AC 7

The World Cup, Nationalism, and Aggression

The political scientist Andrew Bertoli argues in a [2017 peer reviewed article](https://academic.oup.com/isq/article/61/4/835/4657195#107206490) that nationalism (identifying with a nation-state and seeing all others as different or *other*) drives international aggression. That is, the greater the nationalist attitudes in a country, the more likely that country is to pick fights with, or engage in aggressive behavior toward, other countries.

This is an empirical question, but not one easily answered. We can’t just go around and randomly assign different countries to different levels of nationalism. Not only would it be unethical, it wouldn’t be possible. But Bertoli came up with a clever research design to make causal inferences. How you ask? By using a RD design based on how close a country’s soccer team got to qualifying for the World Cup as a *running variable* and entrance to the World Cup as the *treatment*.

Bertoli’s clever idea was that qualifying for the World Cup is a point of national pride for citizens of a given country and that making it to the World Cup would lead to a spike in nationalism. That spike would theoretically be as-if-random near the point of the discontinuity surrounding the scoring process by which teams qualify for the World Cup.

I’ve made the data from Bertoli’s study available. The code to read it into R is below. The dataset is fairly self-explanatory, but just to get you started, some key colums in the data are:

* The outcome: change\_agg, which is the change in the rate of conflict initiation in the two years post qualification relative to the two years pre qualification.
* The treatment: treatment, which is 1 if the country’s soccer team qualified for the World Cup.
* The running variable: cutpoint\_margin, which is the point margin relative to the threshold necessary to qualify for the World Cup.

Here’s what you need to read in the data:

## packages you'll need  
library(tidyverse)  
library(estimatr)  
  
## the data  
url <- "https://raw.githubusercontent.com/milesdwilliams15/Teaching/main/DPR%20201/Data/world\_cup.csv"  
data <- read\_csv(url)

For this analysis challenge, I want you to perform the following tasks and answer the posed questions along the way. You’ll do these in a .Rmd file. When you’re done, knit to word or htlm and submit to Canvas.

1. Does this research design seem plausible to you? Why or why not?
2. Visualize the regression discontinuity with geom\_smooth. Set method = lm and level = 0.83, and inside aes() map the group aesthetic to treatment status.
3. Do it again, but this time use method = "loess". Do you notice any difference?
4. Implement the RD design the way we discussed in class and interpret the results. What is your estimate of the LATE? What does it mean in substantive terms? Can you reject the null?
5. Try narrowing the bandwidth to just +/-2 points from qualifying and just calculate the simple difference in means in the outcome given treatment. What is your estimate? What does it mean in substantive terms? Can you reject the null?
6. With the same bandwidth, implement the RD design like in 3. How do the results change?
7. Does this research design still seem plausible to you? Even if you don’t buy the link between the World Cup and nationalism, do you think nationalism can cause international aggression?
8. There are other covariates in the data. Is treatment correlated with any other factors that might influence aggression?