1 The Type I Tobit Model

$$\begin{aligned} y_i^* &= \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i, \\ y_i &= y_i^* \text{ if } y_i^* > 0 \\ y_i &= 0 \text{ if } y_i^* \leqslant 0 \end{aligned}$$

Wooldridge distiguishes between two different applications of this model:

1. Censored data.

"Top coding" typical example:

$$\begin{aligned} E\left(wealth^* \mid \mathbf{x}\right) &= \mathbf{x}\boldsymbol{\beta} \\ wealth &= \min\left(wealth^*, 200\right) \\ wealth^* &= \mathbf{x}\boldsymbol{\beta}+u,\ u \mid \mathbf{x} \sim \mathbf{N}\left(0, \sigma^2\right) \end{aligned}$$

2. Corner solution outcomes

Example: Engel curves for alcohol and tobacco consumption:

$$w_j = \alpha_{ji} + \beta_{ji} \log x_i + \varepsilon_{ji},$$

where w_j is the budget share of commodity j, β_{ji} is the income elasticity that can depend on household characteristics, x is income.

$$\begin{aligned} w_{ji}^* &= \alpha_{ji} + \beta_{ji} \log x_i + \varepsilon_{ji}, \\ w_{ji} &= w_{ji}^* \text{ if } w_{ji}^* > 0 \\ &= 0 \text{ otherwise.} \end{aligned}$$

Different outcomes from this model can be interesting, e.g. $P(y=0)$, $E(y \mid \mathbf{x}, y>0)$, $E(y \mid \mathbf{x})$.

Remember from the threshold model

$$\begin{aligned} P(y_i = 0) &= P(y_i^* \leqslant 0) = P(\varepsilon_i \leqslant -\mathbf{x}_i'\boldsymbol{\beta}) = \\ P\left(\frac{\varepsilon_i}{\sigma} \leqslant \frac{\mathbf{x}_i'\boldsymbol{\beta}}{\sigma}\right) &= \Phi\left(-\frac{\mathbf{x}_i'\boldsymbol{\beta}}{\sigma}\right) = 1 - \Phi\left(\frac{\mathbf{x}_i'\boldsymbol{\beta}}{\sigma}\right) \end{aligned}$$

Implies that

$$P(y_i > 0) = \Phi\left(\frac{\mathbf{x}_i'\boldsymbol{\beta}}{\sigma}\right)$$

2 Truncated normal distributions

Censored random variables:

$$y_i = y_i^* \text{ if } y_i^* > c$$
$$y_i = c \text{ if } y_i^* \leqslant c$$

Truncated random variables:

$$y_i = y_i^* \text{ if } y_i^* > c$$
$$\text{nothing if } y_i^* \leqslant c$$

Truncated normal

If the distribution is truncated from below.

$$\begin{aligned} f(y|Y \geqslant c) &= \frac{f(y)}{P(Y \geqslant c)} \text{ if } y \geqslant c \\ & \quad 0 \text{ otherwise} \end{aligned}$$

If $Y \sim N(\mu, \sigma^2)$ then

$$E(Y|Y \geqslant c) = \mu + \sigma\lambda_1(c^*) \geqslant \mu,$$

where $\lambda_1(c) = \frac{\phi(c)}{1-\Phi(c)}$ and $c^* = \frac{c-\mu}{\sigma}$.

The corresponding result when the distribution is truncated from above:

$$E(Y|Y \geqslant c) = \mu + \sigma\lambda_2(c^*) \leqslant \mu,$$

where $\lambda_2(c) = \frac{-\phi(c)}{\Phi(c)}$.

For a bivariate distribution (Y, X) :

$$\begin{aligned} E(Y|X \geqslant c) &= \mu_y + \left(\sigma_{yx}/\sigma_x^2\right)[E(X|X \geqslant c) - \mu_x] \\ &= \mu_y + (\sigma_{yx}/\sigma_x)\lambda_1(c^*) \end{aligned}$$

Applying this result to the Tobit model:

$$E\left(y_i|\mathbf{x}_i, y_i>0\right)=\mathbf{x}_i'\boldsymbol{\beta}+E\left(\varepsilon_i|\varepsilon_i>-\mathbf{x}_i'\boldsymbol{\beta}\right)=\mathbf{x}_i'\boldsymbol{\beta}+\sigma\frac{\phi\left(-\mathbf{x}_i'\boldsymbol{\beta}/\sigma\right)}{1-\Phi\left(-\mathbf{x}_i'\boldsymbol{\beta}/\sigma\right)}=\mathbf{x}_i'\boldsymbol{\beta}+\sigma\frac{\phi\left(\mathbf{x}_i'\boldsymbol{\beta}/\sigma\right)}{\Phi\left(\mathbf{x}_i'\boldsymbol{\beta}/\sigma\right)}.$$

For any point c, $\frac{\phi(c)}{\Phi(c)}$ is called the inverse of Mill's ratio $\lambda(c)$.

Marginal effect:
It can be shown that

$$\frac{d\lambda(c)}{dc}=-\lambda(c)\left[c+\lambda(c)\right]$$

$$\frac{\partial E\left(y_i \mid \mathbf{x}_i, y_i>0\right)}{\partial x_j}=\beta_j\left[1-\lambda\left(\mathbf{x}_i'\boldsymbol{\beta}/\boldsymbol{\sigma}\right)\left[\mathbf{x}_i'\boldsymbol{\beta}/\boldsymbol{\sigma}+\lambda\left(\mathbf{x}_i'\boldsymbol{\beta}/\boldsymbol{\sigma}\right)\right]\right],$$

where the adjustment factor $\left[1-\lambda\left(\mathbf{x}_i'\boldsymbol{\beta}/\boldsymbol{\sigma}\right)\left[\mathbf{x}_i'\boldsymbol{\beta}/\boldsymbol{\sigma}+\lambda\left(\mathbf{x}_i'\boldsymbol{\beta}/\boldsymbol{\sigma}\right)\right]\right]$ is always between 0 and 1. This means that β_j determines the sign of the partial effect.

$$\begin{aligned} E\left(y_i|\mathbf{x}_i\right) &= E\left(y_i|\mathbf{x}_i, y_i>0\right)P\left(y_i>0\right)+0*P\left(y_i=0\right)= \\ &= \mathbf{x}_i'\boldsymbol{\beta}\Phi\left(\mathbf{x}_i'\boldsymbol{\beta}/\sigma\right)+\sigma\phi\left(\mathbf{x}_i'\boldsymbol{\beta}/\sigma\right) \end{aligned}$$

Partial effect is then calculated as:

$$\frac{\partial E\left(y_i|\mathbf{x}_i\right)}{\partial x_j}=\frac{\partial P\left(y_i>0 \mid \mathbf{x}_i\right)}{\partial x_j}E\left(y_i|\mathbf{x}_i, y_i>0\right)+P\left(y_i>0 \mid \mathbf{x}_i\right)\frac{\partial E\left(y_i|\mathbf{x}_i, y_i>0\right)}{\partial x_j}.$$

This can be simplifyed to the following expression:

$$\frac{\partial E\left(y_i \mid \mathbf{x}_i\right)}{\partial x_{ij}}=\beta_j\Phi\left(\mathbf{x}_i'\boldsymbol{\beta}/\sigma\right)$$

Estimated by maximum likelihood

$$L = \prod_{i \in I_0} P\left(y_i = 0\right) \prod_{i \in I_1} f\left(y_i | y_i > 0\right) P\left(y_i > 0\right)$$

$$\begin{aligned}
\ln L &= \sum_{i \in I_0} \ln P\left(y_i = 0\right) + \sum_{i \in I_1} \left[\ln f\left(y_i | y_i > 0\right)\right] + \ln P\left(y_i > 0\right) \\
&= \sum_{i \in I_0} \ln P\left(y_i = 0\right) + \sum_{i \in I_1} \ln f\left(y_i\right) = \\
&= \sum_{i \in I_0} \ln \left[1 - \Phi\left(\left(\mathbf{x}_i' \boldsymbol{\beta} / \sigma\right)\right)\right] + \sum_{i \in I_1} \ln \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\left(y_i - \mathbf{x}_i' \boldsymbol{\beta}\right)^2 / 2\sigma^2\right)\right]
\end{aligned}$$

3 The Roy Model

The economy consists of two occupations: fishers and hunters. Each individual can choose between these occupations.

$Y_{1i}-$ Earnings if individual i chooses to be a hunter.

$Y_{2i}-$ Earnings if individual i chooses to be a fisher.

The individuals are not equally productive in these occupations and can expect different earnings.

$Y_{1i} = \mu_1 + u_{1i}$

$Y_{2i} = \mu_2 + u_{2i}$

(Y_1, Y_2) joint normal distribution with covariance matrix $\begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$.

$u_1 = Y_1 - \mu_1$

$u_2 = Y_2 - \mu_2$

$Z = \frac{\mu_1 - \mu_2}{\sigma}$ and $u = \frac{u_2 - u_1}{\sigma}$.

$\sigma^2 = V(u_1 - u_2)$

The individual chooses to be a hunter if

$$\begin{aligned} Y_1 &> Y_2 \\ \mu_1 - \mu_2 &> u_2 - u_1 \\ Z &> u \end{aligned}$$

The expected *observed* earnings for hunters will then be

$$E(Y_1 | u < Z) = \mu_1 - \sigma_{1u} \frac{\phi(Z)}{\Phi(Z)},$$

where $\sigma_{1u} = cov(u_1, u)$.

The expected *observed* earnings for fishers is

$$E(Y_2 | u > Z) = \mu_2 + \sigma_{2u} \frac{\phi(Z)}{1 - \Phi(Z)},$$

where $\sigma_{2u} = cov(u_2, u)$.

Case 1:
$\sigma_{1u} < 0$
$\sigma_{2u} > 0.$

Observed incomes for both fishers and hunters are above μ_1 and μ_2 respectively. Individuals choose occupation according to their comparative advantage.

$E(Y_1|u < Z) > \mu_1$
$E(Y_2|u > Z) > \mu_2$

Case 2:
$\sigma_{1u} < 0$
$\sigma_{2u} < 0.$

High u_2 draws corresponds to low $u_2 - u_1$ draws.

This means that hunters are better in both hunting and fishing. Those who choose fishing are below average in both hunting and fishing - but they are better in fishing.

$E(Y_1|u < Z) > \mu_1$
$E(Y_2|u > Z) < \mu_2$

Observed average earnings for hunters is higher than μ_1 and observed average earnings for fishers are below μ_2.

4 The Type II Tobit Model

$$
\begin{aligned}
y_i^* &= \mathbf{x}_{1i}'\boldsymbol{\beta}_1 + \varepsilon_{1i} \\
h_i^* &= \mathbf{x}_{2i}'\boldsymbol{\beta}_2 + \varepsilon_{2i} \\
y_i &= y_i^*,\ h_i = 1 \text{ if } h_i^* > 0 \\
y_i, \text{ not observed } h_i &= 0 \text{ if } h_i^* \leqslant 0
\end{aligned}
$$

$$
\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & 1 \end{pmatrix} \right]
$$

$$
\begin{aligned}
E\left(y_i | h_i = 1\right) &= \mathbf{x}_{1i}'\boldsymbol{\beta}_1 + E\left(\varepsilon_{1i} | h_i = 1\right) \\
&= \mathbf{x}_{1i}'\boldsymbol{\beta}_1 + E\left(\varepsilon_{1i} | \varepsilon_{2i} > -\mathbf{x}_{2i}'\boldsymbol{\beta}_2\right) = \\
&= \mathbf{x}_{1i}'\boldsymbol{\beta}_1 + \frac{\sigma_{12}}{\sigma_2^2} \frac{\phi\left(\mathbf{x}_{2i}'\boldsymbol{\beta}_2\right)}{\Phi\left(\mathbf{x}_{2i}'\boldsymbol{\beta}_2\right)} = \\
&= \mathbf{x}_{1i}'\boldsymbol{\beta}_1 + \sigma_{12}\lambda\left(\mathbf{x}_{2i}'\boldsymbol{\beta}_2\right)
\end{aligned}
$$

Heckman's two stage method:

1. Use a probit and estimate $\lambda_i = \frac{\phi\left(\mathbf{x}_{2i}'\boldsymbol{\beta}_2\right)}{\Phi\left(\mathbf{x}_{2i}'\boldsymbol{\beta}_2\right)}$.

2. Include λ_i in $y_i = \mathbf{x}_{1i}'\boldsymbol{\beta}_1 + \sigma_{12}\lambda_i + \eta_i$, where $\eta_i = \varepsilon_{1i} - E\left(\varepsilon_{1i} | \mathbf{x}_i, h_i = 1\right)$.

Classical paper: Heckman (1979) "Sample selection bias as a specification error" Econometrica, 47, 153-161.