

Midterm 1 - Preparation

ECO 231 - Undergraduate Econometrics

Prof. Carolina Caetano

INSTRUCTIONS

The instructions below are the same that will be given to you in the exam.

1. You have received three booklets. Booklet 1 contains the exam instructions and the exam questions. Booklet 2 contains the numbered pages where you will answer question 1. Booklet 3 contains the numbered pages where you will answer question 2.
2. This exam has 2 questions, each is worth 50 points. Each item inside a question is worth the same. You have until 5 minutes before the end of the regular class time to answer it.
3. You must answer each question exactly in the space provided for it in booklets 2 and 3. You may use the back of the pages if they are empty. If you answer a question out of the order, or otherwise not on the space provided for it in the second booklet, your question will not be graded. If you need more space, you must ask for extra paper from the TA. It is your responsibility at the end of the exam to staple the extra page exactly in the right place in your exam. You may ask for draft paper if you like.
4. You are not allowed the use of notes, cheat sheets, calculators, or electronic devices of any kind. Turn your cell phone off, and put it away. If you did not bring a watch, check the board. The TAs will write down the time in the board every 15 minutes. If your answers are unclear or illegible you may lose points. You may answer in pencil.
5. If you finished your exam until 10 minutes before the end of class time, you may hand it back and leave the room. However, you may not keep booklet 1.
6. If you finished within 10 minutes of the end of class time, you must remain seated. Do not get up when the TA announces the time is up. Follow the TA's instructions about how to hand booklets 2 and 3. You may keep booklet 1 for yourself.
7. Write down your name on booklets 2 and 3. An exam without the name will not be graded.

1 Material Question

In the exam you will receive the actual situation. Here the questions are written in a generic form with variables y , x_1 , etc. There are many questions below, and you must prepare for them all. However, in the exam I will only give you a subset of them, so that you can complete the exam during class time. Notice that you may still be pressed for time during the actual exam, so prepare your answers so that they are concise.

- (a) What is the point of this research question? In other words, who would be concerned with this, and why?
- (b) Describe the ideal experiment to answer this question.
- (c) Is x_2 a confounder? Explain.
- (d) What is the meaning of the regression line of y on x_1 ? Write the formula of the regression line (don't use generic terms, use the names of the specific variable in the situation you were given).
- (e) Suppose that I want to estimate the effect of x_1 on y . Consider the univariate regression line:

$$y = a + b_1x_1$$

how to calculate a and b_1 using OLS?

- (f) Suppose that you estimated the regression line of the previous item, and it is $y = a + b_1x_1$. Comment on the following sentence: if variable x_1 is forced to change Δx units, y will change in $b_1\Delta x$ units.
- (g) Suppose that my data set contains y , x_1 , x_2 and x_3 . Consider the multivariate regression line:

$$y = a + b_1x_1 + b_2x_2 + b_3x_3$$

Express the partialling-out formula of the regression coefficient b_1 in terms of the variables in the model. Interpret it.

- (h) If the model is

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + u$$

where $\mathbb{E}[u|x_1, x_2, x_3] = 0$. What is this model saying about the world?

- (i) Interpret β_0 and β_1 in this model.

- (j) Suppose the model is

$$\log(y) = \beta_0 + \beta_1 \log(x_1) + \beta_2 x_2 + \beta_3 x_3 + u$$

where $\mathbb{E}[u|\log(x_1), x_2, x_3] = 0$. Interpret β_1 in this model.

- (k) This model is far from realistic. At the very least, we should also include information about x_4 . Suppose that our data has that variable. Write the new model, and interpret β_0 . Do you expect it to be higher or lower than β_0 in the model in item (h)?
- (l) If you were looking for an observational data set to answer this question, what would it need to have?
- (m) Suppose that the data set yielded the graph of averages in the following page (in the exam you will be provided with a graph of averages plot). Trace the regression line of y on x_1 . Should a regression line be used to describe this data? Explain your answer.
- (n) Consider the multivariate regression line using the controls from item (h):

$$y = a + b_1 x_1 + b_2 x_2 + b_3 x_3$$

how are a, b_1, b_2, b_3 calculated?

- (o) Write down the formula of the R^2 of the model in item (h), and interpret it. How do you expect the R^2 of the models in items (h) and (k) to compare? Why?

2 Paper Question

This question refers to this year's paper. Find it in the "Download" page in the course website.

- (a) What is the scientific question that the author wants to address? What is the main hypothesis of this study?
- (b) What is the economic rationale behind the hypothesis of customer discrimination in the sales/service sector?
- (c) Describe the ideal experiment to explore the customer-discrimination hypothesis. Discuss its practical applicability.
- (d) What is the drawback of comparing average wages of obese and non-obese women across different occupations?
- (e) Describe the obesity measure employed by the author and how it was constructed from the data. Discuss why the author considers it is a good measure even though it is constructed from different surveys.
- (f) Write the estimated model implicit in the results of Table 3 (it will be provided in the exam). Interpret the results for the coefficients of the interaction between obese and occupation.
- (g) Why does the author include a variable CB as an explanatory variable in Table 4? What does it intent to capture? (Table 4 will be provided in the exam)
- (h) Why does the author include a variable $Self$ as an explanatory variable in Table 4? What does it intent to capture? (Table 4 will be provided in the exam)
- (i) Why does the author include a variable YC as an explanatory variable in Table 4? What does it intent to capture? (Table 4 will be provided in the exam)
- (j) According to the results in Table 3 (it will be provided in the exam), which is the wage penalty (in percentage points) for a 12 years-educated, services sector obese woman, in comparison to a non-obese woman with the same characteristics.
- (k) Discuss what other explanations the author considers to answer the research question.
- (l) Discuss the limitations of the sample of women considered in this study.
- (m) What is the conclusion of the study? Does the methodology allow to a complete understanding of the wage differentials in the sales/services sector? Explain.