

ECON 3161 Homework 5 (Fall 2023)

Due Date: Dec 3 (Sunday)

Instruction: There are 60 points in total. This homework needs computer coding. Stata software is recommended, but you can use other software you prefer. You can either type or handwrite your answer. Then upload your answer file on canvas.

1. (25 pts) In class we discussed the latent variable model

$$y_i^* = \beta_0 + \beta_1 x_i + u_i, \quad (1)$$

and

$$y_i = \mathbb{I}\{y_i^* > 0\} = \begin{cases} 1, & \text{if } y_i^* > 0 \\ 0, & \text{if } y_i^* \leq 0. \end{cases} \quad (2)$$

- (1) (10 pts) Suppose $u_i \sim N(0, 1)$, show that

$$P(y_i = 1|x_i) = \Phi(\beta_0 + \beta_1 x_i), \quad (3)$$

where $\Phi(\cdot)$ is the normal CDF. (Note: you need to show your steps)

- (2) (5 pts) What is the name for model (3)?

- (3) (5 pts) How would you use the maximum likelihood method to estimate the coefficients β_0, β_1 in model (3)? Write down the log-likelihood function.

- (4) (5 pts) If one uses a linear probability model, i.e., a linear regression model with a binary dependent variable as below

$$y_i = \beta_0 + \beta_1 x_i + u_i, \quad (4)$$

what is the disadvantage of the above model (4)?

2. (35 pts) Use the data in LOANAPP for this exercise. The dataset has been posted in the HW5 folder. The binary variable to be explained is *approve*, which is equal to one if a mortgage loan to an individual was approved. The key explanatory variable is *white*, a dummy variable equal to one if the applicant was white. The other applicants in the data set are black and Hispanic. To test for discrimination in the mortgage loan market, a linear probability model can be used:

$$approve = \beta_0 + \beta_1 white + u. \quad (5)$$

(1) (5 pts) If there is discrimination against minorities, and the appropriate factors have been controlled for, what is the sign of β_1 ?

(2) (10 pts) Regress *approve* on *white* and report the results in the usual form. Interpret the coefficient on *white*. Is it statistically significant?

(3) (10 pts) Estimate a probit model of *approve* on *white*. Find the estimated probability of loan approval for both whites and nonwhites. How do these compare with the linear probability estimates?

(4) (10 pts) Now Estimate a logit model of *approve* on *white*. Find the estimated probability of loan approval for both whites and nonwhites.

Hint: If you use stata, you can use command *probit y x* to estimate a probit model of *y* on *x*, and *logit y x* to estimate a logit model of *y* on *x*. If you use R, you can use command *glm(y ~ x, family=binomial(link="probit"))* for probit model, and *glm(y ~ x, family=binomial(link="logit"))* for logit model.