

Advanced Quantitative Methods

Homework 4: limited dependent variables & simulation

Johan A. Elkind
jos.elkind@ucd.ie

Due May 14, 2018, 5pm

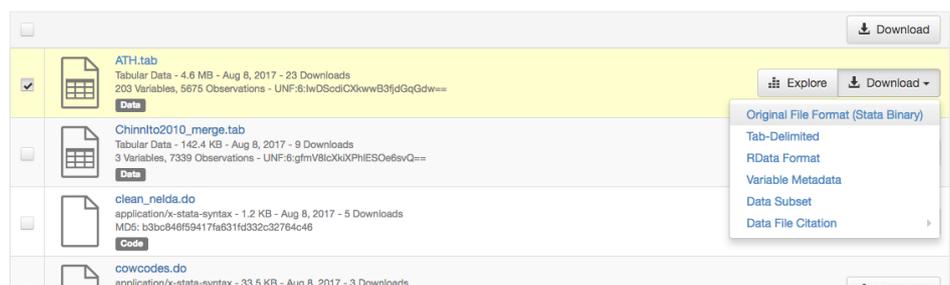
Please submit by email in PDF format. Add R code in a separate .R file, or SPSS code in a separate .sps file, or Stata code in a separate .do file, or the code for any other package you use separately. Alternatively, you can integrate R code and responses in R Markdown and submit both in PDF format (easiest is to use HTML format and then in the web browser save as PDF).

Please note the revised submission date.

Percentages with an asterisk indicate that positive rather than negative marking will be applied.

Data

This homework is based on the replication data for ?, which you can access at their Harvard Dataverse record—probably the first study you find when searching for the keyword “remittance”.¹



The “explore” button might also be very helpful to get a feel for the data.

Questions

We will be investigating the relationship between democracy in neighbouring countries and democracy within a country itself, using the Polity IV democracy score, which runs from -10 for autocracies to +10 for democracies. We will be using some of the variables generated in Homework 3.

¹<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/NFMGFD>

1. Create a variable **democracy** that is one if the Polity IV score (**pol4_polity2**) is greater than or equal to 6 and zero otherwise. Do this by changing the code from Homework 3, so that this precedes the self-merging.
2. (5%) Estimate a linear regression where the dependent variable is **democracy** and the independent variables are the neighbours dummy variable (**neighbor_pol4_polity2.lag**), the lagged dependent variable (**democracy.lag**) and control variables for oil producing countries (**oilproducing.lag**), log of GDP per capita (**mad_lgdppc.lag**), and log of population size (**mad_lpop.lag**).
3. (10%) Estimate the same model using a logistic regression.
4. (15%) Based on the logistic regression, calculate the predicted probability of observing a democracy for oil producing and other countries, keeping other variables at a reasonable value.
5. (10%) Using a bootstrap, calculate a 95% confidence interval for these two proportions.
6. (10%) Estimate the same model using a probit regression.
7. (10%) Based on the probit regression, plot the probability of observing a democracy by average level of democracy in neighbours, separately (two lines in one plot) for oil producing and for other states.
8. (10%) Using simulated parameters, add 95% confidence intervals to the plot.
9. (10%) Produce a regression table with all three regressions.
10. (20*%) Fully interpret the results (approximately 400 words), including model fit, effect sizes, interpretation of statistical tests, and implications for our understanding of international dimensions to democratization.

Grade conversion scheme

Score	Grade		Score	Grade		Score	Grade		Score	Grade	
	UCD	TCD		UCD	TCD		UCD	TCD		UCD	TCD
97-100%	A+	A+	85-87%	B	B	74-76%	C-	C	54-64%	E+	D
94-96%	A	A	83-84%	B-	B	71-73%	D+	C	44-53%	E	D
91-93%	A-	A	80-82%	C+	C+	68-70%	D	C	33-43%	E-	D
88-90%	B+	B+	77-79%	C	C	65-67%	D-	C	0-32%	F	F