

## Exercise Sheet 8

This exercise sheet is to be done prior to your actual Support and Feedback class. The Stata/R file, which you write as a group, should then be submitted by one of your group members via e-mail to your class tutor by Monday 10am of Week 10 (Term 1). You must take the output generated from your Stata/R programme to your tutorial classes.

Use the dataset **condra\_long\_shaver\_wright\_2018.dta/RData**, which is taken from the replication files of “The Logic of Electoral Insurgent Violence” by Condra, Long, Shaver, and Wright, *American Economic Review*, 2018. The study looked at the issue of electoral violence in unstable democracies. Using data from Afghanistan, they examined the extent to which electoral violence discouraged voting at a presidential election.

1. Using OLS, estimate the model:

$$\begin{aligned} total\_v2\_agcho10_i = & \alpha + \beta_1 df\_5to11_i + \beta_2 first_i + \beta_3 temp\_00Z_i + \beta_4 temp\_06Z_i \\ & + \beta_5 rain\_00Z_i + \beta_6 rain\_06Z_i + \varepsilon_i \end{aligned} \quad (1)$$

where *total\_v2\_agcho10* is the voter turnout (%), *df\_5to11* is the number of fire attacks between 5AM and 11AM on the day of election, *first* is the indicator of the first-round election (the presidential election went to a second-round runoff), *temp\_00Z*, *temp\_06Z* are the temperature measures on the day of election at different times (in Kelvin), and *rain\_00Z*, *rain\_06Z* are the amount of rainfall on the day of election at different times (in millimeters). To see variable labels, type “codebook, compact” for Stata, and type “lapply(condra.etal, attr, which=“label”)” for R.

- (a) Save the residuals and fitted values from equation (1) and undertake the RESET(4) test by running a suitably modified version of a regression of residuals on squared, cubed and quartic fitted values.
  - (b) Repeat the test in (a) above, but using *total\_v2\_agcho10* as the dependent variable in place of the residuals.
  - (c) Run a regression to undertake White’s heteroscedasticity test of squared residuals from equation (1) against the level and square of each of the explanatory variables in equation (1). Test the joint significance of all explanatory variables.
  - (d) Compute heteroscedasticity-robust standard errors for (1). Are these robust standard errors different from the OLS standard errors? For R, install “sandwich” package and use “vcovHC” function.
2. Read the introduction (from the begging to right before Section I) and Section III.C (excluding the subsection “IED Deployment during the Preelection Period”) of “The Logic of Electoral Insurgent Violence” by Condra, Long, Shaver, and Wright.

- (a) Now we consider the following estimation equation:

$$\begin{aligned} total\_v2\_agcho10_i = & \gamma_0 + \gamma_1 df\_5to11_i + \gamma_2 windspeed\_06Z_i + \gamma_3 windspeed\_12Z_i \\ & + \gamma_4 windspeed\_06Z_i^2 + \gamma_5 windspeed\_12Z_i^2 + \gamma_6 temp\_00Z_i \\ & + \gamma_7 temp\_06Z_i + \gamma_8 temp\_12Z_i + \gamma_9 rain\_00Z_i + \gamma_{10} rain\_06Z_i \\ & + \gamma_{11} rain\_12Z_i + \gamma_{12} temp\_00Z_i^2 + \gamma_{13} temp\_06Z_i^2 + \gamma_{14} temp\_12Z_i^2 \\ & + \gamma_{15} rain\_00Z_i^2 + \gamma_{16} rain\_06Z_i^2 + \gamma_{17} rain\_12Z_i^2 \\ & + \gamma_{18} population\_2010\_adj_i + \epsilon_i \end{aligned}$$

where *windspeed\_06Z*, *windspeed\_12Z* are wind conditions on the election day at 10:30AM and 4:30PM, *population\_2010\_adj* denotes the district's population, and superscript 2 indicates square of the variable.

Estimate the above equation by two-stage least squares using *plus\_wind\_00Z\_10* as an instrument for *df\_5to11*. Here *plus\_wind\_00Z\_10* is the wind condition on the day of election at 4:30AM.

- (b) Estimate the equation in (a) by OLS (we do not use *plus\_wind\_00Z\_10* for this OLS estimation).
- (c) Using the instrument in (a), test for the endogeneity of *df\_5to11* using the Hausman-Wu test discussed in lectures.

**NOTE:**

Before you e-mail your Stata/R file to your class tutor:

- (i) Check the file runs all the way through from start to end without intervention from the marker.
- (ii) Annotate your Stata/R file with the appropriate question number to which your answer relates to help the marker understand what part of the DO file is answering which question.

All exercise sheets will be marked out of 3 points. Failure to follow the instructions noted above will involve you losing marks.