

# Panel Data Analysis of Microeconomic Decisions

## Assignment Part II

Due Date: December 5 2022

### Assignment

Your assignment should include the relevant Stata code and output, as well as your answers to the questions (maximum 10 pages in total). Note that relevant Stata output should be in the text (not in the Appendix) and counts towards the page limit. Code should be added at the end of the assignment and does not count towards the page limit. Each student must submit individual answers that differ from the answers of the other students. Please submit your assignment on Canvas. If you have any questions, please send them by email to [w.chen@uvvt.nl](mailto:w.chen@uvvt.nl).

### Data

For this assignment we use a US sample from the National Longitudinal Survey of Youth 1979 (NLSY79). The NLSY79 is a nationally representative sample of 12,686 young men and women who were 14-22 years old when they were first surveyed in 1979. These individuals were interviewed annually until 1994 and once every two years after that. The last available round of the survey was conducted in 2016.

The sub-sample you are given comprehends 11,257 individuals ( $N$ ) observed through 12 waves of surveys ( $T$ ), for a total of 135,084 observations ( $N \cdot T$ ). Only the period from 1994 to 2016 is considered. You can find the data in the file `data.dta`. The variables in the data set are given in the Table 1 below.

The students will work on different samples composed of eight consecutive waves. The sample you have to use depends on the last digit of your student number (SNR):

- if the last digit is 0 or 1 you only use waves 1 to 8 (i.e. from 1994 to 2008)

- if the last digit is 2 or 3 you only use waves 2 to 9 (i.e. from 1996 to 2010)
- if the last digit is 4 or 5 you only use waves 3 to 10 (i.e. from 1998 to 2012)
- if the last digit is 6 or 7 you only use waves 4 to 11 (i.e. from 2000 to 2014)
- if the last digit is 8 or 9 you only use waves 5 to 12 (i.e. from 2002 to 2016)

Variable Name	Description
ID	Identification number
YEAR	Year of the survey
WAVE	Survey wave
SEX	1 for women, 0 for men
AGE	Age at the time of the survey
MARRIED	1 for married, 0 otherwise
WHITE	1 for white people, 0 otherwise
INCOME	Total income from wage and salary
SP_INC	Total income from wage and salary of the spouse
EMPL	1 for employed, 0 otherwise
YOUNG_CH	1 for child less than 11, 0 otherwise
NUM_CH	Number of children
EDU_12	1 if highest degree is high school level, 0 otherwise
EDU_13_15	1 if the highest degree is undergraduate level, 0 otherwise
EDU_15	1 if the highest degree is graduate level, 0 otherwise
JOB_SAT	Job satisfaction on a scale from 1 (like it very much) to 4 (dislike it very much)

Table 1: Variables

# Questions

## 1 Binary choice models

1. Use a static random effects logit model to explain employment (EMPL) from age, age squared, marital status, presence of young children, number of children, education, and a dummy for white people. Use people with at most a high school degree as the reference group for the education dummies. Comment on the results.

2. Is the model in question 1 better than a pooled logit model? Motivate your answers.
3. Using the model of question 1, compute the marginal effect of a one year increase in age on the probability to be employed for an observation with probability 0.5 to be employed and age 40 years old. At which age would the marginal effect of age be equal to zero?
4. Consider the specification used in question 2, but now run a static fixed effects model and comment on the results. Carry out a test to choose between the random and the fixed effects specification.
5. Estimate a quasi fixed effects model that allows the individual effects to be correlated with education (you can assume that the parameters that determine the relation between education and the individual effects are the same across years). Comment on the results.
6. Unemployment can be quite persistent over time and you wonder whether state dependence plays any role. Estimate a dynamic version of the random effects logit model using the Wooldridge approach and comment on the results. Discuss the importance of unobserved heterogeneity and state dependence.

## 2 Tobit model

1. Estimate a static random effects tobit model to explain INCOME from the dummy for sex, age, age squared, white, marital status and education dummies. Comment on what you see as the main results.
2. Discuss the importance of unobserved heterogeneity according to this model.
3. Investigate whether it is useful to add year dummies to your model in question 1.
4. Investigate whether the effect of marital status is different for males and females.
5. Estimate the dynamic version of the Tobit model you estimated in question 1 using the Wooldridge approach. Discuss the importance of unobserved heterogeneity and state dependence.
6. Test whether the individual effects are correlated with the initial value of INCOME in the model of question 5.

### 3 Ordered response models

1. Estimate a static random effects ordered logit model to explain job satisfaction (JOB\_SAT) from age, income and education. In this case we are only interested in the sub-sample of employed individuals.
2. Why does the output not include an estimate for the constant term?
3. Comment on the likelihood ratio test reported in the bottom part of the Stata output. What are the null and the alternative hypothesis, the test statistic and the conclusions?
4. Estimate a static fixed effects ordered logit model to explain job satisfaction (JOB\_SAT) from age, income and education. In this case, we are also only interested in the sub-sample of employed individuals. Comment on the difference in estimates compared to the random effects ordered logit. *Hint: you can use the `feologit` function.*
5. Carry out the Hausman test for the coefficient of EDU\_15 in using random effects ordered logit against fixed effects ordered logit. Write down all your calculations as well as the steps of the test.
6. Select your favorite ordered response model from what you have already estimated and other models you consider interesting. Motivate all your steps in the selection process.
7. Compute the marginal effect of EDU\_15 on the probability that JOB\_SAT=1 for an observation with probability of JOB\_SAT=1 equal to 0.5 and compute the probability that JOB\_SAT=4 for an observation with probability of JOB\_SAT=4 equal to 0.5 *for your favorite model selected in question 6.*