**strEcon 4010 Homework 5**

**DUE: Tuesday, October 16 (in class)**

The data in wagexp.csv contains 206 observations for 13 variables. The variables are:

wage = employee hourly wage rate,

ed = years of education from 1st grade,

hsd = 1 if the employee does not have a high school degree, 0 otherwise

hsg = 1 if the employee’s highest educational attainment is a high school degree, 0 otherwise

assc = 1 if the employee’s highest educational attainment is an associate’s degree, 0 otherwise

bach = 1 if the employee has earned a bachelor’s degree, 0 otherwise

age = age of employee,

exp = years of working experience for the employee

fmale = gender of the employee (1 = female, 0 = male)

nonwh = 1 if employee is nonwhite, other than hispanic

hisp = 1 if employee is hispanic

For the first problems, estimate the regression :

1. Perform a t-test at 5% significance with the null hypothesis that . Write out the formula for the test, discuss the critical value, and t-value you get from the formula. What do the results suggest about the model?
2. Write out the VIF (see the function used in class VIF.R) for each coefficient in the above model. Do these results suggest any of these estimates are suffering from multicolliearity?
3. Create a scatter plot of age against experience and print it below. What is the correlation coefficient for age and experience? What does this number tell you?
4. Do you suspect this model suffers from multicollinearity? Support your answer with at least three reasons why or why not.

For the following problems, estimate the equation:

1. Provide the summary statistics of this regression below. What is the interpretation of ?
2. What is the interpretation of ? What does this coefficient suggest about the difference in returns to education between sexes?
3. Run White’s test for heteroskedasticity on the model above. Provide the regression equation for the test, as well as a summary of the results. Explain the findings from this test.
4. Plot the residuals on education below. Does this variable look heteroskedastic? Why or why not?
5. Run the above regression with White’s corrected standard errors and print the summary results below. How do these results differ from the results in question 5?
6. Do you suspect this model suffers from heteroskedasticity? Support your answer with at least two reasons why or why not.