Lab 2 – Dummy, Heteroskedasticity, Specification Issues

1 The scaling issue.

1. Download the datafile "BWGHT.dta" from your EE325 Moodle page.
2. Open the STATA software program
3. type: regress bwghtlbs cigs faminc
4. type: gen faminc_thb = faminc*33200
5. type: regress bwghtlbs cigs faminc_thb

2 Does "beauty" help increase wage?

1. Download the datafile "beauty.xlsx" from your EE325 Moodle page.
2. Open the STATA software program. Click on the "Data Editor" icon.
3. Copy the entire dataset from the excel file and paste it onto the STATA's Data Editor page.

   Choose "Treat first row as variable names".
4. Save the new STATA dataset.

   Choose File -> Save As -> (then name the new dataset "beauty_stata")
5. Open a new do-file and save it.

   Choose "New Do-file Editor" icon

   On the Do–file's top panel, choose File -> Save As -> (then name the new do-file "second_stata_lab")

   - **To explore and understand the data**
6. type: browse
7. type: sum
8. type: tab look
9. type: tab look female

   - **We want to test whether "good look" has a positive impact on wage**
10. type: gen belavg = 0
11. type: replace belavg = 1 if look < 3

12. type: gen log_wage = log(wage)
13. regress log_wage belavg abvavg

    - **Seems like we may have the omitted variable bias. Let's take into account other variables.**

14. type: regress log_wage abvavg belavg educ
15. type: regress log_wage abvavg belavg educ exper expersq
16. type: regress log_wage abvavg belavg educ exper expersq bigcity
17. type: regress log_wage abvavg belavg educ exper expersq bigcity black
18. type: regress log_wage abvavg belavg educ exper expersq bigcity black union
19. type: regress log_wage abvavg belavg educ exper expersq bigcity black union female

    - **Now, let's export the data into a formal format**

20. type: ssc install outreg2

    (This command is to install a command called "outreg2". Once you install this command, your computer will recognize it. So, no need to reinstall in the future.)

21. type: regress log_wage abvavg belavg educ
22. type: outreg2 using statalab2.docx
23. type: regress log_wage abvavg belavg educ exper expersq bigcity
24. type: outreg2 using statalab2.docx, append

    - **Check if we have the heteroskedasticity problem. (Let's use the White Test (special case))**

25. type: regress log_wage abvavg belavg educ exper expersq bigcity black union female
26. type: predict u_hat, resid
27. type: predict y_hat, xb
28. type: gen u_hat_sq = u_hat^2
29. type: gen y_hat_sq = y_hat^2
30. type: regress u_hat_sq y_hat y_hat_sq
31. Calculate $LM = nR^2$
32. Do we reject H_0 : homoskedasticity at 5% level of confidence?

33. Now, try using the ready-made test by STATA
34. type: regress log_wage abvavg belavg educ exper expersq bigcity black union female
35. type: estat hettest
36. Do we reject H_0 : homoskedasticity at 5% level of confidence?

- **Should we believe that the value of β are the same for female and male?** (Chow Test)
- Chow statistic is a type of F-statistic $F = \frac{SSR_p-(SSR_1+SSR_2)}{SSR_1+SSR_2} \cdot \frac{[n-2(k+1)]}{k+1}$

37. To get SSR_p: regress log_wage abvavg belavg educ exper expersq bigcity black union
38. To get SSR_1: regress log_wage abvavg belavg educ exper expersq bigcity black union if female == 0
39. To get SSR_2: regress log_wage abvavg belavg educ exper expersq bigcity black union if female == 1
40. What is the value of the F-statistic? Can we reject H_0 (can use the same model)?

3 Labor Force Participation of Female

1. Download the "MORA.DTA" dataset from your EE325 Moodle page and open it in STATA.

   Choose File -> Open -> (then direct to the location of the file)
2. type: des
3. type: regress inlf nwifeinc educ exper expersq age kidslt6 kidsge6
4. type: estat hettest
5. type: regress inlf nwifeinc educ exper expersq age kidslt6 kidsge6, robust
6. type: predict y_hat, xb
7. type: twoway scatter inlf educ || line y_hat educ
8. type: sort educ

9. type: twoway scatter inlf educ || line y_hat educ

   - **We need to "hold other things constant". Suppose nwifeinc = 30, exper = 10, age = 35, kidslt6 = 0, kidsage6 = 0.**

10. type: gen y_new = 0.585 + 30*(-0.0034) + educ*(0.0379) + 10*(0.0395) + 100*(-0.0006) + 35*(-0.0161)

11. type: twoway scatter inlf educ || line y_new educ