The International College of Economics and Finance
Econometrics – 2020-2021
Final Exam, June 10, 2021.
(One session 2 hours without break)
Solutions

SECTION A

Answer all questions 1 and 2 from this section.

Question 1. Preparing her econometric essay the student uses data on 32 individuals living in London and tries to investigate the relationship between the price of wine and consumption of wine for one week in 2021. She runs the following regression

$$\log(\hat{wine}) = \underset{(0.8911)}{4.2514} - \underset{(0.4100)}{0.8323}\log(price), \quad n = 32, \; R^2 = 0.932 \qquad \textbf{(1)}$$

wine denotes the amount of wine consumed per week in millilitres (a medium glass contains 175ml) and *price* denotes the average price of a selection of wines during the week in GBP (£). The numbers in parentheses are the standard errors.

(a) (10 marks) a1) The student wants to prove the significance of regression coefficients and asks you to show how to do this. What is your conclusion?

Solution:

$t = \dfrac{-0.8323}{0.41} = 2.03 < 2.042 = t_{crit}^{5\%}(30)$, so we cannot reject $H_0: \beta = 0$ against $H_1: \beta \neq 0$.

[4 marks]

a2) Your friend, who is helping the student with you, noticed that the value of the standard error in the regression coefficient is clearly at contradiction with the value of the coefficient of determination, and hypothesized that by mistake the standard error is overestimated by approximately 10 times. Check this consideration and reconstruct the correct value of the standard error of regression coeffitient (and all conclusions derived from it), assuming that the R-squared of equation **(1)** is calculated correctly.

Solution: $F = \dfrac{R^2}{(1-R^2)} \cdot (n-2) = \dfrac{0.932}{(1-0.932)} \cdot 30 = 411.1765 \;\Rightarrow\; t = -\sqrt{F} = -20.277 \;\Rightarrow\; s.e. = \dfrac{-0.8323}{-20.277} = 0.041$.

Now $|t| = 20.277 > 2.75 = t_{crit}^{1\%} \;\Rightarrow$ significant.

[6 marks]

(b) (10 marks) Your friend offered another idea: he noticed that the demand for wine is less elastic for people who have eaten at a restaurant during the week, arguing that eating in a restaurant encourages people to drink wine regardless of the price. Following his advise the student defines a dummy variable D_i that takes the value 1 if individual *i* ate at a restaurant during the week, and 0 otherwise. She obtains the following regression result:

$$\log(\hat{wine}) = \underset{(0.8911)}{4.2133} - \underset{(0.0281)}{0.8218}\log(price) + \underset{(0.011)}{0.0889}D \times \log(price), \quad n = 32, \; R^2 = 0.997. \qquad \textbf{(2)}$$

b1) Conduct a test based on dummy variable approach that may offer support for the friend's claim. Specify the null and the alternative hypothesis. What is your conclusion?

Solution: $H_0: \beta_{D\cdot\log(price)} = 0$, against the one-sided alternative $H_1: \beta_{D\cdot\log(price)} > 0$.

$t = \dfrac{0.0889}{0.0011} = 80.818 > 2.462 = t_{crit.oneside}^{1\%}$ .we reject the null, so the friend's idea is most likely true.

[5 marks]

b2) The student remembered that there is alternative approach to introducing dummy variables, namely the Chow test, and asks your advice on whether it can be applied here (at the same time explaining the essence of this test).

Solution: Chow test is based on the comparison of the *RSS* for the whole sample and for the subsamples of people consuming wine at restaurants and at home $F = \dfrac{(RSS(all) - (RSS(rest) + RSS(home))/2}{(RSS(rest) + RSS(home))/(30 - 2 \cdot 2)}$. But here it is not applicable as any F-test is two sided.

[5 marks]

(c) (10 marks) c1) The student decided to measure the amount of wine consumed per week in number of medium glasses instead of millilitres and to measure prices in dollars instead of pounds (the dollar rate in pounds is 0.71) and asks your advice, can she just adjust the regression coefficients **(1)** without having to do the regression all over again?

Solution: Both numerator and denominator of $\beta = \dfrac{\frac{\Delta wine}{wine}}{\frac{\Delta price}{price}}$ are dimentionless, so the coefficient will not change.

[6 marks]

c2) What about R-square? Why?

Solution: $R^2 = \dfrac{ESS}{RSS}$ also dimentionless, so it also will not change.

[4 marks]

Question 2. A student was instructed by her scientific adviser to construct econometric models explaining the willingness of households to buy ecologically produced apples. She uses data where each family was presented with a description of the attitude to ecologically friendly apples, along with prices (in $) of regular apples (*pr*) and prices of the ecolabeled apple (*pe*).

The variable we want to explain is the binary variable, *ecobuy* which equals 1 if the household buys ecologically friendly apples and 0 otherwise. Additional household variables we have are family income in $1000s, *faminc*, household size, *hhsize*, years of schooling, *educ*, and *age*.

Using a sample of 660 households, the following results were obtained:

	(1)	**(2)**	**(3)**
	OLS A	Probit A	Probit B
C	0.890	1.088	-0.244
	[.068]	(.206)	(.474)
Pr	0.735	2.029	2.030
	[.132]	(.378)	(.381)
Pe	-0.845	-2.344	-2.267
	[.106]	(.318)	(.321)
faminc			0.0014
			(.0015)
Hsize			0.069
			(.037)
Educ			0.071
			(.024)
Age			-0.001
			(.004)
R^2	.086		
$\log L$		-407.60	-399.04

The heteroskedasticity robust standard errors are reported in squared brackets and the (asymptotic) standard errors are reported in parentheses.

(a) (10 marks) a1) The student was confused by the fact that the signs of the coefficients at the prices of the different apples were opposite and thought it was a mistake. Explain to her that the equation is correct and that the signs of the coefficients are as expected.

Solution: If pe rises under pr fixed the chance for $ecobuy = 1$ decreases. If pr rises under pe fixed the difference $pe - pr$ decreases so the chance for $ecobuy = 1$ increases. So both signs are as expected.

[3 marks]

The supervisor advised the student to consider also the regression **(1a)** of the *ecobuy* variable on the excess price of organic apples over the price of regular apples (difference $dif = pe - pr$) and the total price level calculated as the sum of the prices of the two types of apples ($sum = pe + pr$).

a2) Write out the theoretical equations corresponding to regressions **(1)** and **(1a)**. How are the coefficients of models (1) and (1a) related? Is it possible to find numerical values of the coefficients of model (1a) based on (1)? Which equation more accurately answers the question about explaining the willingness of households to buy ecologically produced apples?

Solution:

Consider two regressions

$$ecobuy = \beta_1 + \beta_2 pr + \beta_3 pe + u \qquad \textbf{(1)}$$

$$ecobuy = \alpha_1 + \alpha_2 dif + \alpha_3 sum + u \qquad \textbf{(1A)}$$

or substituting $dif = pe - pr$, $sum = pe + pr$ and rearranging

$$ecobuy = \alpha_1 + (\alpha_3 - \alpha_2) pr + (\alpha_3 + \alpha_2) pe + u \qquad \textbf{(1A)}$$

So $\begin{cases} \alpha_3 + \alpha_2 = \beta_3 \\ \alpha_3 - \alpha_2 = \beta_2 \end{cases}$ and $\begin{cases} \alpha_3 = \dfrac{\beta_2 + \beta_2}{2} \\ \alpha_2 = \dfrac{\beta_3 - \beta_2}{2} \end{cases}$, from here $\begin{cases} \alpha_3 = \dfrac{-0.845 + 0.735}{2} = -0.055 \\ \alpha_2 = \dfrac{-0.845 - 0.735}{2} = -0.79 \end{cases}$

So two regressions are

$$\widehat{ecobuy} = 0.89 + 0.735 pr - 0.845 pe \qquad \textbf{(1)}$$

$$ecobuy = \alpha_1 - 0.79 dif - 0.055 sum + u \qquad \textbf{(1A)}$$

Equation (1A) answers more accurately the question of consumers' willingness to choose eco-apples for purchase. The coefficient of –0.79 with the variable *dif* expresses the combined effect of price changes of the two apple varieties at a fixed overall price level.

[7 marks]

(b) (10 marks) b1) What methods can be used to evaluate each of the LPM (1) and Probit (2, 3) models. Discuss the benefits/drawbacks of using the Probit model when trying to explain a binary variable.

Solution: The LPM model can be estimated both using OLS and ML method. For probit model the most appriate is ML method of estimation.

Advantages of the Probit model over the LPM: predicted probabilities lie in $[0, 1]$; MLE is (asymptotically) efficient whereas OLS (LPM) are inefficient; heteroskedasticity is not a problem for MLE while it is inherent in the LPM; the marginal effect obtained from Probit also is allowed to depend on individual characteristics.

Disadvantages: in the Probit model the coefficients cannot be directly interpreted as the marginal effects of the regressor(s), and so their evaluation is computationally more complicated.

[5 marks]

b2) Test the joint significance of the <u>nonprice</u> variables. Clearly indicate the null and alternative hypothesis, the test statistic and the rejection rule.

Solution:

The Probit model B: $\Pr(ecobuy = 1) = \Phi(\beta_0 + \beta_1 pr + \beta_2 pe + \beta_3 faminc + \beta_4 hhsize + \beta_5 educ + \beta_6 age)$, where $\Phi(z)$ is the standard normal cumulative distribution function. $H_0: \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$ vs. $H_1: \exists i, \beta_i \neq 0$.

$LR = 2(\log L^U - \log L^R) \overset{a}{\sim} \chi^2_4$ under H_0. Test statistic is: $2 \times (-399.04 - (-407.60)) = 17.12 > 13.77 = \chi^2_{0.01,4}$ so we reject the null and state joint significance of non-price variables.

[5 marks]

(c) (10 marks) c1) Indicate how you can obtain the marginal effect of the price of ecological apples *pe* using model **2** (**Probit model A)** holding constant the price of regular apples. (*calculations are not expected here*).

Solution: To get the marginal effect of *pe* we need to evaluate $\phi(z)\hat{\beta}_{ecoprc}$ with $z=\hat{\beta}_0+\hat{\beta}_1\,pr+\hat{\beta}_2\,pe$ where $\phi(\cdot)$ is the pdf of $N(0,1)$ $\phi(z)=\dfrac{1}{\sqrt{2\pi}}\exp\left(-\dfrac{z^2}{2}\right)$.

[3 marks]

(c2) Using model **2 (Probit A)** obtain the marginal effect of a \$0.10 reduction in *pe* when evaluated at the mean of our explanatory variables (*pr* mean equals \$0.884 and *pe* mean equals \$1.082). Compare it with the marginal effect obtained from model **1 (OLS A).** Discuss whether the effect is statistically significant and explain why it is important to use robust standard errors.

Solution: (ii) In this setting, we have: $z=1.088+2.029\times0.884-2.344\times1.082=0.345428$. Now, we have two ways to obtain the marginal effect. First, we can use the fact that the marginal effect of a one-unit change is given by: $\phi(0.345428)\times-2.344=-0.9926$ (using formula from **c1**). To get the effect of a \$0.10 reduction, we will then need to multiply the above by -0.10: $ME=0.099$.

Alternatively, one can use the table of normal cumulative function: in addition to $z=0.345428$ we will use $z+\Delta z=1.088+2.029\times0.884-2.344\times(1.082-0.1)=0.579828$. $ME=\Phi(0.58)-\Phi(0.35)=0.719-0.6368$ $ME=0.0822$. LPM gives close estimate $ME(LPM)=(-0.845)\cdot(-0.01)=0.845$

The key reason to use robust standard errors is to deal with the heteroskedasticity in the error term inherent in the linear probability model (LPM).

[7 marks]

SECTION B

(Question 3 **OR** Question 4**)**.

Question 3. Let us consider the following $ADL(1,1)$ model:

$$Y_t=\beta_0+\beta_1Y_{t-1}+\beta_2X_t+\beta_3X_{t-1}+u_t,\ |\beta_1|<1, \qquad \textbf{(1)}$$

where X_t and Y_t are $I(0)$ variables and u_t is an i.i.d. $(0,\sigma_u^2)$ error term that is uncorrelated with $Y_{t-1},Y_{t-2},Y_{t-3},\ldots$. You may assume that u_t is independent of $X_{t'}$ for all $t,\ t'$.

(a) (10 marks) a1) Provide the short-run and long-run effect of *X* on *Y*.

Solution: The short-run effect is β_2. To obtain long-run effect we solve the model at equilibrium:

$$Y^e=\beta_0+\beta_1Y^e+\beta_2X^e+\beta_3X^e \quad\Rightarrow\quad Y^e=\frac{\beta_0}{1-\beta_1}+\underbrace{\frac{\beta_2+\beta_3}{1-\beta_1}}_{\text{long-run effect}}X^e$$

. The long-run effect is: $\dfrac{\beta_2+\beta_3}{1-\beta_1}$.

[5 mark]

a2) Explain the difference between these effects and comment their economic meaning with reference to equation **(1)**. (*Full mathematical derivation is not required*).

Solution: The short-run effect gives the immediate effect that a change in *X* has on *Y* (temporary change); whereas the long-run effect indicates what in equilibrium the effect would be of a permanent change in *X*. Following equation (1)

$$Y_t=\beta_0+\beta_1Y_{t-1}+\beta_2X_t+\beta_3X_{t-1}+u_t,\ |\beta_1|<1, \qquad \textbf{(1)}$$

at moment t temporary change in X_t happens and so a direct short-run effect of X_t appears (and thus the coefficient β_2 is taken into account). At the next moment, variable X_t has already returned to its original value, and now variable X_{t-1} experiences temporary shock at that moment (and so coefficient β_3 appears), and simultaneously disturbed lagged dependent variable Y_{t-1} contributes to the changes. In the further process, only the lagged dependent variable Y_{t-1} works, providing the multiplier effect of the initial impact.

[5 mark]

(b) (10 marks) Show that when you omit the relevant variable Y_{t-1} in the above model, you will get evidence of autocorrelation. Explain the result.

Solution:
Let us reformulate your model as:

$$Y_t = \beta_0 + \beta_2 X_t + \beta_3 X_{t-1} + v_t, \qquad \textbf{(2)}$$

where $v_t = \beta_1 Y_{t-1} + u_t$, $\beta_1 \neq 0$ (variable Y_{t-1} is supposed to be relevant). Show that $\mathrm{Cov}(v_{t+1}, v_t) \neq 0$.

$\mathrm{Cov}(v_{t+1}, v_t) = \mathrm{Cov}(\beta_1 Y_t + u_{t+1}, \beta_1 Y_{t-1} + u_t) = \beta_1^2 \mathrm{Cov}(Y_t, Y_{t-1}) + \beta_1 \mathrm{Cov}(Y_t, u_t) + \beta_1 \mathrm{Cov}(u_{t+1}, Y_{t-1}) + \mathrm{Cov}(u_{t+1}, u_t)$.

$\mathrm{Cov}(u_{t+1}, u_t) = 0$ and $\mathrm{Cov}(u_{t+1}, Y_{t-1}) = 0$, as the errors are unrelated to anything in the past. From true equation

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + u_t, \ |\beta_1| < 1, \ (\beta_1 \neq 0)$$

follows $\mathrm{Cov}(Y_t, Y_{t-1}) \neq 0$ and $\mathrm{Cov}(Y_t, u_t) \neq 0$. So generally speaking, $\mathrm{Cov}(v_{t+1}, v_t) \neq 0$.

[10 mark]

(c) (10 marks) c1) Discuss what properties your OLS estimators for the equation **(1)** ($ADL(1,1)$) parameters will have in the presence of the lagged dependent variable. In particular, (i) are the estimators unbiased and consistent, and (ii) should we use robust standard errors? Provide supportive arguments for your answers.

Solution: Presence of Y_{t-1} in the model makes the OLS estimator biased. But u_t is uncorrelated with the regressors X_t and X_{t-1}. As long as there is no autocorrelation, Y_{t-1} and u_t are not correlated.

We do not need to use robust standard errors because there is no heteroskedasticity or autocorrelation.

[5 mark]

c2) Will your answer change if we assume that the random term is characterized by a first-order autocorrelation $u_t = \rho u_{t-1} + e_t$, where e_t is an i.i.d. $(0, \sigma_\varepsilon^2)$ error that is uncorrelated with anything in the past (white noise) and $|\rho| < 1$? Explain.

Solution: Y_t correlates with u_t in equation (1). Taking lag, Y_{t-1} correlates with u_{t-1}. Assuming $u_t = \rho u_{t-1} + e_t$

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + \rho u_{t-1} + e_t$$

ρu_{t-1} correlates with lagged explanatory variable Y_{t-1}, what violates Gauss-Markov conditions.

To test the significance of coefficients one should use Heteroscedasticity and Autocorrelation consistent (HAC) standard errors (only Newey-West standard errors are acceptable).

[5 mark]

(d) Discuss how you would proceed to test for the presence of autocorrelation in the model in **(c)** using the Breusch—Godfrey test (use assumptions from c2).

Solution: We should test $H_0: \rho = 0$ against $H_1: \rho \neq 0$ with $u_t = \rho u_{t-1} + e_t$, where e_t is white noise.

Run OLS on the original equation to obtain the residuals $\hat{u}_t$: $\hat{u}_t = Y_t - \hat{\beta}_0 - \hat{\beta}_1 Y_{t-1} - \hat{\beta}_2 X_t - \hat{\beta}_3 X_{t-1}$.

By the assumption that the alternative is a stationary $AR(1)$ process, the test equation is:

$$\hat{u}_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 X_t + \alpha_3 X_{t-1} + \rho \hat{u}_{t-1} + e_t.$$

Test statistic $nR^2 \overset{a}{\sim} \chi_1^2$ where n is sample size and R^2 is goodness-of-fit of the latter regression.

Alternatively, simple t test on coefficient of e_{t-1}, again with asymptotic validity can be used.

[10 mark]

Question 4. (a) (10 marks) a1) What does it mean that time series y_t is stationary? Give two examples of a stationary time series (*no mathematical proof required*). How to test the time series for stationarity (give the general scheme of any test).

Solution: Definition of (weak) stationarity: time series X_t is said to be (weak) stationary if

1) $\mathrm{E}(X_t)$ constant in time

2) $\mathrm{var}(X_t)$ constant in time

3) $\mathrm{cov}(X_t, X_{t+s})$ depends only on s (but not t).

Examples:1) AR(1) process $X_t = \beta_2 X_{t-1} + \varepsilon_t$ with $0 < \beta_2 < 1$, (or $|\rho| < 1$) is stationary for finite samples

2) MA(1) process $X_t = \varepsilon_t + \alpha\varepsilon_{t-1}$ and MA(2) process $X_t = \varepsilon_t + \alpha_1\varepsilon_{t-1} + \alpha_2\varepsilon_{t-2}$ are stationary.

General scheme of DF test for stationarity: to detect non-stationarity (difference type) we consider regression $X_t = \beta_2 X_{t-1} + u_t$; $H_0: \beta_2 = 1$ (random walk) against $H_1: \beta_2 < 1$ (corresponds to AR(1) model). Transform $\Delta X_t = X_t - X_{t-1} = (\beta_2 - 1)X_{t-1} + u_t$ and test significance of $_2 - 1$ using special DF t-statistics table

[6 marks]

a2) Why it is important to know whether time series is stationary? Explain why using non stationary time series can lead to spurious regression.

Solution: Consider $Y_t = Y_{t-1} + \varepsilon_t$, $X_t = X_{t-1} + v_t$, ε_t and v_t are unrelated white noise processes, $Y_t \in I(1)$, $X_t \in I(1)$. Regress Y_t on X_t

$$Y_t = \pi_0 + \pi_1 X_t + u_t,$$

and tests the significance of the slope coefficient: $H_0: \pi_1 = 0$, $H_1: \pi_1 \neq 0$. Under $H_0: \pi_1 = 0$, $Y_t = \pi_0 + u_t$ and since Y_t is $I(1)$) and π_0 is constant, it follows that u_t must be nonstationary $I(1)$. This violates the standard distributional theory based on the assumption that u_t is stationary. Hence although there is no relationship between *Y* and *X*, the regression will produce a *t*-ratio which will reject the null hypothesis $H_0: \pi_1 = 0$.

[4 mark]

Consider the following time series model for $\{y_t\}_{t=1}^T$:

$$y_t = \alpha + \beta t + u_t,\ t = 1, ..., T, \quad \text{with } u_t = \rho u_{t-1} + \varepsilon_t \text{ and } |\rho| \leq 1, \qquad \mathbf{(1)}$$

where ε_t is an i.i.d. $(0, \sigma_\varepsilon^2)$ error that is uncorrelated with anything in the past (white noise).

(b) (10 marks) b1) What does it mean that time series y_t is trend stationary?

Solution: y_t is trend stationary if by detrending (subtracting linear time trend) we can make y_t stationary.

[3 marks]

b2) Show that y_t defined as **(1)** is trend stationary when $|\rho| < 1$.

By detrending y_t, we have:

$$y_t - \beta t = \alpha + u_t.$$

This would be the case provided u_t is stationary. It is assumed that $u_t = \rho u_{t-1} + \varepsilon_t$. If $|\rho| < 1$, $AR(1)$ is a stationary process, hence y_t is trend stationary under this condition.

[7 marks]

(c) (10 marks) c1) What does it mean that time series y_t is difference stationary?

Solution: y_t is difference stationary if by differencing (taking the difference of $y_t - y_{t-1}$) we can make y_t stationary.

[3 marks]

c2) Show that y_t defined as **(1)** is difference stationary when $\rho = 1$.

Solution: Taking the difference of $y_t - y_{t-1}$, we have:

$$\begin{aligned} y_t &= \alpha + \beta t + u_t \\ y_{t-1} &= \alpha + \beta(t-1) + u_{t-1} \\ \hline y_t - y_{t-1} &= \beta + u_t - u_{t-1} \\ \Delta y_t &= \beta + \varepsilon_t , \end{aligned}$$

If $\rho = 1$ then $u_t = \rho u_{t-1} + \varepsilon_t$ converts into $u_t - u_{t-1} = \varepsilon_t$. Since u_t is i.i.d for all t, ε_t is stationary and so $\beta + \varepsilon_t$ also is, hence y_t is difference stationary.

[7 marks]

(d) (10 marks) d1) Discuss the importance of distinguishing between trend stationary and difference stationary processes. What implications regarding properties of the regression estimators this distinction have?

Solution: It is important to know whether a variable is difference- or trend-stationary because for difference-stationary variables shocks have a **permanent effect**

$$X_t = X_0 + \varepsilon_1 + \varepsilon_2 + ... + \varepsilon_{t-1} + \varepsilon_t$$

(one can see that all shocks are present in any $X_t, \forall t$), whereas for trend-stationary variables **shocks are transitory** $X_t = \alpha + \beta t + \varepsilon_t$ (here temporary shock is present only in X_t).

For difference stationary time series there are also statistical implications: estimators become inconsistent if the assumptions are not satisfied (it was shown in a2).

[7 marks]

d2) Which of time series of these two types exhibit strong dependence or persistence of the effect and which one exhibits weak dependence? Explain?

Solution: Implication for the long-run behaviour of a process: Time series that are trend stationary always revert to the trend in the long run (the effects of shocks are eventually eliminated). That is, **they exhibit weak dependence**. Time series which are difference stationary never recover from shocks to the system (the effects of shocks are permanent). **That is, they exhibit strong dependence or persistence of the effect** (it was shown in d1)**.**

[3 marks]