Econometrics 2

Difference in differences

Sergey Lychagin

CEU

Winter 2020

John Snow's study of cholera infection in London

- Lambeth Company sourced water downstream in 1849, upstream in 1854.
- Southwark & Vauxhall sources water downstream in both years.

Sub-Districts.	Deaths from Cholera in 1849.	Deaths from Cholera in 1854.	Water Supply.
First 12 sub-districts .	2261	2458	Southwk. & Vauxhall.
Next 16 sub-districts .	3905	2547	Both Companies.
Last 4 sub-districts .	162	37	Lambeth Company.

Effect of minimum wage on teenage employment

- Theory: Lower demand $\Rightarrow$ unemployment
- New Jersey: raised minimum wage from \$4.25 to \$5.05 on April 1, 1992.

Card and Krueger (1994): "Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania"

Effects of minimum wage change on teenage employment

1. Before-After Design

$$Y_{it} = \alpha + \rho D_{it} + u_{it}$$

- Y_{it} employment of teenage workers in New Jersey fast food restaurants at $t = \text{Feb}, \text{Nov}1992$, and

$$D_{it} = \begin{cases} 1 & \text{if } t = Nov \\ 0 & \text{if } t = Feb \end{cases}$$

- Is ρ causal effect of the minimum wage change? Assumption:
  - In absence of minimum wage change ρ would be zero.
  - No other factors change employment over time.
  - Change in employment in NJ between February and November only due to change in minimum wage.

Very brave assumptions! How about seasonal changes? Business cycle? Other factors?

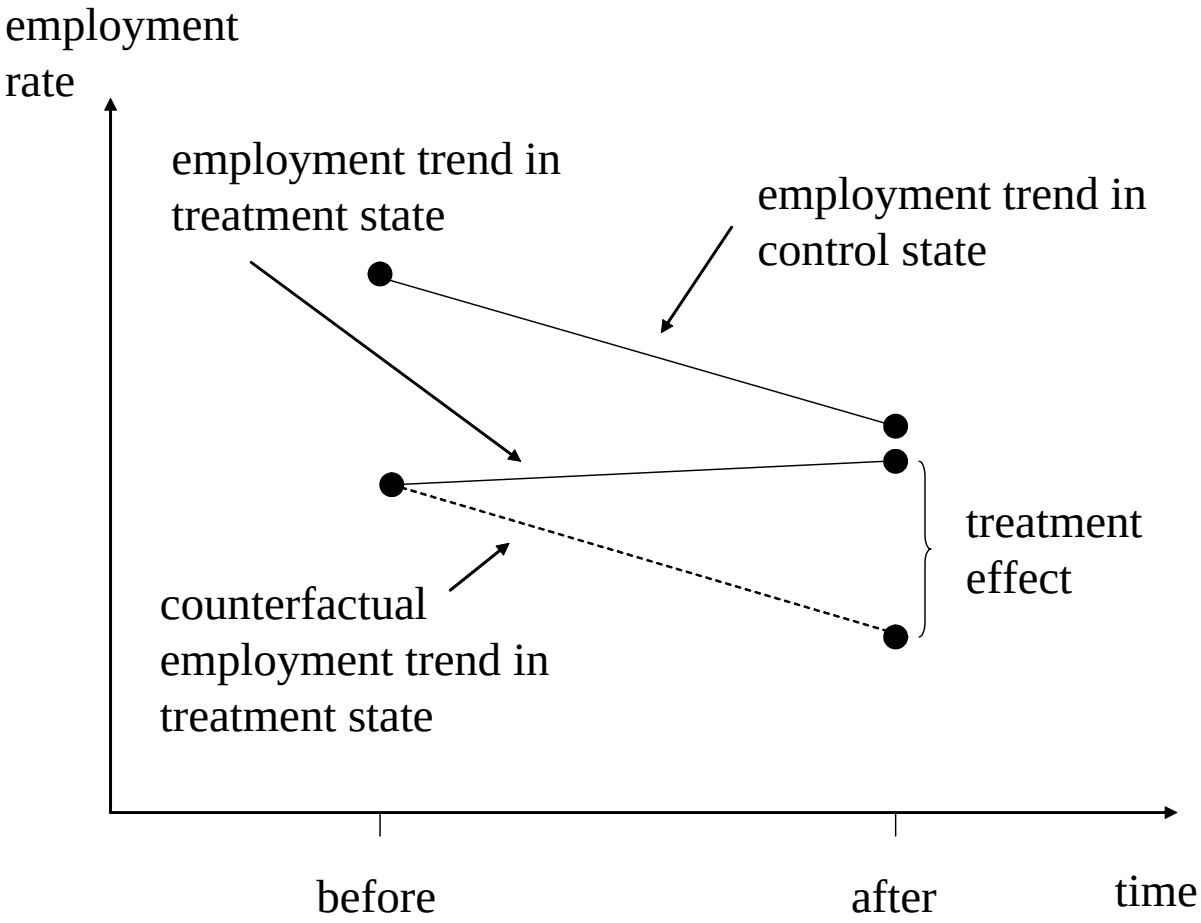

Figure 5.2.1: Causal effects in the differences-in-differences model

Differences in Differences

- Panel/repeated CS design where the regressor of interest varies at a more aggregate level
- Policy changes over time affecting certain population groups, regions, ...
- OVB: unobserved variables at the state/year/. . . level: *group fixed effects*.

Differences in Differences

2. Before-After Design with Untreated Comparison Group

- Pennsylvania left the minimum wage unchanged in 1992
- Compare employment change in PA fast food restaurants
- Potential Outcome:
  - Y_{1ist} restaurant i, state s, time t with high min. wage
  - Y_{0ist} with low min wage

$$\begin{aligned} E[Y_{0ist}|s,t] &= \gamma_s + \lambda_t \\ E[Y_{1ist} - Y_{0ist}|s,t] &= \rho \end{aligned}$$

Let D_{st} be a dummy equal one if minimum wage is high

$$Y_{ist} = \gamma_s + \lambda_t + \rho D_{st} + \epsilon_{ist}$$

Differences in Differences

$$
\begin{aligned}
E[Y_{ist}|s = PA, t = Nov] &- E[Y_{ist}|s = PA, t = Feb] \\
&= \lambda_{Nov} - \lambda_{Feb} \qquad (1) \\
E[Y_{ist}|s = NJ, t = Nov] &- E[Y_{ist}|s = NJ, t = Feb] \\
&= \lambda_{Nov} - \lambda_{Feb} + \rho \qquad (2)
\end{aligned}
$$

Differences in Differences:

$$(2) - (1) = \rho$$

Parallel trends assumption:

- In absence of minimum wage change employment in NJ would have developed the same way as in PA
- No other factors only affect NJ and not PA

Regression Diff in Diff

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \rho(NJ_s \times d_t) + \epsilon_{ist}$$

$$\begin{aligned} NJ_s \times d_t &= D_{st} \\ NJ_s &= \begin{cases} 1 & \text{if } s = NJ \\ 0 & \text{if } s = PA \end{cases} \\ d_t &= \begin{cases} 1 & \text{if } t = Nov \\ 0 & \text{if } t = Feb \end{cases} \end{aligned}$$

Saturated Model

$$\begin{aligned} \alpha &= E[Y_{ist}|s = PA, t = Feb] = \gamma_{PA} + \lambda_{Feb} \\ \gamma &= E[Y_{ist}|s = NJ, t = Feb] - E[Y_{ist}|s = PA, t = Feb] = \gamma_{NJ} - \gamma_{PA} \\ \lambda &= E[Y_{ist}|s = PA, t = Nov] - E[Y_{ist}|s = PA, t = Feb] = \lambda_{Nov} - \lambda_{Feb} \\ \rho &= E[Y_{ist}|s = NJ, t = Nov] - E[Y_{ist}|s = NJ, t = Feb] \\ &\quad - \{E[Y_{ist}|s = PA, t = Nov] - E[Y_{ist}|s = PA, t = Feb]\} \end{aligned}$$

Table 5.2.1: Average employment per store 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 (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)

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 stores with data on employment. Employment at six closed stores is set to zero. Employment at four temporarily closed stores is treated as missing. Standard errors are reported in parentheses

Remarks

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \rho NJ_s \times d_t + \epsilon_{ist}$$

- Main Effects: NJ_s, d_t
- Interaction $NJ_s \times d_t$
- Ideally, we want γ, λ small and ρ big (most convincing).
- Good to add additional states, time periods
- We don't need a panel, a repeated cross section would do just fine

More data from the pre-treatment period?

Time $t = -\underline{T}, \ldots, 0, 1$. Treatment: between $t = 0$ and $t = 1$.
Check if the trends are indeed parallel:

$$Y_{ist} = \gamma_0 NJ_s + \sum_{\tau=-\underline{T}}^{1} \lambda_\tau d_{\tau t} + \sum_{\tau=-\underline{T}}^{-1} \beta_\tau d_{\tau t} \times NJ_s + \rho d_{1t} \times NJ_s + \varepsilon_{ist}$$

Test $H_0 : \beta_{-1} = \ldots = \beta_{-\underline{T}} = 0$.

No visible co-movement – bad sign

In Oct 1996, a federal minimum-wage increase affects PA.

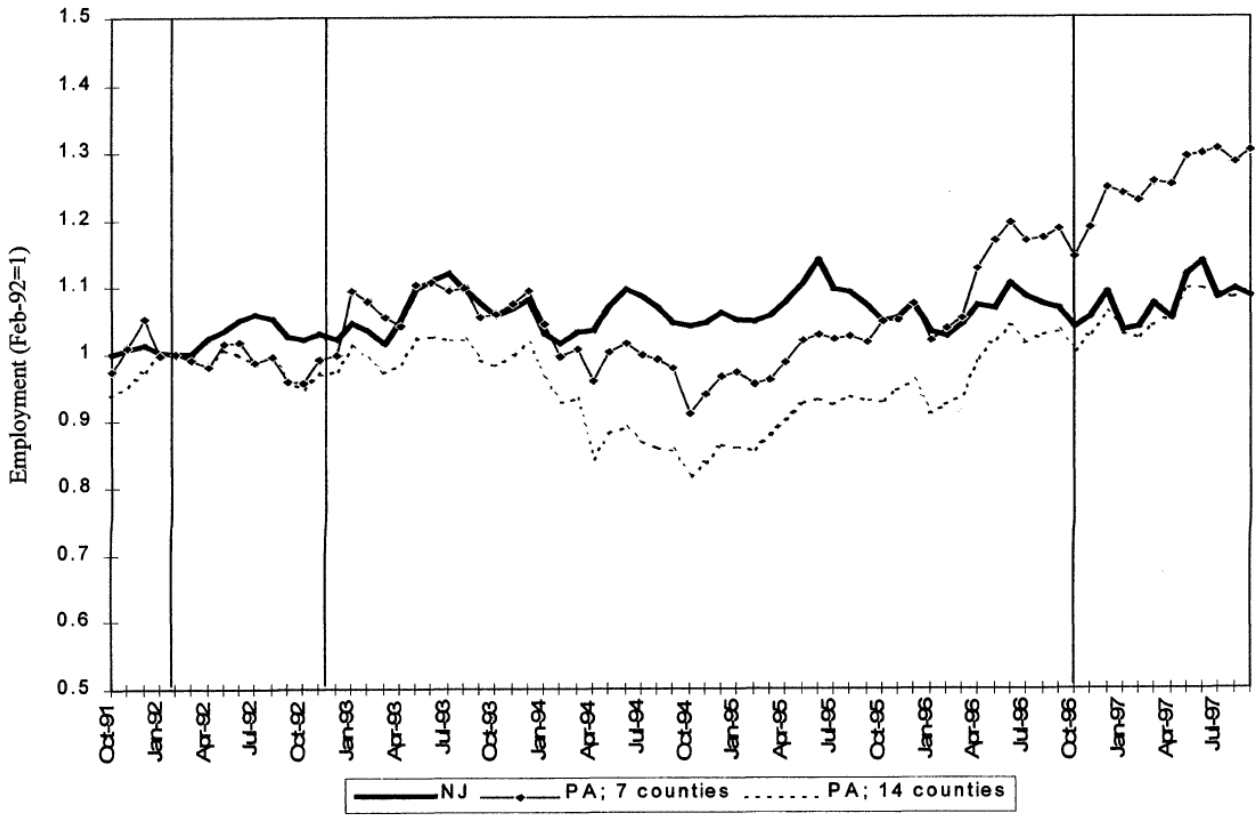

FIGURE 2. EMPLOYMENT IN NEW JERSEY AND PENNSYLVANIA FAST-FOOD RESTAURANTS, OCTOBER 1991 TO SEPTEMBER 1997

Note: Vertical lines indicate dates of original Card-Krueger survey and the October 1996 federal minimum-wage increase.
Source: Authors' calculations based on BLS ES-202 data.

PA and NJ employment aren't really moving in parallel between '92 and '96.

Ashenfelter's dip – another bad sign

Selection based on outcome prior to treatment. Example: effect of a training program on worker earnings.

- Treatment group: workers who enrolled into training
- Control: workers who did not.
- But how do workers self-select into enrollment? Temporary wage decline $\rightarrow$ more likely to enroll? Then we have a problem.

Test if $\beta_{-1} = 0$.

Alternative test: try predicting treatment using past outcomes and **same set of controls** as in the main regression. Are past outcomes significant?

More data from the post-treatment period?

Time $t = -\underline{T}, \ldots, \overline{T}$. Consider delayed effects, set up a fully saturated model. In Card & Krueger's case, we would estimate

$$Y_{ist} = \gamma_0 NJ_s + \sum_{\tau=-\underline{T}}^{\overline{T}} \lambda_\tau d_{\tau t} + \sum_{\tau=-\underline{T}}^{-1} \beta_\tau d_{\tau t} \times NJ_s + \sum_{\tau=1}^{\overline{T}} \rho_\tau d_{\tau t} \times NJ_s + \varepsilon_{ist}$$

Labor market frictions $\rightarrow$ effect of treatment is gradual, $|\rho_2| > |\rho_1|$.

Multiple periods: beware of serial correlation!

Consider an absurd variation of Card & Krueger:

- We start with 4 observations: 1 month before/after wage increase, for PA and NJ. Too small of a sample!
- We obtain employment in PA and NJ at an *hourly* frequency. Large sample, hourly variation is small $\rightarrow$ very tight estimates!
- This must feel wrong, though!

Employment in PA now is likely to be approx. the same as one hour ago. Serial correlation $\rightarrow$ need to adjust std. errors.

- Quick fix: cluster at the level of treatment unit. But then we are left with 2 clusters: NJ and PA.
- Cluster by county? What about spatial correlation?

To learn more, check Bertrand et al. "How much should we trust diff-in-diff estimates?"

Treatment in waves (staggered treatment)

Example: access to broadband internet. Does it affect economic outcomes (jobs, wages, sales)?

- Every unit i (a city or a village) is treated at some point. c – treatment cohort: all units in c get internet access at $t = c$.
- Idea: use untreated cohorts as a control group for treated cohorts.
- Assume same effect across cohorts, parallel trends; estimate

$$Y_{ict} = \gamma_c + \lambda_t + \sum_{\tau \geq c} \rho_{\tau - c} d_{\tau t} + \varepsilon_{it}$$

  Control for γ_c and λ_t using dummies.
- γ_c — cohort assignment can be based on the level of Y

Continuous treatment

Example (a term paper ≈5 years ago):

- Construction of a radar in Pécs. The locals are unhappy about the project (health, safety concerns).
- Any effect on real estate prices? Mechanism: lower demand for properties close to the site $\rightarrow$ prices drop.
- Treatment = distance to the radar site. $t = 0, 1$ — before/after the announcement, i – sales ad

$$Y_{it} = \alpha + \gamma DIST_i + \lambda d_{1t} + \rho d_{1t} \times DIST_i + \varepsilon_{it}$$

- Any obvious issues?

Diff-in-diff and controls

Selection: properties close to the radar site are different from those in the city center. Different market segments $\rightarrow$ different price trends. Control for X_{it} – house type, area, distance to schools, etc:

$$Y_{it} = \alpha + \gamma DIST_i + \lambda d_{1t} + \rho d_{1t} \times DIST_i + \beta X_{it} + \varepsilon_{it}$$

βX_{it} isn't causal; it helps soak up differences in trends between treatment/control group.

Triple differences

Example: Bisztray "The effect of FDI on local suppliers"

- Treatment: Audi opens a plant in Győr. Do local suppliers become more productive?
- Treated firms: (a) produce car parts, (b) are located in Győr.
- Same industry, other locations — control for industry-specific shocks.
- Same location, other industries — shocks specific to Győr.
- Other locations, other industries — overall economy trends.

$$\begin{aligned} TFP_{it} = \alpha + \gamma GYOR_i + \beta SUPPLIER_i + \lambda POST_t \\ + \phi SUPPLIER_i \times POST_t + \psi GYOR_i \times POST_t \\ + \delta SUPPLIER_i \times GYOR_i \\ + \rho SUPPLIER_i \times GYOR_i \times POST_t + \varepsilon_{it} \end{aligned}$$

- Effect of interest: ρ