Assignment 5: Non-Constant Variance

PSCI 8357, Spring 2016
February 18, 2016

This assignment must be turned in by the start of class on **Thursday, February 25**. You must follow the instructions for submitting an assignment.

Main Task

This week's assignment is an extension of last week's. You will once again use the Neumayer (2005) data; see last week's handout for details. You will begin by fitting the same model you ended up selecting in last week's assignment, this time on the full dataset instead of just half of it. From there:

1. **Before** running any tests, think about your model. Do you think the errors are heteroskedastic? If so, what variables do you think are associated with the variance of the error term?

2. Conduct the following tests for heteroskedasticity:

    1. A Breusch-Pagan test using the variables in your model.
    2. A White test using the variables in your model.
    3. A Breush-Pagan test using all the independent variables in the dataset. (Ignore any variable not listed on last week's handout.)

    Interpret the results of the tests. If you reject the null hypothesis of homoskedasticity, what seems to be the source of the heteroskedasticity? Does it line up with your expectations?

3. Estimate Huber-White standard errors for your model. Re-run the composite test(s) you performed last week with the new estimated variance matrix. Do your main conclusions change?

4. Based on your findings from step 2, pick a reasonable set of weights ω and obtain the corresponding weighted least squares estimates. How do the WLS coefficient estimates compare to the OLS estimates? Re-run

your composite test(s) using the WLS estimates. Once again, do your main conclusions change?

5. Think about what you just did in step 4. You ran one model (the original OLS), ran some tests on it (the Breusch-Pagan and White tests), used the results of those tests to set up a new model (the WLS), and then finally reported some test results from it (of the relevant composite hypothesis). Is this p-hacking? Can we trust the p-value you reported alongside your final composite hypothesis test? Why or why not?

   For the purposes of this question, ignore the fact that the variable specification itself was "hacked" in last week's assignment—imagine the specification was handed down from on high. I only want you to focus on the dimensions related to the testing and modeling of heteroskedasticity.

6. Which of your sets of estimates do you trust most: OLS with ordinary standard errors, OLS with Huber-White standard errors, or WLS? Why?

Weekly Visualization Challenge

This is also an extension of last week's task. Compare the estimated relationship of interest and the associated measures of uncertainty across the three sets of estimates you obtain here. Highlight the values of the covariates for which our inferences (about the null hypothesis that MCCE = 0) would change, if any.

References

Neumayer, Eric. 2005. "Bogus Refugees? The Determinants of Asylum Migration to Western Europe." *International Studies Quarterly* 49 (3): 389–410. http://dx.doi.org/10.1111/j.1468-2478.2005.00370.x.