Limited Dependent Variable Models

Semester 2/2013

Part 4: Tobit model for corner solution responses, Censored and Truncated regression models

Chayanee Chawanote

Corner solution response

- A dependent variable is zero for a nontrivial fraction of the population, but is continuously distributed over positive values
- Example: the amount of expenditure an individual spends on alcohol in a given month
  - For some significant fraction of the population, the amount spent on alcohol is 0
- What if we use a linear model for this dependent variable?
  - We could possbily obtain negative fitted values, similar to the problems with LPM for binary outcomes
  - The distribution of y piles up at zero. So, y cannot have a conditional normal distribution as in a linear model

Tobit model

- Define an underlying latent variable:
  $y^* = \beta_0 + \mathbf{x}\beta + u,\ u|x \sim Normal(0, \sigma^2)$ (1)
- Our observed dependent variable:
  $y = max(0, y^*)$ (2)
- y* has a normal, homoskedastic distribution with a linear conditional mean (the classical linear model assumptions)
- y has a continuous distribution over strictly positive values

Tobit model

- The density of y given x is the same as the density of y* given x for positive values
- For y=0, $P(y=0|x) = P(y^* < 0|x) = P(\frac{u}{\sigma} < -\frac{x\beta}{\sigma}|x)$
  $= \Phi(-\frac{x\beta}{\sigma}) = 1 - \Phi(\frac{x\beta}{\sigma})$ (3)
- For y>0, the density of y_i given x_i is
  $\frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{(y-x_i\beta)^2}{2\sigma^2}\right] = \frac{1}{\sigma}\phi\left[\frac{y-x_i\beta}{\sigma}\right]$ (4)
- Log likelihood is

Interpreting the Tobit estimates

- In Tobit models, we have 2 expections of interest:
- $E(y|y > 0, \mathbf{x})$: 'conditional expection' as it is conditional on y>0
- $E(y|\mathbf{x})$: 'unconditional expection'
- Given $E(y|y > 0, \mathbf{x})$, we can fine $E(y|\mathbf{x}) = P(y > 0|x)E[y|y > 0, x]$ (5)
- Note: if $z \sim N(0, 1)$, then $E[z|z > c] = \frac{\phi(c)}{1-\Phi(c)}$ for any constant c

Interpreting the Tobit estimates

- $E(y|y>0,\mathbf{x}) = \mathbf{x}\beta + \sigma\lambda\left(\frac{\mathbf{x}\beta}{\sigma}\right)$, $\lambda\left(\frac{\mathbf{x}\beta}{\sigma}\right) = \frac{\phi(\mathbf{x}\beta/\sigma)}{\Phi(\mathbf{x}\beta/\sigma)}$ "inverse Mills ratio" (6)
- $E(y|\mathbf{x}) = P(y>0|x)E[y|y>0,x] = \Phi\left(\frac{\mathbf{x}\beta}{\sigma}\right)\mathbf{x}\beta + \sigma\phi\left(\frac{\mathbf{x}\beta}{\sigma}\right)$ (7)
- Partial effects of x_j on $E[y|y>0,\mathbf{x}]$:
  $\frac{\partial E[y|y>0,x]}{\partial x_j} = \beta_j\left\{1-\lambda\left(\frac{\mathbf{x}\beta}{\sigma}\right)\left[\frac{\mathbf{x}\beta}{\sigma}-\lambda\left(\frac{\mathbf{x}\beta}{\sigma}\right)\right]\right\}$ (8)
    - same sign with β_j
- Partial derivative of $E[y|x]$ with respect to x_j:
  $\frac{\partial E[y|\mathbf{x}]}{\partial x_j} = \beta_j\Phi\left(\frac{\mathbf{x}\beta}{\sigma}\right)$ (9)

Interpreting the Tobit estimates

- The partial effect at the average: evaluating the adjustment (scale) factor, $\Phi(\bar{x}\hat{\beta}/\hat{\sigma})$
- The average partial effect: compute the scale factor as $\frac{1}{n}\sum_{i=1}^{n}\Phi(\mathbf{x_i}\hat{\beta}/\hat{\sigma})$
- Notice that the scale factor is always between zero and one: $0 < \Phi(x\hat{\beta}/\hat{\sigma}) < 1$ since $\hat{P}(y_i > 0|\mathbf{x_i}) = \Phi(\bar{x}\hat{\beta}/\hat{\sigma})$
  - If there are few observations with $y_i = 0$, the scale factor will be closer to one.

Censored and Truncated Regression Models

- Censoring: missing data on the response variable
- Truncation: missing data on both the response variable and explanatory variables

Censored Regression Models

- Let y follows the classical linear model:
  $y_i = \mathbf{x}_i\beta + u_i,\ u_i|x_i,\ c_i \sim N(0, \sigma^2)$ (10)
  $w_i = min(y_i, c_i)$ (11)
  - We observe y_i only if it is less than a censoring value, c_i
  - Assume u_i is independent of c_i
  - This setup is right censoring, e.g. top coding data
- If using OLS regression, it will use only the uncensored observations ($y_i < c_i$), producing inconsistent estimators of β_j
- This is similar to corner solution response (in Tobit model with economic behavior resulting in zero outcomes), but here it is a data collection problem. For some reason, the data are censored!

Censored Regression Models

- Estimating by using maximum likelihood
- For uncensored observations ($w_i = y_i$), the density of w_i is the same as that of y_i : $N(\mathbf{x}\beta, \sigma^2)$
- For censored observations ($w_i = c_i$), $P(w_i = c_i|\mathbf{x}_i) = P(y_i \geq c_i|\mathbf{x}_i) = P(u_i \geq c_i - \mathbf{x}_i\beta) = 1 - \Phi[(c_i - \mathbf{x}_i\beta)/\sigma]$ (12)
- Then combine these two parts, maximize the sum of these across i, with respect to β_j and σ to obtain the MLEs.

Truncated Regression Models

- Let y follows the classical linear model:
  $y = \mathbf{x}\beta + u,\ u|\mathbf{x} \sim N(0, \sigma^2)$ (13)
- A random draw $(\mathbf{x}_i, y_i)$ is observed only if $y_i \leq c_i$, where c_i is the truncation threshold that can depend on exogenous variables.
- To estimate β_j and σ, we need the distribution of y_i, given that $y_i \leq c_i$, and $\mathbf{x_i}$:
  $g(y|\mathbf{x}_i, c_i) = \frac{f(y|\mathbf{x}_i\beta,\sigma^2)}{F(c_i|\mathbf{x}_i\beta,\sigma^2)}$ (14)
  - $f(\cdot)$ the normal desity with mean $\mathbf{x}_i\beta$ and variance σ^2
  - $F(\cdot)$ the normal cdf with the same mean and variance, evaluated at c_i
  - We renormalize the density by dividing by the area under $f(y|\mathbf{x}_i\beta, \sigma^2)$ that is to the left of c_i