ECONOMETRICS FINAL EXAM

UNIVERSIDAD CARLOS III DE MADRID

JUNE 3, 2022

NAME:
NIA:
GROUP:

INSTRUCTIONS:

1. Write your name and group clearly in all sheets.

2. Leave an ID card with your picture on the desk.

3. Each of the four questions will be answered on (both sides of) the sheet where it is written. You cannot use the space on other sheets, or additional sheets.

4. You can use both sides of this sheet only for calculations which will not be evaluated.

5. All parts in each question have the same value.

6. The exam lasts 120 minutes:

   (a) Questions 1, 2 and 3 will be answered in the first 80 minutes using only a pen or pencil.

   (b) Then you can take your personal computer, where the Wooldridge database will have been downloaded in advance. 5 minutes will be given to boot the computer. Question 4 will be answered in 35 minutes using GRETL. Only GRETL can be visible on the computer screen, no other programs can be running. A personal calculator can be used if GRETL's one is not working. Critical values and p-values can be obtained in GRETL.

QUESTION 1 (20%): Consider a linear model with an explanatory variable that may be endogenous. We have a random sample to estimate the causal relationship between the explained and explanatory variable using an instrument that satisfies the conditions of exogeneity and relevance.

a. With the information available, derive an expression for the slope parameter of the model in terms of the covariances of the instrumental variable with the explained and explanatory variable (70%). From this expression, provide a consistent estimator of the slope parameter of the model (30%).

b. Explain how the two-stage least squares estimator of the slope is obtained (50%), and show that the estimator obtained is algebraically identical to the estimator obtained in a. (50%).

c. Express the parameters of the structural form in terms of the parameters of the reduced forms of the explained and explanatory variable (50%). From the relation obtained, obtain a consistent estimator of the slope parameter of the model (10%), and show that it is identical to the estimator in b. (40%).

ANSWER:

<u>QUESTION 2 (20%)</u>: Consider a linear regression model with a dependent variable, Y, and two explanatory variables, X_1 and X_2.

a. Suppose that you test for the global significance of the model at a $100\alpha\%$ significance level by means of a sequential test: we reject the null hypothesis of global significance when the individual significance hypothesis for some of the slopes is rejected at the $100\alpha\%$ significance level. Use a confidence ellipse at the $100(1-\alpha)\%$ and the corresponding individual confidence intervals for the two coefficients to show that, for a given significance level, the conclusion of the sequential test might be opposite to that of the test based on the F statistic.

b. Suppose that the conditional variance of the model errors is not constant and that we test the joint significance of the model coefficients using an F statistic that imposes the homoskedasticity assumption. What consequences would it have on the decisions resulting from this test?

c. Suppose we want to test whether the partial/marginal effect between Y and X_1 depends on the value that the variable X_2 takes. Explain how you would perform this test at a $100\alpha\%$ significance level.

<u>ANSWER</u>:

QUESTION 3 (20%): Consider a linear model with one endogenous explanatory variable and r exogenous explanatory variables.

a. Suppose that we have m possible instrumental variables. Explain which conditions have to satisfy to be valid instruments.

b. Explain which it means that the m instruments are weak (25%). Which are the consequences of using weak instruments on the inferences performed? (25%) How would you test that these instruments are weak? (50%)

c. How many potential instruments do you need to be able to test that they are exogenous? (25%) Explain how would you execute the test. (75%)

ANSWER:

QUESTION 4 WITH GRETL (40%): Use the data in **KIELMC** database of Wooldridge. Data are for houses that were sold during 1981 in North Andover, Massachusetts; 1981 was the year a garbage incinerator was built in that town. To study the effect of the garbage incinerator on house prices, consider a regression model which explains $\ln(price)$, where $price$ is the price of the house in dollars, in terms of $\ln(dist)$, $\ln(intst)$, $\ln^2(intst)$, $\ln(area)$, $\ln(land)$, $rooms$, $baths$, and age, where $dist$ is the distance to the incinerator in feet, $intst$ is the distance to the interstate in feet, $area$ is the area of the house in squared feet, $land$ is the area of the lot in squared feet, $rooms$ is the number of bedrooms, $baths$ is the number of bathrooms and age is the age of the house in years.

a. Provide an estimate of the elasticity of the relation of house prices and distance to the interstate road for houses 1000 feet away of this road and a confidence interval for this elasticity, **using only data for 1981**.

b. Estimate the effect on the house price of transforming a bedroom into a bathroom (25%). Give the appropriate interpretation to your estimates (25%). Is this effect significative? (50%)

c. To evaluate whether the new incinerator had an effect on house prices it is decided to compare the estimates of the coefficient of $\ln(dist)$ obtained in two separated regressions, one with only 1978 data and another one with only 1981 data. Assuming that both samples are independent, perform an statistical test about whether the incinerator construction has an effect on the nearby houses.

d. Explain how to run the previous test using a single OLS regression with the binary variable $y81$.

ANSWER: