EXTRAORDINARY EXAM

Econometrics

Universidad Carlos III de Madrid

23/06/21

Write your name and group in each answer sheet. Answer all the questions in 2:30 hours.

QUESTION 1 (30%)

Consider production data for the year 1994 on 30 US firms in the sector of primary meat industries. For each firm, values are given on production (Y, valued added in millions of dollars), and capital (K, real capital stock in millions of 1987 dollars). A log-linear production function is estimated by OLS with the following result (standard errors assuming homoskedasticity in parenthesis).

$$\ln Y_i = \underset{(0.451)}{0.701} + \underset{(0.091)}{0.756} \ln L_i + \underset{(0.110)}{0.242} \ln K_i + \hat{u}_i, \; RSS = 1.81551, \; R^2 = 0.956888, \qquad (1)$$

with RSS denoting sums of squared residuals. There are also estimated by OLS two alternative specifications,

$$\ln Y_i = \underset{(0.358)}{0.010} + \underset{(0.026)}{0.524} \ln (K_i \cdot L_i) + \hat{u}_{1i}, \; RSS = 2.37214, \; R^2 = 0.94367, \qquad (2)$$

$$\ln \frac{Y_i}{K_i} = \underset{(0.132)}{0.686} + \underset{(0.089)}{0.756} \ln \frac{L_i}{K_i} + \hat{u}_{2i}, \; RSS = 1.825652, \; R^2 = 0.95665, \qquad (3)$$

where RSS is the sum of squares residuals. Critical values of the standard normal Z: $Z_{0.005} = 2.58$, $Z_{0.01} = 2.33$, $Z_{0.025} = 1.96$, $Z_{0.05} = 1.64$, $Z_{0.1} = 1.28$, where $\mathbb{P}(Z > Z_\alpha) = \alpha$. The critical values of the χ^2_q/q for $q = 1, \ldots, 5$ at 5% are $\chi^2_{1,0.05} = 3.84$, $\chi^2_{2,0.05}/2 = 3.00$, $\chi^2_{3,0.05}/3 = 2.60$, $\chi^2_{4,0.05}/4 = 2.37$, $\chi^2_{5,0.05}/5 = 2.21$.

a. (1/3) Test that the output elasticities with respect to capital and labor are identical using the $R^{2\prime}s$ at 5% of significance. Then, show that the test statistic can be expressed in terms of the $RSS's$.

b. (1/3) Test that the production technology exhibits constant returns to scale. Explain whether or not the test can be carried out either using the $R^{2\prime}s$ or the $RSS's$.

c. (1/3) Discuss how you could obtain a 95% confidence region for $\beta_{\ln L}$ and $\beta_{\ln K}$ (confidence ellipse). What additional information do you need? Briefly comment on whether such a confidence region may assist us in testing

$$H_0 : \beta_{\ln L} = 0.9 \text{ and } \beta_{\ln K} = 0.1 \quad \text{vs} \quad H_1 : \beta_{\ln L} \neq 0.9 \text{ and/or } \beta_{\ln K} \neq 0.1.$$

Use a graph to illustrate your explanations.

QUESTION 2. (35%)

A researcher has data for 100 workers in a large organization on hourly earnings ($earn$), skill level of the worker ($skill$), and a measure of the worker's intelligence (IQ). She hypothesizes that the relation between these variables are given by the following two equations:

$$\begin{aligned} \ln earn &= \beta_0 + \beta_1 skills + u, \\ skills &= \alpha_0 + \alpha_1 IQ + v, \end{aligned} \tag{4}$$

where u and v are disturbance terms uncorrelated with IQ. The researcher is not sure whether u and v are correlated.

a. (1/6) Justify whether each variable in the two equations is exogenous or endogenous and derive the reduced form equations for the endogenous variables.

b. (2/6) Demonstrate mathematically under which circumstances the OLS estimator $\hat{\beta}_1$ of β_1 is consistent and under which circumstances is inconsistent.

c. (2/6) Demonstrate mathematically how the researcher could use instrumental variables (IV) estimation to estimate consistently β_1.

d. (1/6) Explain the advantages and disadvantages of using IV rather than OLS, to estimate β_1 when there is no certainty on the consistency of $\hat{\beta}_1$.

QUESTION 3. (35%)

Our goal is to estimate the causal relationship between house prices and pollution. For this, we have a sample of 506 neighborhoods in the Boston area (USA). We estimate a model that relates the median dollar price of houses in each neighborhood ($price$) with the amount of nitrogen oxide in the area, measured in parts per 100 million (nox), controlling for $dist$: the weighted distance from the neighborhood to the five main employment centers, in miles, by $rooms$: the average number of rooms in the houses in the neighborhood, by *crime*: the number of crimes committed per capita (calculated as the number of crimes divided by the number of inhabitants multiplied by 100000), and by *stratio*: the average of the ratio of students per teacher in the neighborhood schools. The population model is

$$\begin{aligned}\ln(price) &= \beta_0 + \beta_1 \ln(nox) + \beta_2 \ln^2(nox) + \beta_3 dist + \beta_4 dist^2 + \beta_5 dist \cdot \ln(nox) \\ &\quad + \beta_6 rooms + \beta_7 stratio + \beta_8 crime + \beta_9 crime \cdot \ln(nox) + u,\end{aligned}$$

where the error u has zero mean, conditional to the explanatory variables considered, and the conditional variance can be a function of the explanatory variables. GRETL output with the OLS estimation of this model with the variance and covariance matrix of the estimated coefficients, as well as the estimation of a transformation, is at the end of the exam. Use the critical values in question 1.

a. (1/3) Provide a 95% confidence interval for the *price* elasticity with respect to nox, for $nox = 5$, $dist = 4$, and $crime = 0.5$.

b. (1/3) Which is the estimated *dist* value such that the relation between *price* and *dist* changes its sign when $nox = 5$?

c. (1/3) Obtain an estimator of the *price* elasticity with respect to *crime* for $nox = 5$ and $crime = 0.5$. Then, test at the 1% of significance whether this elasticity is different from zero.

Model 1: OLS, using observations 1–506

Dependent variable: lprice

Heteroskedasticity-robust standard errors, variant HC1

	Coefficient	Std. Error	t-ratio	p-value
const	18.5445	2.37133	7.8203	0.0000
ln(nox)	−8.35034	2.37062	−3.5224	0.0005
$\ln^2$(nox)	1.84037	0.581296	3.1660	0.0016
dist	−0.813055	0.222751	−3.6501	0.0003
$dist^2$	0.0168214	0.00346286	4.8577	0.0000
dist·lnox	0.382561	0.125710	3.0432	0.0025
rooms	0.242263	0.0236869	10.2277	0.0000
stratio	−0.0461280	0.00490355	−9.4070	0.0000
crime	0.202440	0.0430921	4.6978	0.0000
crime·lnox	−0.113157	0.0227057	−4.9836	0.0000

R^2	0.679808	Adjusted R^2	0.673998
$F(9, 496)$	97.99094	P-value(F)	4.0e–104

Coefficient covariance matrix

const	ln(nox)	$\ln^2$(nox)	dist	dist^2	
5.6232	−5.5742	1.3521	−0.50367	0.0064633	const
	5.6199	−1.3739	0.49179	−0.0060575	ln(nox)
		0.33791	−0.11702	0.0014084	$\ln^2$(nox)
			0.049618	−0.00066546	dist
				1.1991e–05	dist^2

dist·ln(nox)	rooms	stratio	crime	crime·ln(nox)	
0.28267	−2.0012e–05	−0.0033857	0.013530	−0.0061238	const
−0.27820	−0.0053622	0.0023673	−0.015960	0.0074156	ln(nox)
0.066446	0.0015378	−0.00045917	0.0045256	−0.0021542	$\ln^2$(nox)
−0.027784	−0.00028577	0.00042420	−0.00055342	0.00018813	dist
0.00034633	−2.6614e–06	−5.6973e–06	1.1637e–05	−5.1086e–06	dist^2
0.015803	0.00023215	−0.00024123	0.00026781	−7.9769e–05	dis·ln(nox)
	0.00056107	4.1068e–05	5.1632e–05	−2.2403e–05	rooms
		2.4045e–05	1.5567e–05	−9.9102e–06	stratio
			0.0018569	−0.00097728	crime
				0.00051555	crime·ln(nox)

Model 2: OLS, using observations 1–506

Dependent variable: lprice

Heteroskedasticity-robust standard errors, variant HC1

	Coefficient	Std. Error	t-ratio	p-value
const	18.5445	2.37133	7.8203	0.0000
ln(nox)	−0.952759	0.120556	−7.9030	0.0000
dist	−0.813055	0.222751	−3.6501	0.0003
dist^2	0.0168214	0.00346286	4.8577	0.0000
rooms	0.242263	0.0236869	10.2277	0.0000
stratio	−0.0461280	0.00490355	−9.4070	0.0000
crime	0.202440	0.0430921	4.6978	0.0000
$\ln(\text{nox})(\ln(\text{nox}) - 2\cdot\ln(5))$	1.84037	0.581296	3.1660	0.0016
$\ln(\text{nox})(\text{dist} - 4)$	0.382561	0.125710	3.0432	0.0025
$\ln(\text{nox})(\text{crime} - 0.5)$	−0.113157	0.0227057	−4.9836	0.0000

R^2	0.679808	Adjusted R^2	0.673998
$F(9, 496)$	97.99094	P-value(F)	4.0e–104