**Week 4 Lab Handout**

**PA 5032 – Applied Regression**

February 12th, 2021

Wenchen Wang

**PART A: How to Compare Between Models ~10min**

**PART B: Interaction Terms ~20min**

**PART C: Report 1 Questions**

**Data:** WAGE1.dta (On Class Canvas Site)

**Contents:** 526 observations

**Variables:** wage = average hourly earnings

educ = years of education

exper = years of experience

tenure = years with current company

nonwhite = dummy variable where 1 = nonwhite

female = dummy variable where 1 = female

numdep = number of dependents

Also includes several dummies for region of residence and occupation

**PART A: How to Compare Between Models**

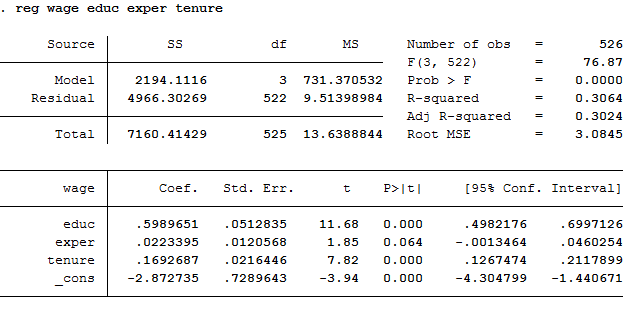
If we want to evaluate what are the factors that might affect wages, what variables would you include in your regression, and what directions would you predict these variables to affect wage?

\*Answer:

If we want to mainly look at how education is affecting wages, and we want to include only two other control variables, what is the best option here?

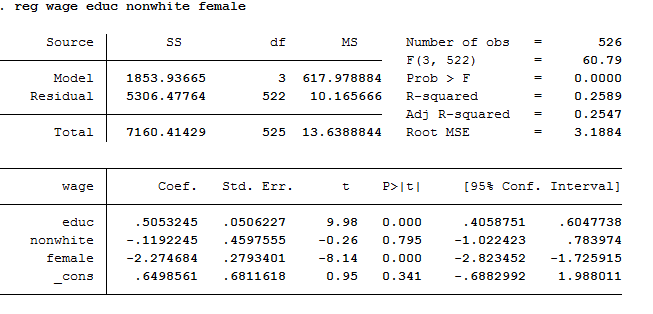
Let’s first look at this regression:

*reg wage educ exper tenure*

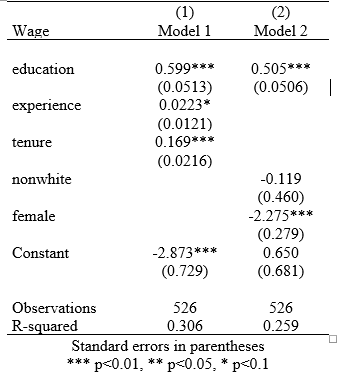


Let’s then look at another regression:

*reg wage educ nonwhite female*

**

Comparison between the two using outreg2 two function:



Which one is a better model?

\*Answer:

**PART B: Interaction Terms**

If we think that the relationship between education and earnings differs for men and women, we can use interaction terms to test that hypothesis.

Let’s first look at a regression of education on wage adding female as a control variable:

*Wage = B0+B1\*educ+B2\*female*

B1 here is the unique effect of education on wages, holding female constant. **Adding a control variable is estimating several regression lines, each with a different intercept but the same slope.**

Let’s add an interaction terms between education and female:

*Wage = B0+B1\*educ+B2\*female+B3\*educ\*female*

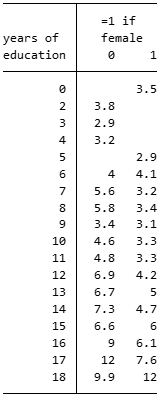
The interaction term here means that the effect for education on wages is different for being a male or being a female, so unique effect of education on wage not only depends on B1 but also depends on the values of B3 and female (since female is a categorical variable, it is either 0 or 1). The unique effect of education is represented by everything that is multiplied by education in the model, which is B1+B3\*female. **Interaction terms allow us to test for whether slopes are different for different groups.** The slope for female’s return to education is B1+B3\*1=B1+B3, and the slope for male’s return to education is B1+B3\*0=B1.

What if we want to look at the effect of being female on wages?

\*Answer:

Before we test for a differential in returns to education on wages between men and women, let’s see what is going on with the data

*table educ female, c(mean wage)*



What can you find from the table?

\*Answer:

Now there exists a difference, let’s use interaction terms to determine whether the slopes for returns on education for men and women are significantly different.

Let’s create our interaction term by multiplying educ and female.

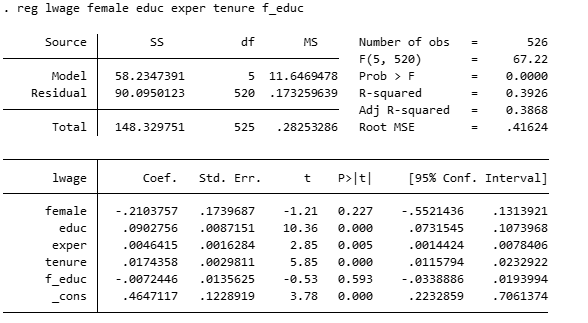
*gen f\_educ = female\*educ*

Next, we can run a regression that interacts education with female to see if returns to schooling are different for men and women. We’ll also control experience and tenure

*Wage = β0 + δ0 education + β1 female + β2 exper + β3tenure + δ1educ\*female + e*

*reg wage educ female exper tenure f\_educ*

H0: δ1=0



Can we reject our null hypothesis?

\*Answer:

How do we interpret the interaction term?

\*Answer:

**PART C: Report 1 Questions**