

Advanced Quantitative Methods

Homework 2: Diagnostics & Time-Series

Johan A. Elkink
jos.elkink@ucd.ie

Due March 8, 2017, 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. Note that if you do not use Latex, there is an “Equation” entry in Microsoft Word under “Insert” that will allow you to include matrices and other mathematical equations.

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

(5*%) of the grade is used for an overall evaluation of the clarity and presentation of your code.

Data

This homework is based on the replication data for Ross and Voeten (2016), which you can access at their Harvard Dataverse record.¹

The screenshot displays a list of data files from the Harvard Dataverse. The file 'ReplicationdataRossVoeten.tab' is highlighted, and its dropdown menu is open, showing various download options. The 'Original File Format (Stata 13 Binary)' option is circled in red.

The “explore” button might also be very helpful to get a feel for the data. It is of course advisable to have a quick glance at their paper as well.

¹<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/BON24T>

Questions

1. Based on Ross and Voeten (2016), we will be investigating the relationship between oil exports and membership of “structured” international organisations using the following model:

$$strucint_i = \beta_1 + \beta_2 oilexp_i + \beta_3 democracy_i + \beta_4 lngdp_i + \beta_5 lnpop_i + \varepsilon_i$$

- (a) Select only cases for the year 2001.
 - (b) Perform the above regression.
 - (c) (5%) Use residual plots against fitted values and against **oilexp** to investigate potential heteroscedasticity.²
 - (d) (5%) Perform a Breusch-Pagan test for heteroscedasticity in the errors. What do you conclude?
 - (e) (5%) Calculate “HC3” robust standard errors and recalculate the *t*- and *p*-values using these. Do the results of the regression change?
 - (f) (5%) Calculate Variance Inflation Factors. What do you conclude?
2. We will continue with the same model, but starting with the original, full data set.

- (a) Select only the Indonesian cases (**ctryname** is “Indonesia”).
- (b) Re-estimate the model using this data set.
- (c) (5%) Perform a Breusch-Godfrey test for autocorrelation in the residuals. What do you conclude?
- (d) (5%) Add a lagged dependent variable to the model and re-estimate.

```
data <- data[order(data$ctryname, data$Year), ]  
T <- dim(data)[1]  
data$Ldemocracy <- c(NA, data$democracy[-T])
```
- (e) (5%) Add the lagged version of the **oilexp** variable to the model and re-estimate (keep the lagged dependent variable), so the model is now:

$$strucint_t = \beta_1 + \beta_2 oilexp_t + \beta_3 democracy_t + \beta_4 lngdp_t + \beta_5 lnpop_t + \beta_6 strucint_{t-1} + \beta_7 oilexp_{t-1} + \varepsilon_i$$

- (f) (10%) Calculate the short-term and long-term impact of an increase in oil exports on membership of a “structured” IGO.
- (g) (5%) Work out the what this means if oil exports increases by 0.22 (i.e. by one standard deviation).
- (h) (5%) Perform an Adjusted Dickey-Fuller test on the residuals of the regression. What do you conclude?

²It will help to use `na.action = na.exclude` as an option to the `lm()` command, so that the `residuals()` command will return a vector of the same length as the original data vectors. Alas, if you want to use the `sandwich` package to calculate the “HC3” standard errors, you need to estimate the model without the `na.action` parameter as well.

- (i) (5%) Calculate first differences for all variables and re-estimate the model using these first differenced variables (without the lagged variables added earlier).
`data$Ddemocracy <- data$democracy - data$Ldemocracy`
- (j) (5%) Perform a Durbin-Watson test for autocorrelation in the residuals of the first difference model. What do you conclude?
3. This question is about bringing the above together and interpreting the results fully. Interpret both statistical results and their relation to political science / international relations generally, i.e. what can we substantively conclude?
- (a) (10*%) Produce a publishable regression table, with one column for each of the above six regressions.
- (b) (20*%) Write a one page report on your findings, discussing this study of the relation between oil exports and membership of “structured” IGOs, including a clear interpretation of the short- and long-term impacts.

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

Ross, Michael L. and Erik Voeten. 2016. “Oil and international cooperation.” *International Studies Quarterly* 60(1):85–97.

URL: <https://www.sscnet.ucla.edu/polisci/faculty/ross/publications.html>