Econometrics
Lab Hour – Session 3

Agustín Bénétrix
benetria@tcd.ie
Office hour:
Wednesday 4-5
Room 3021

Martin Schmitz
schmitzm@tcd.ie
Office hour:
Monday 5-6
Room 3021

Outline

- Importing the dataset
- Regression analysis in Microfit
  - Today: Heteroscedasticity
- Project time

Import Dataset

- Import to Microfit (using Sleep75.fit)
- Or using Sleep75.csv
  - Make sure you know how to import .csv files

Tasks

- Estimate the model

$$sleep = \beta_0 + \beta_1 totwrk + \beta_2 educ + \beta_3 age + \beta_4 age^2 + \beta_5 yngkid + \beta_6 male + u,$$

Output

```
Dependent variable is SLEEP
 706 observations used for estimation from    1 to  706
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 CONSTANT                  3840.9           239.4139           16.0427[.000]
 TOTWRK                   -.16342            .018163           -8.9974[.000]
 EDUC                    -11.7133            5.8720           -1.9948[.046]
 AGE                      -8.6974           11.3291           -.76771[.443]
 AGESQ                     .12844            .13467            .95375[.341]
 YNGKID                  -.022801           50.2764          -.4535E-3[1.00]
 MALE                     87.7546           34.6679            2.5313[.012]
*******************************************************************************
 R-Squared                    .12275   R-Bar-Squared                   .11522
 S.E. of Regression         418.0269   F-stat.    F(  6, 699)   16.3018[.000]
 Mean of Dependent Variable  3266.4   S.D. of Dependent Variable    444.4134
 Residual Sum of Squares   1.22E+08   Equation Log-likelihood        -5259.3
 Akaike Info. Criterion     -5266.3   Schwarz Bayesian Criterion     -5282.3
 DW-statistic                1.9385
*******************************************************************************
                            Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version        *         F Version          *
*******************************************************************************
*                     *                          *                            *
* A:Serial Correlation*CHSQ(   1)=   .66603[.414]*F(   1, 698)=   .65911[.417]*
*                     *                          *                            *
* B:Functional Form   *CHSQ(   1)=   1.4660[.226]*F(   1, 698)=   1.4524[.229]*
*                     *                          *                            *
* C:Normality         *CHSQ(   2)= 184.0682[.000]*       Not applicable       *
*                     *                          *                            *
* D:Heteroscedasticity*CHSQ(   1)=   .82601[.363]*F(   1, 704)=   .82463[.364]*
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

Note on Heteroscedasticity test

- White LM Test included in Microfit
- H_0: Homoscedasticity
  - p-value above 10%
  - Thus, we can't reject the H_0

Tasks

- Calculate heteroscedasticity robust standard errors. Are these s.e. different from the usual ones?

Solution

```
        Ordinary Least Squares Estimation
             Based on White's Heteroscedasticity adjusted S.E.'s
*******************************************************************************
 Dependent variable is SLEEP
 706 observations used for estimation from    1 to  706
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 CONSTANT                   3840.9           259.1258            14.8223[.000]
 TOTWRK                    -.16342            .020683            -7.9014[.000]
 EDUC                     -11.7133             5.7475            -2.0380[.042]
 AGE                       -8.6974            11.7869            -.73789[.461]
 AGESQ                      .12844             .13602             .94426[.345]
 YNGKID                   -.022801            53.9053          -.4230E-3[1.00]
 MALE                      87.7546            35.5425             2.4690[.014]
*******************************************************************************
```

Task

- Find a model that allows the variance of u to differ between men and women (and not on any other factors)!

Solution

The assumption that the variance of *u* given all explanatory variables depends only on gender is

$$Var(u \mid totwrk, educ, age, yngkid, male) = Var(u \mid male) = \delta_0 + \delta_1 male$$

Then the variance for women is simply δ_0 and that for men is $\delta_0 + \delta_1$; the difference in variances is δ_1.

Task

- Estimate the model for heteroskedasticty!
  - Take residuals from OLS equation.
  - Square the residuals
    - Residsq = RESIDUALS^2
- Is the variance higher for men or women?

Solution

After estimating the above equation by OLS, we regress $\hat{u}_i^2$ on $male_i$, $i = 1,2, \ldots, 706$ (including, of course, an intercept).
We can write the results as

$$\hat{u}^2 = \underset{(20,546.4)}{189,359.2} - \underset{(27,296.5)}{28,849.6\ male} + residual$$

$$n = 706, \quad R^2 = .0016.$$

Because the coefficient on *male* is negative, the estimated variance is higher for women.

Output

```
 Ordinary Least Squares Estimation
*******************************************************************************
 Dependent variable is RESIDSQ
 706 observations used for estimation from    1 to  706
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 CONSTANT                 189359.2           20546.4             9.2162[.000]
 MALE                     -28849.6           27296.5            -1.0569[.291]
*******************************************************************************
 R-Squared                     .0015842   R-Bar-Squared                .1660E-3
 S.E. of Regression          359414.4   F-stat.    F(  1, 704)    1.1170[.291]
 Mean of Dependent Variable  173013.8   S.D. of Dependent Variable    359444.3
 Residual Sum of Squares     9.09E+13   Equation Log-likelihood       -10032.1
 Akaike Info. Criterion      -10034.1   Schwarz Bayesian Criterion    -10038.6
 DW-statistic                  2.0385
*******************************************************************************
```

Task

- Is the variance of *u* statistically different for men and women?

Solution

- No. The *t* statistic on *male* is only about –1.06, which is not significant at even the 20% level against a two-sided alternative.

- Project time….