# Advanced Econometrics

姓名：宋其来　　学号：502024740012

(Due Monday, October 28 in class)

**Important:** The empirical exercises should be done in Stata and a copy of your Stata code should be turned in along with your homework. Your homework should be typed/handwritten (you will grab results from the computer output and edit them), it should not contain irrelevant material (when you edit strip out things that are not useful), and it should be a clear and neat presentation of the results (insert comments where necessary in your codes).

1. The data in the file Koop-Tobias.csv are used in Koop and Tobias' s (2004) study of the relationship between wages and education, ability, and family characteristics. Their data set is a panel of 2,178 individuals with a total of 17,919 observations. The variables are defined in the article. Let $X_1$ equal a constant, education, experience, and ability (the individual' s own characteristics). Let $X_2$ contain the mother' s education, the father' s education, broken home, and the number of siblings (the household characteristics). Let $Y$ be the log of the hourly wage.

   i. Compute the least squares regression coefficients in the regression of $Y$ on $X_1$ . Report the coefficients in a table (same for the following regressions)

   ii. Let' s verify the Frish-Waugh-Lowell Theorem. Regress each of the three variables in $X_2$ (excluding Broken home) on all the variables in $X_1$ and save the residuals into a matrix $X_2^*$ . Then regress y on $X_1$ and $X_2^*$ . How do your results compare to the results of the regression of y on $X_1$ and $X_2$ (excluding Broken home)?

   iii. Run the regression of $y$ on $X_1$ and $X_2$ . Test for the presence of heteroscedasticity using both White' s general test and the Breusch-Pagan (1980) and Godfrey (1988) Lagrange multiplier test. Do your results suggest the presence of the heteroscedasticity? Report the result along with the correct SE. How do the results change from part i? Based on your results, what is the estimate of the marginal value, in \$/hour, of an additional year of education, for someone who has 12 years of education when all other variables are at their means and $Brokenhome = 0$ ?

   iv. Use an F testv. Use an F statistic to test the joint hypothesis that the coefficients on the four household variables in $X_2$ are zero.

   vi. We are interested in possible nonlinearities in the effect of education on log wage. (Koop and Tobias focused on experience. As before, we are not attempting to replicate their results.) Plot a histogram of the education variable, which will show values from 9 to 20, a spike at 12 years (high school graduation), and a second at 15.Consider aggregating the education variable into a set of dummy variables:

$$HS = 1 \ if \ Educ \leq 12, 0 \ otherwise$$
$$Col = 1 \ if \ Educ > 12 \ and \ Educ \leq 16, 0 \ otherwise$$

$$Grad = 1 \ if \ Educ > 16, 0 \ otherwise.$$

Replace Educ in the model with (Col, Grad), making high school (HS) the base category, and recompute the model. Report all results. How do the results change? Based on your results, what is the marginal value of a college degree? What is the marginal impact on log wage of a graduate degree?

vii. The aggregation in part vi actually loses quite a bit of information. Another way to introduce nonlinearity in education is through the function itself. Add Educ2 to the regresssion in part iii and recompute the model. Again, report all results. What changes are suggested? Test the hypothesis that the quadratic term in the equation is not needed - that is, that its coefficient is zero.

viii. One might suspect that the value of education is enhanced by greater ability. We could examine this effect by introducing an interaction of the two variables in the equation. Add the variable

$Educ_A bility = Educ \times Ability$ to the base model in part iii. Now, what is the marginal value of an additional year of education?

解. 证毕

2. Critical Summary Writing. Pick a published paper from top journals (AER, Econometrica, JPE, QJE, RES, JF, JFE, RFS, JFQA, MS) that uses field/quasi experiments for causal inference. The paper should be different from those in the presentation list. Write a summary of the paper in your own words in no more than 500 words (in English). Do not copy and paste the abstract of the paper. Pay particular attention to the experimental design, empirical methodology, and how the authors address internal and external validity concerns.

解.