

ECN 3972
Lab 10

Download data from D2L: DRUGS5

Then, call up

In R:

```
library(mosaic)  
library(sandwich)  
library(lmtest)
```

Install data from Excel file: DRUGS5

Check out the variable names.

In R:

```
head(DRUGS5)
```

We are going to estimate the following model:

$$CV_i = \beta_0 + \beta_1 N_i + \beta_2 P_i + \beta_3 IPC_i + \varepsilon_i$$

where CV_i = consumption of pharmaceuticals in i th country divided by the consumption in the U.S.
 N_i = population of the i th country divided by the population in the U.S.
 P_i = average pharmaceutical price in country i
 IPC_i = 1 if country i encourages price competition; = 0 otherwise

In R:

```
model1 = lm(CV~N + P + IPC, data=DRUGS5)  
summary(model1)
```

Get the residuals and square them.

In R:

```
r2 = resid(model1)^2
```

Then, we can plot the residuals against each of our independent variables.

In R: (Make sure you look at the results after each line here!)

```
xyplot(r2~N, data=DRUGS5)  
xyplot(r2~P, data=DRUGS5)  
xyplot(r2~IPC, data=DRUGS5)
```

Let's test for heteroskedasticity using the Breusch-Pagan test. You need to estimate this model:

$$residuals_i^2 = \alpha_0 + \alpha_1 N_i + \alpha_2 P_i + \alpha_3 IPC_i + u_i$$

In R:

```
model2 = lm(r2~N + P + IPC, data=DRUGS5)
summary(model2)
```

You can now calculate the test statistic. It is equal to $N \times R^2$ and is distributed χ^2 with degrees of freedom equal to the number of independent variables in this regression. Next, find the critical value from your χ^2 table (assume the 5% level of significance).

Now, let's test for heteroskedasticity using the White test. You need to estimate this model:

$$residuals_i^2 = \alpha_0 + \alpha_1 N_i + \alpha_2 P_i + \alpha_3 IPC_i + \alpha_4 N_i^2 + \alpha_5 P_i^2 + \alpha_6 IPC_i^2 + \alpha_7 (N_i \times P_i) + \alpha_8 (N_i \times IPC_i) + \alpha_9 (P_i \times IPC_i) + u_i$$

In R:

```
model3 = lm(r2~N + P + IPC + N^2 + P^2 + IPC^2 + N*P + N*IPC + P*IPC, data=DRUGS5)
summary(model3)
```

You can now calculate the test statistic. It is equal to $N \times R^2$ and is distributed χ^2 with degrees of freedom equal to the number of independent variables in this regression. Next, find the critical value from your χ^2 table (assume the 5% level of significance).

IF we have heteroskedasticity, we should use White-corrected standard errors for our regression:

In R:

```
coeftest(model1, vcov = vcovHC(model1,type="HC1"))
```

These would be the standard errors and t-statistics you would use for your hypothesis tests.