Econometrics 2018, Final Exam, Time to complete : 60 minutes
Minimum points required for a positive grade: 20
Name and student ID _____
Signature _____

This exam contains 4 pages (including this cover page) and 2 questions.
Total of points is 40.

Grade Table (for teacher use only)

Question	Points	Score
Problem 1	20	
Problem 2	20	
Total:	40	

You can find the repository for this exam here:
`https://github.com/feb-uni-sofia/econometriccs2018-exam-2018-06-22`

1. The dataset `trips` contains data on 1991 randomly sampled taxi trips that took place during 2015 in New York, USA. For each trip the data contains:

   tip_amount (numeric): The tip given to the driver in USD.

   trip_distance (Character): Trip distance in miles.

   (a) (2 points) Create a new variable in the dataset `trips` called `short_trip` that equals `TRUE` if the trip distance was shorter than average and equals `FALSE` otherwise.

   (b) (2 points) Fit the linear regression model:

   $$\text{tip\_amount}_i = \beta_0 + \beta_1 \text{short\_trip}_i + u_i \tag{1}$$

   with $i = 1, \ldots, n$ and where u_i are independent random terms with zero mean and constant variance. Note that the `lm` function treats logical variables as 0/1 variables: `TRUE` corresponds to 1 and `FALSE` corresponds to 0.

   (c) (2 points) Write down the estimated regression *equation*.

   (d) (3 points) Let μ_{short} denote the expected tip amount for short trips. Estimate μ_{short} using the estimated regression coefficients.

(e) (3 points) Let μ_{long} denote the expected tip amount for long (above average) trips. Estimate μ_{long} using the estimated regression coefficients.

(f) (2 points) Give an approximate 95% confidence interval for the *difference* between the expected tip amount on short trips μ_{short} and the expected tip amount on long trips μ_{long} using the regression output.

(g) (3 points) A taxi driver claims that the expected tip amount on short trips is equal to the expected tip amount on long trips. Formulate this hypothesis in terms of β_1.

(h) (3 points) Write down the t-test statistic for the hypothesis from (g). Write down the p-value of the test and explain your decision to reject or not to reject the hypothesis at a 95% significance level (5% error probability).

2. The dataset `homeCredit` is a random sample of 2000 loan records from Home Credit, a consumer finance provider that lends to people with little or no credit history. Assume that each row in the data corresponds to a single person who has received a loan from Home Credit. The dataset contains the following columns:

   credit (numeric): Amount of credit received (in USD).

   income (numeric): The person's yearly income in USD.

   age (numeric): Person's age in years.

   (a) (2 points) Create two new variables in the dataset `homeCredit` called `income1000` that equals `income` divided by 1000 and `credit1000` that equals `credit` divided by 1000.

   (b) (2 points) Fit the linear regression model :

   $$\text{credit1000}_i = \beta_0 + \beta_1 \text{income1000}_i + u_i \tag{2}$$

   with $i = 1, \ldots, n$ and where u_i are independent random terms with zero mean and constant variance.

   (c) (3 points) Take a look at the estimated regression coefficients. How would you explain the positive association between the level of income and the expected amount of credit received (short answer)?

   (d) (5 points) Estimate the expected amount of credit for persons with 100,000 USD yearly income and give an approximate 95% confidence interval. Explain the meaning of the confidence interval.

   (e) (2 points) What is the meaning of the intercept β_0 in this model (in the context of *this* data)?

(f) (2 points) Estimate the following linear regression model:

$$\text{credit1000}_i = \beta_0 + \beta_1 \text{income1000}_i + \beta_2 \text{age}_i + u_i \tag{3}$$

with the same assumptions about u_i as in model (2).

(g) (2 points) Interpret the intercept in model (3). Is this a meaningful quantity?

(h) (2 points) A 20 year old person with no income applies for a credit. Estimate the amount of credit she can expect to receive: