UNIVERSITY OF OSLO
DEPARTMENT OF ECONOMICS

Exam: ECON3150/4150 – Introductory Econometrics

Date of exam: Friday, 27 May 2022 **Grades are given:** 17 June 2022

Time for exam: 09.00-12.00 (three hours)

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

Resources allowed:

- All written and printed resources – as well as two alternative calculators - are allowed

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

Exam ECON3150/4150: Introductory Econometrics – Spring 2022

1. (80%) Suppose you have the following data from the American "Current Population Survey" from 1992, with average hourly earnings (`ahe`), a dummy variable (`bachelor`) that equals one if a person holds at least a bachelor degree and is zero for those with only a high-school degree, and finally age (`age`):

```
##                   mean        SD         min         max    N
## year     1992.0000000 0.0000000 1992.000000 1992.00000 7612
## ahe        11.6168339 5.6194795    1.242788   46.63414 7612
## bachelor    0.3891224 0.4875832    0.000000    1.00000 7612
## age        29.7104572 2.8063185   25.000000   34.00000 7612
```

   You estimate the following OLS regression:

```
reg = feols(ahe ~ bachelor + age + I(age^2), df, vcov="hetero")
reg
```

```
## OLS estimation, Dep. Var.: ahe
## Observations: 7,612
## Standard-errors: Heteroskedasticity-robust
##                Estimate Std. Error  t value  Pr(>|t|)
## (Intercept) -17.337506   6.990530 -2.48014 0.0131545 *
## bachelor      4.340013   0.128357 33.81212 < 2.2e-16 ***
## age           1.501379   0.479201  3.13309 0.0017363 **
## I(age^2)     -0.019472   0.008155 -2.38772 0.0169773 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 5.13253   Adj. R2: 0.16536
```

```
vcov(reg)
```

```
##             (Intercept)      bachelor         age      I(age^2)
## (Intercept) 48.86750761 -0.0474883117 -3.34648688  5.678874e-02
## bachelor    -0.04748831  0.0164754481  0.00255322 -3.687860e-05
## age         -3.34648688  0.0025532205  0.22963356 -3.904200e-03
## I(age^2)     0.05678874 -0.0000368786 -0.00390420  6.650205e-05
```

   a. Interpret the estimated coefficient on `bachelor`.

   b. Construct and interpret the 68 percent confidence interval for the estimate in 1.a.

c. Can we give the estimate in 1.a a causal interpretation? Motivate your answer.

d. What is the interpretation of the Intercept?

e. How much does a 25-year-old with a bachelor degree earn on average per hour?

f. Compute the average marginal effect of age.

g. Compute the standard error of the estimate in 1.f.

h. Suppose you want to test the joint significance of the age profile at the 5% level. Explain how you would go about testing this and what exact critical value you would use.

A friend suggests to estimate the following regression instead:

```
feols(log(ahe) ~ bachelor + age, df, vcov="hetero")
```

```
## OLS estimation, Dep. Var.: log(ahe)
## Observations: 7,612
## Standard-errors: Heteroskedasticity-robust
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 1.383033   0.055445 24.9440 < 2.2e-16 ***
## bachelor    0.374641   0.010504 35.6653 < 2.2e-16 ***
## age         0.027219   0.001856 14.6683 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 0.453515   Adj. R2: 0.154567
```

i. Your friend claims that this regression is better. What do you reply?

j. Interpret the estimated coefficient on `age`. Is this similar to your results above?

k. In a next step your friend wants to investigate the hypothesis that people with a bachelor degree have *steeper* age profiles than those with only a high-school degree. Explain how to do this.

2. (20%) In the Netherlands applicants to medical school used to be admitted solely based on a random lottery. Below you find data on such lotteries with information on admissions (`admitted`), medical school degrees (`medschool`), and later income when applicants were about 35 years old (`income`):

```
xtabs( ~ medschool + admitted, df)
```

```
##          admitted
## medschool    0    1
##         0 1162  154
##         1  965 2409
```

```
aggregate(income ~ medschool + admitted, df, mean)
```

```
##   medschool admitted   income
## 1         0        0 21.58743
## 2         1        0 20.80111
## 3         0        1 24.04733
## 4         1        1 23.25962
```

Use these data to

a. estimate the causal effect of medical school on income,

b. state your assumptions and

c. discuss their validity in the current setting, providing support from the data where possible, and

d. interpret your findings.