Lecture 12: Fixed Effect Model and Difference in Differences

Introduction to Econometrics,Fall 2018

Zhaopeng Qu

Nanjing University

12/20/2018

Panel Data: What and Why

Introduction

- A panel dataset contains observations on multiple entities, where each entity is observed at two or more points in time.
- If the data set contains observations on the variables X and Y,then the data are denoted

$$(X_{it}, Y_{it}),\ i = 1, ...n\ and\ t = 1, ..., T$$

  - the first subscript,i refers to the entity being observed
  - the second subscript,t refers to the date at which it is observed
- whether some observations are missing
  - balanced panel
  - unbalanced panel

Introduction

TABLE 1.3 Selected Observations on Cigarette Sales, Prices, and Taxes, by State and Year for U.S. States, 1985–1995

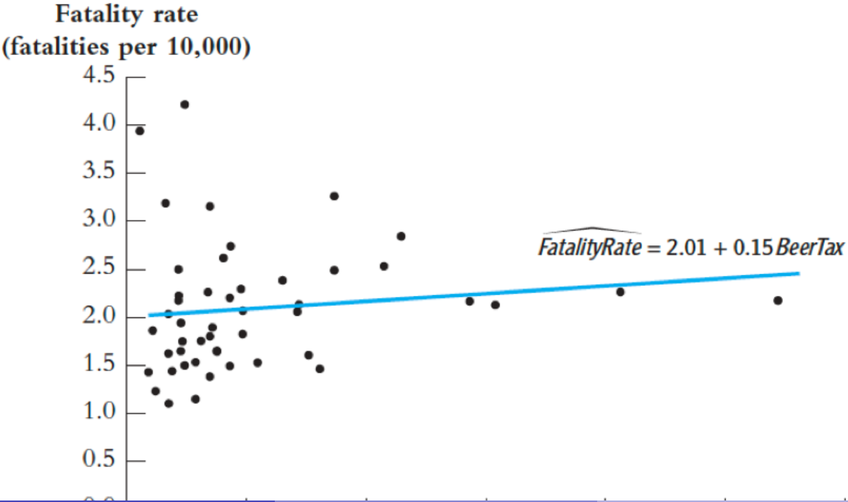
Observation Number	State	Year	Cigarette Sales (packs per capita)	Average Price per Pack (including taxes)	Total Taxes (cigarette excise tax + sales tax)
1	Alabama	1985	116.5	$1.022	$0.333
2	Arkansas	1985	128.5	1.015	0.370
3	Arizona	1985	104.5	1.086	0.362
⋮	⋮	⋮	⋮	⋮	⋮
47	West Virginia	1985	112.8	1.089	0.382
48	Wyoming	1985	129.4	0.935	0.240
49	Alabama	1986	117.2	1.080	0.334
⋮	⋮	⋮	⋮	⋮	⋮
96	Wyoming	1986	127.8	1.007	0.240
97	Alabama	1987	115.8	1.135	0.335
⋮	⋮	⋮	⋮	⋮	⋮

Example: Traffic deaths and alcohol taxes

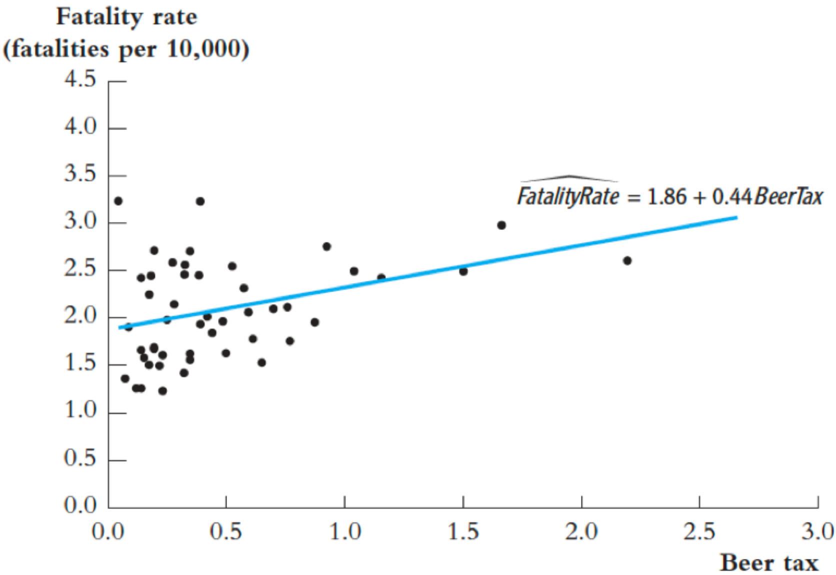
- Observational unit: a year in a U.S. state
- 48 U.S. states, so n = of entities = 48
- 7 years (1982,..., 1988),so T = # of time periods = 7
- Balanced panel, so total # observations $= 7 \times 48 = 336$
- Variables:
  - Dependent Variable: Traffic fatality rate (# traffic deaths in that state in that year, per 10,000 state residents)
  - Independent Variable: Tax on a case of beer
  - Other Controls (legal driving age, drunk driving laws, etc.)

U.S. traffic death data for 1982

- Higher alcohol taxes, more traffic deaths



U.S. traffic death data for 1988



Simple Case: Panel Data with Two Time Periods

- Let Z_i be a factor that determines the fatality rate in the i state but does not change over time.
    - local cultural attitude toward drinking and driving.
- Before and After Model

$$FatalityRate_{it} = \beta_0 + \beta_1 BeerTax_{it} + \beta_2 Z_i + u_{it}$$

where u_{it} is the error term and $i = 1, ...n \; and \; t = 1, ..., T$

- its omission of Z_i might cause omitted variable bias but we don't have data on Z_i.
- The key idea: Any **change** in the fatality rate from 1982 to 1988 cannot be caused by Z_i, because Z_i (by assumption) does not change between 1982 and 1988.

Panel Data with Two Time Periods:Before and After Model

- The math: Consider the regressions for 1982 and 1988...

$$FatalityRate_{i1988} = \beta_0 + \beta_1 BeerTax_{i1988} + \beta_2 Z_i + u_{i1988}$$
$$FatalityRate_{i1982} = \beta_0 + \beta_1 BeerTax_{i1982} + \beta_2 Z_i + u_{i1982}$$

- Then make a difference

$$FatalityRate_{i1988} - FatalityRate_{i1982} =$$
$$\beta_1(BeerTax_{i1988} - BeerTax_{i1982}) + (u_{i1988} - u_{i1982})$$

- Assumption: if $E(u_it|BeerTax_{it}, Z_{it}) = 0$,then $(u_{i1988} - u_{i1982})$ is uncorrelated with either $BeerTax_{i1988}$ or $BeerTax_{i1982}$

Panel Data with Two Time Periods

- Then this “difference” equation can be estimated by OLS, even though Z_i isn’t observed.
- Because the omitted variable Z_i doesn’t change, it cannot be a determinant of the change in Y.

Traffic deaths and beer taxes

1982 data:

$$\widehat{FatalityRate} = \underset{(.11)}{1.86} + \underset{(.13)}{0.44}BeerTax \qquad (n = 48)$$

1988 data:

$$\widehat{FatalityRate} = \underset{(.15)}{2.01} + \underset{(.13)}{0.15}BeerTax \qquad (n = 48)$$

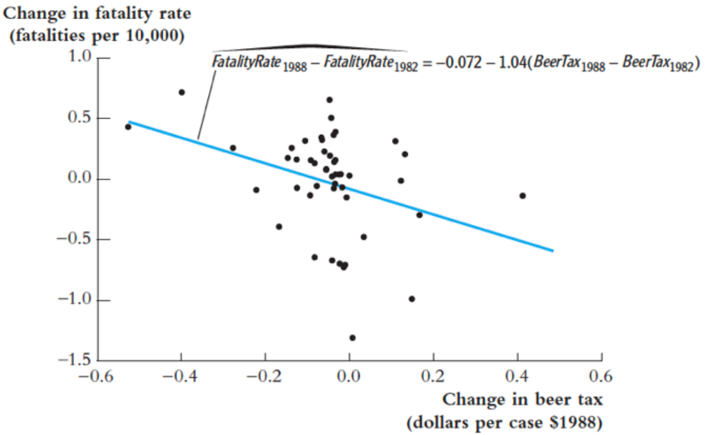
Difference regression ($n = 48$)

$$\widehat{FR_{1988} - FR_{1982}} = \underset{(.065)}{-.072} - \underset{(.36)}{1.04}(BeerTax_{1988} - BeerTax_{1982})$$

Change in traffic deaths and change in beer taxes

FIGURE 10.2 Changes in Fatality Rates and Beer Taxes, 1982–1988

This is a scatterplot of the *change* in the traffic fatality rate and the *change* in real beer taxes between 1982 and 1988 for 48 states. There is a negative relationship between changes in the fatality rate and changes in the beer tax.



Wrap up

- In contrast to the cross-sectional regression results, the estimated effect of a change in the real beer tax is **negative**, as predicted by economic theory.
- By examining changes in the fatality rate over time, the regression in Equation controls for fixed factors such as cultural attitudes toward drinking and driving. But there are many factors that influence traffic safety, and if they change over time and are correlated with the real beer tax, then their omission will produce omitted variable bias.
- This "before and after" analysis works when the data are observed in two different years. Our data set, however, contains observations for seven different years,and it seems foolish to discard those potentially useful additional data. But the "before and after" method does not apply directly when $T > 2$. To analyze all the observations in our panel data set, we use the method of **fixed effects** regression

Fixed Effect Model

Introduction

- Fixed effects regression is a method for controlling for omitted variables in panel data when the omitted variables vary across entities (states) but do not change over time.
- Unlike the "before and after" comparisons,fixed effects regression can be used when there are two or more time observations for each entity.

Fixed Effects Regression Model

- the dependent variable (FatalityRate) and observed regressor (BeerTax) denoted as Y_{it} and X_{it}, respectively:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Z_i + u_{it} \tag{10.9}$$

- where Z_i is an **unobserved variable** that varies from one state to the next but **does not change over time** (for example, Z_i represents cultural attitudes toward drinking and driving).
- We want to estimate β_1, the effect on Y of X holding constant the unobserved state characteristics Z.

Fixed Effects Regression Model

- Because Z_i varies from one state to the next but is constant over time,then let $\alpha_i = \beta_0 + \beta_1 Z_i$,the Equation becomes

$$Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it} \tag{10.10}$$

- Equation (10.10) is the **fixed effects regression model**, in which α are treated as unknown intercepts to be estimated, one for each state. The interpretation of α_i as a state-specific intercept in Equation (10.10).
- Because the intercept α_i in Equation (10.10) can be thought of as the "effect" of being in entity i (in the current application, entities are states), the terms α_i,an are known as **entity fixed effects**.
- The variation in the entity fixed effects comes from omitted variables that, like Z_i in Equation (10.9), vary across entities but not over time.

Alternative : Fixed Effects by using binary variables

- To develop the fixed effects regression model using binary variables, let $D1_i$ be a binary variable that equals 1 when i = 1 and equals 0 otherwise, let $D2_i$ equal 1 when i = 2 and equal 0 otherwise, and so on.
- Arbitrarily omit the binary variable $D1_i$ for the first group. Accordingly, the fixed effects regression model in Equation (10.10) can be written equivalently as

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \gamma_2 D2_i + \gamma_3 D3_i + ... + \gamma_n Dn_i + u_{it} \quad (10.11)$$

Fixed Effects by using binary variables

- Thus there are two equivalent ways to write the fixed effects regression model, Equations (10.10) and (10.11).
- In Equation (10.10), it is written in terms of n state specific intercepts.

$$Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it} \tag{10.10}$$

- In Equation (10.11), the fixed effects regression model has a common intercept and $n-1$ binary regressors

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \gamma_2 D2_i + \gamma_3 D3_i + ... + \gamma_n Dn_i + u_{it} \tag{10.11}$$

- In both formulations, the slope coefficient on X is the same from one state to the next.

Fixed Effects: Extension to multiple X's.

- The fixed effects regression model is

$$Y_{it} = \beta_1 X_{1,it} + ... + \beta_k X_{k,it} + \alpha_i + u_{it} \tag{10.12}$$

- Equivalently, the fixed effects regression can be expressed in terms of a common intercept

$$\begin{aligned} Y_{it} =& \beta_0 + \beta_1 X_{1,it} + ... + \beta_k X_{k,it} \\ &+ \gamma_2 D2_i + \gamma_3 D3_i + ... + \gamma_n Dn_i + u_{it} \end{aligned}$$

Estimation and Inference

- In principle the binary variable specification of the fixed effects regression model [Equation (10.13)] can be estimated by OLS.
- But it is tedious to estimate so many fixed effects.If $n = 1000$, then you have to estimate $1000 - 1 = 999$ fixed effects.
- These special routines are equivalent to using OLS on the full binary variable regression, but are faster because they employ some mathematical simplifications that arise in the algebra of fixed effects regression.

Estimation: The "entity-demeaned"

- Computes the OLS fixed effects estimator in two steps
- The first step:
  - take the average across times t of both sides of Equation (10.10);

$$\bar{Y}_i = \beta_1 \bar{X}_i + \alpha_i + \bar{u}_t$$

  - demeaned: let

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

- The second step: accordingly,estimate

$$\tilde{Y}_{it} = \beta_1 \tilde{X}_{it} + \tilde{u}_{it} \tag{10.14}$$

- In fact, this estimator is identical to the OLS estimator of β_1 obtained by estimation of the fixed effects model in Equation (10.11)

Fixed effect estimator

- Our fixed effects model is

$$Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it} \tag{10.10}$$

- we can think of c_i as fixed effects or "nuisance parameters" to be estimated
- OLS estimation with fixed effects yields

$$(\hat{\beta}, \hat{\alpha}_1, \dots, \hat{\alpha}_n) = \underset{b, a_1, \dots, a_n}{\operatorname{argmin}} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - bX_{it} - a_i)^2$$

this amounts to including n dummies in regression of Y_{it} on X_{it}

Fixed effect estimator

- The first-order conditions (FOC) for this minimization problem are:

$$\sum_{i=1}^{n}\sum_{t=1}^{T}(Y_{it}-\hat{\beta}X_{it}-\hat{\alpha}_i)X_{it}=0$$

- And

$$\sum_{i=1}^{n}\sum_{t=1}^{T}(Y_{it}-\hat{\beta}X_{it}-\hat{\alpha}_i)=0$$

Fixed effect estimator

- Therefore, for $i = 1, \ldots, N$,

$$\hat{\alpha}_i = \frac{1}{T}\sum_{t=1}^{T}(Y_{it} - \hat{\beta}X_{it}) = \overline{Y}_i - \overline{X}_i\hat{\beta},$$

where

$$\bar{X}_i \equiv \frac{1}{T}\sum_{t=1}^{T} X_{it}; \bar{Y}_i \equiv \frac{1}{T}\sum_{t=1}^{T} Y_{it}$$

Fixed effect estimator

- Plug this result into the first FOC to obtain:

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{t=1}^{T}(Y_{it}-\hat{\beta}X_{it}-\hat{\alpha}_i)X_{it} &= \sum_{i=1}^{n}\sum_{t=1}^{T}(Y_{it}-X_{it}\hat{\beta}-\overline{Y}_i+\overline{X}_i\hat{\beta})X_{it} \\
&= \left(\sum_{i=1}^{n}\sum_{t=1}^{T}(Y_{it}-\overline{Y}_i)X_{it}\right) \\
&\quad -\hat{\beta}\left(\sum_{i=1}^{n}\sum_{t=1}^{T}(X_{it}-\bar{X}_i)X_{it}\right)=0
\end{aligned}$$

Fixed effect estimator

- Then we could obtain

$$\hat{\beta} = \left(\sum_{i=1}^{n} \sum_{t=1}^{T} (X_{it} - \bar{X}_i)(X_{it} - \bar{X}_i) \right)^{-1} \left(\sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - \bar{Y})(X_{it} - \bar{X}_i) \right)$$

$$= \frac{\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{X}_{it} \tilde{Y}_{it}}{\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{X}_{it}^2}$$

with time-demeaned variables $\tilde{X}_{it} \equiv X_{it} - \bar{X}, \tilde{Y}_{it} \equiv Y_{it} - \bar{Y}_i$

The sampling distribution, standard errors and statistical inference

- unbiasedness and consistency

$$\begin{aligned}
\hat{\beta}_{fe} &= \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}\tilde{Y}_{it}}{\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}^{2}} \\
&= \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}(\beta_1\tilde{X}_{it}+\tilde{u}_{it})}{\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}^{2}} \\
&= \beta_1 + \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}\tilde{u}_{it}}{\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}^{2}}
\end{aligned}$$

- Paralleling the derivation of OLS estimator, we could prove the estimator of fixed effects model is unbiased and consistent.

Extension: Regression with Time Fixed Effects

Introduction

- Just as fixed effects for each entity can control for variables that are constant over time but differ across entities, so can time fixed effects control for variables that are constant across entities but evolve over time.
- safety improvements in new cars as an omitted variable that changes over time but has the same value for all states.

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Z_i + \beta_3 S_t + u_{it} \quad (10.16)$$

- where S_t is unobserved and where the single t subscript emphasizes that safety changes over time but is constant across states. Because $\beta_3 S_3$ represents variables that determine Y_{it}, if S_t is correlated with X_{it}, then omitting S_t from the regression leads to omitted variable bias.

Time Effects Only

- suppose that the variables Z_i are not present

$$Y_{it} = \beta_1 X_{it} + \lambda_t + u_{it} \tag{10.17}$$

- so the terms λ are known as time fixed effects.The variation in the time fixed effects comes from omitted variables that, like S_t in Equation (10.16), vary over time but not across entities.
- Just as the entity fixed effects regression model can be represented using $n-1$ binary indicators, so, too, can the time fixed effects regression model be represented using $T-1$ binary indicators:

$$Y_{it} = \beta_0 + \beta_1 X_{1,it} + \delta_2 B2_t + ... + \delta_T BT_t + \alpha_i + u_{it} \tag{10.18}$$

Both Entity and Time Fixed Effects

- If some omitted variables are constant over time but vary across states (such as cultural norms) while others are constant across states but vary over time (such as national safety standards),
- The combined entity and time fixed effects regression model is

$$Y_{it} = \beta_1 X_{it} + \alpha_i + \lambda_t + u_{it}$$

Application to traffic deaths

$$\widehat{FatalityRate} = \underset{(0.36)}{-0.64}\,BeerTax + StateFixedEffects + TimeFixedEffects. \quad (10.21)$$

The Fixed Effects Regression Assumptions and Standard Errors

The Fixed Effects Regression Assumptions

KEY CONCEPT 10.3

The Fixed Effects Regression Assumptions

$$Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it}, i = 1, \ldots, n, t = 1, \ldots, T,$$

where

1. u_{it} has conditional mean zero: $E(u_{it} \mid X_{i1}, X_{i2}, \ldots, X_{iT}, \alpha_i) = 0$.
2. $(X_{i1}, X_{i2}, \ldots, X_{iT}, u_{i1}, u_{i2}, \ldots, u_{iT}), i = 1, \ldots, n$ are i.i.d. draws from their joint distribution.
3. Large outliers are unlikely: (X_{it}, u_{it}) have nonzero finite fourth moments.
4. There is no perfect multicollinearity.

For multiple regressors, X_{it} should be replaced by the full list $X_{1,it}, X_{2,it}, \ldots, X_{k,it}$.

Autocorrelated in Panel Data

- An important difference between the panel data assumptions in Key Concept 10.3 and the assumptions for cross-sectional data in Key Concept 6.4 is Assumption 2.
  - **Cross-Section**: Assumption 2 holds: i.i.d sample.
  - **Panel data**: independent across entities but no such restricition **within** an entity.
- if $Cov(X_t, X_s)$ for some $t \neq s$, the X_t is said to be **autocorrelated or serially correlated**.
- In the traffic fatality example, X_{it}, the beer tax in state i in year t, is autocorrelated:
  - Most of the time, the legislature does not change the beer tax, so if it is high one year relative to its mean value for state i,it will tend to be high the next year,too.

Autocorrelated in Panel Data

- Similarly,u_{it} would be also autocorrelated. It consists of time-varying factors that are determinants of Y_{it} but are not included as regressors, and some of these omitted factors might be autocorrelated. It can formally be expressed as

$$Cov(u_{it}, u_{is}|X_{it}, X_{is}, \alpha_i) \neq 0 \; for \; t \neq s$$

  - eg. a downturn in the local economy and a road improvement project.
  - eg. severe winter driving conditions.

- The result: an analogy of heteroskedasticity.

- OLS panel data estimators of β are unbiased and consistent but the standard errors will be wrong

  – usually the OLS standard errors understate the true uncertainty

- This problem can be solved by using **"heteroskedasticity and autocorrelation-consistent(HAC) standard errors"**

Standard Errors for Fixed Effects Regression

- The standard errors used are one type of HAC standard errors, **clustered standard errors**.
- The term **clustered** arises because these standard errors allow the regression errors to have an arbitrary correlation within a cluster, or grouping, but assume that the regression errors are uncorrelated across clusters.
- In the context of panel data,each cluster consists of an entity.Thus **clustered standard errors** allow for heteroskedasticity and for arbitrary autocorrelation *within an entity*, but treat the errors as *uncorrelated across entities*.
- Like **heteroskedasticity-robust standard errors** in regression with cross-sectional data, **clustered standard errors** are valid whether or not there is heteroskedasticity,autocorrelation,or both.

Difference in Differences

Introduction

- A typical RCT design requires a causal studies to do as follow
  1. Randomly assignment of treatment to divide the population into a "treatment" group and a "control" group.
  2. Collecting the data at the time of post-treatment then comparing them.
- It works because *treatment* and *control* are randomized.
- what if we have the treatment group and the control group, but they are not fully randomized?
- If we have observations across two times at least(one before treatment, the other after treatment), then an easy way to make causal inference is **Difference in Differences(DID)** method.

DID with Regression

- Formally, an simple DID regression is

$$Y_{it} = \beta_0 + \beta_1 T_i + \beta_2 D_t + \beta_3 (T_i \times D_t) + u_{it}$$

- T_i: a time dummy denotes pre$(T_i = 0)$ or post treatment$(T_i = 1)$,thus
- D_t: a treatment dummy denotes in treatment$(D_t = 0)$ or control group$(D_t = 1)$
- $T_i \times D_t$: an interaction term

DID with Regression

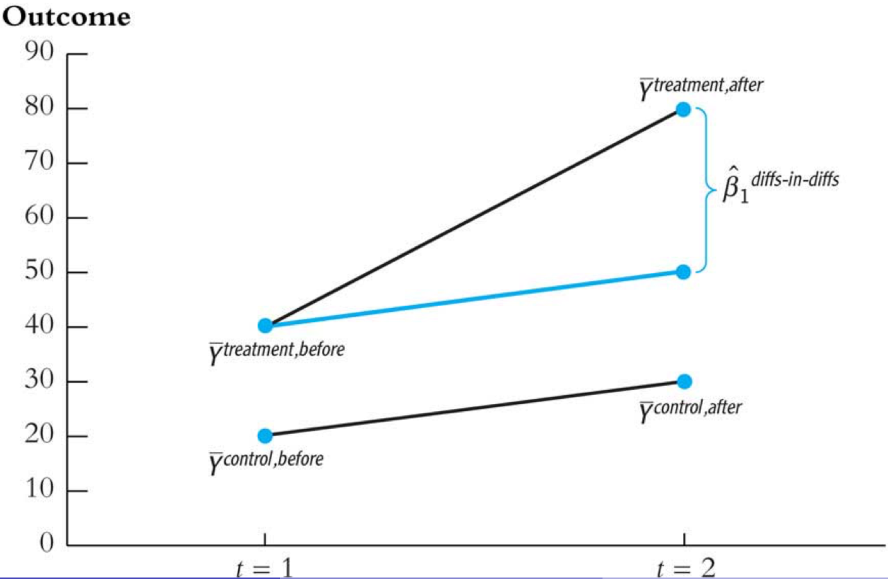
- DID model:

$$Y_{it} = \beta_0 + \beta_1 T_i + \beta_2 D_t + \beta_3 (T_i \times D_t) + u_{it}$$

- then DID estimator

$$\hat{\beta}_{DID} = (\bar{Y}_{treat,after} - \bar{Y}_{treat,before}) - (\bar{Y}_{control,after} - \bar{Y}_{control,before})$$

DID estimator



Key Assumption For DID

- A key identifying assumption for DID is: **Common Trend**
    - Treatment would be the same "trend" in both groups in the absence of treatment.
- There are some unobservable factors affected on outcomes of both group. But as long as the effects have the same trends on both groups, then DID will eliminate the factors.
- Advantage: Omitted variables and other bias are likely reduced by the use of the untreated comparison group
- Other changes are not likely to always influence all groups in the same way.

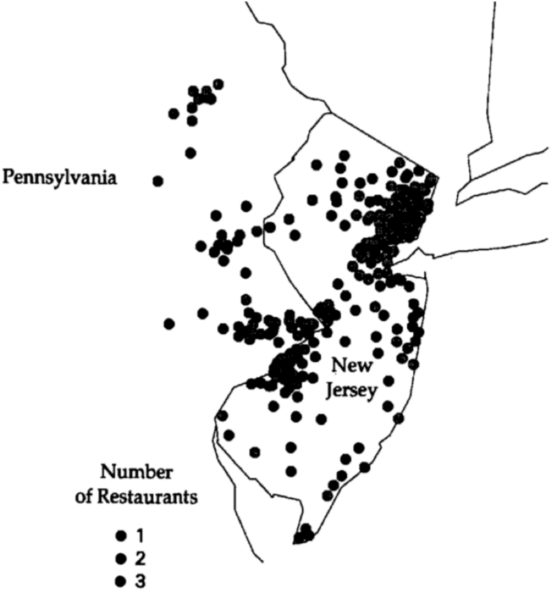
Card and Krueger(1994): minimum wage on employment

- Theoretically,in competitive labor market, increasing binding minimum wage decreases employment.But what about the reality?
- Ideal experiment: randomly assign labor markets to a control group (minimum wage kept constant) and treatment group (minimum wage increased), compare outcomes.
- Policy changes affecting some areas and not others create natural experiments.
    - Unlike ideal experiment, control and treatment groups not randomly assigned.

Card and Krueger(1994): Backgroud

- Policy Change: in April 1992
  - Minimum wage in New Jersey from \$4.25 to \$5.05
  - Minimum wage in Pennsylvania constant at \$4.25
- Research Design:
  - Collecting the data on employment at fast food restaurants in NJ(treatment group) in Feb.1992 (before treatment)and again November 1992(after treatment).
  - Also collecting the data from the same type of restaurants in eastern Pennsylvania(PA) as control group where the minimum wage stayed at \$4.25 throughout this period.

Card & Krueger(1994): Geographic background



Card & Krueger(1994):

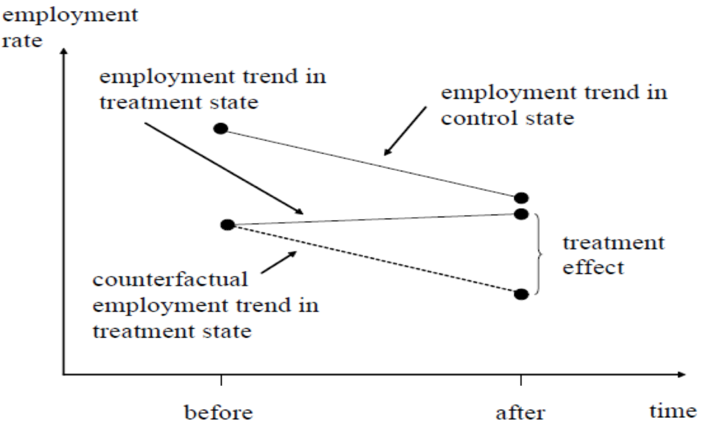


Figure 5.2.1: Causal effects in the differences-in-differences model

Card & Krueger(1994):Result

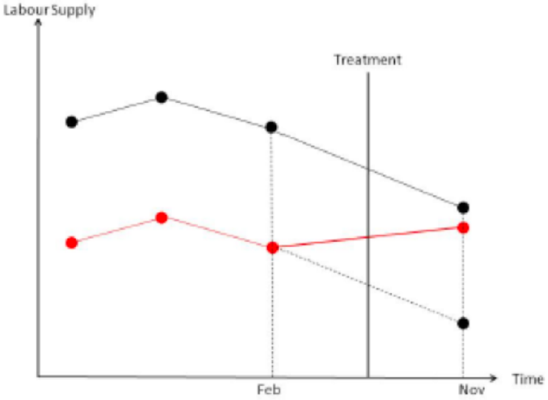
Table 5.2.1: Average employment per store before and after the New Jersey minimum wage increase

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

Notes: Adapted from Card and Krueger (1994), Table 3. The

Assessing natural experiment

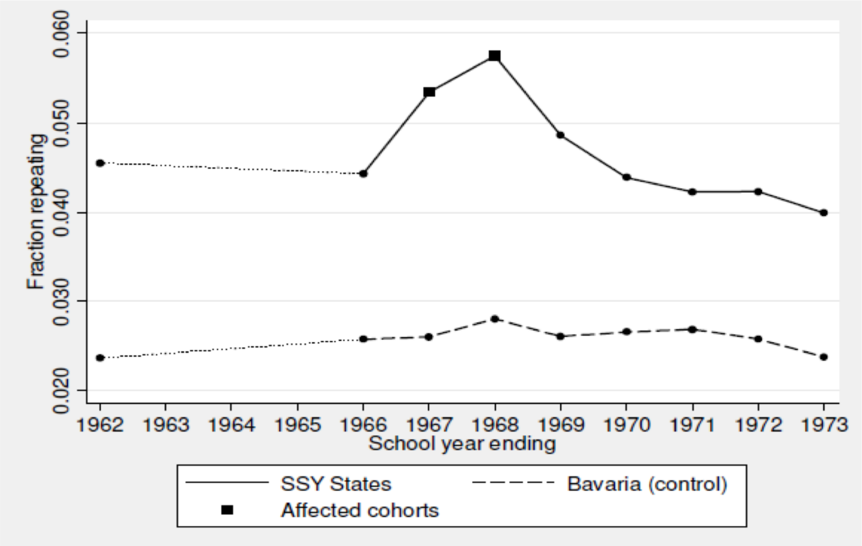
- Common Trend: Common trend assumption is difficult to verify but one often uses pre-treatment data to show that the trends are the same.



An Encouraging Example: Pischeke(2007)

- Topic: the length of school year on student performance
- Backgroud:
  - Until the 1960s, children in all German states except Bavaria started school in the Spring. In 1966-1967 school year, the Spring moved to Fall.
  - It make two shorter school years for affected cohort, 24 weeks long instead of 37.
- Reseach Design:
  - Dependent Variable: Retreating rate
  - Independent Variable: spending time on school
  - Treatment group: Students in the German **States except Bavaria**.
  - Control group: Students in **Bavaria**.

An Encouraging Example: Pischeke(2007)



Pischke(2007)

- This graph provides strong visual evidence of treatment and control states with a common underlying trend.
- A treatment effect that induces a sharp but transitory deviation from this trend.
- It seems to be clear that a short school years have increased repetition rates for affected cohorts.

Extensions of DID: DDD and Synthetic Controls

Extensions of DID: DDD

- More convincing analysis sometime comes from higher-order contrasts: **DDD** or **Triple D** design.
  - Build the third dimension of contrast to eliminate the potential bias.
- e.g: Health Plan for elderly
  - Treatment group: Elderly aged above 65 in treatment place.
  - Control group 1: Elderly aged between 55-65 in treatment place.
  - Assumption 1: the elderly on different age would have the same trends of health status if there were not the plan.
  - Control group 2: Elderly aged above 65 in control place.
  - Assumption 2: the elderly in different places would have the same trends of health status if there were not the plan.
- It can loose the simple *common trend* assumption in simple DID.

Extensions of DID: Synthetic Controls Method

- In some cases, treatment and potential control groups do not follow parallel trends. Then standard DID method would lead to biased estimates.
- Synthetic Control (SC) is a method to evaluate the causal effect of treatment on aggregate outcomes of one (or very few) treated unit.
- The basic idea behind synthetic controls is that a combination of units often provides a better comparison for the unit exposed to the intervention than any single unit alone.
- It is a data-driven procedure to use a small number of non-treated units to build the suitable counterfactuals.

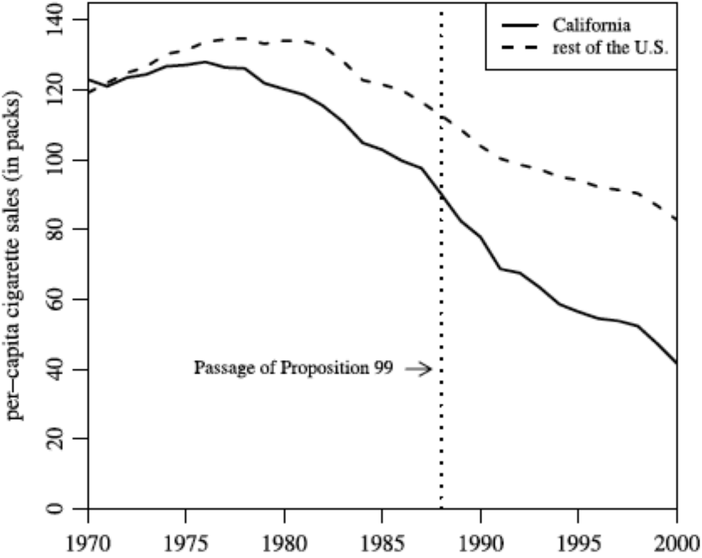
Extensions of DID: Synthetic Controls Method

- Use (long) longitudinal data to build the weighted average of non-treated units that best reproduces characteristics of the treated unit over time in pre-treatment period.
- The weighted average of non-treated units is the synthetic cohort.
- Causal effect of treatment can be quantified by a simple difference after treatment: treated vs synthetic cohort.

Abadie et. al (2010): tobacco tax on cigarette consumption

- In 1988, California passed comprehensive tobacco control legislation: Increased cigarette taxes by $0.25 per pack ordinances.
- estimate the effect of the policy on cigarette consumption in California.

Abadie et. al (2010): tobacco tax on ciga-consumption



Abadie et. al (2010): tobacco tax on ciga-consumption