ECON 4151
Lab Session 5: Multinomial logit/probit models and Tobit model

Wonjin Lee

Washington University in St. Louis

November 13, 2020

Outline

1. Multinomial regression models (Gujarati Ch. 9 and 10)

2. Tobit model (Wooldridge Ch. 17)

Multinomial regression models (MRM)

- Nominal (or unordered) MRM
  1. Chooser or individual-specific data $\implies$ multinomial logit model
     - Regressors vary across individuals.
     - Ex. demographics
  2. Choice-specific data $\implies$ conditional logit model
     - Regressors vary across alternatives.
  3. Both $\implies$ conditional logit model with dummies (mixed logit model)
- Ordered MRM
- Check the data structure!

Exercise (Gujarati Ch. 9 and 10)

- Multinomial logit model: 9.2 School choice
- Conditional logit model: 9.3 Choice of travel mode
- Mixed logit model: 9.4 Choice of travel mode
- Ordered logit model: Exercise 10.3 Firm's rating score

Tobit model

- Tobit model:

$$y_i^* = x_i'\beta + u_i, \quad u|x \sim \mathcal{N}\left(0, \sigma^2\right),$$
$$y_i = \max\left\{0, y_i^*\right\}$$

- Ex. hours of work, earnings, ...
- An alternative to the Tobit model is the Heckit model which is simpler but not as efficient as the ML estimator.
- The density of y given x:
  - For $y > 0$

$$f\left(y_i|x_i\right) = \frac{1}{\sigma}\phi\left(\frac{y_i - x_i'\beta}{\sigma}\right)$$

  - For $y = 0$

$$P\left(y_i = 0|x_i\right) = 1 - \Phi\left(\frac{x_i'\beta}{\sigma}\right)$$

MLE

- Log-likelihood function:

$$\ln \mathcal{L}(\beta, \sigma) = \sum_{i=1}^{N} \left[\mathbb{1}_{\{y_i=0\}} \ln \left(1 - \Phi \left(\frac{x_i'\beta}{\sigma} \right) \right) \right.$$
$$\left. + \mathbb{1}_{\{y_i>0\}} \left[\ln \left(\phi \frac{y_i - x_i'\beta}{\sigma} \right) - \ln \sigma \right] \right]$$

- Maximize it to get $\hat{\beta}$ and $\hat{\sigma}$.

Tobit estimates

- Conditional expectation:

$$\mathbb{E}(y|y>0,x) = x'\beta + \mathbb{E}\left(u|u>-x'\beta\right) = x'\beta + \sigma\frac{\phi\left(\frac{x'\beta}{\sigma}\right)}{\Phi\left(\frac{x'\beta}{\sigma}\right)}$$

  - $\phi(c)/\Phi(c)$ is the inverse Mills ratio.
  - In the OLS, the inverse Mills ratio is an omitted variable $\implies$ β is not consistently estimated.

- Unconditional expectation:

$$\mathbb{E}(y|x) = P\left(y>0|x\right)\mathbb{E}(y|y>0,x) = \Phi\left(\frac{x'\beta}{\sigma}\right)x'\beta + \sigma\phi\left(\frac{x'\beta}{\sigma}\right)$$

- Either case, we have a nonlinear function of x and β.
- Partial effect at the average and average partial effect are defined just like logit/probit models.

One more comment

- $\beta \implies$ the marginal impact of a variable on the mean value of the latent variable, holding other variables constant.
- Multiply β by the probability of observing the latent variable $\implies$ the marginal impact of a variable on the observed value of the latent variable, holding other variables constant.
- The ML estimators are consistent only if the assumptions about the error term are valid; otherwise, inconsistent (ex. heteroskedasticity and non-normal error).

Exercise (Wooldridge Ch. 17)

- Tobit model: Married women's labor supply