UNIVERSITY OF OSLO
DEPARTMENT OF ECONOMICS

Postponed exam: **ECON3150/4150 – Introductory econometrics**

Date of exam: Thursday, June 19, 2014

Time for exam: 09:00 a.m. – 12:00 noon

The problem set covers 5 pages (incl. cover sheet)

Resources allowed:
- Open book exam, where all written and printed resources, as well as calculator, is allowed

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

Postponed Exam ECON4150: Introductory Econometrics.
June 19; 09:00h-12.00h.

This is an open book examination where all printed and written resources, in addition to a calculator, are allowed. If you are asked to derive something, give all intermediate steps. Do not answer questions with a "yes" or "no" only, but carefully motivate your answer. In the grading, questions 1 and 2 will together count for 2/3 and questions 3 and 4 will together count for 1/3.

Question 1

A researcher wants to investigate the wage returns to a job training program. He has set up an experiment where 400 individuals were randomly assigned to a treatment group (the job training program) and to a control group. After the experiment the researcher collected data for the 400 individuals on wages ($Wage$) (in NOK) and on whether the individual has participated in the job training program ($Training$=1) or not ($Training$=0). The researcher decides to estimate the following regression model by OLS

$$Wage_i = \beta_0 + \beta_1 \cdot Training_i + u_i \tag{1}$$

and obtains the following regression results

```
. regress Wage Training, robust

Linear regression                               Number of obs =        400
                                                   F(  1,   398) =   243.25
                                                   Prob > F      =   0.0000
                                                   R-squared     =   0.3794
                                                   Root MSE      =   52.024
```

Wage	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
Training	81.18112	5.205143	15.60	0.000	70.94811	91.41413
_cons	162.1307	3.741819	43.33	0.000	154.7745	169.4869

a) Give an interpretation, in words, of the two estimated coefficients.

b) The researcher wants to analyze whether there is a difference in the returns to participating in a job training program between men and women. Describe in detail how you would extend model (1), such that you can test the null hypothesis that the wage returns to a job training program do not dependent on gender.

c) Some individuals who were assigned to the treatment group did not participate in the job training program while some individuals who were assigned to the control group did participate in the job training program Do you think that the OLS estimator of β_1 in model (1) is a consistent estimator of the effect of participating in the job training program on wages? Explain why or why not.

d) The researcher decides to use an instrumental variable approach to estimate the returns to participating in a job training program. He uses the assignment to the treatment ($Z_i = 1$) or control group ($Z_i = 0$) as an instrument for whether or not an individual participated in the job training program. The researcher obtains the following first stage OLS estimates.

```
. regress Training Z, robust noheader
```

Training	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Z	.155	.0495038	3.13	0.002	.0576783	.2523217
_cons	.435	.0351433	12.38	0.000	.3659103	.5040897

Do you think that the instrument relevance condition holds? Is Z a weak instrument?

e) The researcher also wants to know whether participating in the program increases the likelihood of being promoted to a higher position within a company. The data set contains an additional variable $Promotion_i$ which is equal to one if an individual has been promoted to a higher position and is zero otherwise. The researcher estimates a probit model and obtains the following estimation results

```
Probit regression                          Number of obs   =        400
                                           LR chi2(    1)      =      15.15
                                           Prob > chi2     =     0.0001
Log likelihood =   -269.18127               Pseudo R2       =     0.0274
```

Promotion	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Training	.4916828	.1268232	3.88	0.000	.2431139	.7402516
_cons	-.1874829	.0903272	-2.08	0.038	-.364521	-.0104447

On the basis of these estimation results, what is the change in the probability that an individual is promoted to a higher position that is associated with participating in the job training program?

Question 2

An economist wants to build a forecasting model for the annualized rate of inflation. He has quarterly data on the inflation rate (`inflation`). Let `d_inflation` be the change in the inflation rate from period $t-1$ to period t.

a) The economist estimates the following AR(1) model $\triangle inflation_t = \beta_0 + \beta_1 \triangle inflation_{t-1} + u_t$ and obtains the following estimation results

```
> regress d_inflation L1.d_inflation if tin(1973q1,2004q4), noheader
```

d_inflation	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
d_inflation L1.	-.2227082	.0871244	[redacted]	[redacted]	[redacted]	[redacted]
_cons	-.0074748	.16348	-0.05	0.964	-.330997	.3160474

Compute a 95% confidence interval for β_1.

b) The economist wants to know how many lags of $\triangle inflation_t$ to include in the autoregression. He estimates $AR(p)$ models for $p = 1, 2$ and 3 over the sample period 1973:1 to 2004:4 (first quarter of 1973 through the fourth quarter of 2004) and obtains the following sum of squared residuals (SSR) for each of the estimated models.

p	**1**	**2**	**3**
SSR	431.014	364.635	352.8687

Use the Bayes Information Criterion (BIC) to estimate the number of lags that should be included in the autoregression.

c) The economist augments the AR(1) model of part (a) with three lagged values of the annualized unemployment rate. The economist computes the Granger-causality F-statistic on the three lags of the unemployment rate and obtains the following results.

```
. test L1.unemployment=L2.unemployment=L3.unemployment=0

 ( 1)  L.unemployment - L2.unemployment = 0
 ( 2)  L.unemployment - L3.unemployment = 0
 ( 3)  L.unemployment = 0

       F( [redacted],   123) =    11.78
            Prob > F =     0.0000
```

Do the unemployment rates help to predict the inflation rate?

d) The previous regressions were based on $\triangle inflation_t$, because the economist is worried that $inflation_t$ has a stochastic trend. Use the estimation results below to test for the presence of a stochastic trend in $inflation$, use a 5% significance level.

```
. regress d_inflation L1.inflation L1.d_inflation L2.d_inflation if tin(1973q1,2004q4)
```

d_inflation	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
inflation						
L1.	-.0933371	.0499211	-1.87	0.064	-.1921449	.0054707
d_inflation						
L1.	-.2491508	.0869439	-2.87	0.005	-.4212371	-.0770645
L2.	-.3516491	.0848594	-4.14	0.000	-.5196097	-.1836886
_cons	.4367815	.2789657	1.57	0.120	-.1153697	.9889326

Question 3

Discuss whether each of the following statements is correct or not.

a) The R^2 can never be equal to zero.

b) Omitted variable bias can be solved by computing heteroskedasticity robust standard errors.

c) By including entity fixed effects in a panel data model you control for omitted variables that vary across entities but not over time.

d) A forecast error is the same as an OLS residual.

e) In a regression without explanatory variables the OLS estimate of the constant term is equal to the mean of the dependent variable.

f) If the first stage F-statistic is larger than 10 the instrument exogeneity condition is satisfied.

Question 4

Consider the following population regression model $W_i = \beta_0 + \beta_1 E_i + u_i$ with $E[u_i|E_i] = 0$. The researcher observes E_i but does not observe W_i, instead he observed a noisy measure $W_i^* = W_i + \varepsilon_i$ where $Cov(W_i, \varepsilon_i) = Cov(E_i, \varepsilon_i) = 0$, $E[\varepsilon_i] = 0$ and $Var(\varepsilon_i) = \sigma_\varepsilon^2$. The researcher estimates the following equation by OLS

$$W_i^* = \beta_0 + \beta_1 E_i + v_i$$

b) Is the OLS estimator of β_1 consistent?