体育经济分析：理论与应用

附加实证专题：面板数据+DID

周正卿

02 May 2022

引言

大纲

上次实证方法

- 匹配估计 (*MHE* 3.2 和 C&T 25.4).

今天

- panel data + DID

引言

引言

- 面板数据(Panel data)
    - 同时包含横截面(cross-section)和时间序列(time series)
- 有用的
    - 更多变异（包括横截面和时间序列的变化）
    - 可以处理**非时变的不可观察到的因素**。

引言

- 分类1：短面板和长面板
  - 短面板是指个体维度 N 较大，时间维度 T 较小
  - 长面板是指数据的 N 较小，T 较大
- 分类2：平衡面板与非平衡面板
  - 平衡面板: 每个个体都有相同时间 T 的观测点
  - 非平衡面板: 有部分个体没有相同时间 T 的观测点
  - 若非平衡面数据由随机原因造成的，那么处理方法和平衡面板一样，但是如果数据缺失由非随机原因造成的，则必须考虑缺失的原因: 如**样本选择偏差**

引言

Framework

Framework with Panel Data

- Consider the model

$$y_{it} = \beta' x_{it} + \epsilon_{it}, E[\epsilon_{it}|x_{it}] = 0$$

  where x_{it} is a k-dimensional vector

- If there is no correlation between x_{it} and ϵ_{it}, you can estimate the model by OLS **(pooled OLS)**

- A concern here is the omitted variable bias.

Introducing fixed effect (固定効果)

- Suppose that ϵ_{it} is decomposed as

$$\epsilon_{it} = \alpha_i + u_{it}$$

where α_i is called **unit fixed effect (固定効果)**, which is the time-invariant unobserved heterogeneity.

- With panel data, we can control for the unit fixed effects by incorporating the dummy variable for each unit i!

$$y_{it} = \beta' x_{it} + \gamma_2 D2_i + \cdots + \gamma_n Dn_i + u_{it}$$

where Dl_i takes 1 if $l = i$.

Fixed Effect Model

- Model

$$y_{it} = \beta' x_{it} + \alpha_i + u_{it}$$

- Assumptions:
  1. u_{it} is uncorrelated with $(x_{i1}, \cdots, x_{iT})$, that is $E[u_{it}|x_{i1}, \cdots, x_{iT}] = 0$
  2. (Y_{it}, x_{it}) are independent across individual i.
  3. No outliers
  4. No perfect multicollinarity between explantory variables x_{it} and fixed effects α_i.

Assumption 1: Mean independence

- Assumption 1 is weaker than the assumption in OLS.
- Here, the time-invariant unobserved factor is captured by the fixed effect α_i.

Assumption 4: No Perfect Multicolinearity

- Consider the following model

$$wage_{it} = \beta_0 + \beta_1 experience_{it} + \beta_2 male_i + \beta_3 white_i + \alpha_i + u_{it}$$

  - $experience_{it}$ measures how many years worker i has worked before at time t.

- Multicollinearity issue because of $male_i$ and $white_i$.

- Intuitively, we cannot estimate the coefficient β_2 and β_3 because those **time-invariant variables are captured by the unit fixed effect α_i**.

Estimation

Estimation with Fixed Effects

- Can estimate the model by adding dummy variables for each individual.
  - **least square dummy variables (LSDV) estimator**.
  - Computationally demanding with many cross-sectional units
- We often use the following **within transformation**.

Estimation by within transformation

- Define the new variable $\tilde{Y}_{it}$ as

$$\tilde{Y}_{it} = Y_{it} - \bar{Y}_i$$

  where $\bar{Y}_i = \frac{1}{T}\sum_{t=1}^{T} Y_{it}$.

- Applying the within transformation, can eliminate the unit FE α_i

$$\tilde{Y}_{it} = \beta' \tilde{X}_{it} + \tilde{u}_{it}$$

- Apply the OLS estimator to the above equation!.

Importance of within variation in estimation

- The variation of the explanatory variable is key for precise estimation.

- Within transformation eliminates the time-invariant unobserved factor,
  - a large source of endogeneity in many situations.

- But, within transformation also absorbs the variation of X_{it}.

- Remember that

$$\tilde{X}_{it} = X_{it} - \bar{X}_i$$

  - The transformed variable $\tilde{X}_{it}$ has the variation over time t within unit i.
  - If X_{it} is fixed over time within unit i, $\tilde{X}_{it} = 0$, so that no variation.

Various Fixed Effects

- You can also add **time fixed effects (FE)**

$$y_{it} = \beta' x_{it} + \alpha_i + \gamma_t + u_{it}$$

- The regression above controls for both **time-invariant individual heterogeneity** and **(unobserved) aggregate year shock**.

- Panel data is useful to capture various unobserved shock by including fixed effects.

Cluster-Robust Standard Errors

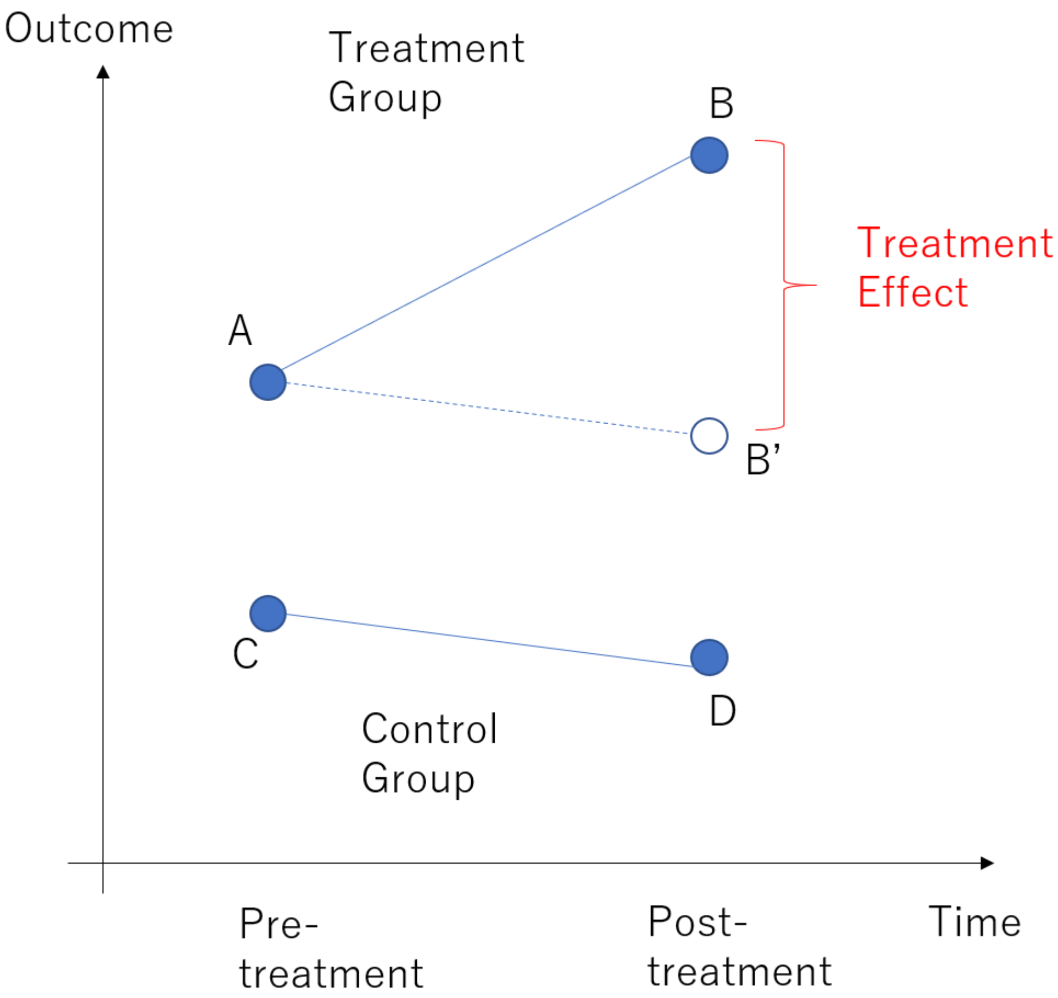
- In OLS, we considered two types of error structures:
  1. Homoskedasticity $Var(u_i) = \sigma^2$
  2. Heteroskedasitcity $Var(u_i|x_i) = \sigma(x_i)$
- They assume the independence between observations, that is $Cov(u_i, u_{i'}) = 0$.
- In the panel data setting, we need to consider the **autocorrelation (自己相関)**.
  - the correlation between u_{it} and $u_{it'}$ across periods for each individual i.
- **Cluster-robust standard error (クラスターに頑健な標準誤差)** considers such autocorrelation.
  - The cluster is unit i. The errors within cluster are allowed to be correlated.

Introduction

Introduction

- **Difference-in-differences (DID, 差の差分法)** exploits the panel structure to estimate the causal effect.
- Two types of simple comparisons
  - Treatment and control group comparison: selection bias
  - Before v.s. after comparison: time trend
- DID combines those two comparisons to draw causal conclusion.

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

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

- 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.

$$\begin{aligned} AT\hat{T} =& \{\bar{y}(t = 2, G = 1) - \bar{y}(t = 1, G = 1)\} \\ & - \{\bar{y}(t = 2, G = 0) - \bar{y}(t = 1, G = 0)\} \end{aligned}$$

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

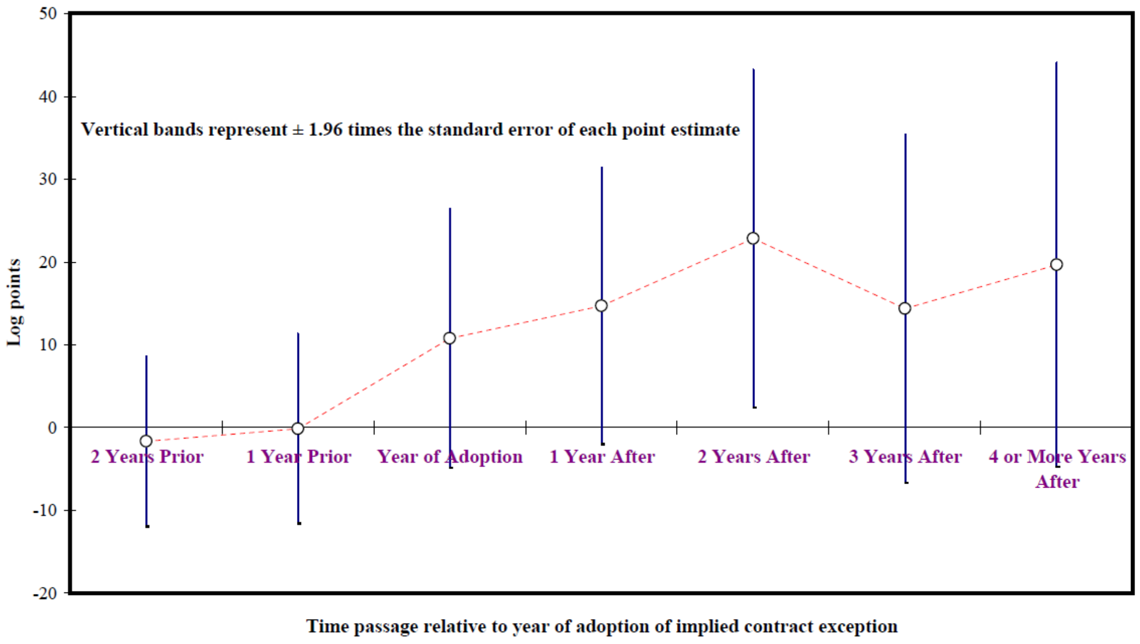
- Idea: If both groups are indeed similar, there should be **no treatment effect in the pre-treatment period.** (placebo outcome)
- We run the following event-study specification

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

- 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 e§ects before, during, and after adoption.

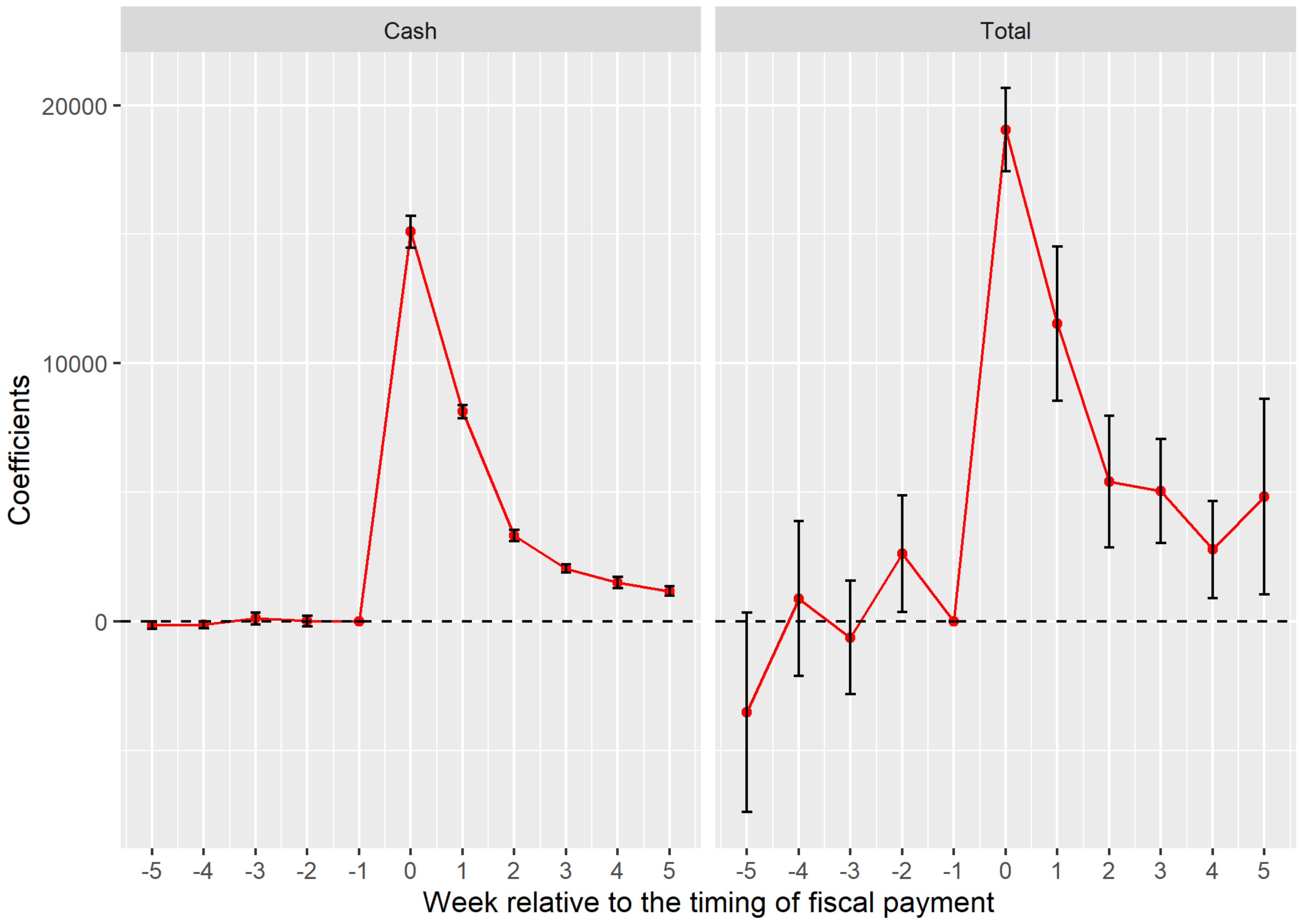


Example: Kubota, Onishi, and Toyama (2021)

- 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)


Cash
Total
20000
10000
0
Coefficients
-5
-4
-3
-2
-1
0
1
2
3
4
5
Week relative to the timing of fiscal payment

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

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

  - 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.