Prof. Dr. Kai Carstensen Professor of Econometrics

Home Assignment - Econometric Methods

Winter Term 2023/24

Deadline: 04 February 2024

Preliminary remarks:

1. Please read these instructions carefully!
2. This is a **voluntary** exercise.
3. You can earn up to six points that are valid for both examination periods of the WS2023/24 and will be added to the points that you obtain in your exam.
4. Use Stata as a software package.
5. You are allowed to work in groups (up to 5 participants).
6. Create all necessary outputs by Stata commands and answer open questions as comments (* in Stata) in your do-file.
7. Please submit a single `do-file` with answers to: l.tran@stat-econ.uni-kiel.de until **Feb 4th, 11:59pm**.
8. Write name, enrolment and stu-numbers of each group member as a comment at the top of your `do-file`!
9. Please name your files `HA_EconMethods_Name.do` with the surname of the participant responsible for submitting them.

Good luck!

This exercise is based on Evans, William, N., Matthew C. Farrelly, and Edward Montgomery (1999): "Do Workplace Smoking Bans Reduce Smoking." *American Economic Review*, 89 (4): 728-747.

Between 1986 and 1993, the time period after the Surgeon General's report on passive smoke, smoking participation rates among workers fell 2.6 percentage points (standard error of 0.010) more than the decline for nonworkers .... One possible explanation for this phenomenon is that the rise in workplace smoking bans over this time period has reduced workers' demand for cigarettes.

In this assignment, you will analyze the impact of workplace smoking bans on smoking behavior. The dataset `HA_smoking.dta` [1] provides information of 10,000 indoor workers on the individuals' firm smoking policies, smoking behavior and characteristics. Below is short description of variables.

Variables	Description
smoker	=1 if current smoker, =0 otherwise
smkban	=1 if there is a work area smoking ban, =0 otherwise
age	age in years
hsdrop	=1 if high school dropout, =0 otherwise
hsgrad	=1 if high school graduate, =0 otherwise
colsome	=1 if some college, =0 otherwise
colgrad	=1 if college graduate, =0 otherwise
female	=1 if female, =0 otherwise

Note: The educational binary indicators refer to the highest level attained and thus are mutually exclusive. An individual with a Master's degree or higher has values of 0 for hsdrop, hsgrad, colsome, and colgrad.

Questions

1. (**4P**) Consider the linear probability model

$$smoker_i = \alpha + \beta_1 smkban_i + \gamma Z_i + e_i \tag{1}$$

where Z are control variables including age, age^2, $female$, $hsdrop$, $hsgrad$, $colsome$ and $colgrad$. Assume there is no omitted variable bias.

   (a) Import and briefly describe the dataset.

   (b) Estimate the regression using the OLS method with heteroscedasticity robust standard errors. Interpret $\hat{\beta}_1$. Is it statistically significant?

[1] Dataset are provided on textbook website 'Introduction to Econometrics' by Stock and Watson.

(c) Based on the regression output, does the probability of smoking increase or decrease with the level of education? Perform the Wald test that high school dropout and high school graduate plays the same role on the smoking probability, i.e. $H_0 : \gamma_{hsdrop} = \gamma_{hsgrad}$.

(d) Calculate the predicted probability of smoking for a woman who is subject to a workplace smoking ban, 70 years old and graduate college. Based on your result, discuss why the linear probability model is not appropriate in this case.

2. (**2P**) Consider the probability model

$$smoker_i = G(\alpha + \beta_1 smkban_i + \gamma Z_i) + e_i \tag{2}$$

Again, we assume there is no omitted variable bias.

(a) Estimate the average partial effects of smoking ban on probability of smoking using i) probit and ii) logit model. Are the effects statistically significant? Compare the results with the one obtained from the linear probability model. To this end, generate an output table that includes all three estimation results.
Hints: To calculate the average partial effect of binary variable *smkban*, you can use syntax *i.variable* in regression for binary variable. For example,
logit smoker i.smkban age, vce(oim)
margins, dydx(i.smkban)

(b) Use the probit model to calculate the effect of smoking ban on the smoking probability for groups:
i) male, 40 years old, college graduate.
ii) female, 20 years old, high school dropout.