# Difference-in-Differences with Modern Advancements

## Applied Econometrics Resources

Kevin Li

# Classic Example

New Jersey increased minimum wage in April 1992. We want to find the causal effect of this on employment rate $Y$.

Within new jersey, we sample fast food restaurants. Each unit $i$ is sampled both before and after the minimum wage increase:

- Before $t = 0$ (March 1992). Average employment rate $\mathbb{E}[Y_{i0}]$
- After $t = 1$ (December 1992). Average employment rate $\mathbb{E}[Y_{i1}]$.

Is the causal effect the difference in average $Y_{it}$ between the two periods $t$?

$$\tau = \mathbb{E}[Y_{i1}] - \mathbb{E}[Y_{i0}]$$

**No**. Why? maybe between $t = 0, 1$ there was a recession that lowered the employment rate. We wouldn't know if the change is due to the minimum wage increase, or the recession.

# Solution: Imputing Trends

So how can we isolate the causal effect if we don't know what the effect of the recession is?

Let us look at another state, Pennsylvania. More formally:

- New Jersey (and its observations i) will be in group $G = 1$.
- Pennsylvania (and its observations i) will be in group $G = 0$.

Pennsylvania $G = 0$ does not have minimum wage change between $t = 0, 1$. But it is affected by recession.

So, we can calculate the effect of recession with the difference between pennsylvania between the two time periods:

$$\text{Recession Effect} = \mathbb{E}[Y_{i1} | G_{it} = 0] - \mathbb{E}[Y_{i0} | G_{it} = 0]$$

# Difference-in-Differences

We know the recession effect now.

So we can take the difference between $t = 0, 1$ in New Jersey, then adjust (minus out) the recession effect to get the causal effect:

$$\tau = \underbrace{(\mathbb{E}[Y_{i1}|G_{it} = 1] - \mathbb{E}[Y_{i0}|G_{it} = 1])}_{\text{New Jersey diff.}} - \underbrace{(\mathbb{E}[Y_{i1}|G_{it} = 0] - \mathbb{E}[Y_{i0}|G_{it} = 0])}_{\text{Pennsylvania Diff.}}$$

- New Jersey diff is the original estimate not accounting for the recession.
- Pennsylvania difference isolates the causal effect.

Notice how it is the difference of two differences, hence, difference-in-differences.

- The first difference is the "correlation" between after and before treatment in the treated group.
- The second difference accounts for confounders/changes in time periods.

# Parallel Trends

Of course, this technique requires a key assumption of **parallel trends**.

$$\mathbb{E}[Y_{i1}(0) - Y_{i0}(0)|G_i = 1] = \mathbb{E}[Y_{i1}(0) - Y_{i0}(0)|G_i = 0]$$

- Read: had the treated group $G = 1$ never gotten treated in $t = 1$, they would have followed the same trend as the untreated $G = 0$ group.
- In other words, had New Jersey not been treated, their difference would be equal to that of pennsylvania.
- Or even more intuitively: if New Jersey had not had a minimum wage change, it would have only been impacted by the recession (and other factors).

  This impact of the recession (and other factors)should be consistent between New Jersey and Pennsylvania.

# Identification Proof

Let us prove that this difference-in-differences estimate really gets us a causal estimand, using potential outcomes:

$$
\begin{aligned}
\mathrm{DID} &= (\mathbb{E}[Y_{i1}|G_{it}=1]-\mathbb{E}[Y_{i0}|G_{it}=1])-(\mathbb{E}[Y_{i1}|G_{it}=0]-\mathbb{E}[Y_{i0}|G_{it}=0]) \\
&= (\mathbb{E}[Y_{i1}(1)|G_{it}=1]-\mathbb{E}[Y_{i0}(0)|G_{it}=1])-(\mathbb{E}[Y_{i1}(0)|G_{it}=0]-\mathbb{E}[Y_{i0}(0)|G_{it}=0]) \\
&= (\mathbb{E}[Y_{i1}(1)|G_{it}=1]-\mathbb{E}[Y_{i0}(0)|G_{it}=1]) \\
&\qquad -(\mathbb{E}[Y_{i1}(0)|G_{it}=0]-\mathbb{E}[Y_{i0}(0)|G_{it}=0])+(\mathbb{E}[Y_{i1}(0)|G_{it}=1]-\mathbb{E}[Y_{i1}(0)|G_{it}=1]) \\
&= \mathbb{E}[Y_{i1}(1)|G_{it}=1]-\mathbb{E}[Y_{i1}(0)|G_{it}=1] \\
&\qquad +\mathbb{E}[Y_{i1}(0)-Y_{i0}(0)|G_{i}=1]-\mathbb{E}[Y_{i1}(0)-Y_{i0}(0)|G_{i}=0] \\
&= \mathbb{E}[Y_{i1}(1)|G_{it}=1]-\mathbb{E}[Y_{i1}(0)|G_{it}=1] \\
&= \mathbb{E}[Y_{i1}(1)|D_{it}=1]-\mathbb{E}[Y_{i1}(0)|D_{it}=1] \\
&= \mathrm{ATT}
\end{aligned}
$$

## More Time Periods

In our example, we currently have only two time periods, t $= 0, 1$.

We can have more time periods. These time periods will be split into two types:

1. Pre-treatment time periods (treated the same as t $= 0$)
2. Post-treatment time periods (treated the same as t $= 1$).

We will use the variable R to describe time periods:

- Pre-treatment time periods are negative values of R.
- Post-treatment time periods are 0 and positive values of R.
- R $= 0$ is the first year that treatment G $= 1$ adopts treatment.