Problem Set 3

POLS 8810

Due 13 March 2025

This problem set will require the same ANES data file (and codebook) available used last week. In case you did not save the link, it is available here.

In Problem Set 2, you looked at the relationship between the income of respondents and their feelings towards the Republican Party. Imagine that after this, you were talking about it to a colleague. That colleague was surprised to learn that your findings were so weak and suggested that perhaps, the relationship between income and Republican Party support is not direct but conditional. Specifically, the colleague suggests that perhaps the relationship is actually much stronger *but only for men.*

You decide that your colleague's suggestion may have merit and formulated the following hypothesis based on their comment:

H_1**:** Among men only, as income increases, support for the Republican Party should also increase

Your goal for this assignment is to test this hypothesis by estimating a linear regression model examining the relationship between respondent income and support for the Republican Party conditional on gender. In addition to clearly and correctly specifying a *single* linear regression model that accounts for the conditional nature of H_1,[1] you will need to present the results from your model in a clear way using table(s) and/or graph(s) as appropriate, and correctly interpret those results in a substantively meaningful way.

Prior to conducting any analysis, it is a good idea to get a sense of what your data looks like. There are a variety of ways to do this using descriptive statistics and/or graphs to look at summaries of individual variables and relationships between variables. Doing this at the beginning of any analysis can save time and trouble later and help to ensure reliability of your estimates. *While you should always do this as a standard step in the process any time*

[1] While subsetting the data and estimating separate models for only male respondents and only female respondents would be a valid approach to testing this hypothesis, that is NOT a valid approach for this assignment.

you are working with data, is not necessary to include any of this in your write up (unless you feel that something you found at this stage is relevant to the primary analysis/analyses).

For the primary analysis, you will need to estimate an appropriately specified regression model that takes into account any relevant information learned from your preliminary examinations of the data you conducted. You will also need to ensure that all Gauss-Markov assumptions are met. If some of the Gauss-Markov assumptions are not met, you will need to discuss why this is not a problem and/or how you are dealing with the problem(s).

Your write-up should contain any information the reader would need to replicate your data manipulations, recoding of variables, variable transformations, model estimation, etc. However, please pay close attention to how you write this up. Your goal is to emulate the way that the data, methods, and results sections of a publishable research paper are written. The handouts I have provide should help with this, but the best way to learn to write professionally is to read things that have been published and emulate their style. (*Note that repetition of the errors in how we discuss our variables and their coding highlighted as Problem 1 in the "Problem Set 1: Most Common Problems" handout (i.e. the use of variables coded in a way that is not suitable for a regression analysis, will negatively impact your grade on this assignment.*)

The primary grade will be based on an evaluation of the student's understanding of how to appropriately estimate a linear regression model to test a conditional relationship and how to clearly and correctly present and interpreted the results of that model. However, quality/clarity of writing, quality of tables and figures, demonstrated understanding and description of the variables used, and other considerations matter as well. The write up should be approximately 5 pages, with 7 pages as an absolute maximum. **Assignments that fail to follow these direction or that cut and paste raw output from the software will receive a grade of 0 for the assignment.**

Finally, while I *strongly* encourage you to work together to develop an understanding of data, coding, model estimation and interpretation, and the collaborative nature of research generally, the final write up must be your own. Moreover, it must be typed and written in the manner of the data, methods, and results sections of an article manuscript complete with any and all professional looking tables and graphs appropriate for the assignment.

Your Problem Set 3 write up is to be submitted *both* via iCollege under /Assessments/ Assignments/Problem Set 3 (in order to allow for automated plagiarism/AI detection) and in physical hard copy following the instructions on the syllabus (for grading). Additionally, you should prepare a replication file with the code you use that should be submitted via email at the same time the problem set is submitted. If the replication code does not fully

replicate all results presented in the write-up, you will be notified and your problem set will not be graded until this is corrected. If this is not resolved within 48 hours of your receiving that notification, you will receive a grade of 0 for the problem set.