

Introductory Econometrics: Computer Problem Sheet 2

You can download the data from the course website
<http://personal.strath.ac.uk/gary.koop/ie.html>

Question 1

This data set has been collected from the UNESCO yearbook and it contains data from 1997 for 38 countries on their educational spending, GDP and population. In particular, it contains three variables:

EDUC = public recurrent expenditure on education (millions of US dollars)
GDP = gross domestic product (millions of US dollars)
POP = population (millions)

1. *Testing for heteroskedasticity.* Sort the data by GDP and perform a Goldfeld-Quandt test, running regressions using the subsamples of fourteen countries with the smallest and largest values for GDP. Repeat this exercise sorting the data by POP.
2. *The White test.* Perform White's test for heteroskedasticity on a regression of EDUC on GDP and POP. Do the results of this test agree with the results of the Goldfeld-Quandt test?
3. *Correcting for heteroskedasticity: weighted least squares.* Based on your results from question 1, investigate whether scaling by POP or GDP corrects the heteroskedasticity problem. That is, run a regression of EDUC/POP on GDP/POP and test for heteroskedasticity using the White test. Run a regression of EDUC/GDP on 1/GDP and test for heteroskedasticity using the White test. Does either of these transformations solve the heteroskedasticity problem. How do you interpret the coefficients (including the intercept) in these transformed regressions?
4. *Correcting for heteroskedasticity by logging.* Investigate whether the heteroskedasticity problem can be corrected by logging the variables. That is, use a White test investigate whether the a heteroskedasticity problem exists in a regression of $\ln(\text{EDUC})$ on $\ln(\text{GDP})$ and $\ln(\text{POP})$. How do you interpret the coefficients in this regression?

Question 2:

In Computer Problem Sheet 1 OLS methods were used in a regression analysis involving the house price data set. Was the use of OLS appropriate? Hint: Use a White test to check if heteroskedasticity is present in the regression you ran in Computer Problem Sheet 1.