Stata Intensive: Workshop 3

Fall 2016, D-LAB

Isabelle Cohen

Oct 26, 2016

PRACTICE EXERCISES

These practice exercises are based off the Stata Test given in 2014 to job applicants at Evidence for Policy Design at the Harvard Kennedy School.

1. Importing Data

Import the file "RE - Test Data.xls" into Stata

2. Reshaping

Reshape the test score data so that each student's data is only one row

3. Labeling Variables

Label the newly created variables for math scores

4. Merging

Merge this data with "RE - Student Info.dta"

5. Data Cleaning with Loops

a. Create a variable showing the difference between the normalized post-test

and normalized pre-test scores for each student and label as appropriate.

b. Create a dummy variable for each value of Third.

Name each as third\_(x), where x represents the third of the distribution

into which the student falls, and label the variable accordingly. Use a loop.

6. Labeling Values

Label the values of the treatment variable. 0 represents the Control Group, and 1 the Treatment Group.

7. Regression Analysis and Outputting Results

Regress the difference between the normalized post-test and pre-test scores on treatment for the entire sample, then separately for each third of the initial distribution. Control for the normalized pre-test scores and cluster the standard errors at the school grade level.

Please output the results from all regressions in Excel in the clearest form possible. Include the number of observations.

CHALLENGE: Use only 4 lines of code (not counting \\\ separations)