

George Johnson

Advanced Policy Research Methods

Professor Atsusaka

4 November 2024

Regression Discontinuity Design

1. The causal estimand is the effect of candidate ethnicity on coethnic contributions.
2. **Treatment assignment** here is based on if the candidate is ethnic and is in a close primary election. They use this because it is the only mechanism that allows for a **deterministic** study to find this correlation pattern and produces the best outcome to maximize for confounders. The **running variable** is coethnic contributions. The **cutoff** would be in non close primary elections, so the election has to be both close and primary.
3. As if random means that in this study the treatment can be assumed random and that the confounders are balanced among the control groups. I do not think this is justifiable because of the difference in cultures of things that don't include ethnicity and race that cannot be balanced here. These include confounders like religion, age, income.
4. $E[Y_i(1)|X_i, D_i = 1] = 1005 + 5x$
5. The single value of the treatment effect at the cutoff point 5
6. The syntax in the provided r code in the submission works because it leverages discontinuity at cutoff to identify local treatment effect. This assumes that units just above and just below are similar in all ways with respect besides the treatment effect.

This controls confounders that affect r by creating quasi random variation due to the cutoff.

Difference-in-Differences

1. The causal estimand is the effect of candidate ethnicity on coethnic contributions. (The same estimand)
2. In the provided syntax in the R code file, the syntax works and is useful because by including fixed time effects, it controls for confounders that are within each district at each year.
3. A time varying confounder that affects candidate race and candidate contributions would be housing cost, this can affect the demographic that lives in the district and also the amount of money that can be contributed.
4. Average education level would be a unit varying confounder. Unit varying since it can be assumed relatively stable over time in each district.
5. This assumption is fair due to the amount of confounders they have accounted for that would most affect the study.
6. Unobserved confounders that varies in years and districts would prevent the study from correctly estimating the treatment effect in DiD since it causes changes that can't be separated from treatment.

Missing Data

1. Litewise has the disadvantages of data loss, bias, and reducing generalizability. Litewise does have justification if the data is MCAR. This means that the missing data is random and does not depend on any variables, observed or unobserved.

2. Partial identification is a process that allows one to make an educated guess for an outcome when there is missing data. This gives a range of possible values for the average outcome based on the data we do have.
3. All done in r
4. The informativity of the bound depends on the amount of missing data, but it provides the true mean within these two bounds. So from the code the upper bound is 5.334231 and the lower bound is 0.5247263 showing that the true mean is between these two values.
5. The means are close with a difference of 0.164527