ECON 4151
Lab Session 7: Heckit model

Wonjin Lee

Washington University in St. Louis

November 20, 2020

Outline

1. Heckit model (Wooldridge Ch. 17)

Selectivity: Heckit model

- A common form of sample selection is incidental truncation.
  - We always observe the explanatory variables x; but we only observe y for a subset of the population.
  - The rule determining whether we observe y does not depend directly on the outcome of y; it depends on another variable (an exogenous variable) $\Longrightarrow$ exogenous sample selection.
- Ex. the wage offer, or the hourly wage
  - If the person is actually working at the time of the survey, then we observe the wage offer b/c we assume it is the observed wage.
  - But for the people out of the workforce, we cannot observe the wage offer.
  - The truncation of wage offer is incidental b/c it depends on another variable, labor force participation.
  - Note that we observe all other information about an individual, such as education, prior experience, gender, marital status, and so on.

Selectivity: Heckit model

- Heckit model:

$$y_i = x_i'\beta + u_i, \quad \mathbb{E}[u|x] = 0$$
$$s_i = \mathbb{1}(z_i'\gamma + \nu_i \geq 0) \;\; : \text{selection eq'n}$$

  where $s = 1$ if we observe y, and $s = 0$ otherwise.
  - x and z are always observed and $x \subset z$
  - ν is an unobserved error and has a standard normal distribution.
  - Assume (u, ν) is independent of z.
- Assume

$$\begin{pmatrix} u \\ \nu \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{pmatrix} \right]$$

- $\text{Corr}(u, \nu) \implies$ a sample selection problem.

Selectivity: Heckit model

- Assuming u and ν are jointly normal (with zero mean), we get

$$\mathbb{E}[y|z, s = 1] = x'\beta + \rho\lambda(z'\gamma)$$

  where ρ is the correlation coefficient b/w u and ν.
  - We have two regressors: x and λ.
  - The expected value of y given z and observability of y is equal to $x'\beta$ plus an additional term that depends on the inverse mills ratio evaluated at $z'\gamma$.
  - If $\rho = 0$, OLS of y on x using the selected sample consistently estimates β.
- Take the expected value of the selection equation, then

$$\Pr(s = 1|z) = \Phi(z'\gamma)$$

  - Estimate γ by probit model of s_i on z_i using the entire sample.

Sample selection correction

- Steps:
  1. Use all the observations. Estimate a probit model of s_i on z_i and get $\hat{\gamma}$. Compute the inverse Mills ratio, $\hat{\lambda}_i = \lambda(z_i'\hat{\gamma})$.
  2. Use the selected sample (observations for which $s_i = 1$). Run the regression of y_i on x_i and $\hat{\lambda}_i$.