

Pol Sci 630: Problem Set 7 - Dummy Variables and Interactions (II) and Short Research Outline

Prepared by: Jan Vogler (jan.vogler@duke.edu)

Due Date: Wednesday, October 19th, 2016, 1.25 PM (Beginning of Class)

Note 1: It is absolutely essential that you show all your work, including intermediary steps, in your (mathematical) calculations and that you comment on your R code to earn full credit (you can comment on your R code both with the use of `#` in the R code and in the \LaTeX code). Showing all steps and commenting on code will also be required in future problem sets.

Note 2: Please submit a PDF file created through knitr containing all your answers to the problem set. knitr allows you to combine R code and \LaTeX code in one document, meaning that you can include both the answers to R programming and math problems. Also submit the source code that generates the PDF file (i.e. the .Rnw file).

Note 3: Make sure that the PDF files you submit do not include any references to your identity. The grading will happen anonymously. You can submit your answer at the following website: <http://ps630-f15.herokuapp.com/>

R Programming

Problem 1 (4 points)

Do the following in R:

a) Load the *VOTE1* dataset that you used in problem set 5. Estimate a linear model with the same variables that you used in problem 2 of problem set 5 (with incumbent vote share as the dependent variable) and add a fixed effect for state. Display the regression results in R.

b) Then, please describe the consequences of using a fixed effect for state. Are there any differences in terms of the direction or significance of the key variables? What is different between a regular regression and one that uses fixed effects, specifically in this case?

Problem 2 (8 points)

a) This assignment is based on the regression you ran in problem 1.

When explaining the vote share that the incumbent party has received, there could be an interaction between the strength of the incumbent party (*prtystrA*) and its expenditures (*expendA*). We might expect that there is a positive feedback between both as parties with a stronger organization can potentially use the money more effectively. (Note: This expectation might or might not be true.) Please run a regression that includes an interaction term of both variables. Then display the regression results.

Note: Please **do not** include state-fixed effects in this regression (other than in problem 1).

b) Please make statements about all coefficients in your model, including the interaction term. Interpret the direction and magnitude of the marginal effects as well as the level of significance of the relationship of all coefficients. Make sure that you explicitly answer the question of how the marginal effect of expenditures varies with different levels of party strength.

c) Please graphically show the interaction between party strength and expenditures of the incumbent party with the following plots:

1. Create a marginal effects plot with confidence intervals of the influence of expenditures on vote share at different levels of party strength.
2. Create a marginal effects plot with confidence intervals of the influence of party strength on vote share at different levels of expenditures.

3. Create a coefficient plot in which you show the coefficients and associated confidence intervals of expenditures, party strength, and their interactions.
 4. Create a predicted values plot with confidence intervals of the influence of party strength on vote share at different levels of expenditures. Choose the 5th and 95th percentile of expenditures to show two different predicted value lines.
- d) Please answer the following questions:
1. Does a lower level of party strength have a positive or negative impact on the marginal effect that expenditures have on vote share?
 2. Does a higher level of expenditures have a positive or negative impact on the marginal effect that party strength has on vote share?

Interactions: Math and Interpretation

Problem 3 (4 points)

Assume that you have the following linear model

$$Y = 10 + (5) * X_1 + (2) * X_2 + (1) * X_1 * X_2 + u$$

- a) Calculate the derivative of Y with respect to X_1 and X_2 .
- b) Use R to plot the marginal effect of X_1 at different levels of X_2 . Assume that X_2 is an integer that varies between -10 and 10.
- c) Please interpret the marginal effect of X_1 based on both the derivative and the plot that you generated.
- d) If there is an interaction between X_1 and X_2 in reality, what would be the potential consequences of omitting this interaction from our model? Explain carefully (no need to show it mathematically).

Short Research Outline

“Problem 4” (ungraded)

Please write a short research outline with a minimum length of 3 paragraphs and a maximum length of 2 pages. One paragraph on each question is sufficient. Please upload your answers as a separate files to the Sakai Drop Box of the course.

a) Write one paragraph on the background of your research idea. What is the question you want to ask? Why is it relevant? Which academic works are most closely related to the specific question you want to answer?

b) Write one paragraph on how the theoretical concepts you are interested in can be measured as variables. What is your unit of analysis? What is your dependent variable? What is your main independent variable? Which control variables would you like to include and why?

c) Write one paragraph on which datasets would be appropriate to use. Which data sets contain the data you need? How many observations are available over which time period? How difficult would it be to obtain the data?