PROBLEM SET 5
Due back by Tuesday May 2.


Name: BONANNO DAVID

Id: 187841

Instructions:

- Make sure you are working on your problem set as each problem set is different.

- **The answers to the questions of this problem set are to be given exclusively in the answer sheet**

- **The answers sheet MUST be printed and not photocopied. Photocopies will not be accepted.**

- Questions marked with the symbol ♣ admit more than one correct answer

- Please fill the boxes in the answer sheet completely using a **black pen** as follows

<p align="center">Question 1: ▨ B C D E</p>

- The answer sheet must not be creased or folded otherwise your problem set won't be graded.

- You can hand back your problem set at the END of class on May 2.
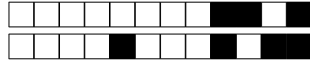
1

With a sample of 706 observations, we estimate the following model:

$$ln(hwage_i) = \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 educ_i + \beta_4 yngkid_i + u_i$$

and obtain these results:

$$ln(hwage_i) = \underset{0.41175}{-1.00179} + \underset{0.01941}{0.07121} age_i - \underset{0.00024}{0.00076} age_i^2 + \underset{0.01116}{0.07131} educ_i + \underset{0.07006}{0.09439} yngkid_i$$

where *lhwage* is the logarithm of the hourly wage in euro, *age* is measured in years, *educ* is years of education and *yngkid* is a variable equal to 1 in case the person has a child younger than three years.

Question 1   What is the interpretation of β_2?

A Increasing age by one year, the hourly wage decreases by 0.077% on average, ceteris paribus.

B Increasing the square of age by one year, the hourly wage decreases by 0.077% on average, ceteris paribus.

C Increasing the square of age by one year, the hourly wage decreases by 0.00077 euros on average, ceteris paribus.

D By itself does not have a proper interpretation.

Question 2   What is the interpretation of β_3?

A One year more of education is associated with a change of about 0.07 euros in hourly wage, on average, ceteris paribus.

B One year more of education is associated with a change of about 7% in hourly wage, on average, ceteris paribus.

C One year more of education is associated with a change of about 0.0007 in hourly wage, on average, ceteris paribus.

D An increase of 1% in education is associated with a change of about 7% euros in hourly wage, on average, ceteris paribus.

E One year more of education is associated with a change of about 0.07% in hourly wage, on average, ceteris paribus.

F An increase of 1% in education is associated with a change of about 0.07 euros in hourly wage, on average, ceteris paribus.

Question 3   Is β_1 statistically significant?

A It is not at 1% level

B We cannot check for this, it makes no sense.

C It is not at 5% level.

D It is at 10% level.

Question 4   What are we testing when we check whether β_2 is significant?

A We check whether the logarithm of hourly wage depends negatively on age.

B We check whether the logarithm of hourly wage depends on age.

C We check whether the logarithm of hourly wage depends positively on age.

D We check whether the logarithm of hourly wage depends linearly on age.

Question 5   Keeping other variables fixed, at what age the logarithm of hourly wage is maximized?

A At about 46.7 years.

B At about 0, but this makes no sense.

C At about 93.3 years.

D At about 56.3 years.

Question 6   Using a subset of the variables in the previous model, we would like to write a new one such that we obtain the elasticity of the hourly wage to education, and that, given in increase of one year in age, it returns a change in hourly wage in percent points. Choose the correct model among these:

A $ln(hwage_i) = \beta_0 + \beta_1 ln(age_i) + \beta_2 educ_i + u_i$

B $ln(hwage_i) = \beta_0 + \beta_1 age + ln(\beta_2 educ_i) + u_i$

C $hwage_i = \beta_0 + \beta_1 ln(age_i) + \beta_2 educ_i + u_i$

D $hwage_i = \beta_0 + \beta_1 age_i + \beta_2 ln(educ_i) + u_i$

E $ln(hwage_i) = \beta_0 + \beta_1 age_i + \beta_2 ln(educ_i) + u_i$

2

Let us define with Y the amount of cholesterol in mlg in the blood and with Med a dummy variable which takes the value of 1 for medication B and 0 for medication A, where A and B are two different medications that lower cholesterol. Female is a dummy variable which takes the value of 1 for females and 0 otherwise.

Consider the following regression:

$$Y = \beta_0 + \beta_1 \times med + \beta_2 \times female + \beta_3 \times med \times female + u.$$

Question 7   What is the effect of using medication B with respect to medication A for men?

A $\beta_0 + \beta_1$

B None of the others.

C $\beta_1 - \beta_0$

D β_1

E β_0

Question 8   What is the average cholesterol value for women using medication A?

A β_0

B None of the others.

C β_2

D $\beta_0 + \beta_2$

E $\beta_0 + \beta_2 + \beta_3$

Question 3

These data are taken from the Medical Expenditure Panel Survey survey conducted in 1996. These data were provided by Professor Harvey Rosen of Princeton University and were used in his paper with Craig Perry "The Self-Employed Are Less Likely Than Wage-Earners to Have Health Insurance. So What?" in Douglas Holtz-Eakin and Harvey S. Rosen, eds., Entrepeneurship and Public Po licy, MIT Press 2004.

Among the variables in the dataset, `ins` is a dummy equal to one if the interviewee has the insurance; `selfemp` is equal to one if the interviewee is a self-employed workers; `gender` is equal to one if the in dividual is a male; `married` is one if the individual is married; `health` is one if the individual reports to be in good health; `educ` is 0 if the person has no education, 1 if he/she achieved middle school diploma, 2 for the high school diploma, 3 for the bachelor degree, 4 for the master degree and 5 for the PhD; `age` is in years and `age2` is the square of age.

We estimate two models:

$$Pr(ins = 1|X) = \beta_0 + \beta_1 \times selfemp + \beta_2 \times married + \beta_3 \times gender + \beta_4 \times health$$
$$+\beta_5 \times gender * health + \beta_6 \times educ + \beta_7 \times age + \beta_8 \times age^2$$

```
Coefficients:
                Estimate Std. Error t value  Pr(>|t|)
(Intercept)    0.2974634  0.0580248    5.13  0.0000003
selfemp       -0.1742361  0.0141740  -12.29  < 2e-16
married        0.1181062  0.0094187   12.54  < 2e-16
gender        -0.0232270  0.0343575   -0.68  0.49903
health         0.0744310  0.0247243    3.01  0.00262
genderxhealth -0.0206248  0.0353131   -0.58  0.55920
educ           0.0529807  0.0029210   18.14  < 2e-16
age            0.0105315  0.0027482    3.83  0.00013
age2          -0.0000788  0.0000333   -2.37  0.01796
---

Heteroskadasticity robust standard errors used
```

$$Pr(ins = 1|X) = \Phi(\beta_0 + \beta_1 \times selfemp + \beta_2 \times married + \beta_3 \times gender + \beta_4 \times health$$
$$+ \beta_5 \times gender * health + \beta_6 \times educ + \beta_7 \times age + \beta_8 \times age^2) \quad \text{(II)}$$

```
Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)    -0.844932   0.195991   -4.31 0.000016
selfemp        -0.651923   0.046842  -13.92  < 2e-16
married         0.455241   0.034845   13.06  < 2e-16
gender         -0.040238   0.111653   -0.36  0.71856
health          0.300503   0.082988    3.62  0.00029
genderxhealth  -0.124880   0.116613   -1.07  0.28422
education       0.226139   0.012852   17.60  < 2e-16
age             0.029150   0.009899    2.94  0.00323
age2           -0.000162   0.000126   -1.29  0.19821
---
```
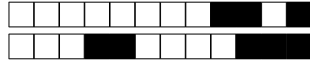
Question 9

Under model (I), what is the effect of being in good health on the probability of having an insurance, for men and women respectively?

A 7.44% for women, -2,06% for men.

B 7.44% for women, 5.38% for men.

C 7.44% for men, 5.38% for women.

D 7.44% for men, -2,06% for women.

Question 10   What is the intepretation of β_2 in model (II)?

A It does not have a proper interpretation in terms of magnitude.

B On average, a married worker has a probability of 45.5% to have an insurance, *ceteris paribus*.

C On average, married individuals are 45.5% less likely than others to have an insurance, controlling for all other factors.

D On average, married individuals are 45.5% more likely than others to have an insurance, controlling for all other factors.

Question 11   Does increasing level of education have a significant impact on the probability to buy an insurance at 5% level under model (II)?

A Yes, since the model includes the variable "educ".

B It depends on the values of all other covariates.

C No, since the coefficient β_6 is not significant.

D Yes, since the coefficient β_6 is significant.

Question 12   Does increasing level of education have a positive impact on the probability to buy an insurance at 5% level under model (II)?

A No, since the coefficient β_6 is not significant.

B Yes, since the t-test on the coefficient β_6 is higher than 1.64.

C Yes, since the model includes the variable "educ".

D It depends on the values of all other covariates.

Question 13 ♣   Under model (II), which of the following statements are true?

A On average, self employment has a negative effect on the probability of having an insurance, *ceteris paribus*.

B We cannot interpret directly the estimated coefficients in terms of magnitude.

C We cannot possibly know whether education level has an effect on the probability of having an insurance.

D The probability of having an insurance for a married person is, on average, 45.5% higher than for other people, *ceteris paribus*.

E *None of these answers are correct.*

Question 14 ♣   Under model (I), which of the following statements are true?

A Statistically speaking, having a good health has the same impact on the probability of having an insurance for men and women.

B On average, self employed people are less likely to have an insurance, controlling for all other factors.

C The probability of having an insurance is linked to age in a linear fashion.

D The older an individual grows, the higher the probability of having an insurance.

E *None of these answers are correct.*

Question 15   In model (I) we used heteroskedastic-robust standard errors. Could have we used homoskedastic standard errors instead?

[A] No, because homoskedasticity rarely occurs in nature.

[B] Only if we knew that the errors were in facts homoskedastic.

[C] No, because the errors are very likely to autocorrelated.

[D] No, because homoskedasticity is not possible when the dependent variable is a dummy.
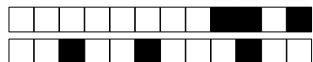
Question 16   What is the estimated probability that a 24 years old non-educated non-self-employed single female who is not in good health and with middle school diploma will buy insurance coverage?

[A] More than 95%.

[B] Less than 5%.

[C] Approximately 0.5.

[D] I cannot say because the model is nonlinear.
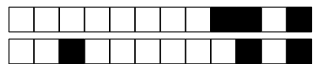
Name: **BONANNO DAVID** Id: **187841**

- Answers must be given exclusively on this sheet: answers given on the other sheets will be ignored.

- **This sheet MUST be printed out and not photocopied. Photocopies will not be accepted.**

- Please fill the boxes below completely using a **black pen**.

- Do not crease or fold.

- Due back: Tuesday, May 2 (right after class)

Question 1: A B C D

Question 2: A B C D E F

Question 3: A B C D

Question 4: A B C D

Question 5: A B C D

Question 6: A B C D E

Question 7: A B C D E

Question 8: A B C D E

Question 9: A B C D

Question 10: A B C D

Question 11: A B C D

Question 12: A B C D

Question 13: A B C D E

Question 14: A B C D E

Question 15: A B C D

Question 16: A B C D