Utrecht University

Utrecht University School of Economics

Endterm exam Econometrics (WISB377)

Thursday, 11 November 2021, 9:00-11:00 CET

For those who have permission from the Board of Examiners for an extension of time, they can hand in the answer sheet on 11:40 CET ultimately.

This exam contains 6 questions on 3 numbered pages

Exam instructions

At the start of the exam

- Candidates who arrive 30 minutes after the time scheduled for the start of the examination will not be permitted entry to the examination room.

During the examination

- Nobody is allowed to leave the room within the first 30 minutes after the start of the exam.
- You are not allowed to go to the restroom unless you have permission of the Examiner or an invigilator.
- **MOBILE PHONES AND OTHER COMMUNICATION DEVICES ARE ONLY ALLOWED WHEN SWITCHED OFF AND STORED IN CLOSED BAGS**.
- *It is a closed book exam. It is **not** allowed to use any study aids such as books, readers, (preprogrammed) calculators*
- You may use a simple calculator and a dictionary (without any [handwritten] notes in it).
- The exam form is **NOT** allowed to be taken home by the candidate

Results/Post-examination regulations:

- The results of the examination will be announced on Blackboard within two weeks of the exam date. At the same time the time & place of the exam inspection will be announced.
- We do not discuss exam results over the phone or by email.
- After the announcement of the exam results in OSIRIS you have four weeks within which to lodge an appeal against your grade.
- Four weeks after the results of this exam are published, the original exam is available to you, when a declaration is signed, stating that no appeal has been made or will be made.

You can request a photocopy of your answers at the Student Desk up and until four weeks after publication of the results

Questions

1) Using a random sample of *n* observations, the parameters of the regression equation

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + u_i \qquad i = 1,\ldots,n$$

are estimated by Ordinary Least Squares. It is assumed that that variance of *u*, σ_u^2, is known and finite. We would like to develop a *Z*-test for hypothesis testing on $\boldsymbol{\beta}$.

Question: is it necessary to assume multivariate normality of the error term $\mathbf{u} \mid \mathbf{X}$ or are we required to assume univariate normality of the error term? Please motivate carefully your answer.

2) We start with the GLS estimator $\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\boldsymbol{\Psi}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\Psi}^{-1}\mathbf{y}$

and the assumption that $\mathbf{u} \mid \mathbf{X} \sim Normal(\mathbf{0}, \boldsymbol{\Psi})$, for which $\boldsymbol{\Psi}$ is known and of full rank. It can be shown that

$$\hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + \left(\mathbf{X}'\boldsymbol{\Psi}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\Psi}^{-1}\mathbf{u}$$

We would like to develop a test for hypothesis testing on $\boldsymbol{\beta}$.

Question: could you please derive the test statistic for $\hat{\boldsymbol{\beta}} \mid \mathbf{X}$? Could you please mention the assumptions that you are taking?

3) By making use of Ordinary Least Squares, a researcher wants to estimate the causal effect of a variable x_j that is included in the vector $\mathbf{x}_i$ of the linear regression equation.

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + u_i \qquad i = 1,\ldots,n$$

There is strong indication that the covariance between the error term *u* and the variable x_j is not equal to zero. The researcher argues that for the purpose of the research it is better to estimate the model with a sample of 100,000 observations than for a sample of 1000 observations, because of the Central Limit Theorem.

Question: do you agree with the researcher? Please motivate your answer.

4) Assume that there is a regression equation, for which there are some outliers in the dependent variable. The researcher has strong indication that the outliers are not due to measurement error of the dependent variable. Furthermore there is indication of heteroskedasticity.

Question: What would be your advice to the researcher?

5) Consider the error structure of a regression equation

$$u_t = \rho u_{t-1} + e_t \quad |\rho| < 1 \; t = 2,\ldots,T$$

It is assumed that the error term e_t is identically and independently distributed, with $Ee_t = 0$ and $Var(e_t) = \sigma_e^2$. The covariance between e_t and u_{t-1} is zero.

Questions:

a) Please compute the variance of the error term u_t

b) Please compute the covariance between u_t and the two-period lagged u_{t-2}

c) Please construct the covariance matrix of the error term u_t

6) Let's consider the panel data model

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

Questions:

a) Suppose the model is estimated with a first-difference estimator. Why is it impossible to estimate the parameter γ on the linear time trend, using the first-difference estimator?

b) Calculate the autocorrelation between the error terms for a pooled OLS estimator.

< end of the exam >