UNIVERSITY OF OSLO
DEPARTMENT OF ECONOMICS

Exam: **ECON3150/4150 – Introductory Econometrics**

Date of exam: Thursday, May 23, 2019   **Grades are given:**   June 14, 2019

Time for exam: 09:00 – 12:00

The problem set covers 4 pages (incl. cover sheet)

Resources allowed:
- Open book examination, where all written and printed resources, in addition to calculator, are allowed.

The grades given: A-F, with A as the best and E as the weakest passing grade. F is fail.

Exam Econ 4150, spring 2019

1. (15%) The zero conditional mean assumption of the Gauss-Markov conditions is often written as $E[u_i|X_i] = 0$

    (a) Does $E[u_i|X_i] = 0$ imply that u_i and X_i are uncorrelated? Explain your answer.

    (b) Give an example of a regression equation where the $E[u_i|X_i] = 0$ is likely to be violated.

    (c) If $E[u_i|X_i] \neq 0$ will more observations of X help?

2. (7,5%) You draw a sample of N individuals from a population and use income data from the sample to estimate the average income in the population. Explain why $\tilde{y} = \frac{1}{N-1}\sum_{i=1}^{N} y_i$ is a biased but consistent estimator of the mean income in the population.

3. (7,5%) You don't have access to your computer and you have to estimate the parameters in this simple OLS model by hand:

$$y = \beta_0 + \beta_1 x + u.$$

    Lucky for you there are only five observations, displayed below. Find the OLS estimates $\hat{\beta}_0$ and $\hat{\beta}_1$

obs.	y	x
1	2	1
2	5	2
3	6	3
4	7	3
5	5	1

4. (20%) A researcher wants to estimate how mothers influence the earnings of their daughters and collects data on the wage of 100 women and data on their mother's education, IQ and reading habits. She obtains the following results (the standard error is in parenthesis below the coefficient)

$$\widehat{lnwage} = \underset{(0.54)}{1.58} + \underset{(0.23)}{1.24}\, mothedu + \underset{(0.56)}{1.60}\, IQ + \underset{(1.09)}{1.19}\, books$$
$$N = 100, \ R^2 = 0.42, \ F(3, 96) = 18.9$$

    where $mothedu$, IQ, $books$ refer to an individual's mother's logged education level, their logged score on a standard IQ test and the logged number of books they have read respectively.

(a) What is the interpretation on the coefficient on books?

(b) Are each of the variables individually significant at the 95% confidence level?

(c) Test whether the coefficient on *mothedu* is significantly different from 1 at a 5% significance level?

(d) It turns out that there is a strong positive correlation between *books* and *mothedu*, is this a problem?

(e) What would happen to the coefficient on *mothedu* if *books* was omitted from the above regression? Why would this happen?

5. (50%) Data from a household survey ($N = 9000$) in Bangladesh contains information on whether or not a household has a member, a migrant, that works abroad. Migrants regularly send money to their household back home. These transfers are called remittances. A researcher wants to use the household survey data to estimate whether receiving remittances from migrants affect the income earned locally at home by the household that receives money from their migrant member.

A priori one can imagine that receiving money from the migrant can either reduce or increase how much the household earns at home. To investigate this question the researcher uses OLS to regress the log of income earned by the household, excluding the money received from the migrant, ($ln(income)$) on a dummy that indicates whether or not the household receives remittances ($rem = 1$ if a household receives remittences (money from the migrant), 0 if not):

$$ln(income)_i = \beta_0 + \beta_1 rem_i + u_i \qquad (1)$$

(a) She obtains $\widehat{\beta}_1 = 0.11$, give an interpretation of this coefficient.

(b) The standard error of $\widehat{\beta}_1$ is 0.03. What is the 95 % confidence interval for $\widehat{\beta}_1$?

(c) The researcher is also interested in estimating how remittances affect the poverty status of a household. To this end she creates a indicator variable *poverty* that is equal to 0 for households with an income above the poverty line and equal to 1 for households with an income at or below the poverty line.

  i. Is it problematic to use *poverty* as the dependent variable in an OLS regression? Explain your answer

  ii. What alternatives to OLS can be used?

(d) Return to the regression model $ln(income)_i = \beta_0 + \beta_1 rem_i + u_i$. Discuss this statement: $\widehat{\beta}_1 = 0.11$ captures the average causal effect on earned family income of having a family member abroad that sends remittences?

(e) The researcher uses distance from Dhaka (capital in Bangladesh) as an instrument for remittances in an IV-estimation. What criteria must distance fulfill to be a valid instrument.

(f) Write down the first stage equation of the IV-regression that uses distance from Dhaka as an instrument. What would you look for in this first stage to determine the validity of the instrument?

(g) Someone suggests that another instrument could be used, namely the ownership of non-agricultural land.

    i. Is it possible to use both instruments simultaneously to predict migration? Explain your answer.

    ii. With two potential instruments, can any tests be performed to check the exogeneity of the instruments? Explain your answer.