Utrecht University
Utrecht University School of Economics

Full retake exam Econometrics (WISB377)

Thursday, 2 February 2023, 11:00 – 13:00 CET full retake

*The students who were given extra time from the Board of Examiners have additional 20 minutes (end time of the exam: 13:40 CET).

Remarks:
- This entrance test consists of 12 sub-questions 6 numbered pages (included front page).
- Write your name and registration number on each page of your exam.
- Please do not post copies of this exam on the Internet.

Question 1

a) The Ordinary Least Squares estimator (OLS estimator) of $\boldsymbol{\beta}$ is obtained by

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\text{argmin}}\ L(\boldsymbol{\beta})$$

for which the loss function is

$$L(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

Question 1. please derive the OLS estimator $\hat{\boldsymbol{\beta}}$. What are the necessary assumptions?

Question 2. why is the Hessian of this minimization procedure a positive definite matrix?

b) For a set of information of *n* firms, consider the following population regression equation.

$$Log(Costs_i) = \beta_0 + \beta_1 \log(Firmsize_i) + \beta_2 Productivity_i + \beta_3 DumOld_i + u_i \qquad i=1,\ldots,n$$

log is the natural logarithm, *Costs* is the costs of a firm in thousands of euros, *Firmsize* is the number of employees and *Productivity* is the value of the production per worker in thousands of Euros. *DumOld* is a 0-1 indicator variable that has the value of 1 if the firm already existed prior to the year 2000 (and zero elsewhere).

Question: please give a precise interpretation of the regression parameters β_1, β_2 and β_3

c) For the bivariate regression equation

$$y_i = \beta_0 + \beta_1 x_i + u_i \qquad i = 1,\ldots,n$$

and a random sample of *n* observations, it is assumed that $E(\mathbf{u} \mid \mathbf{X}) = \mathbf{0}$ (strict exogeneity), for which $\mathbf{X}$ is a (*n* x 2)-dimensional matrix.

Question: Using the Law of Iterated Expectations, please proof that

$$E(\mathbf{u} \mid \mathbf{X}) = \mathbf{0} \Rightarrow Cov(u, x) = 0.$$

d) We formulate $\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$.Let's continue with equation (4) in the previous exercise. The *n* vectors $\begin{pmatrix} 1 \\ x_1 \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ x_n \end{pmatrix}$ are identically and independently distributed 2-dimensional random variables for which $E\left(\begin{pmatrix} 1 \\ x_i \end{pmatrix}\begin{pmatrix} 1 & x_i \end{pmatrix}\right) = \mathbf{C}$ where $\mathbf{C}$ is a finite and non-singular matrix. In addition, the *n* random variables $x_1 u_1, x_2 u_2, \ldots, x_n u_n$, are identically and independently distributed with $E x_i u_i = 0$. Furthermore, $u_1, u_2, \ldots, u_n$, are identically and independently distributed with $E u_i = 0$. We consider the OLS estimator $\hat{\boldsymbol{\beta}}_n$.

Question: demonstrate how you can make use of all of these assumptions to proof that the OLS-estimator converges in probability to $\boldsymbol{\beta}$

$$\hat{\boldsymbol{\beta}}_n \xrightarrow{p} \boldsymbol{\beta} \text{ as } n \to \infty$$

Question 2

For a random sample of *n* observations, we consider the 4-dimensional vector of regression parameters $\boldsymbol{\beta}$ of the linear regression model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

$\boldsymbol{\beta}$ is estimated by the OLS estimator. It is assumed that the column rank of $\mathbf{X}$ is 4, and that the conditional distribution $\mathbf{u} \,|\, \mathbf{X} \sim Normal(\mathbf{0}, \mathbf{I}_n)$ for which $\mathbf{I}_n$ is an identity matrix of dimension *n*, and $\mathbf{0}$ an *n*-dimensional vector of zeros.

A researcher formulates a null hypothesis: $\beta_1 = 0$ and $\beta_2 = \beta_3$.
Question: Please show for which matrix $\mathbf{R}$ and vector $\mathbf{r}$ the null hypothesis can be written as

$$\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$$

a) Question: Please derive the Wald test statistic under the null hypothesis, $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$, and explain how the aforementioned assumptions are required for the derivation.

b) Question: What is the statistical distribution of the test statistic under the null hypothesis?

Question 3

For a sample of *n* observations, for the linear regression model.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

let’s assume that the variance covariance matrix of the error terms contains heteroskedasticity:

$$Var(\mathbf{u} \mid \mathbf{X}) = \text{diag}\left(\sigma_i^2\right) \qquad i = 1,\ldots,n$$

a) Please, discuss the consequences of heteroskedasticity for the consistency of the OLS estimator

b) It can be demonstrated that $Var(\hat{\boldsymbol{\beta}} \mid \mathbf{X}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Psi}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$, for which $\mathbf{X}'\boldsymbol{\Psi}\mathbf{X} = \sum_{i=1}^{n} \mathbf{x}_i \sigma_i^2 \mathbf{x}_i'$. Please carefully describe both stages of the estimation procedure to calculate the White robust standard errors.

Question 4

Let's consider the moving average model. The MA(1) model is:

$$u_t = e_t + \alpha_1 e_{t-1} \qquad t = 2,...,T$$

Subscript *t* refers to the *t*-th period. The error term e_t is i.i.d. (identically and independently distributed), with expected value of zero and constant variance:

$Ee_t = 0$ and $Var(e_t) = \sigma_e^2$.

<u>Question:</u> Show that the (*T*-1) x (*T*-1) covariance matrix of the error term *u* is

$$Var(u \mid \mathbf{X}) = \mathbf{\Psi} =$$

$$= \begin{pmatrix} (1+\alpha_1^2)\sigma_e^2 & \alpha_1\sigma_e^2 & 0 & \cdots & & 0 \\ \alpha_1\sigma_e^2 & (1+\alpha_1^2)\sigma_e^2 & \alpha_1\sigma_e^2 & 0 & & \\ 0 & \alpha_1\sigma_e^2 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \alpha_1\sigma_e^2 & 0 \\ & & & \alpha_1\sigma_e^2 & (1+\alpha_1^2)\sigma_e^2 & \alpha_1\sigma_e^2 \\ 0 & & \cdots & 0 & \alpha_1\sigma_e^2 & (1+\alpha_1^2)\sigma_e^2 \end{pmatrix}$$

Question 5

We consider the panel data model

$$y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \alpha_i + u_{it} \qquad i = 1,\dots,n; t = 1,\dots,T$$

for which α_i is the individual-specific effect (random variable), and u_{it} is the identically and independently distributed error term with expected value zero and constant variance.

a) Question: Please calculate the autocorrelation for the random effects estimator. What are the major assumptions of the random effects estimator?

b) Question: For the model at the level of the individual

$$\mathbf{y}_i = \alpha_i \boldsymbol{\iota} + \mathbf{X}_i \boldsymbol{\beta} + \mathbf{u}_i$$

$$\mathbf{y}_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{pmatrix}; \ \mathbf{X}_i = \begin{pmatrix} x_{1i1} & x_{2i1} & \cdots & x_{ki1} \\ x_{1i2} & x_{2i2} & & x_{ki2} \\ \vdots & \vdots & & \vdots \\ x_{1iT} & x_{2iT} & & x_{kiT} \end{pmatrix} = \begin{pmatrix} x_{i1}' \\ x_{i2}' \\ \vdots \\ x_{iT}' \end{pmatrix}; \ \mathbf{u}_i = \begin{pmatrix} u_{i1} \\ u_{i2} \\ \vdots \\ u_{iT} \end{pmatrix}$$

$\mathbf{y}_i$ and $\mathbf{u}_i$: T x 1 vectors for individual i;
$\mathbf{X}_i$: T x k matrix for individual i;
$\boldsymbol{\iota}$ is a T x 1 vector of ones.

We want to derive the first-difference estimator. Show how we can make use of the ((T-1) x T)-matrix $\mathbf{D}$ to obtain the first-difference estimator of $\boldsymbol{\beta}$.

$$\mathbf{D} = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & & -1 & 1 \end{pmatrix}$$

c) Question: What are the essential assumptions for the first-difference estimator? Using the derivation of the previous sub-question, please give a careful motivation for your answer.

< End of the exam >