Notes on Lab Session 6

Giulia Tani
https://tanigiulia.github.io/

TSE - MSc course in Program Evaluation
March 2026

Lee (2008): Motivation

- **Does incumbency affect the probability of winning an election?**

  Idea: elected representatives may use performance in office, visibility, resources,... to gain advantage for their party in the next elections.

- What if we regress election success on incumbency status?

  $$Y_i = \alpha + \rho D_i + \epsilon_i$$

  Problem: incumbency D_i is not randomly assigned. Winning past elections may indicate better political skills, greater popularity, etc. These unobserved factors ϵ_i are correlated with D_i, which leads to endogeneity.

- Solution: **Regression Discontinuity Design (RD)**

RD intuition

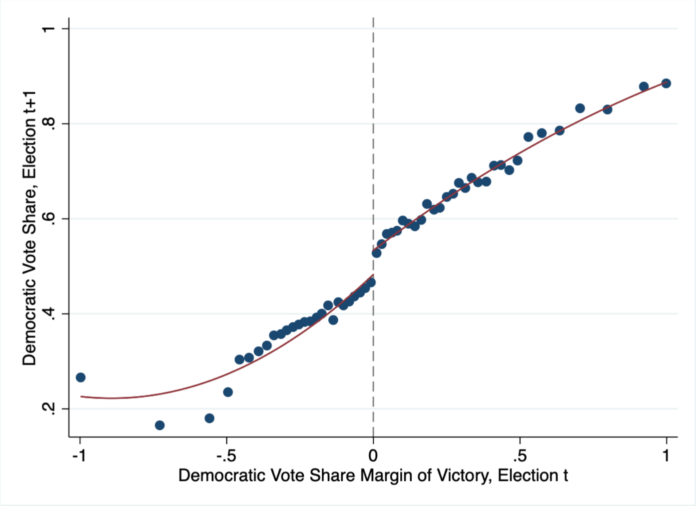
- Treatment (Democratic incumbent) is a **deterministic** and **discontinuous** function of victory in the previous election:

$$D_i = \begin{cases} 1, & \text{if } mov_i \geq 0 \\ 0, & \text{if } mov_i < 0 \end{cases}$$

where mov_i = % Democratic votes - % Republican votes in district i.

- Idea: local randomization at the cutoff.
  - » Small shifts in votes near $mov_i = 0$ are essentially random (e.g. last-minute swings).
  - » Districts where democrats barely won (mov_i just above 0) should be very similar to those were they barely lost (mov_i just below 0), **except** for the incumbency status.
  - » If there is a **sharp discontinuity** in future election success exactly at $mov_i = 0$, it suggests a causal incumbency advantage.

Binscatter plot: evidence of party incumbency advantage



RD estimation

- Control for the gap between Democratic and Republican votes in the previous election (mov_i):

$$Y_i = \alpha + \beta mov_i + \rho D_i + \epsilon_i$$

  Districts where mov_i was high (low) are more likely to lean Democratic (Republican) in the next elections. If we don't control for this, we incorrectly estimate the incumbency effect (ρ).

- Treatment D_i is a **deterministic** function of mov_i.

  $\longrightarrow$ **Assumption**: all other determinants of Y_i vary**smoothly** with mov_i at the cutoff.

- **Common support assumption is violated** by construction: there is no value of mov_i at which we get to observe both treatment and control observations.

  $\longrightarrow$ Validity hinges on our willingness to **extrapolate** at the cutoff point.

RD estimation (cont'd)

- We can be more general:

$$E(Y_i(0) \mid mov_i) = f_0(mov_i) = \alpha + \beta_{01} mov_i + \ldots + \beta_{0p} mov_i^p$$
$$E(Y_i(1) \mid mov_i) = f_1(mov_i) = \alpha + \rho + \beta_{11} mov_i + \ldots + \beta_{1p} mov_i^p$$

Then:

$$\begin{aligned} E(Y_i \mid mov_i) &= E(Y_i(0) + D_i(Y_i(1) - Y_i(0)) \mid mov_i) \\ &= f_0(mov_i) + D_i(f_1(mov_i) - f_0(mov_i)) \\ &= \alpha + \beta_{01} mov_i + \ldots + \beta_{0p} mov_i^p + \rho D_i + \beta_1^* D_i mov_i + \ldots + \beta_p^* D_i mov_i^p \end{aligned}$$

where $\beta_1^* = \beta_{11} - \beta_{01}, \ldots, \beta_p^* = \beta_{1p} - \beta_{0p}$.

- So we regress:

$$Y_i = \alpha + \beta_{01} mov_i + \ldots + \beta_{0p} mov_i^p + \rho D_i + \beta_1^* D_i mov_i + \ldots + \beta_p^* D_i mov_i^p + \epsilon_i$$

where ρ is the treatment effect at $mov_i = 0$.

RD estimation (cont'd)

- Problem: Validity of the previous model hinges on whether we correctly specified the conditional mean functions as p-th order polynomials that are continuous at the cutoff. If not, the jump at the cutoff might be due to nonlinearities we did not account for.

- Solution: Look only at data in a neighborhood around the discontinuity $(-h, h)$. We assume:

$$\lim_{h \to 0} E(Y_i \mid 0 < mov_i < h) - E(Y_i \mid -h < mov_i < 0) = E(Y_i(1) - Y_i(0) \mid mov_i = 0)$$

  Then comparisons of average outcomes in the neighborhood to the left/right of the cutoff provide an estimate of the treatment effect that is less sensitive to the specification of the conditional mean functions.

- We run a **local linear regression**:

$$Y_i = \alpha + \beta_{01} mov_i + \rho D_i + \beta_1^* D_i mov_i + \epsilon_i, \quad \text{using only } |mov_i| < h.$$

- In choosing h we face the usual trade-off between bias and efficiency.

Local linear regression