First and Last Name:
Student ID No.:

Applied Microeconometrics using Stata

Empirical Part

Prof. Dr. Uwe Jirjahn
M. Sc. Mehrzad B. Baktash
M. Sc. Thi Xuan Thu Le

Summer Semester 2022
60 Points

Task 1: Please choose the correct answer(s) for the following questions, **without the help of Stata**. You may use Stata and move to Task 2 only after completing and handing over the Task 1.
*(**Note:** Multiple correct answers is possible. 2 Points/Question)*

Q1. Which of the following criteria is the most optimal for assessing the goodness of the fit of a multiple linear regression model?
a. Adjusted R^2 **b.** R^2 **c.** The intercept **d.** The coefficient

Q2. What is the post estimation command that you can use after the regress command in Stata to compute the predicted mean-Y values of interest?
a. pcorr **b.** esttab **c.** margins **d.** marginsplot

Q3. What is the name of the statistical test that can help us determine whether to choose a fixed effects or a random effects model?
a. Hausman test **b.** z-test **c.** chi square test **d.** Link-Wallace test

Q4. Which of the following is the command that will give us the mean of a variable?
a. describe **b.** codebook **c.** sum **d.** mean

Q5. Which of the following do we use to type in commands in Stata?
a. Command window **b.** Review window **c.** Variable window **d.** Do-file editor

Q6. Which variant of logistic regression is recommended when you have a categorical dependent variable with more than two values?
a. Logistic regression **b.** Multinomial logistic regression **c.** Ordered logit regression
d. Poisson regression

Q7. What does it mean when we say that our panel is balanced?
a. When we have more time periods than units
b. When we have a large sample of units
c. When we have an equal number of time periods per unit
d. When we have an unequal number of time periods per unit

Q8. Which of the following do we use to create new variable?
a. generate **b.** replace **c.** create **d.** egen

Q9. Which of the following do we use to estimate the determinants of the 'Number of Alcoholic Beverages a Worker Drunk in the Last Month'?

a. Logit   **b.** Probit   **c.** Poisson   **d.** Negative Binomial

Q10. Which of the following codes are wrong?

a. gen age2=age*age
b. gen age2==age*age
c. keep if gender==2
d. keep if gender=2

First and Last Name:
Student ID No.:

Task 2: Please solve the following question **using Stata**. Create a **do-file** and start a **log-file** with your full name and matriculation number in the file names. Do not use the menu or the command line to create your solutions, as the results from these options are not reproducible. Instead, always use and store the appropriate commands in the do-file. Do not forget to close your log-file at the end of your do-file. After the exam, send me your do-file and log-file to **baktash@uni-trier.de**, I will let you know immediately whether I received your files.

Load the **exam_task_2.dta** dataset.

The dataset contains a nationally representative cohort of adults from country X, with the following variables:

Variable name	Definition
id	Person identifier
gender	Male=1, Female=2, Other=3, Missing= -1
t	Survey year
income	Income in Euros
stress	Score of stress (1= Lowest, 10= Highest)
vaccine	Dummy equals 1 if the worker is vaccinated by the time of interview at least once and equals 0 if not.
exp	Work experience in years
birth	The worker's year of birth

Individuals in country X were randomly vaccinated from time t=2021 on. Solve the following:

    a. Identify the treatment and the control groups. (*5p*)
    b. Generate a variable for the age of the workers. (*4p*)
    c. Keep only the male and female workers. (*2p*)
    d. What is the share of male workers in the dataset? (*3p*)
The share of male workers in the dataset is _____ %.
    e. Create a new dummy variable 'female' that identifies females as 1 and males as 0. (*3p*)
    f. Remove all the observations for the workers that report a zero income. (*2p*)
    g. Label the variable 'exp' as 'Work Experience'. (*2p*)
    h. Label the value '1' of stress variable as 'Low Stress' and value '10' as 'High Stress'. (*3p*)
    i. Estimate the influence of getting vaccination on the stress level of the workers. (*6p*)
Interpret the results of (h.) shortly here:

    j. Test for the common trends assumption graphically. (*5p*)
Interpret the results of (i.) shortly here:

    k. Estimate the influence of getting vaccination on the stress level of the female workers who are older than 30 and have at least 5 years of work experience. (*5p*)
Interpret the results of (j.) shortly here: