Exercise - Week 2

Part 1: Application of Simple and Multiple Regression Models

1. Housing Prices and Air Pollution

- Load the HPRICE2.DTA data set which describes housing prices in a sample of 506 communites in the Boston area, USA. (type **use HPRICE2**).
- We are interested in correlating housing prices with the amount of air pollution in the community which is measured by nitrogen oxide. Estimate the elasticity of housing prices with respect to the amount of nitrogen oxide.
- Now, try to estimate the elasticity of housing prices with respect to the property tax.
- What is the estimated % change of housing prices when an average number of rooms increases one unit?
- We want to estimate the elasticity of housing prices with respect to the amount of nitrogen oxide, while holding the log property tax constant. Is the estimated elasticity quite different from the one without holding the log property tax fixed?
- Could you connect this finding with omitted variable bias? If so, how big is the size of estimated bias?

2. Global Warming

- Load the global.dta dta set (**use global**).
- Construct the log of the following variables: temperature, CO2 concentration. For example, **gen lnt = ln(temp)**.
- Construct the dummy variable before (including) and after 1900. To this end, type the following:
  gen dperiod = year
  recode dperiod min/1900 = 0 1900/max = 1
  tab dperiod
- Perform the following regression:
  ln(temp) = beta0 + beta1*ln(Co2) + beta2*dperiod + u.
- What do you expect from the regression equation? What do you find from the regression?
- Now, perform the following regression:
  ln(temp) = beta0 + beta1*ln(Co2) + beta2*dperiod*ln(Co2) + u.
- What do you expect from the regression equation? What do you find from the regression?

Part 2: Simulation of Unbiasedness for OLS estimators by using Do-Files in STATA

- First, start a new do-file in Stata (Just click a small "envelop" icon saying "Bring Do-File Editor to Front")
- In order to set up, type the following:
  clear
  capture log close
  log using unbiased.txt, text replace
  di "unbiased.txt by Katrien Stevens: Simulation of Unbiasedness for OLS"
  version 7
- If you want to save the log file (unbiased.txt) in a directory, say, C:\mydirectory, type **cd "C:\mydirectory"** after clear.
- In the beginning, we would like to generate one random sample with 20 observations according to the true model: y = 1 + 2*x + u, where u and x are iid from N(0,1). Type the following:
  set obs 20
  gen x = invnorm(uniform())
  gen u = invnorm(uniform())
  gen y = 1 + 2*x + u
  summarize
- Type **reg y x**.
- From the log file that you will open later, you can see the regression estimation results from one random sample with 20 observations.
- Now we want to generate a 1000 number of random samples, at each of which we will regress y on the constant and x and get estimates of the intercept and the slope. By stacking 1000 estimates of the intercept and the slope across random samples, we will get the sample averages of OLS estimates across random samples and check how close these sample averages are to the true parameter.
- We first define the program that will be used at each time of 1000 random samples. Type the following:

```
capture program drop regwrs
program define regwrs
if "`1'" == "?" {
        global S_1 "b_const se_const b_slope se_slope"
        exit
        }
        drop all
        set obs 20
```

```
    gen x = invnorm(uniform())
    gen u = invnorm(uniform())
    gen y = 1 + 2*x + u
    reg y x
    post `1' (_b[_cons]) (_se[_cons]) (_b[x]) (_se[x])
end
```

- Now we instruct our do-file to run the simulations with 1000 repetition using the program regwrs. Type
  simul regwrs, reps(1000)
  summarize
- Finally, type **log close**.
- Save the do-file, called "unbiased.do". And press the icon saying "Do Current File".
- Go to the directory where the log file "unbiased.txt" is saved. Compare the OLS estimates from one random sample and 1000 number of random samples and check whether the idea of unbiasedness for OLS estimators is confirmed.