EC3303 Econometrics I Tutorial Problem Set 4

1. A researcher is interested in studying how years of education affects hourly earnings and whether the effect of another year of education is the same for males and females. He runs a regression of *log hourly earnings (AHE)* on *Years of Education*, *Female* (a binary variable =1 if person is female and =0 if person is male), an *interaction term where Female and Years of Education are interacted*, *Potential Experience*, *Potential Experience- Squared*, and *location dummy variables* (ie. Midwest is a binary variable = 1 if person lives in the Midwest and = 0 otherwise; South is a binary variable =1 if person lives in the South and =0 otherwise; West is a binary variable =1 if person lives in the West and =0 otherwise. The omitted region is Northeast. These 4 locations are mutually exclusive and exhaustive). Answer the questions below using the following regression results:

Regressor	**(1)**
Years of Education	0.1032** (0.0012)
Female	-0.451** (0.024)
Female × Years of Education	0.0134** (0.0017)
Experience	0.0143** (0.0012)
Experience^2	-0.000211** (0.000023)
Midwest	-0.095** (0.006)
South	-0.092** (0.006)
West	-0.023** (0.007)
Intercept	1.503** (0.023)
Adjusted R^2	0.267
Observations	52,790

(a) Consider a man with 16 years of education and 2 years of experience who is from a Western state in the United States. Estimate the expected change in the ***logarithm of*** average hourly earnings (AHE) associated with an additional year of experience.

Answer:
With 2 years of experience, the man's expected *ln(AHE)* is:

$$\begin{aligned} ln\widehat{(AHE)} &= (0.1032 \times 16) - (0.451 \times 0) + (0.0134 \times 0 \times 16) + (0.0143 \times 2) \\ &\quad - (0.000211 \times 2^2) - (0.095 \times 0) - (0.092 \times 0) - (0.023 \times 1) + 1.503 \\ &= 3.159 \end{aligned}$$

With 3 years of experience, the man's expected *ln(AHE)* is:

$$\begin{aligned} ln\widehat{(AHE)} &= (0.1032 \times 16) - (0.451 \times 0) + (0.0134 \times 0 \times 16) + (0.0143 \times 3) \\ &\quad - (0.000211 \times 3^2) - (0.095 \times 0) - (0.092 \times 0) - (0.023 \times 1) + 1.503 \\ &= 3.172 \end{aligned}$$

Difference = 3.172 - 3.159 = 0.013

(b) Consider a man with 16 years of education and 2 years of experience who is from a Western state in the United States. Estimate the expected change in the average hourly earnings (AHE) associated with an additional year of experience.

Answer: (i.e. average hourly earnings (*AHE*) are expected to increase by 1.3% when this man's experience increases by 1 year, from 2 to 3 years)

(c) Repeat (a), assuming 10 years of experience.

Answer: With 10 years of experience, the man's expected *ln(AHE)* is:

$$\widehat{ln(AHE)} = (0.1032 \times 16) - (0.451 \times 0) + (0.0134 \times 0 \times 16) + (0.0143 \times 10) - (0.000211 \times 10^2) - (0.095 \times 0) - (0.092 \times 0) - (0.023 \times 1) + 1.503 = 3.253$$

With 11 years of experience, the man's expected *ln(AHE)* is

$$\widehat{ln(AHE)} = (0.1032 \times 16) - (0.451 \times 0) + (0.0134 \times 0 \times 16) + (0.0143 \times 11) - (0.000211 \times 11^2) - (0.095 \times 0) - (0.092 \times 0) - (0.023 \times 1) + 1.503 = 3.263$$

Difference = 3.263 - 3.253 = 0.010

(d) Repeat (b), assuming 10 years of experience.

Answer: (i.e. average hourly earnings (*AHE*) are expected to increase by 1.0% when this man's experience increases by 1 year, from 10 to 11 years)

(e) More generally, what is the effect of an additional year of experience on hourly earnings?

Answer: The regression in nonlinear in experience (it includes *Potential experience*^{2}). Since the regression is a polynomial (quadratic, to be precise) in experience, we cannot interpret the coefficient on $exper^2$, for example, as the effect on $\Delta\ ln(AHE)$ from a unit change in $exper^2$. This does not make sense. Instead, because the polynomial is quadratic in experience, with the coefficient on $exper^2$ being negative and the coefficient on $exper$ being positive, this tells us that the effect of an additional year of experience on $\Delta\ ln(AHE)$ depends on the starting value of years of experience (i.e. where the change in experience is evaluated at). As can be seen from parts (a) and (c), the effect of an additional year of experience on $ln(AHE)$ (and also AHE) is estimated to be lower as a person's year of experience becomes greater.

(f) Interpret the coefficients on the location dummy variables

Answer:
The best thing to do in this case would first be to write out the estimated regression function. The estimated regression function is:

$$\widehat{ln(AHE)} = 1.503 + 0.1032educ - 0.451Female + 0.0134(Female \times educ) + 0.0143exper - 0.000211exper^2 - 0.095Midwest - 0.092South - 0.023West$$

Holding education, gender, and experience constant:

a. Compared to a person living in the Northeast (base category), a person living in the Midwest is predicted to earn (0.095 × 100)% = 9.5% less.
b. Compared to a person living in the Northeast (base category), a person living in the South is predicted to earn (0.092 × 100)% = 9.2% less.
c. Compared to a person living in the Northeast (base category), a person living in the West is predicted to earn (0.023 × 100)% = 2.3% less.

(g) What is the effect of an additional year of education for males? What is the effect of an additional year of education for females? Is the effect of an additional year of education different for males and for females?

Answer:
We want to know how $ln(AHE)$ responds to change in education (so that we can know, eventually, how AHE itself responds to a change in education). Since, there is a Female-Education interaction, the effect of a change in education on $ln(AHE)$ is hypothesized to be different for males and for females.

For males (Female=0),

$$\frac{\Delta\, ln(AHE)}{\Delta educ} = 0.1032$$

The above is obtained either by using the "before and after method" or differentiating $ln(AHE)$ w.r.t *educ*, setting *Female* =0.

For females (Female=1),

$$\frac{\Delta\, ln(AHE)}{\Delta educ} = 0.1032 + 0.0134(1) = 0.1166$$

The above is obtained either by using the "before and after method" or differentiating $ln(AHE)$ w.r.t *educ*, setting *Female* =1.

So, holding location and experience constant, 1 additional year of education is predicted to increase AHE by (0.1032×100)% = 10.32% for males. Similarly, holding location and experience constant, 1 additional year of education is predicted to increase AHE by (0.1166×100)% = 11.66% for females.

Notice that the coefficient on the interaction, $Female \times educ$, is statistically significant. This implies that there is evidence that the effect of an additional year of education is different for males and for females.

<u>Stata Exercise (to be done in tutorial with the tutor)</u>

2. The California Standardized Testing and Reporting data set (caschool.dta) contains data on test performance, school characteristics, and student demographic backgrounds for 420 Californian school districts. Today, we will learn how to perform non-linear regressions.