

## Final Exam Sample

ECON 340: Economic Research Methods

Instructor: Div Bhagia

Print Name: \_\_\_\_\_

This is a closed-book test. You may not use a phone or a computer.

Time allotted: 90 minutes

Total points: 20

Please show sufficient work so that the instructor can follow your work.

*I understand and will uphold the ideals of academic honesty as stated in the honor code.*

Signature: \_\_\_\_\_

## 1. (6 pts) Multiple-choice questions

(a) Why do we use the *Adjusted- $R^2$*  instead of the  $R^2$  when comparing models with different number of variables?

- $R^2$  is always higher for a model with more variables even if some variables don't make sense.
- It's not possible to calculate  $R^2$  for models with more than one variable.
- Adjusted- $R^2$*  is concerned with whether estimates are causal.
- All of the above

(b) Which of the following is OLS trying to achieve?

- Choose  $\beta_0$  and  $\beta_1$  to minimize  $\sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X)$
- Choose  $\beta_0$  and  $\beta_1$  to minimize  $\sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X)^2$
- Choose  $\beta_0$  and  $\beta_1$  to minimize  $\sum_{i=1}^n (Y_i - (\beta_0 - \beta_1 X))^2$
- None of the above

(c) The regression model:

$$sales = \beta_0 + \beta_1 advertising\_expenditure + u$$

yields an  $R^2$  of 0.24, then

- Advertising expenditure explains 76% of the variation in sales.
- Sales explain 24% of the variation in advertising expenditure.
- Advertising expenditure explains 24% of the variation in sales.
- None of the above

(d) Suppose I estimate the following model:

$$\ln Y = 100 - 12.45 \ln X$$

Then according to my model:

- If  $X$  increases by 1%,  $Y$  decreases by 12.45%
- If  $X$  increases by 1,  $Y$  decreases by 12.45%
- If  $X$  increases by 1,  $Y$  decreases by 12.45
- If  $X$  increases by 1%,  $Y$  decreases by 12.45

(e) Suppose I estimate the following model:

$$\widehat{wages} = 500 + 2000age - 10age^2$$

According to this model, predicted wages for a 40-year-old individual are:

- \$80,500
- \$64,500
- \$500
- \$112,500

(f) Consider the following model:

$$Y = \beta_1 + \beta_2(D_1 \times D_2) + \beta_3D_1 + \beta_4D_2$$

where  $D_1$  and  $D_2$  are dummy variables that only take values 1 and 0. Then  $E(Y|D_1 = 1, D_2 = 1) - E(Y|D_1 = 1, D_2 = 0)$  is given by:

- $\beta_2$
- $\beta_2 + \beta_3$
- $\beta_2 + \beta_4$
- $\beta_1 + \beta_2$

2. (8 pts) A researcher runs an experiment to determine whether small class sizes improve kindergarten school performance. To this end, kindergarten students are assigned at random to either a *regular* or a *small* class. Based on this experiment, the researcher writes down the following model for test scores:

$$\text{testscore} = \beta_0 + \beta_1 \text{smallclass} + u$$

where

testscore: test score on a standardized test

smallclass: takes value 1 if the student is assigned to the small class, and 0 otherwise

The researcher estimates the equation by OLS and finds:

$$\hat{\beta}_0 = 904.5 \quad S_{\hat{\beta}_0} = 1.55$$

$$\hat{\beta}_1 = 14.35 \quad S_{\hat{\beta}_1} = 2.63$$

The  $R^2$  from the regression is 0.01.

- (a) (2 pts) Given that the students were randomly assigned to the different types of classes, do you think we can interpret  $\beta_1$  as the causal effect of a small class size on kindergarten performance? Why or why not?
- (b) (2 pts) Interpret each coefficient.
- (c) (1 pt) How do I interpret  $\hat{\beta}_0 + \hat{\beta}_1$ ?
- (d) (1 pt) Does the low  $R^2$  imply that the model is incorrect?
- (e) (2 pts) You want to test the hypothesis that a *small* class is as effective as a *regular* class.
- Formally state the null and alternative hypothesis.
  - Calculate the  $t$ -value associated with this test. Can you reject your null at 5% level of significance?

3. (3 pts) Suppose that wages are determined as follows:

$$wages = \beta_0 + \beta_1 education + \beta_2 ability + u$$

But instead we estimate the following regression model:

$$Y = \tilde{\beta}_0 + \tilde{\beta}_1 education + \tilde{u}$$

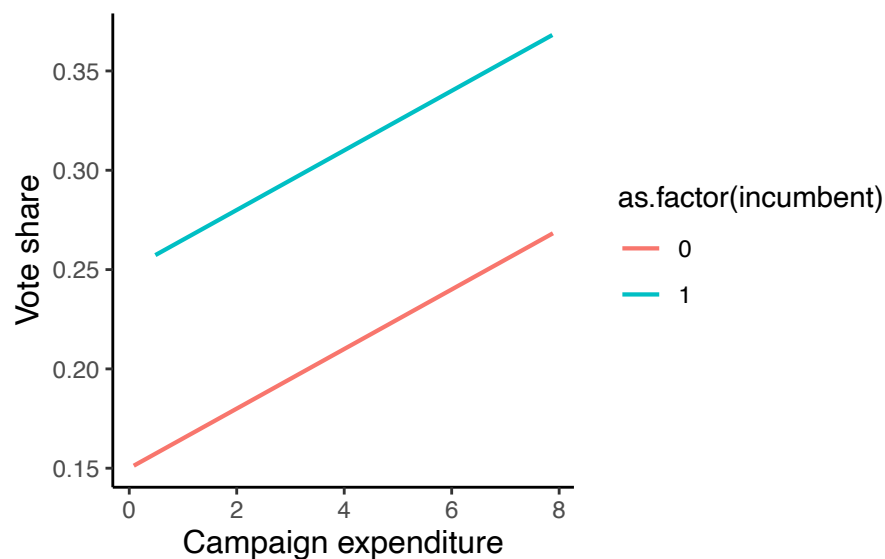
Do you think there will be an upward bias ( $\tilde{\beta}_1 > \beta_1$ ) or a downward bias ( $\tilde{\beta}_1 < \beta_1$ )? Explain your reasoning.

4. (3 pts) I estimated the following model:

$$vote\_share = \beta_0 + \beta_1 incumbent + \beta_2 campaign\_exp + u$$

Here, *incumbent* is a binary variable that takes the value 1 if the political candidate is an incumbent (i.e., won the previous election) and 0 if not. *campaign\_exp* measures money spent on the candidate's campaign leading up to the election in millions of dollars.

The fitted data from this model is presented below. Given the fitted line answer the following.



- (a) (1 pt) Is  $\hat{\beta}_1 > 0$  or  $\hat{\beta}_1 < 0$ , or is it not possible to determine the sign of  $\hat{\beta}_1$ ?  
Explain your reasoning.
- (b) (1 pt) Is  $\hat{\beta}_2 > 0$  or  $\hat{\beta}_2 < 0$ , or is it not possible to determine the sign of  $\hat{\beta}_2$ ?  
Explain your reasoning.
- (c) (1 pt) What is the interpretation of  $\beta_2$ ?