**SECTION A**

Answer ANY FOUR of the following six questions.

1. Explain what is meant by an error term. What assumptions do we make about the error term when estimating an ordinary least squares (OLS) regression. [5 marks]
2. Give at least three examples from economics where you expect some nonlinearity in the relationship between variables. Interpret the slope in each case. [5 marks]
3. How is the slope coefficient interpreted in a log-linear model, where the dependent variable is (i) in logarithms but the independent variable is not (i.e. a log-linear model), (ii) in a linear-log model and (iii) in a log-log model? [5 marks]
4. Formulate the Gauss-Markov theorem. Discuss briefly its assumptions and the consequences for the OLS in case of their violation. Illustrate your answers with examples and graphs when appropriate. [5 marks]
5. Explain the concept of non-stationarity of time series. Give examples of economic time series which are (1) stationary and (2) non-stationary. How does the autocorrelation function (ACF) help in detecting non-stationarity? [5 marks]
6. Carefully discuss the advantages of using heteroskedasticity-robust standard errors over standard errors calculated under the assumption of homoskedasticity. Give at least two examples where it is very plausible to assume that the errors display heteroskedasticity. [5 marks]

**SECTION B**

Answer ANY TWO of the following three questions.

Each question in this section carries 30 marks.

**Question 1**

Using a sample of 200 countries a researcher investigates the long-run determinants of growth. He knows that economic growth is determined by, among other factors, the rate of investment and the population growth. To test these predictions he estimated the following regression:

where is the real GDP per capita growth rate, is the average population growth rate and is the average share of investment in the GDP. All variables are averages for the period 1960-2010. The numbers in parentheses are standard errors.

1. Interpret the results and comment on the signs of the coefficients. Calculate the t-statistics. Do you think the coefficients are significantly different from zero? Explain the meaning of the standard error of regression (SER).
2. The overall F-statistic for the regression is 79.11. The critical values at the 5% and 1% level are respectively 3.00 and 4.61. What is your decision on the null hypothesis?

In a second regression the researcher also includes for each country the average educational attainment (Educ) in the regression. He obtains the following result:

Does the inclusion of the level of education affect the previous results? Do you think the first regression (without controlling for the lelvel of education) suffers from omitted variable bias?

1. One country in the sample has the following values: , , and . Does the regression in point (c) overpredict or underpredict the real GDP growth rate for this country?

**Question 2**

A study, published in 1993, used U.S. state panel data to investigate the relationship between minimum wages and employment of teenagers. The sample period was 1977 to 1989 for all 50 states. The author estimated a model of the following type:

u

where E is the employment to population ratio of teenagers, M is the nominal minimum wage, and W is average hourly earnings in manufacturing. In addition, other explanatory variables, such as the adult unemployment rate, the teenage population share, and the teenage enrolment rate in school, were included.

1. Name some of the factors that might be picked up by time and state fixed effects.
2. The author decided to use eight regional dummy variables instead of the 49 state dummy variables. What is the implicit assumption made by the author? Could you test for its validity? How?
3. The results, using time and region fixed effects only, were as follows:

Interpret the result briefly.

**Question 3**

Having learned in macroeconomics that consumption depends on disposable income, you want to determine if disposable income helps predict future consumption. You collect data for the sample period 1962:I to 1995:IV and plot the two variables.

(a) To determine if past values of personal disposable income growth rates help to predict consumption growth rates, you estimate the following relationship.

+

The Granger causality test for the exclusion on all four lags of the GDP growth rate is 0.98. The critical values for the 1%, the 5%, and the 10% level of F4,∞ are 3.32, 2.37, and 1.94 respectively. Make a decision on whether or not these additional variables Granger cause the change in the growth rate of consumption.

(b) You are somewhat surprised about the result in the previous question and wonder, how sensitive it is with regard to the lag length in the model. As a result, you calculate BIC and AIC of p and q from 0 to 6. The results are displayed in the accompanying table:

p,q BIC AIC

0 5.061 5.039

1 5.052 4.988

2 5.095 4.989

3 5.110 4.960

4 5.165 4.972

5 5.206 4.973

6 5.270 4.992

Which values for p and q should you choose?

(c) Estimating an model gives you a t-statistic of 1.28 on the coefficient of lagged disposable income growth. What does the Granger causality test suggest about the inclusion of lagged income growth as a predictor of consumption growth?

**SECTION C**

Answer both question in this section.

1. Explain the concept of omitted variable bias and discuss: [10 marks]
   1. What are the two conditions for omitted variable bias to occur?
   2. What is the direction of the bias?
   3. How, if at all, is the bias related to the size of the sample?
   4. What practical advice for selecting explanatory variables can you offer in light of your answer above. Illustrate your answers with economic examples.
2. [10 marks]
   1. What is the forecast error and how does it differ from the OLS predicted values?
   2. If a regression has a stochastic trend, then the OLS estimator of its coefficient and its OLS t-statistic can have non-standard distributions: briefly discuss the problems associated with this.
   3. A researcher estimated an AR(1) with an intercept and finds that the OLS estimate of β1 is 0.88. Does a 95% confidence interval include β1=1? Explain.