DIFFERENCES-IN-DIFFERENCES & FIXED EFFECTS

Quantitative Analysis
Week 11

Thinking Longitudinally

- Up to now, most the data we have been working with has been **cross-sectional** – which means that it consists of observations of many cases at the same moment in time, i.e., one "slice" through the timeline.
- Time series data, also called **longitudinal data**, instead has observations of cases at many different moments in time.
- This allows us to analyse how the relationships between different phenomena develop over time and can reveal things we would never observe in cross-sectional data.
  - For example, how people's attitudes and behaviour respond to specific events;
  - Or situations where there is a time gap between cause and effect, like an election leading to a policy change.

Time series notation

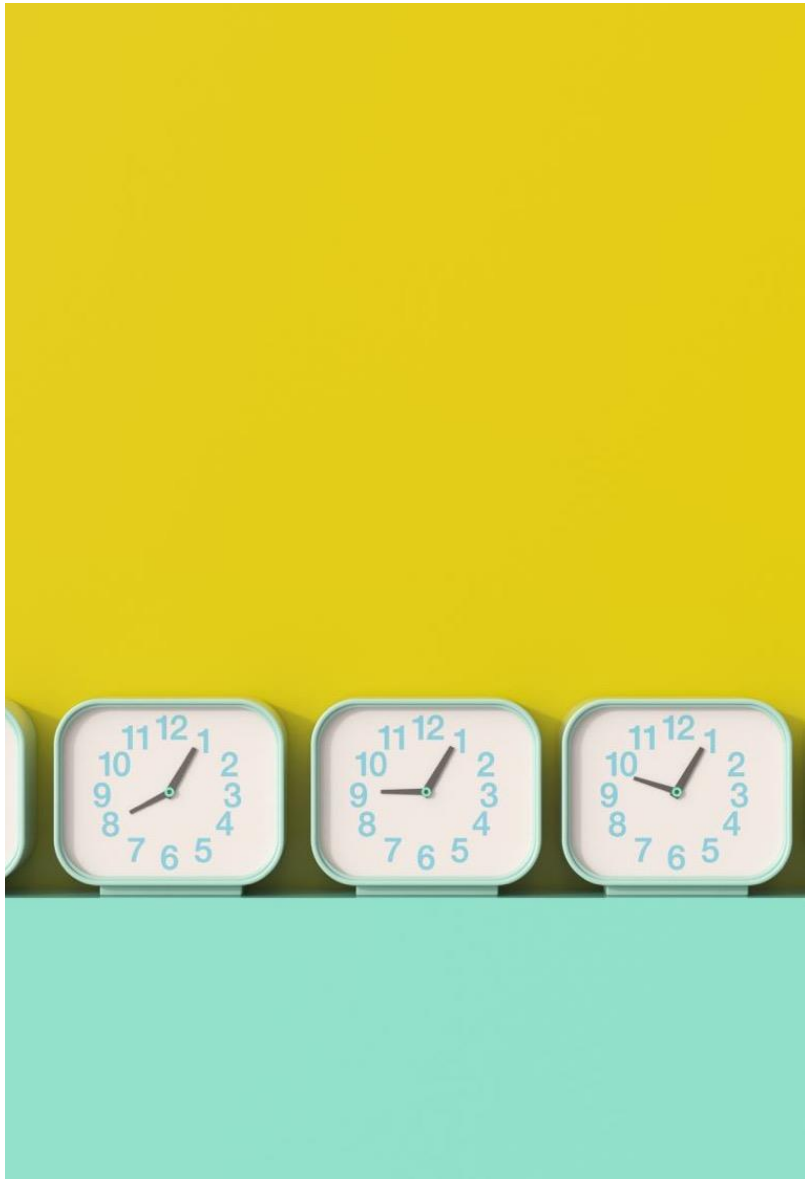
When we talk about time series data, we usually use the variable t to denote the time period we're looking at.

t or t_0 is the present, or the time period of interest (for example, the point where the event we're interested in studying happened).

We then denote earlier time periods by counting backwards in negative numbers: $t_{-1}, t_{-2}, t_{-3} \cdots t_{-n}$

Later time periods: $t_{+1}, t_{+2}, t_{+3} \cdots t_{+n}$

Variable X at time period t_{-3} would be written as X_{t-3}



Is my data really a time series?

When we talk about **time series analysis**, we're usually referring to a specific group of methods that are designed to deal with data that has a lot of observations at regular intervals.

- Daily stock closing prices;
- Monthly opinion polls;
- Annual demographic figures; etc.

Time series analysis provides us with methods for studying this kind of data by controlling for things like **seasonal cycles** and **long-term trends**, which are problems we don't encounter with cross-sectional data.

We'll discuss how to do this in the next class; in this class, we'll be focusing on **panel data**, which are observations from a more limited set of time periods.

Before & After Models

- In many cases, you only have observations for two or three time periods – a common example is an electoral survey which has **waves** conducted both before and after the election.
- With **panel data** like this, there is not enough data for things like seasonality or long-term trends to become an issue.
- In cases where you have observations of the same cases before and after the event, you can use normal regression analysis techniques, using variables from both time periods
- For example, attitudes to a political leader at t_0 might be predicted by a combination of their views of the leader at t_{-1} and their current (t_0) view on the nation's economic performance.
- **Be careful about causality!**
  - A variable from a later time period must not be used to predict an outcome in an earlier time period.

Differences-in-Differences

- The **before-and-after** approach is based on the difference between samples at t_{-1} and t_0 - but what if there are trends that would have caused our measurements to change over time anyway?
- A common approach to this problem is a research design called **differences-in-differences**, or **DiD**.
- DiD is ideal for situations where we are studying a "treatment" – a causal factor applied to one group of observations but not to others, who become the control group.
- It assumes that without the treatment, the measures of the treatment group would have changed in a similar way to the measures of the control group.

Cholera in London

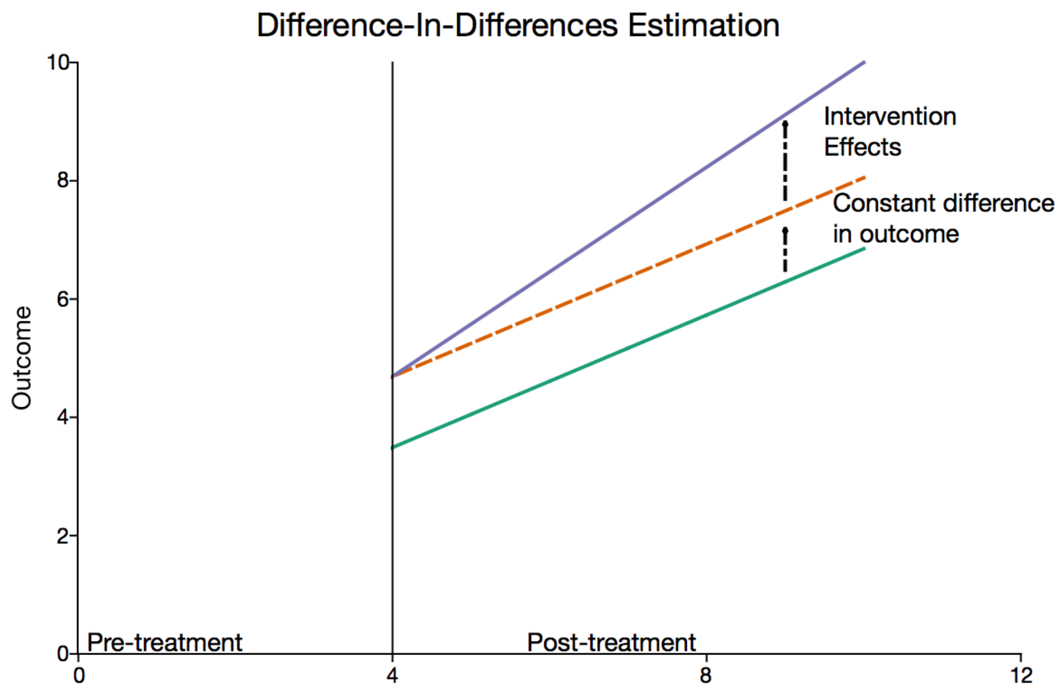
- In 1854, there was an outbreak of cholera in London that killed over 600 people in Soho. At the time, it was believed that disease was spread by ‘*miasma*’ – i.e., ‘bad air’.
- A doctor called John Snow believed that the real problem was drinking dirty water. He measured cholera cases in boroughs of London in 1849 and 1854. His “treatment” was a change to the water supply in the borough of Lambeth in 1852, which moved it up the River Thames and away from London’s pollution.
- His analysis method was, essentially, differences-in-differences; he proved that Lambeth’s cholera cases were lower than would have been expected if the borough had followed the same trend as other regions.
- He used this evidence to convince authorities to remove the pump handle from the water pump in Soho. People stopped using the polluted water it had provided, the cholera outbreak stopped – and the foundations for our modern understanding of disease were laid.

Parallel Trend Assumption

Differences-in-Differences' underlying idea is called the **parallel trend assumption**, and it looks like the graph opposite.

The blue line is the measured outcome for the treatment group; the green line is the measured outcome for the control group.

By measuring the change in the control group, we can predict what the measurement for the treatment group would have been *without* the treatment – and therefore calculate the **treatment effect**.



Differences-in-Differences Regression

- One of the simplest ways to incorporate a Differences-in-Differences design in a regression model is to create **dummy variables** for the time periods (e.g., 0 for the pre-treatment period, 1 for the post-treatment period).
- Then we create an **interaction term** in the regression model between the time period and the treatment effect. For example:
- $y \sim \alpha + \beta_1 treat + \beta_2 time + \beta_3 (treat \times time) + \epsilon$
- This would then give us a coefficient β_3 (i.e., the coefficient for the interaction term) which tells us the effect of the treatment on the outcome compared to the *predicted* outcome, assuming parallel trends.

D-I-D Regression Over Longer Terms

- DiD Regression doesn't just work for situations with two time periods (pre- and post-treatment). We can also create dummy variables for several time periods and see the treatment effects over time.
- This is especially helpful if we have data for a few periods prior to the treatment, because it lets us confirm the parallel trends assumption.
  - If that assumption holds, the regression coefficients should be around zero in the pre-treatment periods.
  - Then you should see a **step change** at the treatment point, with the treated samples being different from the control samples in all post-treatment periods.

FIXED EFFECTS

Building on Differences in Differences

- Think again for a moment about the Differences-in-Differences model, and the parallel trends assumption.
- The strength of this model is that it doesn't assume the treated case (e.g., Lambeth) would be **identical** to the other cases (e.g., the other London boroughs) without the treatment.
- If we could just assume that they would be identical, we could simply show that Lambeth had lower per-capita cholera cases than the other boroughs, and be done with it.
- However, we know intuitively that things are usually more complex than that.
  - There could be many factors which impact on Lambeth's per-capita cholera cases – population density, poverty, age demographics, etc. etc. – so they wouldn't be identical to the other boroughs even regardless of the treatment.

Missing Information

- We've talked several times about **omitted variable bias** – which arises when there's an important factor missing from your model, perhaps because there's no available data for it.
- In this case, the factors which made Lambeth different from the other boroughs before the policy change are examples of omitted variables.
- However, we know one other useful thing: there was no other major change, apart from the water policy, that would alter the underlying factors about Lambeth relative to other boroughs.
- Between 1849 and 1854 (Snow's two measured years), there was no major upheaval that would change Lambeth's population density, poverty, demographics etc. in a significantly different way to the rest of London.

Unit Level Fixed Effects

- What this all means is that we can describe all of those factors related to Lambeth (and the other boroughs) as being **unit-unique and time-invariant**.
- This means that every unit (borough) has its own unique characteristics, and that while we have not measured these characteristics (they are **unobserved**), we know that they do not change over time – at least not in a way that diverges dramatically from unit to unit.
- We can therefore construct a regression model using **unit level fixed effects** – basically saying, "every observation of Unit X, no matter which time period it occurs in, will be explained to some degree just by being similar to all other observations of Unit X".
  - Mathematically, this is generally calculated by **de-meaning** the variables in the regression – subtracting their mean values from each observation, so you are regressing based on the variation around the mean.

When to use Fixed Effects

- Unit level fixed effects are used very commonly in social science research.
- In almost any situation where you're observing specific units (people, businesses, regions, countries, etc.), it will be impossible to observe every variable related to them – so fixed effects are used to reduce omitted variable bias.
- The most important assumption of fixed effects is that there is **no unobserved time-varying confounder** – in other words, the differences between units are **time-invariant**, apart from the differences caused by the treatment. If you have reason to believe that there will be a major time-variant difference between units, you cannot use a fixed effects model.

Fixed Effects at Other Levels

- The most common kind of fixed effects model is the unit-level model – but it's also possible to include fixed effects at other levels.
- For example, you might be studying the experiences of immigrants, and know that their country of origin impacts their opportunities in various ways – so you would set a fixed effect at the country level.
- In this case, all of the people you study from country X would share a single fixed effect parameter for their national origin.
- This could be calculated either through the de-meaning process mentioned before (this is also called the **within transformation**), or by including **dummy variables** for each country.

OVER TO RSTUDIO