

# ECO 5350 Exam2

Seung Ki Lee

All values are rounded up to 2 significant figures below one

## Q1) Use STATA dataset, hprice1 to answer following questions (5pt)

1. Write an equation for regressing price on lotsize, sqrft, bdrms, assess

$$\text{price} = \beta_0 + \beta_1 \text{lotsize} + \beta_2 \text{sqrft} + \beta_3 \text{bdrms} + \beta_4 \text{assess}$$

$$\text{price} = -38.89 + 0.00059 \text{lotsize} - 0.00052 \text{sqrft} + 11.60 \text{bdrms} + 0.91 \text{assess}$$

(21.50)      (0.00050)      (0.017)      (6.55)      (0.10)

2. Write above equation using beta coefficients.

$$Z_{\text{price}} = 0.058 Z_{\text{lotsize}} - 0.0029 Z_{\text{sqrft}} + 0.095 Z_{\text{bdrms}} + 0.84 Z_{\text{assess}}$$

(0.00050)      (0.017)      (6.55)      (0.10)

3. Please mention which of the above variables have the greatest absolute impact on Price. Which has the least?

assess has the greatest absolute impact of 0.84, and sqrft has the least significant impact

4. For the equation in (1), add interaction of sqrft and bdrms. Write the new equation below.

Let interaction term =  $\text{sqrftbdrms}$

$$\text{price} = 65.28 + 0.00059 \text{lotsize} - 0.046 \text{sqrft} - 14.66 \text{bdrms} + 0.88 \text{assess} + 0.012 \text{sqrftbdrms}$$

(65.64)      (0.00049)      (0.032)      (16.95)      (0.10)      (0.0072)

5. What is the effect of bdrms on price, holding other variable constant, when  $\text{sqrft} = 1000$ .

The below function was written rounded to 2 significant figures below zero for convinience, but calculated with whole original values. The final answer was rounded up.

$$\frac{d\text{price}}{d\text{bdrms}} = 0.012 \text{sqrft} - 14.66$$

$$\frac{d\text{price}}{d\text{bdrms}} (\text{sqrft} = 1000) = 0.012(1000) - 14.66 = -2.58$$

For a 1000 square foot house an additional bedroom actually decreases the price by -2.58 units.

## Q2) Using the STATA dataset, hprice1, please answer the following questions (3pt)

1. Write the STATA command to regress log(price) on log(assess), log(lotsize), log(sqrft), and dummy variable for the number of bedrooms. You need to use the name of variables used in STATA dataset, or that created by STATA software.

```
1 | xi : regress lprice lassess llotsize lsqrft i.bdrms
```

2. Holding all variables constant, please mention if there are significant effect of 4 bedrooms as compared to 2-bedroom house on price. What is the magnitude of such effect on price of the house?

For  $P > |t|$  value is greater than 0.05, and at 5% significance level it is statistically insignificant. The magnitude of effect is 0.042. The price of house increase by 4.2% from 2 bedrooms to 4 bedrooms.

3. How many rooms have significant effects on house prices, as compared to two-bedroom houses? (please justify why/why not)

The P-values for number of bedrooms are:

bdrms	P-value
bdrm3	0.313
bdrm4	0.612
bdrm5	0.077
bdrm6	0.037
bdrm7	0.272
bdrm8	0.811

As we need P value less than 0.05 for the effect to be statistically significant, we can see that bdrm6 is the only statistically significant effect.

### Q3) Use the STATA dataset, Charity, to answer the following questions (3pt)

1. Write the equation of regressing respond on avggift, propresp, and resplast. Use Heteroskedastic White correction for the above regression.

$$\text{respond} = \beta_0 + \beta_1 \text{avggift} + \beta_2 \text{propresp} + \beta_3 \text{resplast}$$

Without Correction

$$\text{respond} = \underset{(0.015)}{0.0023} + \underset{(0.000085)}{0.00018} \text{avggift} + \underset{(0.034)}{0.74} \text{propresp} + \underset{(0.018)}{0.095} \text{resplast}$$

With Heteroskedstic White Correction

$$\text{respond} = \underset{(0.013)}{0.0023} + \underset{(0.000030)}{0.00018} \text{avggift} + \underset{(0.034)}{0.74} \text{propresp} + \underset{(0.020)}{0.095} \text{resplast}$$

2. Write the STATA command for marginal effect of the above regression

```
1 regress respond avggift propresp resplast
2 mfx
```

3. **What is the probability of a positive response to for an increase in avggift, holding other variables constant?**

For per unit increase in avggift, the probability of positive reponse increase by 0.018%

**Q4) Use the STATA dataset, Fertil2, for answering the following questions**

1. **Write down the equation for the regression of children on agefbrth, heduc, electric, urban, and frsthalf. (1pt)**

$$\text{children} = \beta_0 + \beta_1 \text{agefbrth} + \beta_2 \text{heduc} + \beta_3 \text{electric} + \beta_4 \text{urban} + \beta_5 \text{frsthalf}$$

$$\text{children} = \underset{(0.030)}{6.11} - \underset{(0.015)}{0.087} \text{agefbrth} - \underset{(0.011)}{0.11} \text{heduc} + \underset{(0.15)}{0.18} \text{electric} - \underset{(0.11)}{0.61} \text{urban} + \underset{(0.098)}{0.16} \text{frsthalf}$$

2. **Test for Heteroskedasticity, and write the test statistic for**

o **LM test:**

$$\chi^2(1) = 70.16$$

$$\text{Prob} > \chi^2 = 0.0000$$

with STATA command

```
1 estat hettest
```

o **F-Test**

$$F(1, 1827) = 64.08$$

$$\text{Prob} > F = 0.0000$$

with STATA command

```
1 estat hettest, fstat
```

**Do you accept or reject Null Hypothesis? Interpret what accepting or rejecting will imply in this situation (3pt)**

I reject the null hypothesis in both LM and F test because the p-value in both cases are less than 0.05, the threshold at 95% confidence interval. This means that under both LM and F test we have heteroskedasticity.

3. **Estimate and write an equation for the above model using FGLS (2pt)**

$$\text{children} = \underset{(0.22)}{5.4} - \underset{(0.011)}{0.059} \text{agefbrth} - \underset{(0.010)}{0.090} \text{heduc} + \underset{(0.13)}{0.18} \text{electric} - \underset{(0.10)}{0.61} \text{urban} + \underset{(0.088)}{0.16} \text{frsthalf}$$

4. **Does FGLS estimation remove heteroskedasticity for the above model? Explain your answer (1pt)**

o **LM test:**

$$\chi^2(1) = 246.77$$

$$\text{Prob} > \chi^2 = 0.0000$$

o **F-Test:**

$$F(1, 1827) = 187.21$$

$$\text{Prob} > F = 0.0000$$

It does not remove heteroskedasticity. Even after the FGLS estimation the p values of both LM and F test is lower than 0.05 at 95% confidence interval, indicating it rejects null hypothesis.

**5. If there is still heteroskedasticity in the above model, can we modify the data to get rid of the issue? If so, how? (1pt)**

We can modify the data by log-transforming any continuous variable. First we would first take log of the dependent variable to see if it is fixed. If not, we can take the log of the independent variable.

**6. Write down the White robust standard errors for equation (1) (1pt)**

$$\text{children} = \underset{(0.28)}{6.11} - \underset{(0.014)}{0.087} \text{agebrth} - \underset{(0.011)}{0.11} \text{heduc} + \underset{(0.13)}{0.18} \text{electric} - \underset{(0.11)}{0.61} \text{urban} + \underset{(0.098)}{0.16} \text{frsthalf}$$

Independent Variable	Robust Std. Err
agebrth	0.014
heduc	0.011
electric	0.13
urban	0.106
frsthalf	0.097
_cons	0.28