

## Co-ethnic Candidates and Voter Turnout

For this problem set, 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. He shows that the association between minority voter turnout and co-ethnic candidates disappears once we take into account district-level racial composition. In particular, he demonstrates that in districts where blacks make up a greater share of the voting-age population, blacks in that district are more likely to vote in elections *regardless* of candidate race. Although the main analyses he carries out are a bit more complicated than what we cover in *QSS*, we can replicate his substantive findings using the multiple regression approach we’ve learned.

A description of the variables is listed below:

Name	Description
<code>year</code>	Year the election was held
<code>state</code>	State in which the election was held
<code>district</code>	District in which the election was held (unique within state but not across states)
<code>turnout</code>	The proportion of the black voting-age population in a district that votes in the general election
<code>CVAP</code>	The proportion of a district’s voting-age population that is black
<code>candidate</code>	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 `blackturnout.csv`. Which years are included in the dataset? How many different states are included in the dataset?

### Question 2

Create a boxplot that compares turnout in elections with and without a co-ethnic candidate. Be sure to use informative labels. Interpret the resulting graph.

### Question 3

Run 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 these coefficients. Do not merely comment on the direction of the association (i.e., whether the slope is positive or negative). Explain what the value of the coefficients mean in terms of the units in which each variable is measured. Based on these coefficients, what would you conclude about blacks voter turnout and co-ethnic candidates?

#### Question 4

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 where the x-axis is the proportion of co-ethnic voting-age population and the y-axis is black voter turnout. Color 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. Interpret the graph.

#### Question 5

Run a linear regression with black turnout as your outcome variable and with candidate co-ethnicity and co-ethnic voting-age population as your predictors. Report the coefficients, including the intercept. 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?

#### Question 7

Based on the regression model with one predictor, what do you conclude about the relationship between co-ethnic candidates and black voter turnout? Based on the regression model with two predictors, what do you conclude about the relationship between co-ethnic candidates and black voter turnout? Please ignore issues of statistical significance for this question given that it will be covered in Chapter 7 of *QSS*.

#### Question 8

Repeat the analysis in the Question 5, but fit the linear regression by year. Which year has a positive coefficient of candidate co-ethnicity? Use `for loop` to check each year. In the loop, use `if()` to check the coefficient and print out the year if the coefficient of candidate co-ethnicity (`candidate`) is positive.