Treatment effects

Christoph Hanck

Summer 2024

UDE

Heterogeneity of treatment effects

Definition: Heterogeneous treatment effect

If the effect of a treatment varies across individuals of the population, we speak of a **heterogeneous treatment effect.**

Heterogeneity of treatment

Heterogeneity

Individuals are heterogeneous. If their characteristics impact the effect of treatment, the effects will be heterogeneous across individuals

→ It is often useful to think of each individual as having their own treatment effect.

How do deal with heterogeneous treatment effects?

- We can try to **estimate the heterogeneous treatment effects** − often it difficult to do so we will instead focus on the other thing we can do with the concept of heterogeneous treatment effects like getting answer to the question that of what is being identified if effects are so heterogeneous.

Average treatment effect

> ## Definition: Average treatment effect (ATE)
>
> The *mean* of the treatment effect distribution is called the average treatment effect: If we impose the treatment on everyone, the ATE is the effect we will see for an **average individual**.

Average treatment effect

Example: Drug on cervical cancer

Assume a drug against Cervical cancer will reduce Terry's chances of cervical cancer by 2% and Angela's by 1%. Andrew and Mark do not have cervices so it will reduce their chances by 0.

What is the average treatment effect?

The drug company does tests only on females, they get

$$ATE = \frac{0.02 + 0.01}{2} = 0.015.$$

Note that 1.5% is probably the better answer for a research question (which one?), but it *cannot* be considered as the average treatment effect among the population.

Categories of treatment effect averages

Main categories of treatment effect averages

1. Treatment effect averages where we only count *the treatment effects of some people but not others*, i.e., treatment effect averages *conditional on something*

2. Treatment effect averages where we count *everyone*, but we count some individuals more than others

Isolating the average effect

> ## Example: Treatment effect for each person
>
> | Name | Gender | Untreated Outcome | Treated Outcome | Treatment Effect |
> |---|---|---|---|---|
> | Alfred | Male | 1 | 2 | 1 |
> | Brianna | Female | 1 | 5 | 4 |
> | Chizue | Female | 2 | 5 | 3 |
> | Diego | Male | 2 | 4 | 2 |
>
> The individuals have *different* treatment effects. The *average* treatment effect is
>
> $$\frac{1 + 4 + 3 + 2}{4} = 2.5.$$

Conditional average treatment effects

Example: Conditional ATE

To get an average effect for a certain group we literally pick the corresponding individuals. E.g., we can run an experiment but only on men, taking guys *like* Alfred and Diego and randomly assign them to get treatment or not.

Name	Treated	Outcome
Alfred	Treated	2
Alfred	Untreated	1
Diego	Treated	4
Diego	Untreated	2

Q: What is the conditional ATE?

Average treatment effect on the (un)treated

- Another common way in which the average effect is taken among just one group is based on **who gets treated**.

- We might end up with the **average treatment on the treated (ATT)** or **the average treatment on the untreated (ATUT)**, which averages the treatment effects among those who actually got treated (or not).

- The distinction between ATT and ATUT (and to know which one we end up with) is an important one in nearly all social science contexts.

- Treated and untreated people are often different in quite a few ways (people who choose to do stuff are generally quite different from those who do not!), and we might expect the treatment effect to be different for them, too.

Marginal treatment effect

Another way in which a treatment effect can focus on just a particular group is with the marginal treatment effect.

> ## Definition: Marginal treatment effect
>
> The **marginal treatment effect** is the treatment effect of a person who is just *on the margin of either being treated or not being treated*.

Weighted average treatment effect

We may also include everyone but weight some people differently than others to get a **weighted average treatment effect**.

> ## Definition: Weighted average treatment effect
>
> The **Weighted Average Treatment Effect** is the treatment effect obtained by *placing distinct weights* on different individuals.

Weighted average treatment effect

> ### Example: Treatment effect for each person − ctd.
>
> Consider the table *Isolating the Average Effect* again:
>
> - The mean of 1, 2, 3, and 4 is 2.5
>
> - Assume that Brianna should count twice as much as everyone else, and Diego should count half as much. Then our weighted average treatment effect is
>
> $$\frac{1 \cdot 1 + 4 \cdot 2 + 3 \cdot 1 + 2 \cdot 0.5}{1 + 2 + 1 + 0.5} = 2.89.$$
>
> - In the context of treatment effects, we rarely get to literally pick weights. Instead, there is something about the study design that weights some people more than others.
>
> - A common way this shows up is as *variance-weighted average treatment effects*.

Variance-weighted average treatment effect

- There may be groups of people having lots of variation in treatment while others have not

- We may weight the treatment effect estimate of those with high variation in treatment more heavily, simply because they can be seen both with and without treatment a lot

> ## Definition: Variance-weighted average treatment effect
>
> A treatment effect average where each individual's treatment effect (after closing back doors!) is weighted based on the *variation in their treatment variable* is called **variance-weighted average treatment effect**.

Variance-weighted average treatment effect

Example: Treatment effect for each person – ctd.

Let us consider the following table:

Name	Number of Treated	Outcome
Brianna	500 treated	5
Brianna	500 untreated	1
Diego	900 treated	4
Diego	100 untreated	2

- Here, we can close the back door, '**being a Brianna' / 'being a Diego**' by subtracting out mean differences between Brianna and Diego, both for the outcome and the treatment

- By doing so we get the treatment effect 3.47

Variance-weighted average treatment effect

Example: Drug on cervical cancer − ctd.

3.47 is closer to Brianna's treatment effect of about 4 than to Diego's treatment effect of 2.

- The variance in treatment among Briannas is $0.5 \cdot 0.5 = 0.25$ and among Diegos is $0.9 \cdot 0.1 = 0.09$

- The **variance-weighted average**, then, is $\frac{0.25 \cdot 4 + 0.09 \cdot 2}{0.25 + 0.09} = 3.47$

Distribution-weighted average treatment effect

> ## Definition: Distribution-weighted average treatment effect
>
> In **Distribution-weighted average treatment effect**, individuals with **common values of the matching variable** are weighted more heavily.

Exogenous variation in treatment

- Another form of weighted treatment effect that pops up often is based upon the **responsiveness** of treatment

- Heterogeneous treatment effects not only apply to the effect of treatment on an outcome but also apply to the effect of exogenous variation on treatment

- There may be heterogeneous treatment effects due to (undesired) effects of treatment assignment *on* treatment, rather than the effect of treatment on outcome

Intent-to-treat estimate

If we have exogenous variation but not everybody follows it, and we limit our data to just the people in our experiment, and we look at the relationship between treatment assignment and the outcome, we then get the intent-to-treat estimate.

> ## Definition: Intent-to-treat
>
> Intent-to-treat is **the effect of assigning treatment, although not the effect of treatment itself**, since not everybody follows the assignment.

Intent-to-treat estimate

Note

- Intent-to-treat gives us the **average treatment effect of assignment**

- It weighs each person's treatment effect by **the proportion of their treatment effect they received**

- It is not exactly a weighted treatment effect as instead of dividing by the *sum of the weights*, we divide by the *number of individuals*

Local average treatment effect

Definition: Local average treatment effect

The local average treatment effect (**LATE**) is a **weighted average treatment effect** where **stronger response to exogenous variation** gets weighted more heavily.

Local average treatment effect

- For LATE we use some **source of exogenous variation to predict treatment**, and then use those predictions instead of our actual data on treatment

- This approach does not just say *were you assigned treatment or not?* but rather *how much more treatment do we think you got due to assignment?*

- If we replace the *number of people* denominator with a *how much more treatment was there?* denominator, LATE becomes very similar to the intent-to-treat

- Specifically, the weights in LATE reflect how much additional treatment each individual would get if assigned to treatment

Importance of ATE

Why ATE is better than other treatment effects

- The treatment effect we get is almost entirely determined by the source of treatment variation we use.

- Usually, we want the average treatment effect—the effect we would see *on average* if we took a single individual and applied the treatment to them.

Can we always get an ATE?

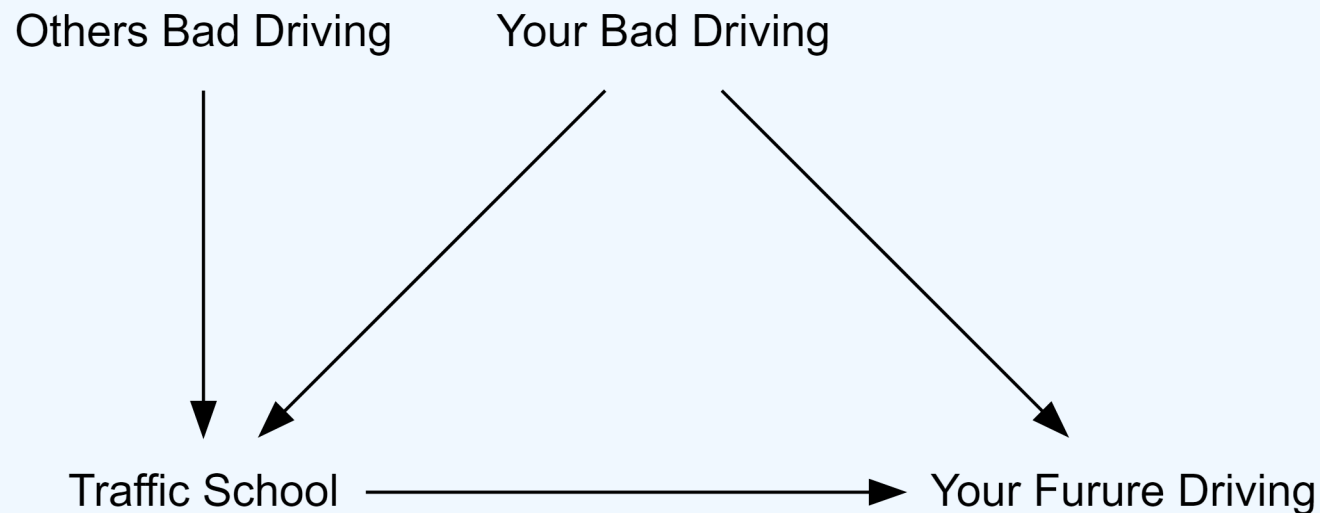Example: Effect of traffic school on future driving performance

Others Bad Driving   Your Bad Driving

Traffic School ——————————————→ Your Furure Driving

Figure 2: Causal diagram of Effect of traffic school on future driving performance

Can we always get an ATE?

> ## Example: Effect of traffic school on future driving performance – ctd.
>
> - There are only two reasons anyone goes to traffic school: **making a terrible driving mistake**, or **having someone else make a terrible driving mistake that you are somehow punished for.**
> - Back Door: *TrafficSchool ← YourBadDriving → YourFutureDriving*
> - We want to identify the effect by measuring and controlling for your own bad driving skills
> - This will identify the effect, but it will also shut out any variation in *TrafficSchool* that is driven by *YourBadDriving*

Can we always get an ATE?

Example: Effect of traffic school on future driving performance

Consider Rodney and Richard:

- Rodney has a 50% chance of not going to traffic school, a 10% chance of going because of someone else's bad driving, and a 40% chance of going because of his own bad driving.

- Richard has a 50% chance of not going to traffic school, a 30% chance of going because of someone else's bad driving, and a 20% chance of going because of his own bad driving.

Can we always get an ATE?

> ## Example: Effect of traffic school on future driving performance
>
> **Outcome**
>
> - We are tossing out that 40% for Rodney and 20% for Richard chances of going because of their own bad driving.
>
> - There is only a 10% chance that Rodney goes to traffic school for the reason we still allow to count, and similarly a 30% chance for Richard.
>
> - That means there is more remaining variation in treatment for Richard than for Rodney, so Richard's treatment effect will be weighted more heavily than Rodney's will.
>
>   **A weighted average treatment effect!**

Which treatment effect do we get?

> ## Thumb rules
>
> 1. If you have *true randomization* in a representative sample and do not need to do any adjustment, then you have an average treatment effect (ATE).
>
> 2. If you have true randomization only within a certain group, and you isolate that group so you can take advantage of that randomization, you have a conditional average treatment effect (CTE)
>
> 3. If you know that some variation in treatment is connected to back doors and so you close those back doors, using only the remaining variation, you have a weighted average treatment effect (WATE)

Which treatment effect do we get?

Thumb rules – ctd.

4. If you are identifying your effect by assuming that some untreated group is what the treated group would look like if they had not been treated, then we have the average treatment on the treated (ATT)

5. If part of the variation in treatment is driven by an exogenous variable, and you isolate just the part driven by that exogenous variable, then you have a local average treatment effect (LATE)