# Limited Dependent Variables

1. Truncated Sample

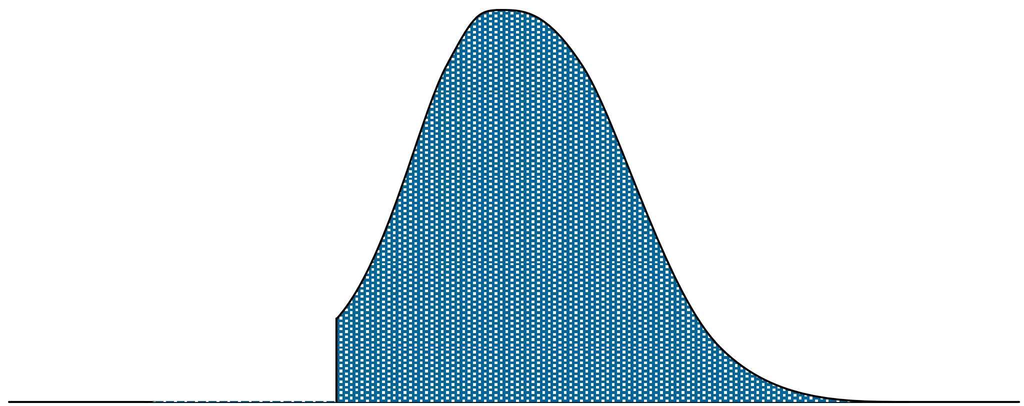
   Some observations are not available for observe. However, those observations do exist but weren't observed.

2. Censored Samples

   The observations can all be observed. However, some observations have no value for observe, hence, report as 0.

# Truncated Sample

## Truncated Normal Distribution

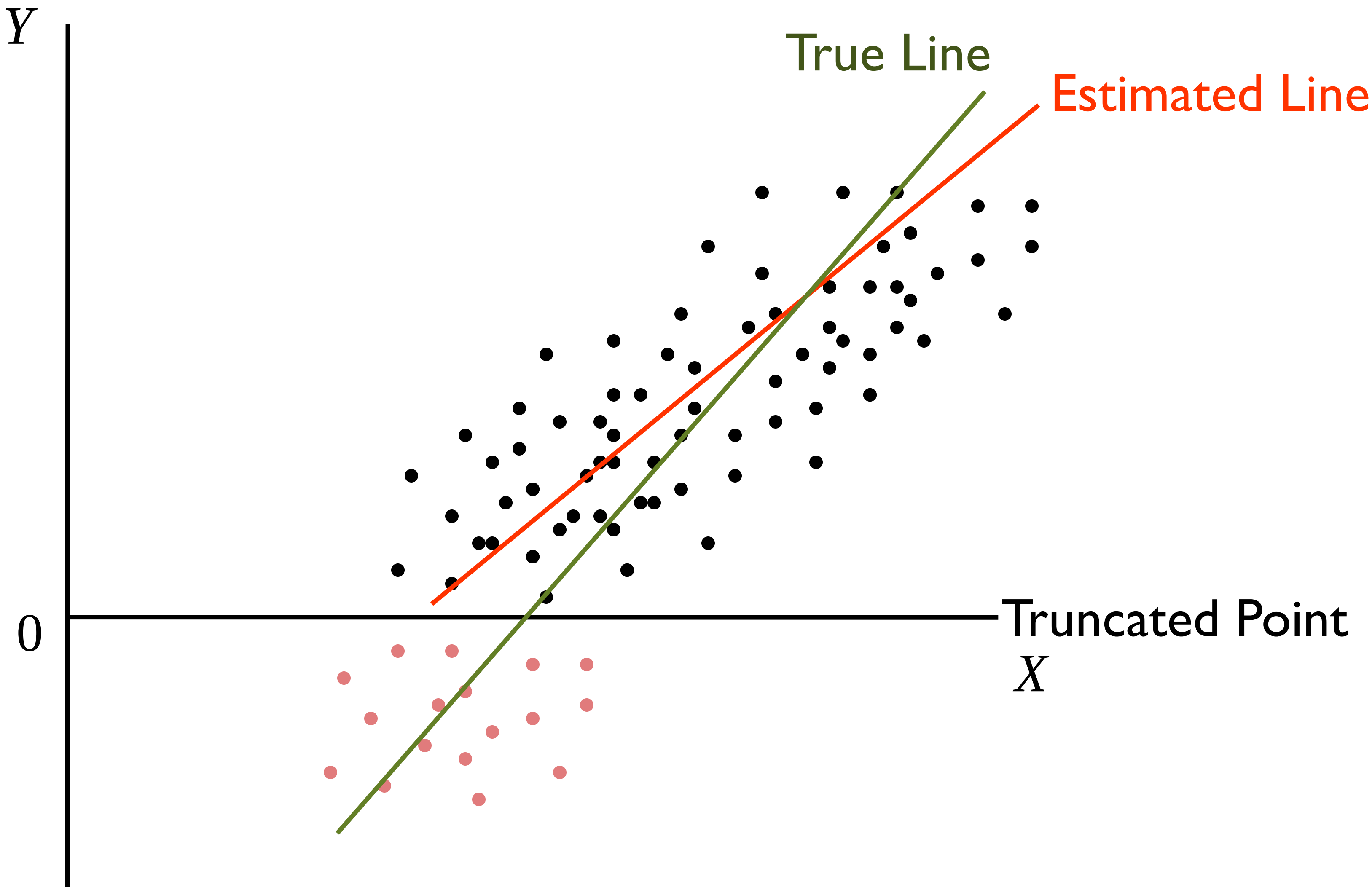



## Probability Function:

$$\Pr(x > T) = 1 - \Phi\left(\frac{T-\mu}{\sigma}\right) = 1 - \Phi(\alpha)$$

$$f(x | x > T) = \frac{f(x)}{1-\Phi(\alpha)} = \frac{\frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)}{1-\Phi(\alpha)}$$

# Biased in Truncated Sample



# Model for Truncated Data

Truncated model from below, we can only observe $y = y^* \quad if \; y^* > T_L$

For example, consumption of durable goods where $T_L = 0$

Truncated model from above, we can only observe $y = y^* \quad if \; y^* < T_U$

For example, only low-income individuals may be observed.

# Model for Truncated Data

Model for truncated data from below can be stated as:

$$y_i = x_i'\beta + \varepsilon_i \quad if\ y^* > T_L \quad \text{and} \quad \varepsilon_i \left| x_i \sim N[0, \sigma^2] \right.$$

Then, $$E\left[ y_i \middle| y_i^* > T_L \right] = x_i'\beta + \sigma \frac{\phi\left[(T - x_i'\beta)/\sigma\right]}{1 - \Phi\left[(T - x_i'\beta)/\sigma\right]}$$

$$= x_i'\beta + \sigma\lambda\left(\alpha_i\right)$$

Marginal Effect of $x_i$

$$\frac{\partial E\left[ y_i \middle| y_i^* > T \right]}{\partial x_i} = \beta\left(1 - \lambda_i^2 + \alpha_i \lambda_i\right)$$



By Tatre Jantarakolica

# Estimation Method for Model for Truncated Data

Model for Truncated data can be estimated by maximizing the truncated log-likelihood function of the model.

e.g., truncated normal probability distribution can be stated as:

$$f(y) = f^*(y|y > T_L) = f^*(y) \big/ \Pr\left[y|y > T_L\right]$$

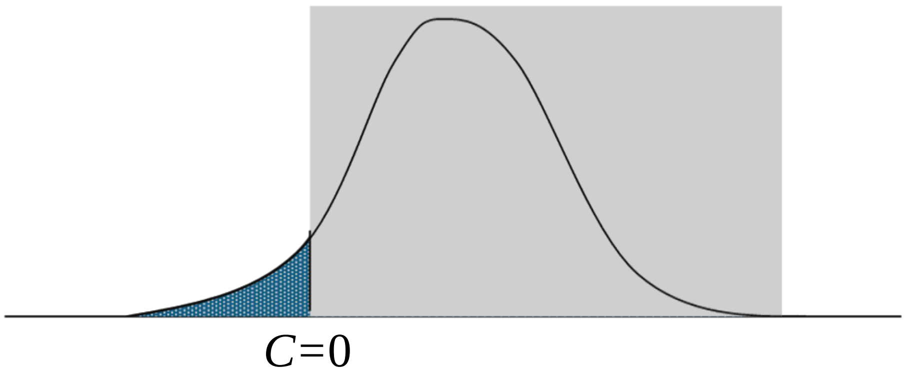
$$= f^*(y) \big/ \left[1 - F^*(T_L)\right]$$

Then, log-likelihood function can be stated as:

$$\ln L = \sum_{i=1}^{N} \left\{ \ln f^*\left(y_i | x_i, \beta\right) - \ln\left[1 - F^*\left(T_{Li} | x_i, \beta\right)\right] \right\}$$

# Censored Sample

## Censored Normal Distribution




$$y=0 \quad if \;\; y^* \leq 0$$

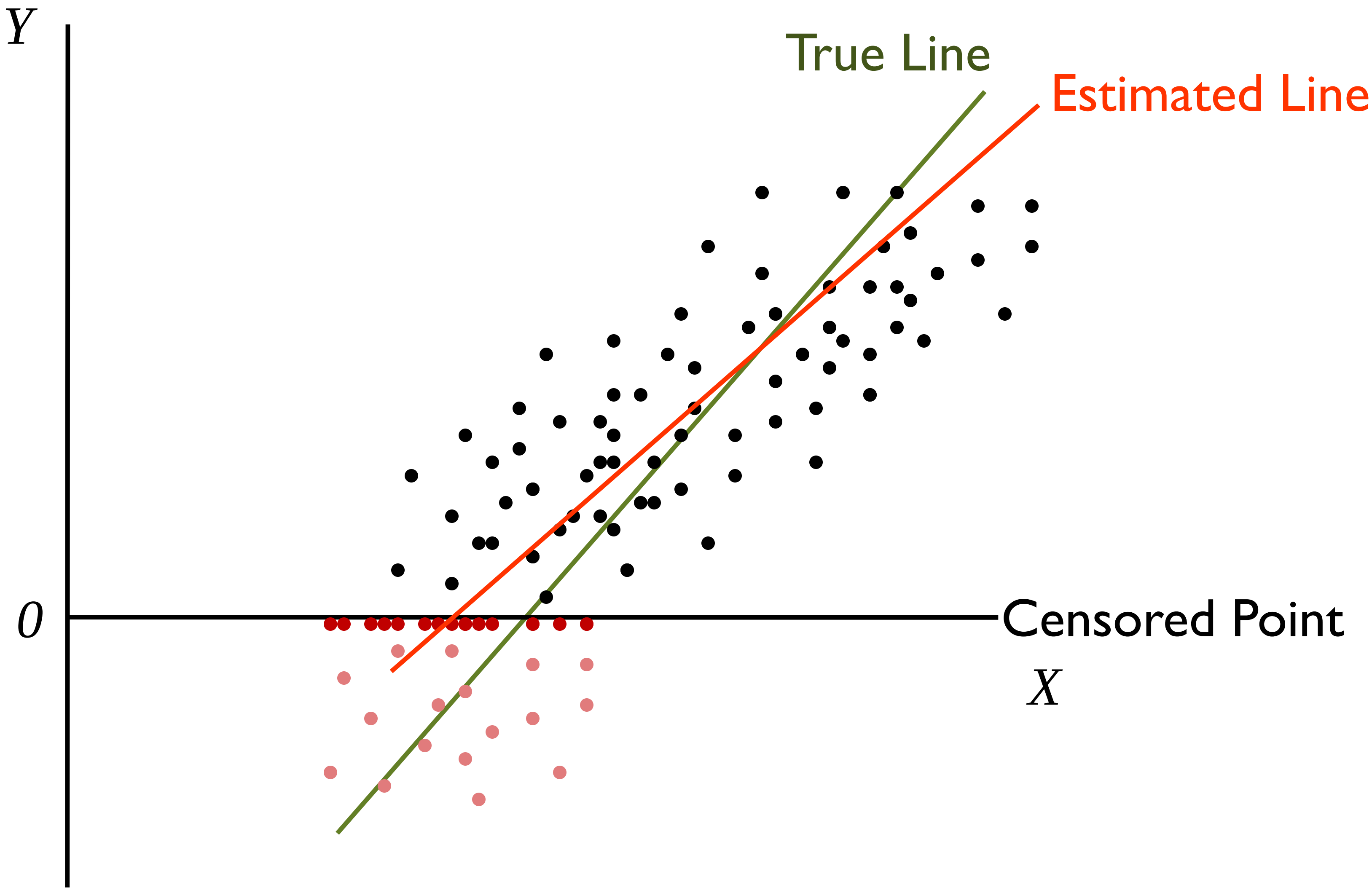
$$y=y^* \; if \;\; y^* > 0$$

## Probability Function:

$$\Pr(y=0)=\Pr(y^* \leq 0)=1-\Phi\left(\frac{\mu}{\sigma}\right)$$

$$E\left[y \middle| c=0\right]=\Phi\left(\mu/\sigma\right)\left(\mu+\sigma\lambda\right) \qquad where\ \lambda=\frac{\phi\left(\mu/\sigma\right)}{\Phi\left(\mu/\sigma\right)}$$

# Biased in Censored Sample






By Tatre Jantarakolica

# Model for Censored Data

Censoring from below, we can observe

$$y = \begin{cases} y^* & if \ \ y^* > C_L \\ C_L & if \ \ y^* \leq C_L \end{cases}$$

For example, consumption of durable goods

Censoring from above, we can observe

$$y = \begin{cases} y^* & if \ \ y^* < C_U \\ C_U & if \ \ y^* \geq C_U \end{cases}$$

For example, maximum income of individuals may only be observed as $>1$ million.

# Tobit Model for Censored Data

Model for censored data from below and $C_L = 0$ can be stated as: $y_i^* = x_i\beta + \varepsilon_i$

where $y = 0$ $if$ $y^* \leq 0$ or $y = y^*$ $if$ $y^* > 0$

Then, $E\left[y_i \middle| x_i\right] = \Phi\left(\frac{x_i\beta}{\sigma}\right)\left(x_i\beta + \sigma\lambda_i\right)$

where $\lambda_i = \frac{\phi\left[(0 - x_i\beta)/\sigma\right]}{1 - \Phi\left[(0 - x_i\beta)/\sigma\right]} = \frac{\phi\left(x_i\beta/\sigma\right)}{\Phi\left(x_i\beta/\sigma\right)}$

Marginal Effect of $x_i$

$\frac{\partial E\left[y_i^* \middle| x_i\right]}{\partial x_i} = \beta$ and $\frac{\partial E\left[y_i \middle| x_i\right]}{\partial x_i} = \beta\Phi\left(\frac{x_i\beta}{\sigma}\right)$

# Estimation Method for Tobit Model for Censored Data

Tobit Model can be estimated by maximizing the censored log-likelihood function of the model. For example, censored normal probability distribution when $C_L = 0$ can be stated as:

$$F^*(0) = \Pr\left[y^* \le 0\right] = \Pr\left[x_i\beta + \varepsilon_i\right] = \Phi\left(-x_i\beta/\sigma\right) = 1 - \Phi\left(x_i\beta/\sigma\right)$$

Censored density can be expressed as:

$$f(y) = \left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2\sigma^2}\left(y - x'\beta\right)^2\right\}\right]^d \left[1 - \Phi\left(\frac{x'\beta}{\sigma}\right)\right]^{1-d}$$

where: $d = 1\ if\ \ y^* > 0\ and\ d = 0\ if\ \ y^* \le 0$

# Estimation Method for Tobit Model for Censored Data

Then, log-likelihood function for Tobit Models assuming censored normal distribution can be stated as:

$$\ln L = \sum_{i=1}^{N} \left\{ \begin{aligned} & d_i \left( -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln \sigma^2 - \frac{1}{2\sigma^2}\left(y_i - x_i\beta\right)^2 \right) \\ & + \left(1 - d_i\right) \ln\left(1 - \Phi y_i \middle| x_i, \beta\right) \end{aligned} \right\}$$

# Tobit Model using Panel Data

Tobit model with random effects can be stated as: $y_{it}^{*} = x_{it}\beta + \alpha_i + \varepsilon_{it}$

where $y_{it} = 0 \;\; if \;\; y_{it}^{*} \le 0$ or $y_{it} = y_{it}^{*} \;\; if \;\; y_{it}^{*} > 0$

$\alpha_i =$ Cross-sectional Random Effects

Joint density function: $\prod_{t=1}^{T}\left[\frac{1}{\sigma_\varepsilon}\phi_{it}\right]^{d_{it}}\left[1-\Phi_{it}\right]^{1-d_{it}}$

where

$$\phi_{it} = \phi\left(\left(y_{it} - \alpha_i - x_{it}\beta\right)/\sigma_\varepsilon\right), \quad \Phi_{it} = \Phi\left(\left(\alpha_i + x_{it}\beta\right)/\sigma_\varepsilon\right)$$

Log-likelihood $\sum_{i=1}^{N} \ln f\left(y_i \middle| x_{it}, \beta, \sigma_\varepsilon^2, \sigma_\alpha^2\right)$

where $f = \int f\left(y_i \middle| x_i, \alpha_i, \beta, \sigma_\varepsilon^2\right)\frac{1}{\sqrt{2\pi\sigma_\alpha^2}}\exp\left(\frac{-\alpha_i}{2\sigma_\alpha^2}\right)^2 d\alpha_i$

# Test Pooled vs Panel Tobit Estimators

To test whether the pooled estimator or the panel estimator should be employed, the test whether $\rho = 0$ should be performed.

where: $$\rho = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\varepsilon^2}$$

If $\rho = 0$, it means that panel-level variance component $\sigma_\alpha^2$ is unimportant.

Then, panel Tobit estimator is no different from the pooled Tobit estimator.