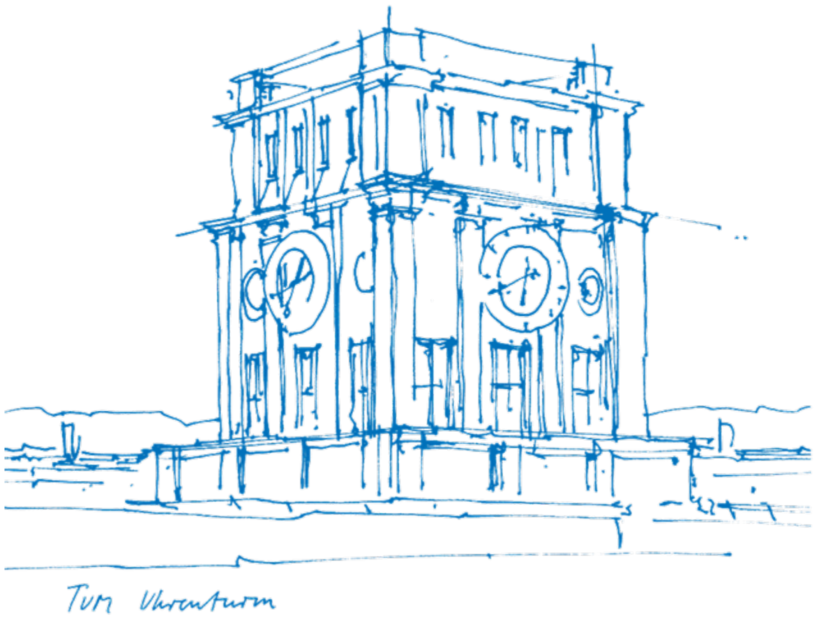
Empirical Research Methods - Lecture 8

Prof. Dr. Helmut Farbmacher
Technical University of Munich
TUM School of Management
Munich, April 9, 2024



Program Evaluation & Identification Strategies

Outline of today's lecture

Program Evaluation

- Potential Outcome Framework
- Treatment Effect

Identification Strategies

- Randomization
- Selection on (Un)Observables
- Differences-in-Differences Estimation

Program Evaluation

Program Evaluation

In economics we are often interested in the **causal effect of a treatment** (e.g., a policy intervention) on an outcome variable. For example,...

- the causal effect of attending college on wages;
- the causal effect of a training program on the probability of finding a job;
- the causal effect of copayments on health care demand.

Standard econometric arguments: Is the binary treatment indicator exogenous?

Arguments in the evaluation literature look slightly different (although the underlying problem is the same).

Helps us to think about causality from a different perspective.

Program Evaluation

Potential Outcome Framework

Potential Outcome Framework

So far, we assumed that each individual outcome is characterized by a data vector $\{y_i, x_i, u_i\}_{i=1}^{n}$.

However, in the evaluation literature we think about two **potential outcomes**

$$Y_i(0): \text{potential outcome if individual } i \text{ not treated}$$
$$Y_i(1): \text{potential outcome if individual } i \text{ treated}$$

Potential outcomes are defined for each individual BUT they will never be realized at the same time (since an individual i is either treated or not).

Define the **treatment status** as

$$D_i = \begin{cases} 0 & \text{if individual } i \text{ is not treated} \\ 1 & \text{if individual } i \text{ is treated} \end{cases}$$

Note: Capitalization of letters is standard in the program evaluation literature.

Potential Outcome Framework

The following **observational rule** gives us the **realized outcome** Y_i

$$Y_i = Y_i(1) \cdot D_i + Y_i(0) \cdot (1 - D_i)$$

We only observe the realized outcome Y and the treatment status D in our dataset. That is, we observe

- the potential outcome $Y(0)$ for those being untreated $(D = 0)$ and
- the potential outcome $Y(1)$ for those being treated $(D = 1)$.

However, we do not oberve the **counterfactual** outcomes:

- $Y(1)$ is the counterfactual outcome for those being untreated $(D = 0)$ and
- $Y(0)$ is the counterfactual outcome for those being treated $(D = 1)$.

Potential Outcome Framework

Y_0	Y_1	D	Y
6.862733	7.819566	0	6.862733
8.364915	8.818250	0	8.364915
7.226931	7.904501	0	7.226931
8.649052	9.221686	0	8.649052
8.821402	8.924327	0	8.821402
6.136669	7.036494	1	7.036494
7.584316	7.830404	1	7.830404
8.677257	8.719317	1	8.719317
7.654305	7.982226	1	7.982226
7.369844	8.324348	1	8.324348

Potential Outcome Framework

Y_0	Y_1	D	Y
6.862733	NA	0	6.862733
8.364915	NA	0	8.364915
7.226931	NA	0	7.226931
8.649052	NA	0	8.649052
8.821402	NA	0	8.821402
NA	7.036494	1	7.036494
NA	7.830404	1	7.830404
NA	8.719317	1	8.719317
NA	7.982226	1	7.982226
NA	8.324348	1	8.324348

Potential Outcome Framework

D	Y
0	6.862733
0	8.364915
0	7.226931
0	8.649052
0	8.821402
1	7.036494
1	7.830404
1	8.719317
1	7.982226
1	8.324348

Program Evaluation

Treatment Effect

Treatment Effect

The **individual treatment effect** is the difference between the two potential outcomes:

$$Y_i(1) - Y_i(0)$$

We cannot observe this quantity since we either observe $Y_i(1)$ or $Y_i(0)$ but never both quantities for the same individual i at the same time.

However, we may be able to identify the **Average Treatment Effect**:

$$\mathrm{ATE} = \mathrm{E}\big(Y_i(1) - Y_i(0)\big)$$

The ATE is the expected effect of the treatment when we randomly draw an individual from the population.

Treatment Effect

Another population quantity of interest is the **Average Treatment Effect on the Treated**:

$$\mathrm{ATET} = \mathrm{E}\big(Y_i(1) - Y_i(0) | D_i = 1\big)$$

The ATET is the expected effect of the treatment when we randomly draw an individual from the sub-population that received the treatment.

Neither the ATE **nor** the ATET can be estimated without additional assumptions because we never observe $Y_i(0)$ and $Y_i(1)$ for the same individual i at the same time.

Treatment Effect

Of course, we can estimate the expected difference in the **realized outcome** between untreated and treated individuals, that is

$$\begin{aligned}\Delta &= \mathrm{E}(Y_i|D_i=1)-\mathrm{E}(Y_i|D_i=0)\\ &= \mathrm{E}\big(Y_i(1)|D_i=1\big)-\mathrm{E}\big(Y_i(0)|D_i=0\big)\end{aligned}$$

In many economic applications it is, however, unlikely that Δ equals the ATE.

Why?

Treatment Effect

Selection into treatment is a fundamental problem if we want to evaluate a program.

Example: Let $Y(0)$ and $Y(1)$ be an individual's wages without and with training, then selection into treatment could mean that individuals with higher education

- are more likely (or willing) to do the training and
- have higher values of $Y(0)$ anyway.

Trick to see this:

$$
\begin{aligned}
\Delta &= \mathrm{E}\big(Y_i(1)|D_i=1\big) - \mathrm{E}\big(Y_i(0)|D_i=0\big) \overbrace{- \mathrm{E}\big(Y_i(0)|D_i=1\big) + \mathrm{E}\big(Y_i(0)|D_i=1\big)}^{\text{expanding by counterfactual term}} \\
&= \mathrm{E}\big(Y_i(1)|D_i=1\big) - \mathrm{E}\big(Y_i(0)|D_i=1\big) + \mathrm{E}\big(Y_i(0)|D_i=1\big) - \mathrm{E}\big(Y_i(0)|D_i=0\big) \\
&= \mathrm{E}\big(Y_i(1) - Y_i(0)|D_i=1\big) + \mathrm{E}\big(Y_i(0)|D_i=1\big) - \mathrm{E}\big(Y_i(0)|D_i=0\big)
\end{aligned}
$$

Treatment Effect

Selection into treatment is a fundamental problem if we want to evaluate a program.

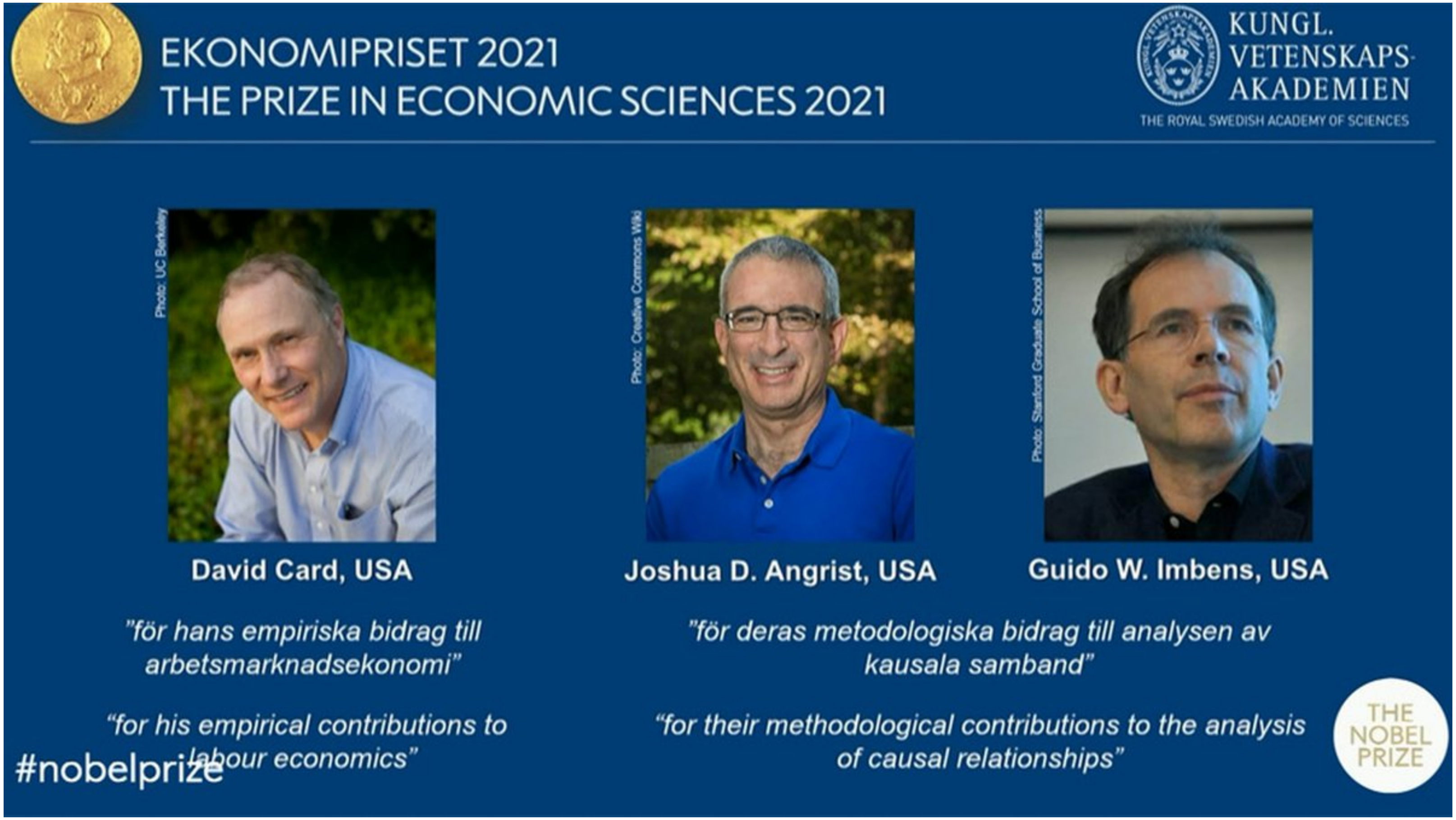
Example: Let $Y(0)$ and $Y(1)$ be an individual's wages without and with training, then selection into treatment could mean that individuals with higher education

- are more likely (or willing) to do the training and
- have higher values of $Y(0)$ anyway.

Trick to see this:

$$\begin{aligned}
\Delta &= \mathrm{E}\big(Y_i(1)|D_i=1\big) - \mathrm{E}\big(Y_i(0)|D_i=0\big) \overbrace{- \mathrm{E}\big(Y_i(0)|D_i=1\big) + \mathrm{E}\big(Y_i(0)|D_i=1\big)}^{\text{expanding by counterfactual term}} \\
&= \mathrm{E}\big(Y_i(1)|D_i=1\big) - \mathrm{E}\big(Y_i(0)|D_i=1\big) + \mathrm{E}\big(Y_i(0)|D_i=1\big) - \mathrm{E}\big(Y_i(0)|D_i=0\big) \\
&= \underbrace{\mathrm{E}\big(Y_i(1) - Y_i(0)|D_i=1\big)}_{\mathrm{ATET}} + \underbrace{\mathrm{E}\big(Y_i(0)|D_i=1\big) - \mathrm{E}\big(Y_i(0)|D_i=0\big)}_{\text{Selection Effect}}
\end{aligned}$$

Treatment Effect



Suppose we have access to observational data for a large number of individuals on their earnings, Y_i, and whether they have completed high school, $D_i = 1$, or not, $D_i = 0$. The observed average difference in earnings across the two groups, is an estimate of $\Delta = E[Y_i|D_i = 1] - E[Y_i|D_i = 0] = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 0]$. It is unlikely that Δ equals the causal effect of interest, however. To see this formally, subtract and add $E[Y_i(0)|D_i = 1]$, a counterfactual term, which yields:

$$\boxed{\underbrace{\Delta}_{\text{Difference in means}} = \underbrace{E\big[\big(Y_i(1) - Y_i(0)\big)\big|D_i = 1\big]}_{\text{causal effect on high school completers}} + \underbrace{E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0]}_{\text{selection bias}}}$$

The first term in this expression, $E\big[\big(Y_i(1) - Y_i(0)\big)\big|D_i = 1\big]$, is the causal effect of interest. In our high school example, it provides the answer to the question: What is the effect on wage earnings of completing high school, among those that did so? The second term, $E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0]$, represents selection bias. It measures how different the earnings would have been in the two populations — completers and non-completers — had they not completed high school. For reasons described above, high school completers would likely earn more than non-completers, even without high-school, implying that the selection effect is positive in this example. In such a case, the comparison of means, Δ, provides an upward biased estimate of the causal effect of high school education on earnings.

Identification Strategies

Identification Strategies

$Y_i(0)$: potential outcome if individual i is not treated
$Y_i(1)$: potential outcome if individual i is treated

Observational rule gives us the realized outcome:

$$Y_i = Y_i(1) \cdot D_i + Y_i(0) \cdot (1 - D_i)$$

Realized outcome mixes two effects:

$$\begin{aligned} \Delta &= \mathrm{E}\big(Y_i(1)|D_i = 1\big) - \mathrm{E}\big(Y_i(0)|D_i = 0\big) \\ &= \underbrace{\mathrm{E}\big(Y_i(1) - Y_i(0)|D_i = 1\big)}_{\text{ATET}} + \underbrace{\mathrm{E}\big(Y_i(0)|D_i = 1\big) - \mathrm{E}\big(Y_i(0)|D_i = 0\big)}_{\text{Selection Effect}} \end{aligned}$$

Now, we are going to discuss several **strategies to disentangle both effects**.

Identification Strategies

Randomization

Randomization

Random assignment of treatment (say, in a real experiment) is the **gold standard** to learn treatment effects.

Randomized Controlled Trials (RCTs) are often used in economics nowadays.

Banerjee, Duflo and Kremer received the Nobel Prize in Economics 2019 more or less for their RCTs to alleviate global poverty.

Analyzing experimental data is (almost) always very easy. Why is this the case?

Randomization

Say, we want to estimate the ATET:

$$\begin{aligned}\mathrm{ATET} &= \mathrm{E}\big(Y_i(1) - Y_i(0) | D_i = 1\big) \\ &= \mathrm{E}\big(Y_i(1) | D_i = 1\big) - \mathrm{E}\big(Y_i(0) | D_i = 1\big)\end{aligned}$$

where the second equality just uses the linearity of the expectation operator.

We can only estimate the difference in the realized outcome:

$$\begin{aligned}\Delta &= \mathrm{E}(Y_i | D_i = 1) - \mathrm{E}(Y_i | D_i = 0) \\ &= \mathrm{E}\big(Y_i(1) | D_i = 1\big) - \mathrm{E}\big(Y_i(0) | D_i = 0\big) \quad \text{[holds by definition of } Y_i\text{]}\end{aligned}$$

Randomization

Say, we want to estimate the ATET:

$$\begin{aligned} \text{ATET} &= \mathrm{E}\big(Y_i(1) - Y_i(0)|D_i = 1\big) \\ &= \mathrm{E}\big(Y_i(1)|D_i = 1\big) - \mathrm{E}\big(Y_i(0)|D_i = 1\big) \end{aligned}$$

where the second equality just uses the linearity of the expectation operator.

We can only estimate the difference in the realized outcome:

$$\begin{aligned} \Delta &= \mathrm{E}(Y_i|D_i = 1) - \mathrm{E}(Y_i|D_i = 0) \\ &= \mathrm{E}\big(Y_i(1)|D_i = 1\big) - \mathrm{E}\big(Y_i(0)|D_i = 0\big) \quad \text{[holds by definition of } Y_i\text{]} \end{aligned}$$

So far, ATET and Δ are <u>not</u> equal.

Let's see how randomization of treatment affects the program evaluation...

Randomization

Randomization of treatment implies that D is **independent** of any other variable. This means that D is also independent of the potential outcomes, $Y(0)$ and $Y(1)$.

We can use the independence of D from the potential outcomes in the following way:

$$\underbrace{\mathrm{E}\big(Y_i(0)|D_i = 1\big)}_{\text{counterfactual}} = \mathrm{E}\big(Y_i(0)|D_i = 0\big) = \underbrace{\mathrm{E}\big(Y_i|D_i = 0\big)}_{\text{can be estimated}}$$

The first equality holds by independence of D and $Y(0)$, which follows from the randomization of the treatment. The second equality holds by the definition of Y.

That is under randomization we have

$$\begin{aligned}\Delta &= \mathrm{E}\big(Y_i(1)|D_i = 1\big) - \mathrm{E}\big(Y_i(0)|D_i = 0\big)\\ &= \mathrm{E}\big(Y_i(1)|D_i = 1\big) - \mathrm{E}\big(Y_i(0)|D_i = 1\big)\\ &= \mathrm{ATET}\end{aligned}$$

Randomization

Randomization enables us to estimate the ATET from the realized outcome ($\Delta = \mathrm{ATET}$). The following regression is an easy way to do so:

$$Y_i = \beta_0 + \beta_1 D_i + U_i$$

Because D is a (binary) dummy regressor, we get as marginal effect the mean difference between treated and nontreated outcomes, i.e.,

$$\beta_1 = \mathrm{E}(Y_i | D_i = 1) - \mathrm{E}(Y_i | D_i = 0) = \Delta$$

which under randomization equals the ATET.

An alternative way to think about consistency of $\widehat{\beta}_1$ is that $\mathrm{E}(D_i U_i) = 0$, which means that OLS.1 is fulfilled, because D is independent of U due to randomization.

What does OLS.2 mean here?

Randomization

Of course, in this setting we also know that $\mathrm{ATET} = \mathrm{ATE}$ because D is independent of any other random variable (due to randomization).

That is we also have

$$\mathrm{ATET} = \mathrm{E}\big(Y_i(1) - Y_i(0) | D_i = 1\big) = \mathrm{E}\big(Y_i(1) - Y_i(0)\big) = \mathrm{ATE}$$

The last result is a unique characteristic of randomization.

In a general setting (i.e. **no** randomization of treatment), there are <u>not</u> many reasons to assume that the average treatment effect in the entire population is equal to the average treatment effect in the treated population.

People with larger **individual** treatment effects have a higher incentive to participate.

Identification Strategies

Selection on (Un)Observables

Selection on (Un)Observables

Selection on observables and/or selection on unobservables is a huge problem in applied economics.

Suppose that $Y(1)$ and $Y(0)$ are an individual's wages with and without training and D is an indicator whether this person did or did not participate in the training.

If people can choose to participate, then it could be that individuals with higher education

1. are more likely (or more willing) to do the training
2. have already a higher value of $Y(0)$

In this case D and $Y(0)$ are not independent.

Selection on (Un)Observables

If selection only depends on **observable variables**, we can simply add these variables as additional exogenous regressors $\mathbf{x}$, i.e. we estimate

$$Y_i = \beta_0 + \beta_1 D_i + \mathbf{x}_i'\gamma + U_i$$

If selection also depends on **unobservable variables**, differences-in-differences or instrumental variables estimation could help to estimate the treatment effect.

Identification Strategies

Differences-in-Differences Estimation

Differences-in-Differences Estimation

Differences-in-Differences (DiD) estimation is a method that is often used to evaluate programs.

It is particularly well suited for the analysis of natural experiments. In a **natural experiment**, individuals are (as good as) randomly assigned to a treatment and a control group.

In principle DiD estimation also works if there is selection on (un)observables.

<u>Data requirement:</u> We can observe individuals of both groups for two periods (before and after the treatment). A repeated cross-section is sufficient (we do <u>not</u> need a real panel dataset).

Differences-in-Differences Estimation

Treatment group: Individuals that are affected by the treatment.
Control group: Individuals that are not affected by the treatment.

Note that both groups need to be unaffected by the treatment **before** the reform. That's why DiD estimation is often used to evaluate policy changes.

Now, let's partition the data into four groups:

	Before	After
Control	$\bar{y}_{00}$	$\bar{y}_{01}$
Treatment	$\bar{y}_{10}$	$\bar{y}_{11}$

where $\bar{y}_{ta}$ denotes the mean outcome in the respective group.

Differences-in-Differences Estimation

From this table we can now already calculate the differences-in-differences:

	Before	After	Difference over time
Control	$\bar{y}_{00}$	$\bar{y}_{01}$	$\bar{y}_{01} - \bar{y}_{00}$
Treatment	$\bar{y}_{10}$	$\bar{y}_{11}$	$\bar{y}_{11} - \bar{y}_{10}$
Differences over groups	$\bar{y}_{10} - \bar{y}_{00}$	$\bar{y}_{11} - \bar{y}_{01}$	DiD

$$\mathrm{DiD}_v = \bar{y}_{11} - \bar{y}_{10} - [\bar{y}_{01} - \bar{y}_{00}]$$
$$\mathrm{DiD}_h = \bar{y}_{11} - \bar{y}_{01} - [\bar{y}_{10} - \bar{y}_{00}]$$

Of course, $\mathrm{DiD}_v = \mathrm{DiD}_h = \mathrm{DiD}$.

While this calculation gives us the point estimate, it does <u>not</u> provide an easy way to calculate standard errors. Therefore, it is better to employ a linear model representation of the DiD...

Differences-in-Differences Estimation

Linear model representation of DiD estimation:

$$y = \alpha + \beta \text{After} + \gamma \text{Treat} + \delta \text{After} \cdot \text{Treat} + u$$

with

$$\text{After} = \begin{cases} 1 & \text{if observations after the treatment} \\ 0 & \text{otherwise} \end{cases}$$

and

$$\text{Treat} = \begin{cases} 1 & \text{if observations belongs to treatment group} \\ 0 & \text{otherwise} \end{cases}$$

Note that Treat $= 1$ also for observations **before** the treatment (if they would belong to the treated group after the treatment).

Differences-in-Differences Estimation

$$y = \alpha + \beta \text{After} + \gamma \text{Treat} + \delta \text{After} \cdot \text{Treat} + u$$

Interpretation of the parameters in this regression:

	Before	After	Difference over time
Control	α	$\alpha+\beta$	β
Treatment	$\alpha+\gamma$	$\alpha+\beta+\gamma+\delta$	$\beta+\delta$
Differences over groups	γ	$\gamma+\delta$	δ

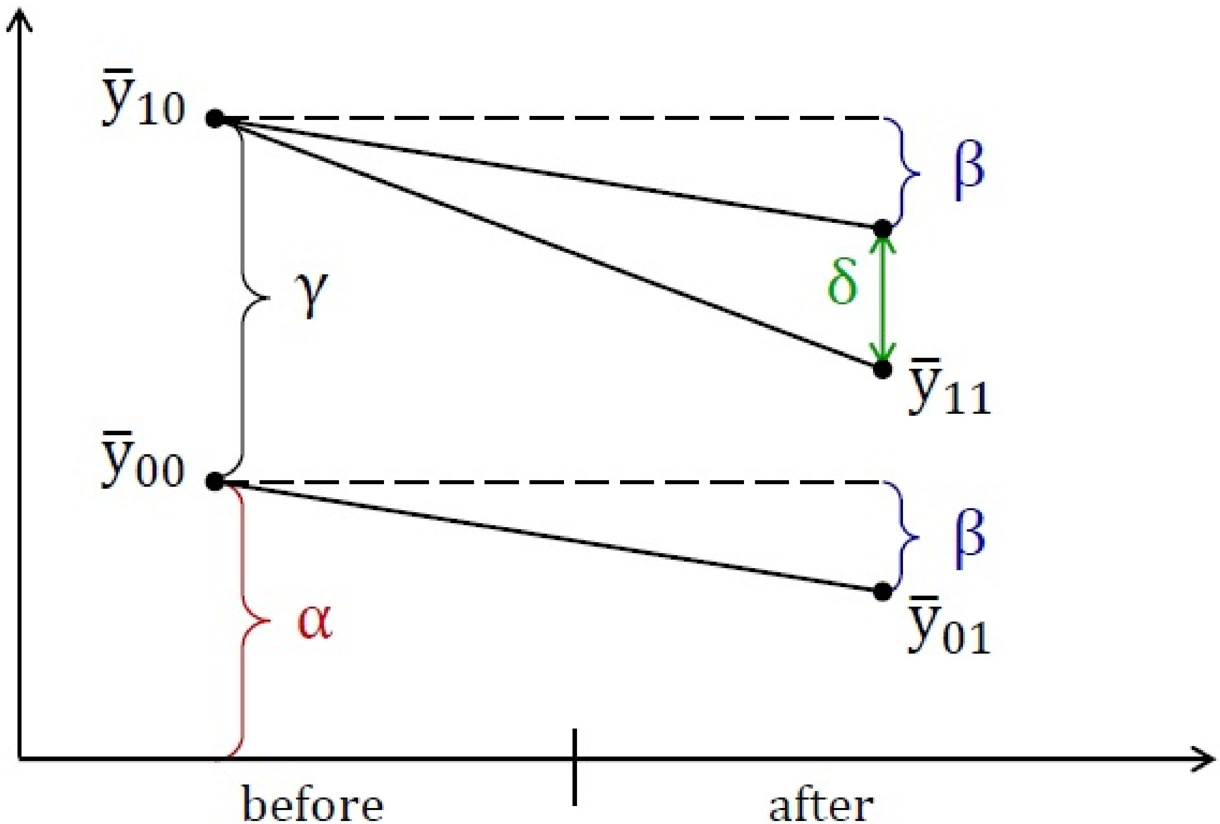
That is, we can interpret $\widehat{\delta}$ as the DiD estimate and additionally get valid standard errors of this estimate.

Moreover, we can easily include control variables in the linear model.

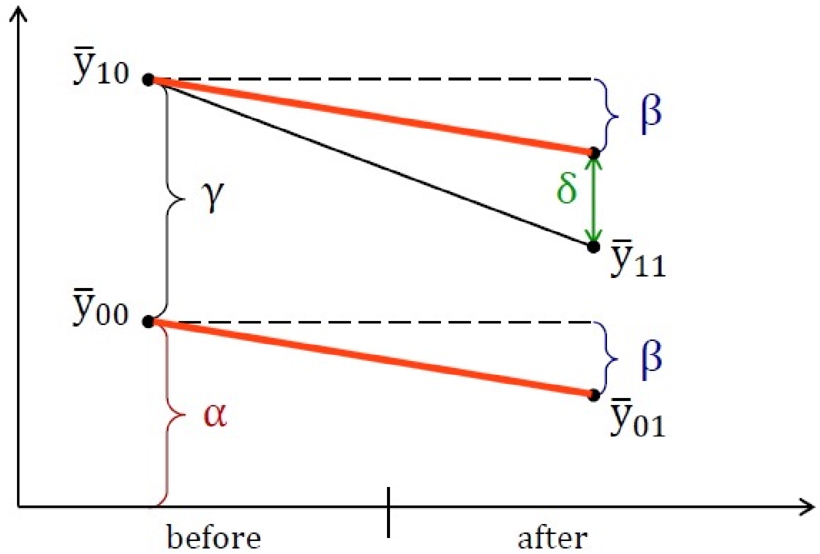
Applied Econometrics

Differences-in-Differences Estimation

Graphical illustration



Differences-in-Differences Estimation



Crucial assumption: Treated observations would follow **the same trend** as control observations in the absence of the treatment (**common trend assumption**).

This is an identifying assumption, which in general cannot be tested with the data at hand.

Differences-in-Differences Estimation: Application

Example: Causal effect of minimum wage on employment.

Theoretical considerations: employment falls due to downward-sloping labour demand curve.

Card and Krueger (1994) "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania" *American Economic Review* Vol. 84 No.4

On April 1st, New Jersey (NJ) increased its minimum wage from $4.25 to $5.05 .

(Adjacent) Pennsylvania (PA) did not change its minimum wage.

Differences-in-Differences Estimation: Application

Data:

They surveyed 410 fast-food stores in NJ and PA before and after the minimum wage increase in NJ.

Fast-food industry: most employees are paid close to the minimum wage; looking at average income earners does not make sense.

Research Strategy:

Differences-in-differences estimation with NJ as treatment group and PA as control group.

Differences-in-Differences Estimation: Application

Figure: Part of Table 3 in Card and Krueger (1994)

Variable	PA (i)	NJ (ii)	Difference, NJ − PA (iii)
	Stores by state		
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)

Outcome: Full-time-equivalent (FTE) employment counts per store

Unexpected results: Employment actually rose in NJ relative to PA after the minimum wage increase.

Contact

Helmut Farbmacher
office.econometrics@mgt.tum.de