

Quantitative and Formal Methods in IR Assignment 3

This assignment is due at the beginning of lecture (Thursday) the seventh week of Hilary Term (05/03/2009). You will hand me a printed copy in class and e-mail the assignment to me (david.armstrong@politics.ox.ac.uk) *before* the Thursday lecture when the assignment is due. The assignments will be graded on the usual scale where 60 is considered passing. For those from outside the Oxford system who have yet to be formally assessed here, scores ranging from 60 to 70 are considered perfectly acceptable. On rare cases, scores of above 70 will be awarded for exceptional work. Scores less than 60 are considered failing. You will be allowed to resubmit any failed assignments by the end of the term as per the course syllabus. Changes to the deadline for assignments will be entertained on a case-by-case basis in the presence of exceptional circumstances when I am contacted *before* the deadline. Assignments not turned in by the deadline will receive a score of zero.

Please type your responses to the questions. I leave the formatting of the responses up to you, but it would be to your benefit to make it easy for me to navigate through the document. This could include, clearly marking sections to which the answers apply and putting all copied Stata output in a fixed-width font (I recommend 8 or 9pt Courier New). I do not require you to distill the Stata output into tables, though if you want, you are certainly welcomed to do so. You may work in groups to generate the results if you wish, but the write-up and interpretation of the results should be your own.

The questions below will require you to use the `assignment3.dta` dataset that can be downloaded from the course website: <http://www.quantoid.net/IRStats.PHP>. Notes pertaining to the dataset and individual variables can be seen by typing `notes` in the Stata command window and hitting “enter.” Below is a description of the variables in the dataset:

abbrev State abbreviation

ccode Correlates of War numeric state code.

cname Country name

new_civlibs Civil Liberties variable created from Freedom House sub-indicator scores.
Higher values mean more respect of civil liberties.

gdppc GDP/capita from Kristian Gleditsch’s dataset

civ_small Civilizations code from Henderson and Tucker with small civilizations coded into “other”.

Questions

1. Run a simple linear regression with `new_civlibs` as the dependent variable and `gdppc` as the independent variable. Interpret the slope and intercept and tell me whether you would reject the null that slope is equal to zero and the null that the intercept is equal to zero when considering the two-sided alternatives for both. From the R^2 , do you think this is a good model?
2. Assess the model for problems with linearity. If you find some, run a new model that accounts for the non-linearity.
3. Assess this new model for problems with the residuals. What do you see?
4. Include in the model, all dummy variables representing the civilization codes. Explain how to interpret the coefficients on the dummy variables.
5. Look at the residuals again. Are they “better” than before? Has accounting for the civilization codes made a difference here? If so, what kind of difference?
6. Include interaction effects between each of the civilization dummy variables and the variable you’re now using to represent GDP/capita (i.e., perhaps some transformation of the original variable if you found non-linearity above).
 - (a) Examine the residuals in a residuals versus fitted plot. What do you see here? Do there appear to be any problems? Propose and execute a solution to deal with problems that exist.
 - (b) Test the null hypotheses that each conditional coefficient for your GDP variable (conditional on the civilization codes) is equal to zero against the two-sided alternative. What do you find?
7. Given the results found in the last section, can you offer a substantive explanation for why we might find something like this?