

POL 681 Problem Set 5: Statistical Interactions and Graphs

Objective: The objective of this problem set is to give you experience in working with interaction terms in a linear regression model. You will be expected to correctly estimate the parameters of interest and correctly interpret them (including interpreting the uncertainty about the parameter estimates.)

Directions: Please use the dataset `hispanic.dta` to address the following questions. Please turn in `Stata` or `R` log files and all graphs. Your answers should be wordprocessed and you should interpret your results using substantive language. Do not use the variable name stubs in your interpretation (i.e. don't refer to the dependent variable as "HISPTHR").

In the dataset `hispanic.dta`, there are four five variables of interest. The first, denoted in the data as `hispthrm` is the NES 101-point "feeling thermometer" question for Hispanics. The scale ranges from 0, denoting least favorably evaluated, to 100, denoting most favorably evaluated. For this problem set, this variable will be the dependent variable.

The variable denoted as `immthrm` is an NES 101-point feeling thermometer question for immigrants. The scale ranges from 0, denoting least favorably evaluated, to 100, denoting most favorably evaluated.

The variable denoted as `prej` is a scale derived from the NES measuring prejudicial attitudes. The scale ranges from 7 (denoting most prejudicial attitudes) to 30 (denoting least prejudicial attitudes).

The variable denoted as `hispa` denotes the percentage of Hispanics residing in the congressional district in which the survey respondent lives (data are from U.S. Census Bureau).

The variable `white` is a dummy variable denoting whether or not the respondent is white and non-Hispanic. A zero denotes the respondent is either Hispanic, African-American, Asian, Native American, or is in some other non-white racial or ethnic group. Hence, a 0 implies the respondent is from a racial or ethnic minority group.

A researcher is interested in the following questions. Please answer them by providing analysis of the data.

1. Do white attitudes toward Hispanics significantly differ from minority respondent attitudes toward Hispanics? Why or not? (8 points)
2. Does there seem to be a connection between attitudes toward immigrants and attitudes toward Hispanics after controlling for race/ethnicity of the survey

respondent? Why or why not? (8 points)

3. Provide a graphical summary of the relationship found in question 2. How is this figure interpreted? How does this graph illustrate the additivity property? (6 points)

4. Suppose that instead of additivity, the researcher conjectured that the impact of the evaluation of immigrants was conditional on whether or not the respondent was white (non-Hispanic) or non-white (or white Hispanic). Construct and estimate a model that tests this proposition. How is this model substantively interpreted? (10 points)

5. Provide a graphical summary of the relationship found in question 4. How is this figure interpreted? How does this graph illustrate *non*-additivity? (6 points)

For the next set of questions, omit the `white` and `immthrm` variables from your analysis.

6. Suppose that our researcher believes that one's level of prejudicial attitudes is multiplicatively related to the number of hispanics residing in close proximity to the respondent? The theory leading to this expectation is that as one is surrounded by more members of the minority group, negative stereotypes may be counteracted due to exposure. Estimate and interpret a model that tests this proposition. How do you substantively interpret this model? Provide me with any graphs or tables of predicted values you think you need in order to interpret this model. (20 points)

7. Because the slope estimates are conditional in this model, so are the standard errors. Are these slope estimates statistically different from 0 across the full range of the interaction term? Provide concise evidence to answer this question. (15 points)

(Note that concise evidence could come in the form of computing the *t*-ratios for the slope of `hispthrm` on `prej` conditional on `hispa`; compute the *t*-ratios for the slope of `hispthrm` on `hispa` conditional on `prej` and then graphing them. Also, graphing the 95 percent confidence intervals might be useful. How you choose to report your evidence is up to you; just make sure I can follow it!)