

Econ 390, Fall 2024: Midterm Exam I

October 3, 2024

Instructor: Maria Ponomareva

INSTRUCTIONS:

- You have 2 hours to complete and upload your work.
- This is an open notes, open books exam. However, collaboration is not allowed (this also includes posting this exam on Chegg and similar resources).
- The total is 100 points.
- **There are 5 questions:**

Question 1:	20
Question 2:	15
Question 3:	35
Question 4:	20
Question 5:	10
<hr/>	
Total:	100

- **Show all your work:** only fully explained answers will earn full credit.
- **Write legibly:** if we can't read it we won't grade it. **This also applies to corrupt files and wrongly uploaded files.**
- GOOD LUCK!

Question 1: ($4 \times 5 = 20$ pts) TRUE/FALSE. State if the following statements are true or false. **Only clearly and correctly explained answers will earn full credit.**

- (a) (5 pts) In real life economic applications, error term can be correlated with the explanatory variable, thus rendering the OLS unable to consistently estimate the causal effect.
- (b) (5 pts) When estimating a linear model, one would prefer to work with a dataset where the explanatory variable varies as little as possible.
- (c) (5 pts) The following linear regression model that explains hourly wage ($wage$, in dollars) in terms of IQ score (IQ) was estimated:

$$\widehat{wage} = 0.613 + 0.087IQ$$

This means that, keeping other things fixed, a 10 point increase in the IQ score is associated with a wage increase of 87 cents per hour.

- (d) (5 pts) If one wants to estimate the average percentage difference in prices between homes with central AC and homes without central AC, one can estimate the following linear regression model:

$$price = \beta_0 + \beta_1 \log(AC) + u$$

where $price$ is the price of a house in the sample, and AC is equal to 1 for all homes that have central AC, and equal to zero for all homes without central AC in the sample.

...continue to Question 2.

Question 2: ($3 \times 5 = 15$ pts) Below you are asked to decide on whether or not to use a one-sided alternative or a two-sided alternative hypothesis for the slope coefficient. **Briefly justify your decision and state the corresponding null and alternative hypotheses.**

- (a) (5 pts) $Crime = \beta_0 + \beta_1 Police + u$, where $Crime$ is the crime rate in a county, and $Police$ is a county's per capita police force. You seek evidence that more policing helps prevent crimes.
- (b) (5 pts) $Kids = \beta_0 + \beta_1 Income + u$, where $Kids$ is the number of kids in a family, and $Income$ is a family's income. You seek evidence that income is an important determinant of how many kids a family has.
- (c) (5 pts) $Price = \beta_0 + \beta_1 NoHighway + u$, where $Price$ is the price of a house, and $NoHighway$ is a dummy variable equal to 1 if that house is located far from a major highway, and 0 otherwise. You seek evidence that homes located far from major highways sell for more money than homes that are close to major highways.

...continue to Question 3.

Question 3: ($7 \times 5 = 35$ pts) Does having a better access to health care make people happier? Researchers at the Gallup World Poll surveyed thousands of respondents in 155 countries, between 2006 and 2009, in order to measure the overall satisfaction with life: each respondent was asked to report on the overall satisfaction with their lives, and their answers were ranked using a “life evaluation” score from 1 (the worst) to 10 (the best). The average of these individual scores is then used to represent the overall life satisfaction in the country, for each of 155 countries. To explain the overall life satisfaction the following model was proposed:

$$satisf_i = \beta_0 + \beta_1 life_exp_i + u_i$$

where $satisf_i$ was the average life evaluation score in country i , and $life_exp_i$ is country i 's life expectancy at birth (number of years a newborn infant is expected to live) and is used as a measure of health care access and quality in the country.

- (a) (5 pts) What are some other factors that can affect people's satisfaction with their lives (other than health care)? List at least 3.
- (b) (5 pts) If having a better access and better quality health care (hence higher life expectancy at birth) does make people happier and more satisfied with their lives, what sign do you expect for the slope coefficient in this model, and why?
- (c) Using the Gallup World Poll data for 155 countries, that model was estimated by OLS and here are the results:

$$\widehat{satisf} = -1.499 + 0.1062 life_exp$$

(0.0913) (0.0068)

$$n = 155, R^2 = 0.6135$$

- (i) (5 pts) Interpret the R^2 for this regression.
- (ii) (5 pts) Interpret the slope coefficient for this regression, keeping in mind that $satisf$ variable is the **country's** average (not the individual's level of satisfaction).
- (iii) (5 pts) In the United States, life expectancy at birth currently is 79 years. What would be the predicted average satisfaction with life for the United States?
- (iv) (5 pts) The answer you gave in part (b) is the theory that you seek evidence in favor of. Conduct the corresponding test at the 5% significance level. What is your conclusion? **Justify the use of a one-sided or two-sided test.**
- (v) (5 pts) Is the slope coefficient significant at the 1% significance level? Explain.

...continue to Question 4.

Question 4: ($4 \times 5 = 20$ pts) The Association for the Education of Young Children (AYEC) wants to evaluate the effect of school recess on the performance of kids in elementary school. In order to do this, at the beginning of a school year each of 300 elementary schools in the state was randomly assigned (based on a roll of a die) to have either a 30 min, a 40 min or a 50 min recess. At the end of the year, the average test score for a standardized math test is computed for each school. The researchers at AEYC plan to use this data to estimate the following model:

$$TestScore = \beta_0 + \beta_1 recess + u$$

where $TestScore$ is the average test score for a school, and $recess$ is the duration of recess in that school (so either $recess = 30$ or $recess = 40$ or $recess = 50$).

- (a) (5 pts) Explain what the u variable represents here. Give at least **3 (three)** specific examples.
- (b) (5 pts) Explain why in this model $E(u|recess) = 0$.
- (c) You seek evidence in favor of a theory that the duration of recess does have an impact on students' test scores. So, you estimate your model by OLS in order to test that theory, and here are the results:

$$\widehat{TestScore} = \underset{(8.13)}{638.3} + \underset{(0.61)}{1.32}recess$$

- (i) (5 pts) Use these regression results to test the theory that the duration of recess does have an impact on students' test scores. Use the 5% significance level. What is your conclusion? **Carefully state your null and alternative hypotheses in terms of the model's parameters (so, β_0 and/or β_1).**
- (ii) (5 pts) Use regression results to test the same theory at the 1% significance level. What is your conclusion now?

...continue to Question 5.

Question 5: ($2 \times 5 = 10$ pts) A researcher wants to see if time spent online (web browsing, TikTok, Facebook, Instagram, etc.) can explain weight gain of teenagers, so he asks the participants (age 14-19) to record their time spent online over a 1-year period (*TimeOnline*) and also their average weight (*AvgWeight*) throughout the year. He plans to use this data to estimate the following model:

$$AvgWeight = \beta_0 + \beta_1 TimeOnline + u$$

- (a) (5 pts) What are other factor that can explain affect someone's weight? List at least three, and briefly explain why you think these factors affect someone's weight?
- (b) (5 pts) Can any of those factor be correlated with the time spent online? Explain.