Jorge Eduardo Pérez Pérez

Advances in Differences in Differences and Bartik instruments: Class Notes

Banco de México

October 8, 2024

Chapter 1
The Basic Difference in Differences Model

In this first chapter, we will review the basics of the difference in differences model. This review will serve as a building block for discussion of recent topics. We follow de Chaisemartin and D'Haultfœuille (2023) and Roth et al. (2023) for this review.

1.1 The simple case

We start with the simple case to introduce notation and assumptions. Units are indexed by i. There are $t = 1,2$ periods. In $t = 2$, treated units receive a treatment, while control units remain untreated in both periods. We will denote treated units by those for which $D_i = 1$, and control units by $D_i = 0$. We observe a panel of units $i = 1,...,N$ for both periods. For each unit, we observe an outcome $Y_{i,t}$ and a treatment status indicator D_i.

Each unit has potential outcomes that would be observed had the unit received treatment or not. For each period t, each unit has two potential outcomes: $Y_{i,t}(0,0)$ is the potential outcome if the unit is not treated and $Y_{i,t}(0,1)$ is the potential outcome if the unit is treated. To simplify, we write $Y_{i,t}(0) = Y_{i,t}(0,0)$ and $Y_{i,t}(1) = Y_{i,t}(0,1)$.

The fundamental problem of causal inference is that we do not observe the potential outcomes. We only observe the realized outcome $Y_{i,t} = D_i Y_{i,t}(1) + (1-D_i)Y_{i,t}(0)$.

The **individual treatment effect** $\tau_{i,t}$ is the difference between the potential outcomes in the treated and untreated states for every individual i. For $t = 2$:

$$\tau_{i,2} := Y_{i,2}(1) - Y_{i,2}(0) = Y_{i,2}(0,1) - Y_{i,2}(0,0) \tag{1.1}$$

We do not try to estimate these individual treatment effects because we do not observe individuals repeatedly. Instead, we try to estimate the **average treatment effect on the treated (ATT)**, that is, the average treatment effect conditional on belonging to the set of treated units.

$$\tau_2 := \mathbb{E}[\tau_{i,2}|D_i = 1] = \mathbb{E}[Y_{i,2}(1) - Y_{i,2}(0)|D_i = 1] \tag{1.2}$$

We do not observe untreated outcomes $Y_{i,2}(0)$ for the treated group, so the ATT is not identified only from data. To identify it, we are going to *impute* the untreated outcomes using the baseline values of the outcome on the control group, and assuming that in absence of treatment, the untreated outcomes would evolve in parallel across treated and control units. This is the parallel trends assumption:

Assumption 1.1 *(Parallel trends)*

$$\mathbb{E}[Y_{i,2}(0) - Y_{i,1}(0)|D_i = 1] = \mathbb{E}[Y_{i,2}(0) - Y_{i,1}(0)|D_i = 0] \tag{1.3}$$

Parallel trends can be justified by a fixed-effects model for the potential untreated outcome: $Y_{i,t}(0) = \alpha_i + \phi_t + \varepsilon_{i,t}$ with $\text{Cov}(D_i, \varepsilon_{i,t}) = 0$. In this model, treatment need not be random: it can be correlated with individual characteristics α_i that are constant over time. However, it cannot be correlated with characteristics that change over time $\varepsilon_{i,t}$.

We also need a no anticipation assumption: treatment status in t only affects the outcome in t:

Assumption 1.2 *(No anticipation)* $Y_{i,1}(0) = Y_{i,1}(1)$ *for all i with* $D_i = 1$

An additional hidden assumption is the stable unit treatment value assumption (SUTVA). This assumption is embedded in how we wrote the outcome equation: $Y_{i,t} = D_i Y_{i,t}(1) + (1-D_i)Y_{i,t}(0)$. Here, the outcome for unit i does not depend on the treatment status of other units. This rules out spillover effects. We do not define it formally since it is embeded in the potential outcomes framework, but list it as in Cunningham (2021).

Assumption 1.3 *(Stable unit treatment value)*

1. *A treated unit cannot impact a control unit such that their potential outcomes change*
2. *Units that are assigned to treatment receive the same treatment value*

With these assumptions we are ready to identify τ_2. We want to impute the average value of the potential outcomes for the untreated units $\mathbb{E}[Y_{i,2}(0)|D_i = 1]$. By parallel trends:

$$\mathbb{E}[Y_{i,2}(0)|D_i = 1] = \mathbb{E}[Y_{i,1}(0)|D_i = 1] + \mathbb{E}[Y_{i,2}(0) - Y_{i,1}(0)|D_i = 0]$$

The second term cannot be obtained from data because it is a potential outcome: It is the **potential** outcome for treated units in time 0 had they not been treated in time 1. However, because of no anticipation, this equals the **observed** outcome for treated units in time 0.

$$\mathbb{E}[Y_{i,1}(0)|D_i = 1] = \mathbb{E}[Y_{i,1}(1)|D_i = 1]$$

Therefore the imputed value of $\mathbb{E}[Y_{i,2}(0)|D_i = 1]$ is:

$$\mathbb{E}[Y_{i,2}(0)|D_i = 1] = \mathbb{E}[Y_{i,1}(1)|D_i = 1] + \mathbb{E}[Y_{i,2}(0) - Y_{i,1}(0)|D_i = 0]$$

which is the baseline value for the treated units plus the change in time of the outcome for control units. Under these assumptions the ATT is:

$$\begin{aligned}\tau_2 &= \mathbb{E}[Y_{i,2}(1) - Y_{i,2}(0)|D_i = 1] \\ &= \mathbb{E}[Y_{i,2}(1)|D_i = 1] - \mathbb{E}[Y_{i,1}(1)|D_i = 1] - \mathbb{E}[Y_{i,2}(0) - Y_{i,1}(0)|D_i = 0]\end{aligned} \tag{1.4}$$

which is the traditional difference in differences expression.

The imputation of the potential outcomes for the treated group can be easily visualized in a graph:

[Cunningham Mixtape II 1.44]

For estimation, equation (1.4) gives us a natural way of estimating the TT from data, simply replacing the expectations by their sample analogs. A more practical way is to write (1.4) as a linear model:

$$Y_{i,t} = \alpha_i + \phi_t + \mathbf{1}(t = 2) \times D_i \times \beta + \varepsilon_{i,t} \tag{1.5}$$

and estimate the parameters, via OLS for example. We call this specification the **two way fixed effects (TWFE)** specification. The OLS estimate of β, $\hat{\beta}$, is equivalent to $\hat{\tau}_2$, an estimate of τ_2.

Equation (1.5) also provides a simple way to conduct inference on $\hat{\beta}$. Under a sampling assumption:

Assumption 1.4 *(Sampling)*
We observe a random sample of N i.i.d draws of $(Y_{i,2}, Y_{i,1}, D_i)'$ *with joint distribution F satisfying assumptions 1-3.*

Under assumptions 1-4 we have that the estimator $\hat{\beta}$ is consistent and asymptotically normal

$$\sqrt{n}(\hat{\beta} - \tau_2) \underset{d}{\rightarrow} \mathcal{N}(0, \sigma^2)$$

as $N \to \infty$. The variance σ^2 can be estimated with standard linear regression methods. We will talk about usual inference issues in this setup later.

1.2 Triple differences

It is worthwhile to also review the triple-difference estimator here. A nice recent paper (Olden and Møen, 2022) summarizes this estimator.

Suppose there are two states, one of which adopts a policy $(T = 1)$ while the other does not $C = 1$. In both states, there are two groups, A and B. Only group B benefits from the policy. A typical double difference estimator could compare individuals in group B between states T and C, under parallel trends in the potential outcomes of group B across states. However, if this assumption is unlikely to hold, we can take advantage of having an untreated group. We can use the untreated group A to measure the difference in trends in potential outcomes between states T and C. Then, we can adjust the difference-in-differences estimator that compares group B across states by the difference in trends in potential outcomes across estates to arrive at an unbiased estimate of the ATT of the policy.

The diff-in-diff across states for group B is:

$$\beta^{DD,B} = \mathbb{E}[Y|T,B,Post] - \mathbb{E}[Y|T,B,Pre] - (E[Y|C,B,Post] - E[Y|C,B,Pre]) \tag{1.6}$$

In terms of potential outcomes, this equals

$$\begin{aligned}
\beta^{DD,B} &= \mathbb{E}[Y(1)|T,B,Post] - \mathbb{E}[Y(0)|T,B,Pre] - (\mathbb{E}[Y(0)|C,B,Post] - \mathbb{E}[Y(0)|C,B,Pre]) \\
&= \mathbb{E}[Y(1)|T,B,Post] - \mathbb{E}[Y(0)|T,B,Post] + \mathbb{E}[Y(0)|T,B,Post] \\
&- \mathbb{E}[Y(0)|T,B,Pre] - (\mathbb{E}[Y(0)|C,B,Post] - \mathbb{E}[Y(0)|C,B,Pre]) \\
&= ATT + \mathbb{E}[Y(0)|T,B,Post] - \mathbb{E}[Y(0)|T,B,Pre] - (\mathbb{E}[Y(0)|C,B,Post] - \mathbb{E}[Y(0)|C,B,Pre])
\end{aligned}$$

Under parallel trends the terms after ATT cancel out. Without parallel trends, we can define the analog diff-in-diff across states for group A:

$$\beta^{DD,A} = \mathbb{E}[Y|T,A,Post] - \mathbb{E}[Y|T,A,Pre] - (E[Y|C,B,Post] - E[Y|C,B,Pre])$$

In terms of potential outcomes, this equals

$$\beta^{DD,A} = \mathbb{E}[Y(0)|T,A,Post] - \mathbb{E}[Y(0)|T,A,Pre] - (E[Y(0)|C,B,Post] - E[Y(0)|C,B,Pre])$$

Assuming that the difference in trends for potential outcomes across states is the same across groups:

$$\begin{aligned}
&\mathbb{E}[Y(0)|T,B,Post] - \mathbb{E}[Y(0)|T,B,Pre] - (\mathbb{E}[Y(0)|C,B,Post] - \mathbb{E}[Y(0)|C,B,Pre]) \qquad (1.7)\\
&= \mathbb{E}[Y(0)|T,A,Post] - \mathbb{E}[Y(0)|T,A,Pre] - (E[Y(0)|C,B,Post] - E[Y(0)|C,B,Pre])
\end{aligned}$$

Then $\beta^{DDD} = \beta^{DD,B} - \beta^{DD,A} = ATT$
To estimate β^{DDD} we can use the following fixed-effects specification:

$$\begin{aligned} Y_{sit} = &\alpha + \beta_1 T + \beta_2 B + \beta_3 Post \\ &+ \beta_4 T \times B + \beta_5 T \times Post + \beta_6 B \times Post \\ &+ \beta^{DDD} + \varepsilon_{sit} \end{aligned} \tag{1.8}$$

Most of the recommendations regarding inference and the issues that we will point out in the section for the difference in differences estimator carry over to the triple difference estimator.

References

Cunningham, Scott (2021), *Causal inference: The mixtape*. Yale university press.

de Chaisemartin, Clément and Xavier D'Haultfœuille (2023), "Two-way fixed effects and difference-in-differences estimators with heterogeneous treatment effects and imperfect parallel trends." *Available at SSRN*.

Olden, Andreas and Jarle Møen (2022), "The triple difference estimator." *The Econometrics Journal*, 25, 531–553.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe (2023), "What's trending in difference-in-differences? a synthesis of the recent econometrics literature." *Journal of Econometrics*.