Homework Set 2, ECO 311

Due Date: check syllabus

Instruction: There are 7 questions. Each question is worth 2 points. You need to submit the answers of only five questions which you choose. The maximum point you can get is 10 points. For the purpose of preparing exam, you need to understand all questions.

I will discuss the homework on the due date. Please do not ask me to go through the homework before the due date. However, you can discuss the homework with your classmates. You need to submit the homework individually though.

Q1: Quadratic and Interaction Terms

Use the House data. You may read example 6.2 and 6.3 in the textbook.

1. Report the regression with a quadratic term $\texttt{rprice} = \beta_0 + \beta_1 \texttt{age} + \beta_2 \texttt{age}^2 + u$. Explain whether $\texttt{age}^2$ should be included. When a house gets older, does the house price fall at increasing or decreasing rate? Hint: generate the quadratic term before running regression.

2. Report the regression with an interaction term $\texttt{rprice} = \beta_0 + \beta_1 \texttt{age} + \beta_2 \texttt{age} * \texttt{bath} + u$. Compare two houses, one with one bathroom and one with two bathrooms. Which one's value falls faster when it gets older? Why?

Q2: log and adjusted R squared

Use the House data. You may read Table 2.3, Section 6.2, and Section6.3 of the textbook.

1. Report regression $\log(\texttt{rprice}) = \beta_0 + \beta_1 \texttt{age} + u$, and regression $\log(\texttt{rprice}) = \beta_0 + \beta_1 \log(\texttt{age}) + u$. Interpret $\hat{\beta}_1$ in each regression

2. How to tell which regression fits data better? (Hint: consider adjusted R^2)

Q3: Two Sample T Test and Dummy Variable

Use the Wage data, and run the following R functions

```
summary(lm(wage~female))
t.test(wage[female==1], wage[female==0], alternative = "two.sided", var.equal = T)
```

1. (1 point) Interpret the coefficient of female produced by the command `lm`. Explain why the absolute value of t value for coefficient of female in command lm is the same as the absolute value of two sample t test in command `t.test`.

2. (1 point) Find the standard deviation of wage for male and female workers, respectively. Can you modify the `t.test` command in order to allow for unequal variances?

Q4: Interpreting Coefficients of Dummy Variables

Use the Wage data

1. (1 point) Run the following regression $\log(\texttt{wage}) = \beta_0 + \beta_1 \texttt{female} + \beta_2 \texttt{married} + u$ Please interpret $\hat{\beta}_1$ and $\hat{\beta}_2$.

2. (1 point) Run a new regression $\log(\texttt{wage}) = c_0 + c_1 \texttt{female} + c_2 \texttt{married} + c_3 \texttt{married*female} + u$. Please interpret $\hat{c}_3$, and explain what the interaction term `married*female` can capture.

Q5: Testing Difference in Regression Functions across Groups

Use the Wage data. You may read section 7-4c of the textbook

1. (1 point) Report two separate (group-wise) regressions, one for female and one for male. The regression is $\log(\texttt{wage}) = \beta_0 + \beta_1 \texttt{educ} + u$. Please interpret β_0 and β_1 in each regression.

2. (1 point) Test the null hypothesis at 5% level that the relation between log wage and educ does NOT depend on gender. Put differently, you need to run the Chow test for the null hypothesis of no structural change.

Q6: Empirical DID Research

Please read the Gun paper posted on my webpage, and explain in details how the authors apply the difference-in-difference (DID) technique to show the effect of Gun Buyback on crimes.

Q7: See last page of DID (L5) slides