Econ 140 - Spring 2016
Section 8

GSI: Fenella Carpena

March 17, 2016

Additional Exercises

Question 1. For each of the following functions, state whether it can be linearized. If yes, write the resulting regression function in a form that can be estimated using OLS. If no, explain why.

1. $Y_i = \beta_0 X_{1i}^{\beta_1} e^{\beta_2 X_2 i}$
2. $Y_i = \frac{X_i}{\beta_0 + \beta_1 X_i}$
3. $Y_i = \frac{e^{\beta_0 + \beta_1 X_i}}{1 + e^{\beta_0 + \beta_1 X_i}}$
4. $Y_i = \beta_0 X_{1i}^{\beta_1} X_{2i}^{\beta_2} + u_i$

Question 2. Consider the following two regressions of wages on age and gender.

$$\widehat{Earn} = \underset{(21.18)}{323.70} + \underset{(0.55)}{5.15} \cdot Age - \underset{(13.06)}{169.78} \cdot Female$$

$$R^2 = 0.13,\ SER = 274.75$$

and

$$ln(\widehat{Earn}) = \underset{(0.08)}{5.44} + \underset{(0.002)}{0.015} \cdot Age - \underset{(0.036)}{0.421} \cdot Female$$

$$R^2 = 0.17,\ SER = 0.75$$

where $Earn$ is weekly earnings in dollars, Age is measured in years, and $Female$ is a dummy variable equal to 1 if the individual is female, and 0 otherwise.

(a) Interpret each regression carefully. For a given age, how much less do females earn on average? Should you choose the second specification on grounds of the higher regression R^2?

(b) Suppose that your professor points out to you that age and ln(earn) profiles typically take on an inverted U-shape. How would you extend the previous regression to test this idea?

(c) Now, consider the regression where you add the square of age to your log-linear regression in part (a).

$$ln(\widehat{Earn}) = \underset{(0.18)}{3.04} + \underset{(0.009)}{0.147} \cdot Age - \underset{(0.033)}{0.42} \cdot Female - \underset{(0.0001)}{0.0016} Age^2$$

$$R^2 = 0.28,\ SER = 0.68$$

Interpret the results from the above regression. Why is the Age coefficient here so large relative to its value in the regression in part (a)?

Question 3. Suppose you have data on weight, age, height and gender for 100 male and female children, between the ages of 9 and 12, who are all at least 4 feet tall. Using this data, you estimate the following relationship

$$\widehat{Weight} = \underset{(3.81)}{45.59} + \underset{(0.46)}{4.32} \cdot Height4$$

$$R^2 = 0.55,\ SER = 15.69$$

where $Weight$ is in pounds, and the $Height4$ variable is inches above 4 feet (so for a child who is 4 feet tall, $Height4$ takes on the value 0, while for a child who is 4 feet and 5 inches tall, $Height4$ takes on the value 5).

(a) Interpret the results of this regression.

(b) You remember from the medical literature that females in the adult population are, on average, shorter than males and weigh less. You also seem to have heard that females, controlling for height, are supposed to weigh less than males. To see if this relationship holds for children, you add a binary variable (DFY) that takes on the value one for girls and is zero otherwise. You estimate the following regression function:

$$\widehat{Weight} = \underset{(5.99)}{36.27} + \underset{(7.36)}{17.33} \cdot DFY + \underset{(0.80)}{5.32} \cdot Height4 - \underset{(0.90)}{1.83} \cdot DFY \cdot Height4$$

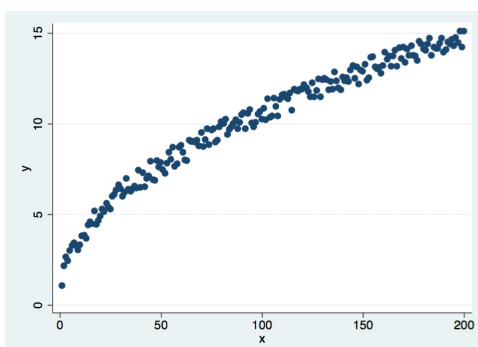
$$R^2 = 0.58,\ SER = 15.41$$

Are the signs on the new coefficients as expected? Are the new coefficients individually statistically significant? Write down and sketch the regression function for boys and girls separately.

(c) Using the regression in part (b), state the hypothesis that the regression function is identical for boys and girls. What test statistic would you use to this hypothesis?

(d) Consider the regression in part (b) but now assume that *in addition* to testing whether the relationship between height and weight changes by gender, you also wanted to test if the relationship between height and weight changes by age. Briefly outline how you would specify the regression to test this relationship, where the regression includes the gender binary variable (DFY) and an age binary variable (call it $Older$) that takes on a value of one for eleven to twelve year olds and is zero otherwise. How would the estimated relationship vary between the following four groups: younger girls, older girls, younger boys, and older boys?

Question 4. Consider the scatterplot of y and x below. Explain what transformation you would use, and what regression you would estimate to model this pattern. Can you think of two variables that might have an economic relationship shaped like this?



Question 5. A regression of *wage* (hourly wage, measured in dollars per hour) and *educ* (years of schooling) using data from a random sample of 526 American workers yields the following:

$$\widehat{wage} = -0.90 + 0.54 \cdot educ$$

(a) Interpret the intercept of this regression.

Suppose that using $ln(wage)$ instead as the response variable, we obtain the following regression:

$$\widehat{ln(wage)} = 0.584 + 0.083 \cdot educ$$

$$n = 526,\ R^2 = 0.186$$

(b) Interpret the slope. Compare the interpretation of the slope in the two regressions when the response variable is $ln(wage)$ vs. $wage$.

(c) Interpret the R^2 in the regression where $ln(wage)$ is the dependent variable.