John Chao
Spring 2015
Due Date: 5/5/15

Economics 422

Econometrics I

Problem Set 4

1. Data were collected from a random sample of 220 home sales from a community in 2003. Let *Price* denote the selling price (in \$1000), *BDR* denote the number of bedrooms, *Bath* denote the number of bathrooms, *Hsize* denote the size of the house (in square feet), *Lsize* denote the lot size (in square feet), *Age* denote the age of the house (in years), and *Poor* denote a binary variable that is equal to 1 if the condition of the house is reported as "poor." An estimated regression yields

$$\begin{aligned} \widehat{Price} &= 119.2 + 0.485BDR + 23.4Bath + 0.156Hsize + 0.002Lsize \\ &\quad +0.090Age - 48.8Poor, \\ \overline{R}^2 &= 0.72,\ SER = 41.5 \end{aligned}$$

(a) Suppose that a homeowner converts part of an existing family room in her house into a new bathroom. What is the expected increase in the value of the house?

(b) Suppose that a homeowner adds a new bathroom to her house, which increases the size of the house by 100 square feet. What is the expected increase in the value of the house?

(c) What is the loss in value if a homeowner lets his housee run down so that its condition becomes "poor"?

(d) Compute the (unadjusted) R^2 for the regression.

2. Consider the regression model

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i.$$

Transform the regression so that you can use a $t-$statistic to test the following restrictions

(a) $\beta_1 = \beta_2$;

(b) $\beta_1 + a\beta_2 = 0$, where a is a constant;

(c) $\beta_1 + \beta_2 = 1$. (*Hint*: You must redefine the dependent variable in the regression.)

3. Consider the regression reported in Question 1 above (where standard errors have now been added):

$$\begin{aligned} \widehat{Price} &= \underset{(23.9)}{119.2} + \underset{(2.61)}{0.485} BDR + \underset{(8.94)}{23.4} Bath + \underset{(0.011)}{0.156} Hsize + \underset{(0.00048)}{0.002}\ Lsize \\ &\quad + \underset{(0.311)}{0.090} Age - \underset{(10.5)}{48.8} Poor, \\ \overline{R}^2 &= 0.72,\ SER = 41.5 \end{aligned}$$

(a) Is the coefficient on BDR statistically significantly different from zero?

(b) Typically five-bedroom houses sell for much more than two-bedroom houses. Is this consistent with your answer to (a) and with the regression more generally?

(c) A homeowner purchases 2000 square feet from an adjacent lot. Construct a 99% confidence interval for the change in the value of her house.

(d) Lot size is measured in square feet. Do you think that another scale might be more appropriate? Why or why not?

(e) The F-statistic for omitting BDR and Age from the regression is $F = 0.08$. Are the coefficients on BDR and Age statistically different from zero at the 10% level?

4. This problem is inspired by a study of the "gender gap" in earnings in top corporate jobs [Bertrand and Hallock (2001)]. The study compares total compensation among top executives in a large set of U.S. public corporations in the 1990s (Each year these publicly traded corporations must report total compensation levels for their top five executives.)

(a) Let $Female$ be an indicator variable that is equal to 1 for females and 0 for males. A regression of the logarithm of earnings onto $Female$ yields

$$\widehat{\ln(Earnings)} = \underset{(0.01)}{6.48} - \underset{(0.05)}{0.44}Female, \quad SER = 2.65.$$

i. The estimated coefficient on $Female$ is -0.44. Explain what this value means.

ii. The SER is 2.65. Explain what this value means.

iii. Does this regression suggest that female top executives earn less than top male executives? Explain.

iv. Does this regression suggest that there is gender discrimination? Explain.

(b) Two new variables, the market value of the firm (a measure of firm size, in millions of dollars) and stock return (a measure of firm performance, in percentage points) are added to the regression:

$$\begin{aligned} \widehat{\ln(Earnings)} &= \underset{(0.03)}{3.86} - \underset{(0.04)}{0.28}Female + \underset{(0.004)}{0.37}\ln(MarketValue) + \underset{(0.003)}{0.004}Return, \\ n &= 46,670, \ \overline{R}^2 = 0.345. \end{aligned}$$

i. The coefficient on $\ln(MarketValue)$ is 0.37. Explain what this value means.

ii. The coefficient on $Female$ is now -0.28. Explain why it has changed from the regression in (a).

(c) Are large firms more likely to have female top executives than small firms? Explain.