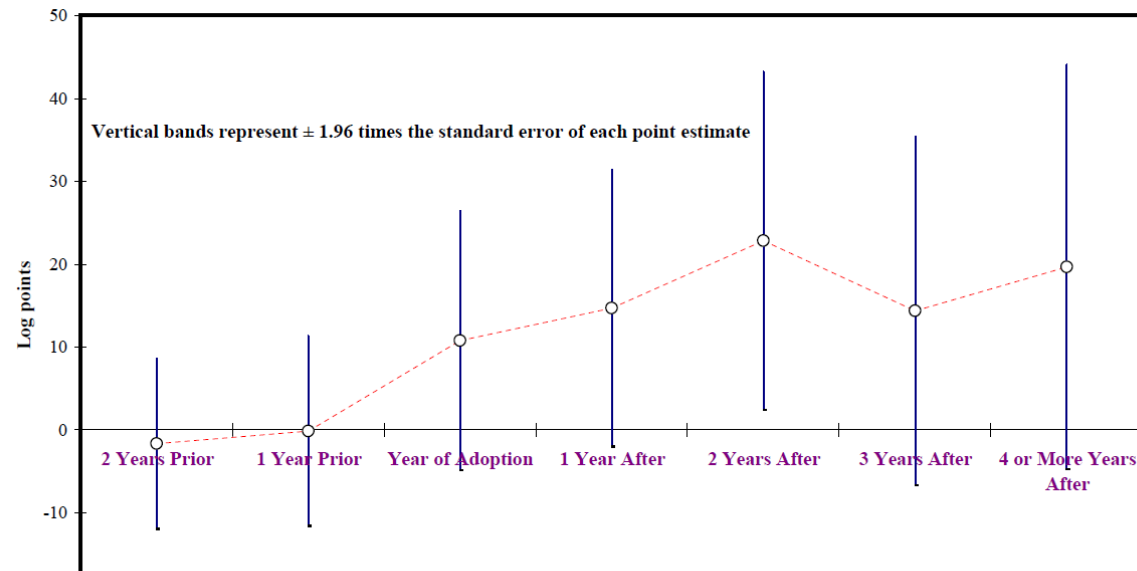
Example: Kubota, Onishi, and Toyama (2021)

- To estimate the effect of receiving the SCP (特別定額給付金) on consumption, we run

$$y_{itw} = \alpha_i + \alpha_{iw} + \alpha_{tpw} + \sum_{k=-5}^{5} \gamma^k D_{itw}^k + u_{itw}$$

  - Household i in week w in year $t(= 2019, 2020)$.
  - $D_{itw}^k = 1\{T_i - w = k\}$, T_i is the timing of receiving the SCP (i.e., treatment)
  - y_{itw}: consumption measure (total withdrawal, ATM withdrawal)
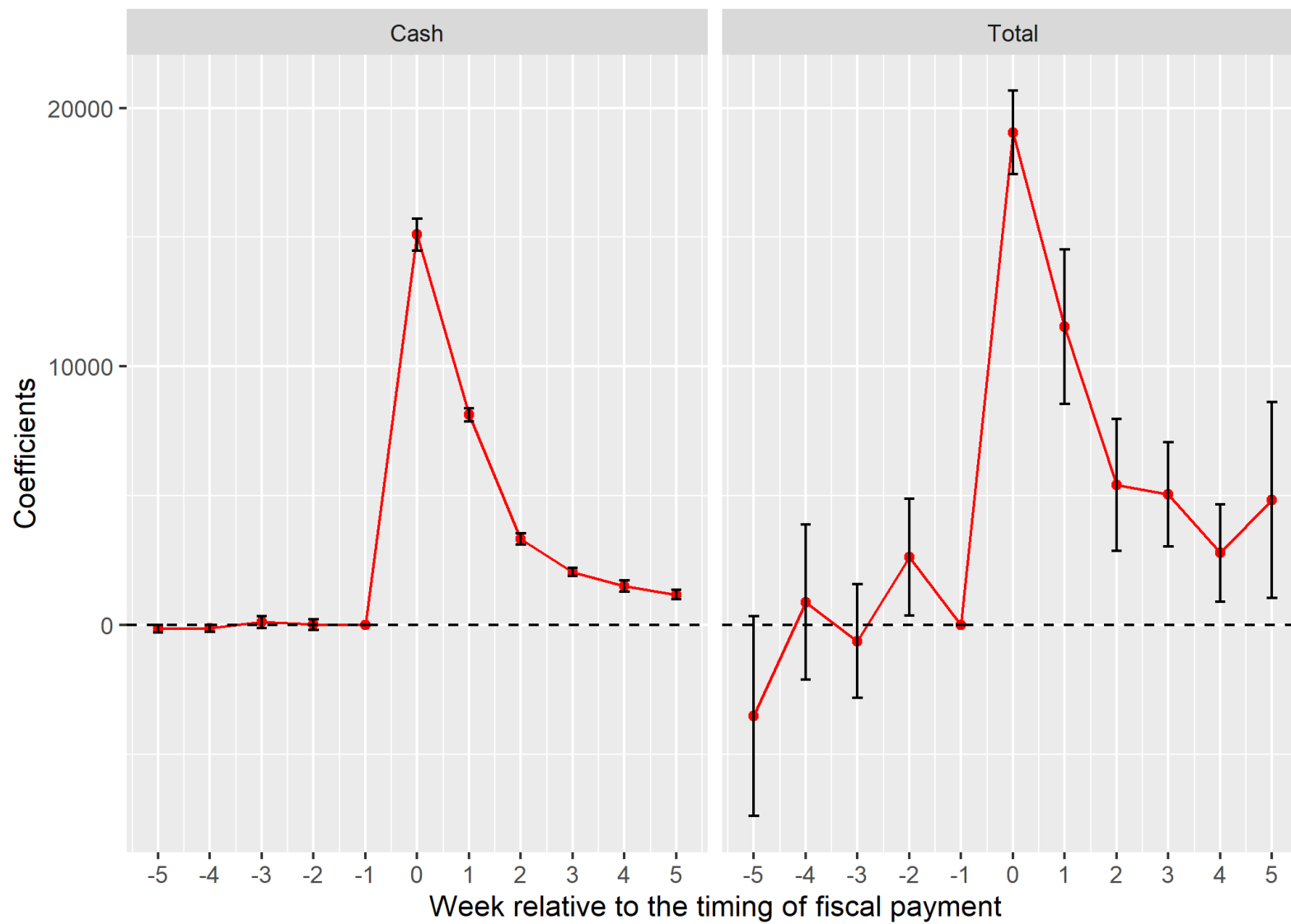
Application

Application: JAL-JAS Merger (Doi and Ohashi 2019)

- Question: What is the effects of JAL-JAS merger on market outcome in airline industry.

- They use both DID and structural estimation (構造推定) approach to answer the question.

- Here, I only discuss their DID analysis.

Background of JAL-JASmerger

- Around 2000, JAL and JAS tried to create a holding company (持株会社) to merge.
    - Market share: JAL 25%, JAS 24%, ANA 48%

- The JFTC (公正取引委員会) issued an interim report in March 2002 and claimed that it was likely to be **a substantial restraint of competition in the domestic air passenger market**.

- The merging party proposed remedial measures.
    - Set the pre-merger price level as the price ceiling for the post-merger period.
    - Release 9 slots at the most congested airport, and be assigned to new entrants.

- The JFTC approved the merger with the remedies and it was consummated on Oct 2002.

Theoretical Effects of a Merger

- **Anticompetitive effect (反競争効果)**: The market becomes more concentrated, less competition, higher price, lower welfare.

- **Efficiency effect (効率性効果)**: The merged party may become more efficient through scale effect, synergy, knowledge transfer, etc.

- These two have different implications on economic welfare (Williamson's tradeoff)

- The JFTC (公取委) reviews the merge case and decides whether approve it.

Descriptive Statistics

Table 2
Monthly relative price and flight frequency.

	Number of Routes [Number of slots available by remedy]		Average shares (%)	2000		2002		2004		(Post-merger averaged outcomes) - (pre-merger averaged outcomes)	
				Price	Frequency	Price	Frequency	Price	Frequency	Price	Frequency
All routes	274			1.3	2.2	1.8	2.3	1.9	2.7	1.2	1.2***
				(14.8)	(1.8)	(14.9)	(1.9)	(15.3)	(2.6)		
By market structure											
Merger-to-monopoly	6		1.5	11.5	2.0	11.5	2.3	15.8	3.3	1.3***	1.5***
				(7.5)	(0.6)	(4.8)	(0.6)	(4.3)	(0.7)		
Merger-to-oligopoly	27	[3]	58.7	−9.8	4.6	−9.7	4.8	−10.1	6.5	1.0	1.3***
				(11.6)	(2.6)	(10.7)	(2.9)	(10.2)	(4.2)		
Monopoly	209		20.3	2.8	1.7	3.4	1.8	3.9	1.9	1.3	1.1*
				(15.3)	(1.2)	(15.5)	(1.3)	(15.9)	(1.6)		
Oligopoly	32	[2]	19.4	0.2	3.4	0.9	3.5	−0.7	3.7	−2.6	1.1
				(9.9)	(1.9)	(10.0)	(1.9)	(11.0)	(2.2)		

Notes: The numbers are means for July of each year with standard deviations in the parenthesis. The numbers in the squared brackets are the numbers of slots transferred under the structural remedy. Price is the distance-adjusted price level (in constant 2005 JPY), as a ratio to the industry average. Frequency is the number of round trips per day. The superscripts, ***, **, and * indicate significance at the 99-, 95-, and 90% confidence levels, respectively. The pre-merger period ranges from April 2000–July 2002, whereas the post-merger period is from October 2002 to October 2005.

DID Analysis

- Outcome: Price, flight frequency, # of domestic passengers

- 274 routes in 2000–2005

- Regression equation

$$\ln\left(y_{jmt}\right) = \gamma_1^A \cdot JJ_{jmt} + \gamma_2^A \cdot post_t + \gamma_3^A \cdot JJ_{jmt} \cdot post_t + \mathbf{x}'_{jmt} \cdot \lambda^A + \kappa_{jmt}^A$$

  - j: firm, m: market (route), t: time (monthly)
  - $post_t = 1$ if t is after October 2002.
  - $JJ_{jmt} = 1$ if a firm j in market m is either JAL or JAS (JAL group).
  - $\mathbf{x}'_{jmt}$: covariates.

- Control group: Non-merged domestic airlines (ANA and others)
  - Either in the same market as JAL–JAL or in a market without JAL-JAS.

- Allow heterogeneous effects of a merger

$$
\begin{aligned}
\ln\left(y_{jmt}\right) =& \gamma_1^B \cdot JJ_{jmt} + \gamma_2^B \cdot post_t + \gamma_3^B \cdot JJ_{jmt} \cdot post_t \\
& + MtM_{jmt} \cdot \left(\gamma_4^B \cdot JJ_{jmt} + \gamma_5^B \cdot post_t + \gamma_6^B \cdot JJ_{jmt} \cdot post_t\right) \\
& + MtO_{jmt} \cdot \left(\gamma_7^B \cdot JJ_{jmt} + \gamma_8^B \cdot post_t + \gamma_9^B \cdot JJ_{jmt} \cdot post_t\right) \\
& + x'_{jmt} \cdot \lambda^B + \kappa_{jmt}^B
\end{aligned}
$$

  - $MtM_{jmt} = 1$ if the route became a monopoly after the JJ merger

  - $MtO_{jmt} = 1$ if it became oligopoly.

Results

Table 3
DID estimates.

	Prices		Flight frequency		Number of passengers	
JJ* post	−0.01**	−0.013**	0.285***	0.082***	0.385***	0.185***
	(0.004)	(0.005)	(0.016)	(0.021)	(0.020)	(0.027)
JJ * post * MTM		0.051***		0.578***		0.966***
		(0.012)		(0.048)		(0.062)
JJ * post * MTO		−0.004		0.287***		0.246***
		(0.008)		(0.033)		(0.042)
R-squared	0.97	0.97	0.93	0.94	0.96	0.96
Number of obs.	5329		5329		5329	

Notes: Each dependent variable is a logarithmic form. The numbers in the parentheses are the standard errors. The superscripts, ***, **, and * indicate significance at the 99-, 95-, and 90% confidence levels, respectively.

Robustness Check

- Pretrend: Regress the outcome on route-FE, time FE, and time FE interacted with the dummy for control routes.

- Placebo test: Use the period b.w. April 2000 to July 2002. Consider placebo merger in July 2001 and see its effects using DID.

Limitations of DID in this context

- Is the choice of control group appropriate?

- How to interpret the results?

- Further analysis using structural estimation approach.

课程计划

- 框架

- R 语言的实现

- **差分法 (DID, 差の差分法)**

选择性偏差

- **选择性偏差**（selection bias）：样本选择过程中的非随机而导致的结论偏差，包括**自选择偏差**（self-selection bias）和**样本选择偏差**（sample-selection bias）
  - 自选择偏差：实验组的非随机分配导致的 → 是否进入实验组是个体选择的结果或是与个体相关的因素（包括可观测和不可观测两类）导致的
  - 例子
  - 样本选择偏差：非随机选择的样本不能够反映总体的特征，从而使结果估计量产生偏差
  - 例子
- 评估政策处理效应的方法就是尽可能的降低选择性偏差

选择性偏差的表达

- 直接面对数据时，通常使用处理样本结果变量的均值与未处理样本结果变量的均值之差来估计 ATE
    - differen-in-mean，简称 DIM
    - simple-difference-in-outcome，简称 SDO
    - Naiive comprison

$$DIM = E\left[Y^1 \mid D = 1\right] - E\left[Y^0 \mid D = 0\right]$$
$$= \frac{1}{N_T}\sum_{i=1}^{n}\left(y_i \mid d_i = 1\right) - \frac{1}{N_C}\sum_{i=1}^{n}\left(y_i \mid d_i = 0\right)$$

- 当满足**独立性假设**（independence assumption，简称 IA）时：$(Y_1; Y_0) \perp D$
    - DIM = ATE = ATT = ATU