# 8: Difference-in-Differences

## Quasi-Experiments

* External validity is generally great as it’s occurred in the real-world, so generalisation is possible.
* Could be some self-selection built into the study as subjects will choose what to do based on what suits them best.
* Assignment to treatment is “as if” random can be a big assumption here and a lot of a research report can be devoted to explaining why it’s a reasonable assumption to make.
* With matching/IPW, you don’t need a specific research design situation, whereas with a DiD experiment, you need a control and treatment group and a before and after.
  + Very situational-based thing.
* The goal of quasi-experiments is to isolate the pathway between treatment and control, but the isolation is achieved by the context/situation rather than actively stopping confounding.
* In matching/IPW you statistically adjust for the confounding and the identification comes from closing all the backdoors. Don’t necessarily need a treatment and control group.
* There’s a specific DAG structure that comes from a DiD, RCT and IV etc.
* Can still do all the adjustments with a natural experiment – DAGs can enhance the experiment.

## Interactions & Regression

* Indicator variables shift the entire regression line up or down.
* Interactions get interpreted slightly differently to continuous or indicator variables i.e. the slope is the global slope PLUS the interaction term when the indicator variable takes a value of 1.
* Often the global intercept won’t mean anything as it’ll be outside the range of reasonable explanatory terms.
* Interaction terms can’t be included in isolation, the component terms must also be present.
* Interaction terms are all shifts in slope relative to the baseline, which doesn’t change if the categorical variable has multiple values instead of just being binary.

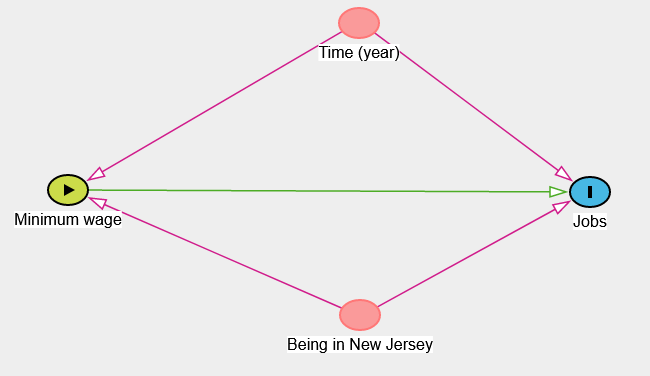
### Example

* A plain hotdog is $2, getting cheese is an additional $0.35… this is the cheese effect.
* The chilli effect is also $0.35.
* The chilli-cheese effect is $0.00 – there’s no additional change in price from combining the two. If the chilli-cheese dog was $2.50, then the chilli-cheese effect would be -$0.20.
* Interaction times are trying to determine the extra effect of combining the two terms.

## Two Wrongs Make a Right

* Relies on a natural counterfactual group existing that is very similar in characteristics to the treatment group.

### Raising the minimum wage

* Raising the minimum wage caused fast food restaurants to employ 0.59 extra workers on average in New Jersey.
* This isn’t the causal effect as we can’t compare this figure to anything i.e. comparing before and after states within a group isn’t enough to say anything about a causal effect.
* The average number of jobs in New Jersey was 21.03 after raising the minimum wage vs 21.17 in Pennsylvania (where the minimum wage didn’t change) or a decrease of 0.14 on average.
* Again, this isn’t a causal effect as there no pre-treatment state to compare the difference to i.e. it’s a single point in time.
* Need to compare both pre- and post-treatment for a treatment and control group simultaneously for a causal effect statement.
* More clear if a DAG is drawn:
  + 
  + The relationship between a change in minimum wage and jobs is confounded by the year that the change occurred and geographical location.
  + Need to do something statistically to both of these to isolate the relationship of interest.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Pre-mean | Post-mean |  |
| Pennsylvania | 23.33 | 21.17 | -2.16 |
| New Jersey | 20.44 | 21.03 | 0.59 |
|  | -2.89 | -0.14 | 2.76 (column or row total) |

* In NJ, there was an increase in the average number of jobs, but comparing NJ and Pennsylvania after the program, it looked as though the program caused a job loss… conflicting conclusions. BUT, this is because neither is a causal effect.
* In NJ, the minimum wage increase caused fast food restaurants to hire an addition 2.76 workers on average.
  + The policy change reversed the economic decline seen in the control group.
* The causal effect is the difference between the actual and hypothetical end-points in the treatment group.
* Plotting the analysis is great because a chart is reasonably intuitive for most people.

### Is there an easier way?

* It’s easy to make a mistake (not to mention tedious) when calculating the means and then subtracting them.
* You only get point estimates – there’s no confidence intervals to provide an indication of statistical uncertainty.
* Near on impossible if there are more than two confounders.
* Could stick all the confounders in a regression. Not always ideal as you then have to assume the relationship is linear i.e. matching and IPW.
* Here it’s ok as the confounders are all binary and including them is like adjusting for those nodes in the DAG.
  + Each of these coefficients describes a part of the 2x2 table.
  + is just treatment minus control
  + is post- minus pre-treatment (after minus before)
* Can build the 2x2 table from the regression coefficients if you really wanted to.
* is the causal effect!

## Difference-in-Differences Assumptions

* Very tricky to find comparable control groups.

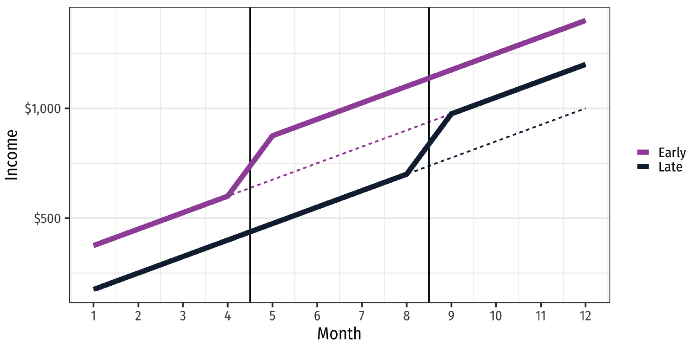
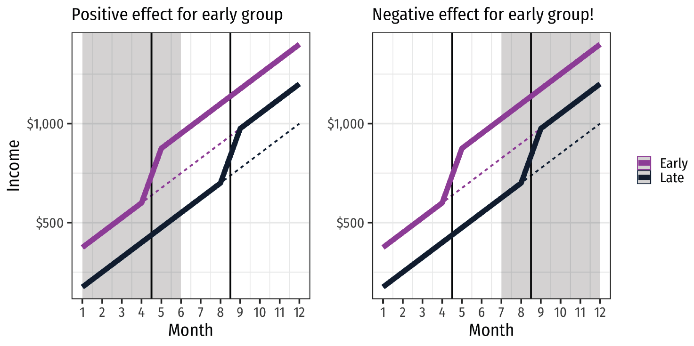
### Parallel trends

* Both groups need to move in a similar way pre-treatment.
* Can be checked graphically using a simple line chart (for a univariate case).
* Credible causal effect (red line) only if it’s plausible that the parallel trends would have continued in the absence of treatment.
* Get as much pre-treatment data as possible to argue that the control group is appropriate (because this assumption is reasonable).

### Causal Effect

* If a DiD effect is identified well before the time of treatment/policy change, it’s likely due to the two groups having different trends rather than the ‘fake’ policy. Therefore, the parallel trends assumption is violated.

### Treatment Timing

* 
* Thus far only been considering the period around months 4-5.
* Suppose that the late group raises their minimum wage some months later and the early group doesn’t. This means there’s a different causal effect around months 8-9, which makes it look worse for the early group.
* 
* It appears as though the second treatment is reducing the positive effect the early group had, because the late group is now getting that positive effect.
* Don’t know which line you’re looking at in real-life. Always looking at one of them, but groups are constantly changing their policies at different times, which will distort the findings.