Co-ethnic Candidates and Voter Turnout

Autumn 2019

For this practical, we will analyze data from the following article:

Fraga, Bernard. (2015) "Candidates or Districts? Reevaluating the Role of Race in Voter Turnout," *American Journal of Political Science*, Vol. 60, No. 1, pp. 97–122.

Fraga assesses the theory that minority voters are more likely to vote in elections featuring co-ethnic candidates (i.e., candidates that share the same race of minority voters).

A description of the variables is listed below:

Name	Description
`year`	Year the election was held
`state`	State in which the election was held
`district`	District in which the election was held (unique within state but not across states)
`turnout`	The proportion of the black voting-age population in a district that votes in the general election
`CVAP`	The proportion of a district's voting-age population that is black
`candidate`	Binary variable coded "1" when the election includes a black candidate; "0" when the election does not include a black candidate

Question 1

Fraga analyzes turnout data for four different racial and ethnic groups, but for this analysis we will focus on the data for black voters. Load the `blackturnout.csv` dataset into R. Which years are included in the dataset? How many different states are included in the dataset?

Question 2

Create a boxplot using the `boxplot()` function that compares turnout in elections with and without a co-ethnic candidate. Be sure to use informative labels. Interpret your figure.

Question 3

Now let's conduct a linear regression with black turnout as your outcome variable and candidate co-ethnicity as your predictor. Report the coefficient on your predictor and the intercept. Interpret the coefficients from your regression model. Based on these coefficients, what would you conclude about blacks voter turnout and co-ethnic candidates?

Question 4

Now let's say you decide to investigate the results of the previous question a bit more carefully because the elections with co-ethnic candidates may differ from the elections without co-ethnic candidates in other ways. Create a scatter plot using the `plot()` function where the x-axis is the proportion of co-ethnic voting-age population and the y-axis is black voter turnout. It would be helpful to colour your points according to candidate co-ethnicity. That is, make the points for elections featuring co-ethnic candidates one color, and

make the points for elections featuring no co-ethnic candidates a different color (Hint: you can do this by using the `ifelse()` function to create a colour variable based upon the values of the `candidate` variable: `df$col <- ifelse(df$candidate == 1, "red", "blue")`, assuming that your dataframe is called `df`). Interpret the graph.

Question 5

Conduct a linear regression with black turnout as your outcome variable and with candidate co-ethnicity and co-ethnic voting-age population as your predictors. Interpret the coefficients on the two predictors, ignoring the intercept for now (you will interpret the intercept in the next question). Explain what each coefficient represents in terms of the units of the relevant variables.

Question 6

Now interpret the intercept from the regression model with two predictors. Is this intercept a substantively important or interesting quantity? Why or why not?