

Midterm: OLS results

Ida Hjermitslev

10 Mar 2015

The datasets that will be used for this project are in the attached .rda file. This .rda file contains two objects, data and dataMiss. The dependent variable for the models you will be running is income inequality at the country level (gini_net_std, higher values indicate greater income inequality). The independent variables are the level of ethnic fractionalization (ELF_ethnic, higher values indicate greater levels of ethnic fractionalization) and a country's democracy score (polity2, higher levels indicate more democratic regimes).

% latex table generated in R 3.1.2 by xtable 1.7-4 package % Wed Mar 11 11:19:57 2015

OLS model on complete dataset						
	Estimate	Std. Error	T-Statistic	P-Value	Lower 95% CI	Upper 95% CI
(Intercept)	-0.22	0.85	-0.26	0.80	-0.47	0.03
Ethnic Fractionalization	2.01	0.63	3.18	0.00	1.82	2.20
Democracy	-0.10	0.08	-1.20	0.23	-0.12	-0.07
F-statistic: 0.09 on 1 and 43 DF, p-value: 0.914.						
OLS model with listwise deletion						
	Estimate	Std. Error	T-Statistic	P-Value	Lower 95% CI	Upper 95% CI
(Intercept)	0.67	0.57	1.18	0.25	0.49	0.85
Ethnic Fractionalization	0.81	0.44	1.82	0.08	0.66	0.95
Democracy	-0.16	0.05	-3.16	0.00	-0.18	-0.15
F-statistic: 0.048 on 1 and 37 DF, p-value: 0.953.						
OLS model with imputed data						
	Estimate	Std. Error	T-Statistic	P-Value	Lower 95% CI	Upper 95% CI
(Intercept)	1.94	0.46	4.22	0.00	1.80	2.07
Ethnic Fractionalization	0.33	0.45	0.74	0.46	0.20	0.47
Democracy	-0.29	0.04	-7.49	0.00	-0.30	-0.28
F-statistic: 0.02 on 1 and 43 DF, p-value: 0.98.						

The last part of the midterm will simply involve a discussion of how the coefficient estimates differ between model, modelListDel, and modelAmelia. As can be seen in the table, the coefficient estimates differ considerably.

For all of the models, a greater level of ethnic fractionalization is associated with higher income inequality and the relationship is statistically significant at the 95% confidence level. However, the magnitude of the effect differs widely. If coefficients are computed on the full data set the income inequality will increase by two units for every unit increase in ethnic fractionalization. The effect is much smaller with listwise deletion, and even smaller with the imputed data.

For all of the models, a higher democracy score is associated with lower income inequality and the relationship is statistically significant at the 95% confidence level. If coefficients are computed on the full data set, the income inequality will decrease by one unit for every ten unit change in democracy. The effect is stronger/more negative for the listwise deletion, and strongest for the imputed data.