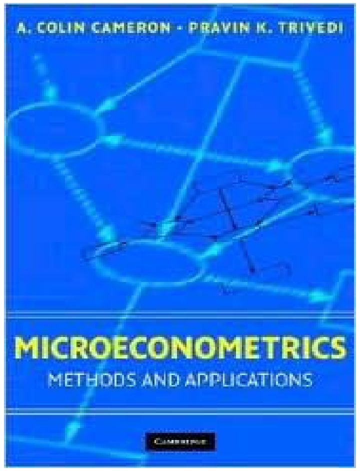
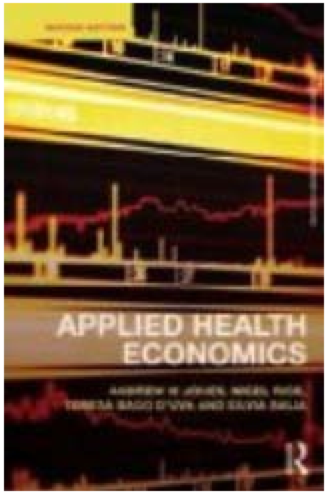
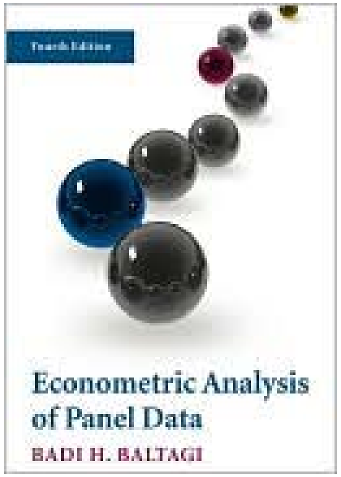
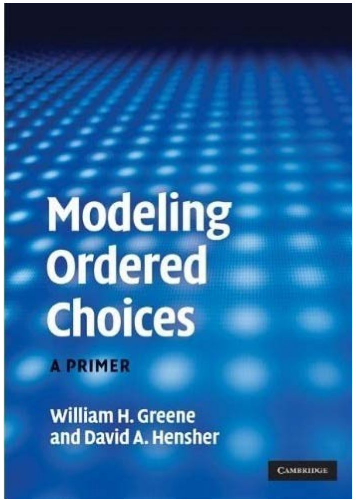
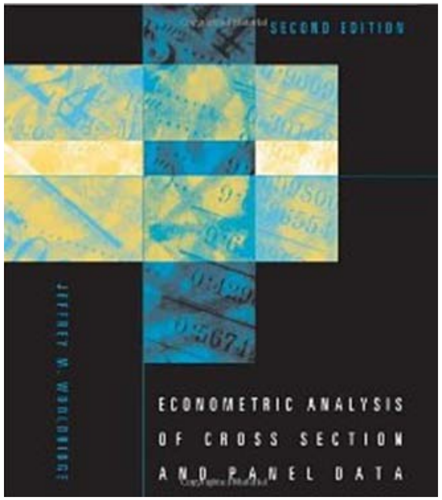


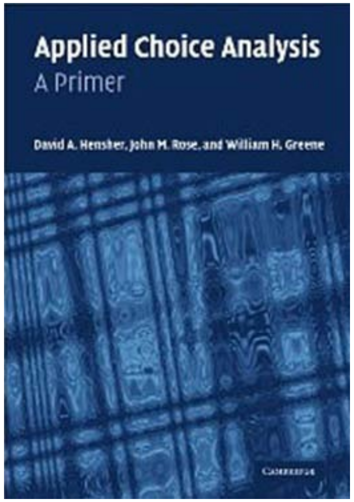




10. Censoring, Tobit and Two Part Models







Censoring and Corner Solution Models

- Censoring model:

  $y = T(y^*) = 0$ if $y^* \le 0$

  $y = T(y^*) = y^*$ if $y^* > 0$.

- Corner solution:

  $y = 0$ if some exogenous condition is met;

  $y = g(x)+e$ if the condition is not met.

  We then model $P(y=0)$ and $E[y|x,y>0]$.

- Hurdle Model:

  $y = 0$ with $P(y=0|z)$

  Model $E[y|x,y > 0]$

The Tobit Model

$y^* = \mathbf{x}'\beta + \varepsilon$,
$y = \text{Max}(0, y^*)$,
$\varepsilon \sim N[0, \sigma^2]$

Variation: Nonzero lower limit
Upper limit
Both tails censored

Easy to accommodate. (Already done in major software.)

Log likelihood and Estimation. See Appendix.

(Tobin: "Estimation of Relationships for Limited Dependent Variables," *Econometrica*, 1958. Tobin's probit?)

Conditional Mean Functions

$$y^* = \mathbf{x}\boldsymbol{\beta} + \varepsilon,\ \ \varepsilon \sim N[0,\sigma^2],\ \ y = \text{Max}(0, y^*)$$

$$E[y^* \mid \mathbf{x}] = \mathbf{x}'\beta$$

$$E[y|\mathbf{x}] = \text{Prob}[y=0|\mathbf{x}] \times 0 + \text{Prob}[y>0|\mathbf{x}]E[y|y>0,\mathbf{x}]$$

$$= \text{Prob}[y^*>0|\mathbf{x}]\ \ E[y^*|y^*>0,\mathbf{x}]$$

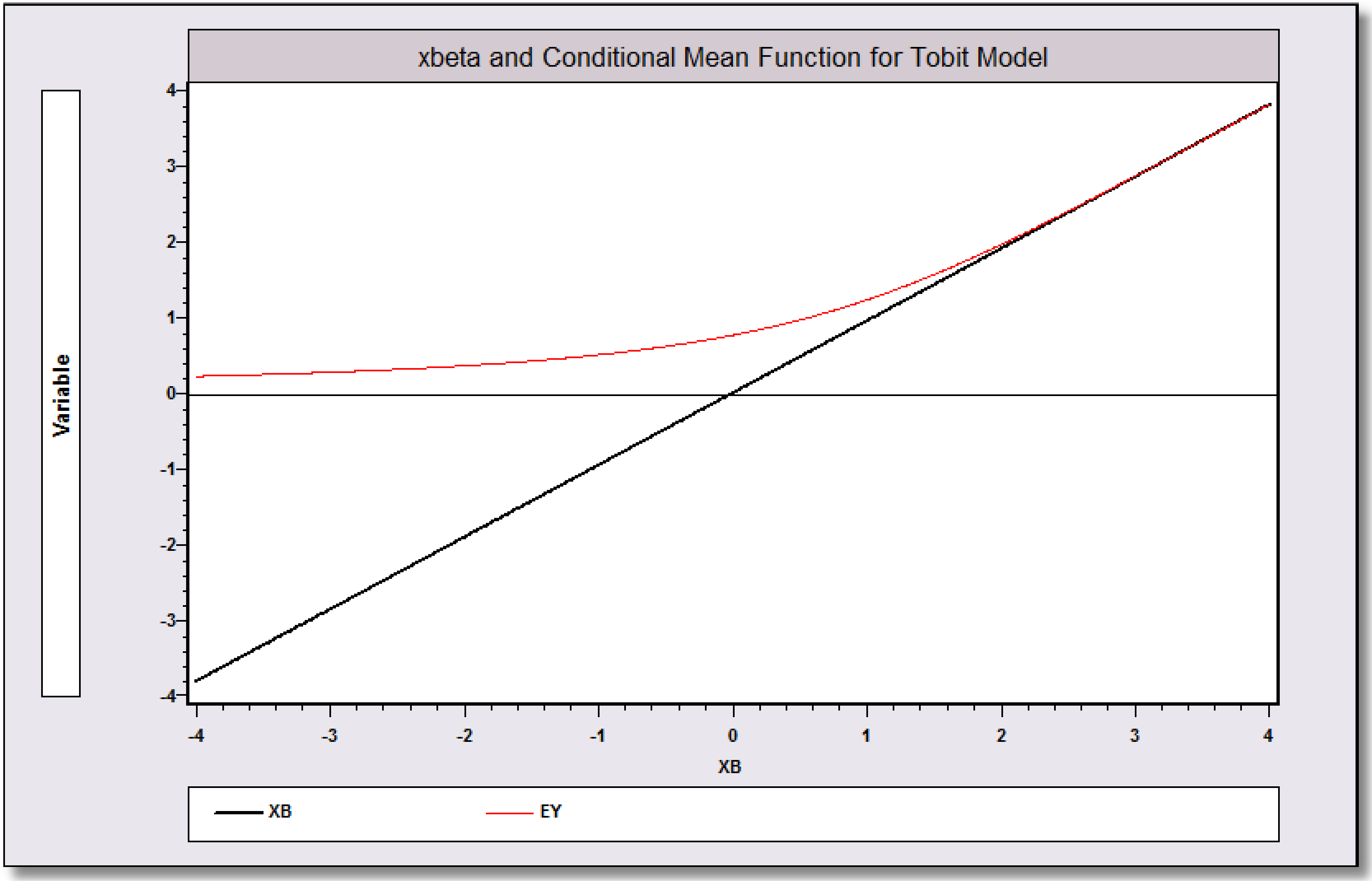
$$= \boxed{\Phi\left[\frac{\mathbf{x}\boldsymbol{\beta}}{\sigma}\right]} \times \boxed{\left(\mathbf{x}\boldsymbol{\beta} + \sigma\frac{\phi(\mathbf{x}\boldsymbol{\beta}/\sigma)}{\Phi(\mathbf{x}\boldsymbol{\beta}/\sigma)}\right)}$$

$$= \Phi(\mathbf{x}\boldsymbol{\beta}/\sigma)\mathbf{x}'\boldsymbol{\beta} + \sigma\phi(\mathbf{x}'\boldsymbol{\beta}/\sigma)$$

$$\boxed{\lambda = \frac{\phi(\mathbf{x}\boldsymbol{\beta}/\sigma)}{\Phi(\mathbf{x}\boldsymbol{\beta}/\sigma)} = \text{"Inverse Mills ratio"}}$$

$$\boxed{E[y|\mathbf{x},y>0] = \mathbf{x}'\boldsymbol{\beta} + \sigma\frac{\phi(\mathbf{x}\boldsymbol{\beta}/\sigma)}{\Phi(\mathbf{x}\boldsymbol{\beta}/\sigma)}}$$

Conditional Means



Predictions and Residuals

- What variable do we want to predict?
  - y*? Probably not – not relevant
  - y? Randomly drawn observation from the population
  - y | y>0? Maybe. Depends on the desired function
- What is the residual?
  - y – prediction? Probably not. What do you do with the zeros?
  - Anything - **x**β? Probably not. **x**β is not the mean.
- What are the partial effects? Which conditional mean?

OLS is Inconsistent - Attenuation

$E[y\,|\,\mathbf{x}] = \Phi(\mathbf{x}'\boldsymbol{\beta}/\sigma)\,\mathbf{x}'\boldsymbol{\beta} + \sigma\phi(\mathbf{x}'\boldsymbol{\beta}/\sigma)$

Nonlinear function of **x**.

What is estimated by OLS regression of y on **x**?

Slopes of the linear projection are approximately equal to the derivatives of the conditional mean evaluated at the means of the data.

$$\frac{\partial E[y\,|\,\mathbf{x}]}{\partial \mathbf{x}} = \Phi(\mathbf{x}\boldsymbol{\beta}\ /\ \sigma) \times \boldsymbol{\beta}$$

Note the **attenuation**; $0 < \Phi(\mathbf{x}\boldsymbol{\beta}\ /\ \sigma) < 1$

Partial Effects in Censored Regressions

For the latent regression:

$$E[y^*|\mathbf{x}] = \mathbf{x}'\boldsymbol{\beta} \;;\quad \boxed{\frac{\partial E[y^*|\mathbf{x}]}{\partial \mathbf{x}} = \boldsymbol{\beta}}$$

$$E[y|\mathbf{x}] = \Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}}{\sigma}\right)\mathbf{x}'\boldsymbol{\beta} + \sigma\phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}}{\sigma}\right);\quad \boxed{\frac{\partial E[y|\mathbf{x}]}{\partial \mathbf{x}} = \Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}}{\sigma}\right)\boldsymbol{\beta}}$$

$$E[y|\mathbf{x}, \mathbf{y} > 0] = \mathbf{x}'\boldsymbol{\beta} + \sigma\left[\phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}}{\sigma}\right) / \Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}}{\sigma}\right)\right]$$

$$= \mathbf{x}'\boldsymbol{\beta} + \sigma\lambda\left(\frac{\mathbf{x}'\boldsymbol{\beta}}{\sigma}\right) = \mathbf{x}'\boldsymbol{\beta} + \sigma\lambda(a)$$

$$\boxed{\frac{\partial E[y|\mathbf{x}, \mathbf{y} > 0]}{\partial \mathbf{x}} = \boldsymbol{\beta}[1\text{-}\lambda(a)(a + \lambda(a))]}$$

Application: Fair's Data

Fair's (1977) Extramarital Affairs Data, 601 observations. Psychology Today.
Source: Fair (1977) and http://fairmodel.econ.yale.edu/rayfair/pdf/1978ADAT.ZIP.
Several variables not used are denoted X1, ..., X5.
y = Number of affairs in the past year,
(0,1,2,3,4-10=7, more=12, mean = 1.46. (Frequencies 451, 34, 17, 19, 42, 38)
z1 = Sex, 0=female; mean=.476
z2 = Age, mean=32.5
z3 = Number of years married, mean=8.18
z4 = Children, 0=no; mean=.715
z5 = Religiousness, 1=anti, …,5=very. Mean=3.12
z6 = Education, years, 9, 12, 16, 17, 18, 20; mean=16.2
z7 = Occupation, Hollingshead scale, 1,…,7; mean=4.19
z8 = Self rating of marriage. 1=very unhappy; 5=very happy

Fair, R., "A Theory of Extramarital Affairs," *Journal of Political Economy*, 1978.
Fair, R., "A Note on Estimation of the Tobit Model," Econometrica, 1977.

Fair's Study

- Corner solution model
- Discovered the EM method in the *Econometrica* paper
- Used the tobit instead of the Poisson (or some other) count model
- Did not account for the censoring at the high end of the data

Estimated Tobit Model

		Tobit			Probit	Truncated Regression	
Variable	***Least Squares (1)***	***Estimate (2)***	***Marginal Effect (3)***	***Scaled by 1/σ (4)***	***Estimate (5)***	***Estimate (6)***	***Marginal Effect (7)***
Constant	5.61	8.18	—	0.991	0.997	8.32	—
	(0.797)	(2.74)	—	(0.336)	(0.361)	(3.96)	—
z_2	–0.0504	–0.179	–0.042	–0.022	–0.022	–0.0841	–0.0407
	(0.0221)	(0.079)	(0.184)	(0.010)	(0.102)	(0.119)	(0.0578)
z_3	0.162	0.554	0.130	0.0672	0.0599	0.560	0.271
	(0.0369)	(0.135)	(0.0312)	(0.0161)	(0.0171)	(0.219)	(0.106)
z_5	–0.476	–1.69	–0.394	–0.2004	–0.184	–1.502	–0.728
	(0.111)	(0.404)	(0.093)	(0.484)	(0.0515)	(0.617)	(0.299)
z_7	0.106	0.326	0.0762	0.0395	0.0375	0.189	0.0916
	(0.0711)	(0.254)	(0.0595)	(0.0308)	(0.0328)	(0.377)	(0.182)
z_8	–0.712	–2.29	–0.534	–0.277	–0.273	–1.35	–0.653
	(0.118)	(0.408)	(0.0949)	(0.0483)	(0.0525)	(0.565)	(0.273)
σ	3.09	8.25				5.53	
log L			–705.5762		–307.2955	–329.7103	

Discarding the Limit Data

$y^* = \mathbf{x}\boldsymbol{\beta} + \varepsilon$

$y = y^*$ if $y^* > 0$, y is unobserved if $y^* \leq 0$.

$f(y|\mathbf{x}) = f(y^*|\mathbf{x}, y^*>0) =$ Truncated (at zero) normal

$$f(y^*|\mathbf{x},y^*>0) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-(y-\mathbf{x}\boldsymbol{\beta})^2}{2\sigma^2}\right)\frac{1}{\text{Prob}[y^*>0|x]}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-(y-\mathbf{x}\boldsymbol{\beta})^2}{2\sigma^2}\right)\frac{1}{\Phi(\mathbf{x}\boldsymbol{\beta}/\sigma)}$$

ML is not OLS

Regression with the Truncated Distribution

$$E[y_i|\mathbf{x}_i, y_i > 0] = \mathbf{x}_i'\boldsymbol{\beta} + \sigma\frac{\phi(\mathbf{x}_i'\boldsymbol{\beta}/\sigma)}{\Phi(\mathbf{x}_i'\boldsymbol{\beta}/\sigma)}$$

$$= \mathbf{x}_i'\boldsymbol{\beta} + \sigma\lambda_i$$

OLS will be inconsistent:

$$\text{Plim } \mathbf{b} = \boldsymbol{\beta} + \text{plim}\left(\frac{\mathbf{X'X}}{n}\right)^{-1} \times \text{plim}\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i\lambda_i$$

A left out variable problem.

Approximately: $\text{plim } \mathbf{b} \approx \boldsymbol{\beta}\ \text{plim}(1 - \bar{a}\bar{\lambda} - \bar{\lambda}^2)$

$$\bar{a} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i'\boldsymbol{\beta}/\sigma, \quad \bar{\lambda} = \lambda(\bar{a})$$

General result: Attenuation

Estimated Tobit Model

		Tobit			Probit	Truncated Regression	
Variable	***Least Squares (1)***	***Estimate (2)***	***Marginal Effect (3)***	***Scaled by 1/σ (4)***	***Estimate (5)***	***Estimate (6)***	***Marginal Effect (7)***
Constant	5.61	8.18	—	0.991	0.997	8.32	—
	(0.797)	(2.74)	—	(0.336)	(0.361)	(3.96)	—
z_2	–0.0504	–0.179	–0.042	–0.022	–0.022	–0.0841	–0.0407
	(0.0221)	(0.079)	(0.184)	(0.010)	(0.102)	(0.119)	(0.0578)
z_3	0.162	0.554	0.130	0.0672	0.0599	0.560	0.271
	(0.0369)	(0.135)	(0.0312)	(0.0161)	(0.0171)	(0.219)	(0.106)
z_5	–0.476	–1.69	–0.394	–0.2004	–0.184	–1.502	–0.728
	(0.111)	(0.404)	(0.093)	(0.484)	(0.0515)	(0.617)	(0.299)
z_7	0.106	0.326	0.0762	0.0395	0.0375	0.189	0.0916
	(0.0711)	(0.254)	(0.0595)	(0.0308)	(0.0328)	(0.377)	(0.182)
z_8	–0.712	–2.29	–0.534	–0.277	–0.273	–1.35	–0.653
	(0.118)	(0.408)	(0.0949)	(0.0483)	(0.0525)	(0.565)	(0.273)
σ	3.09	8.25				5.53	
$\log L$			–705.5762		–307.2955	–329.7103	

Two Part Specifications

Tobit Log Likelihood

$$\log L = \sum_{i=1}^{n} d \log f(y \mid y > 0) + (1-d) \log P(y=0)$$

$$= \sum_{i=1}^{n} d \log f(y \mid y > 0) - d \log P(y > 0)$$

$$+ d \log P(y>0) + (1-d) \log P(y=0)$$

$$= \boxed{\log L_{\text{Truncated Regression}}} + \boxed{\log L_{\text{Probit}}}$$

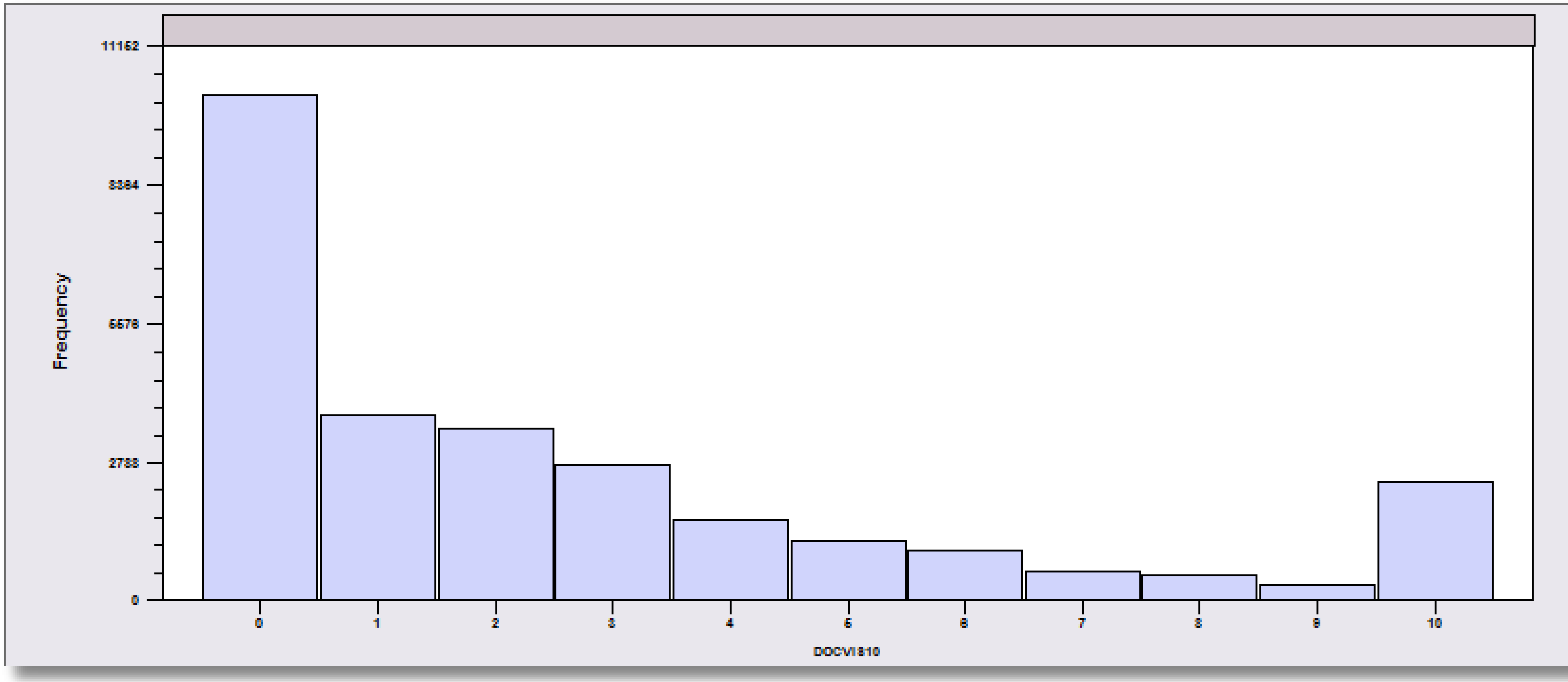
= logL for the model for y when y > 0 +

logL for the model for whether y is = 0 or > 0.

A two part model: Treat the model for quantity differently from the model for whether quantity is positive or not.

==> A probit model and a separate truncated regression.

Doctor Visits (Censored at 10)



Two Part Hurdle Model

Variable	TOBIT		PROBIT		TRUNC	
	Estimate	t ratio	Estimate	t ratio	Estimate	t ratio
Constant	-1.84453	-7.487	-.46071	-6.836	-2.40442	-4.313
AGE	.06969	27.503	.01470	20.326	.08958	17.189
EDUC	-.04965	-3.631	-.00918	-2.451	-.08582	-2.910
INCOME	-.83239	-4.937	-.08445	-1.803	-2.06645	-5.707
FEMALE	1.28128	22.491	.35068	21.911	.89914	7.750
MARRIED	-.09756	-1.431	.00727	.382	-.38544	-2.801
PUBLIC	.61614	6.440	.13814	5.329	.77983	3.667
Sigma	4.33603	171.556			4.67560	72.744
Log-L	-57668.0		-17443.0		-40166.4	
Log-L(0)			-18019.6			

```
|-> calc;list;spectest=2*(ltrunc+lprobit-ltobit)$
[CALC] SPECTEST=      117.2424688
```

Critical chi squared [7] = 14.1. The tobit model is rejected.

Panel Data Application

- Pooling: Standard results, incuding “cluster” estimator(s) for asymptotic covariance matrices
- Random effects
  - Butler and Moffitt – same as for probit
  - Mundlak/Wooldridge extension – group means
  - Extension to random parameters and latent class models
- Fixed effects: Some surprises (Greene, Econometric Reviews, 2005)

Neglected Heterogeneity

$$y_{it}^* = \mathbf{x}_{it}\boldsymbol{\beta} + (c_i + \varepsilon_{it}) \text{ assuming } c_i \perp \mathbf{x}_{it}$$

$$\text{Prob}[y_{it}^* = 0 \mid \mathbf{x}_{it}] = \Phi\left(\frac{\mathbf{x}_{it}\boldsymbol{\beta}}{\sigma_\varepsilon^2 + \sigma_c^2}\right)$$

MLE estimates $\boldsymbol{\beta} / (\sigma_\varepsilon^2 + \sigma_c^2)$ => Attenuated for $\boldsymbol{\beta}$

Partial effects are $\left[\boldsymbol{\beta} / (\sigma_\varepsilon^2 + \sigma_c^2)\right]\Phi\left(\frac{\mathbf{x}_{it}\boldsymbol{\beta}}{\sigma_\varepsilon^2 + \sigma_c^2}\right)$

Consistently estimated by ML even though $\boldsymbol{\beta}$ is not.

(The standard result.)

Fixed Effects MLE for Tobit

Table 3. Tobit Model. Effect of Group Size on Estimates[a]							
Estimate	T=2	T=3	T=5	T=8	T=12	T=15	T=20
β	0.67	0.53	0.50	0.29	0.098	0.082	0.047
δ	0.33	0.90	0.57	0.54	0.32	0.16	0.14
σ	-36.14	-23.54	-13.78	-8.40	-5.54	-4.43	-3.30
ME_x	15.83	8.85	3.65	1.30	0.44	0.22	0.081
ME_d	19.67	11.85	5.08	2.16	0.89	0.46	0.27
S.E.(β)	-32.92	-19.00	-11.30	-8.36	-6.21	-4.98	0.63
S.E.(δ)	-32.87	-22.75	-12.66	-7.39	-5.56	-6.19	0.25
R^2	0.785	0.774	0.761	0.751	0.744	0.740	0.736
Censoring	0.413	0.410	0.408	0.406	0.405	0.404	0.403
Correlation	0.650	0.531	0.411	0.325	0.265	0.237	0.206

No bias in slopes. Large bias in estimator of σ

APPENDIX: TOBIT MATH

Estimating the Tobit Model

Log likelihood for the tobit model for estimation of $\boldsymbol{\beta}$ and σ:

$$\log L = \sum_{i=1}^{n}\left[(1\text{-}d_i)\log\Phi\left(\frac{-\mathbf{x}\boldsymbol{\beta}}{\sigma}\right) + d_i\log\left(\frac{1}{\sigma}\phi\left(\frac{y_i - \mathbf{x}\,\boldsymbol{\beta}}{\sigma}\right)\right)\right]$$

$d_i = 1$ if $y_i > 0$, 0 if $y_i = 0$. Derivatives are very complicated,

Hessian is nightmarish. Consider the Olsen transformation*:

θ=1/σ, γ=-$\boldsymbol{\beta}$/σ. (One to one; σ=1/θ, $\boldsymbol{\beta} = -\gamma / \theta$.)

$$\log L = \sum_{i=1}^{n}\left[(1\text{-}d_i)\log\Phi\left(\mathbf{x}_i'\gamma\right) + d_i\log\left(\theta\phi\left(\theta y_i + \mathbf{x}_i'\gamma\right)\right)\right]$$

$$\sum_{i=1}^{n}\left[(1\text{-}d_i)\log\Phi\left(\mathbf{x}_i'\gamma\right) + d_i(\log\theta + (1/2)\log 2\pi - (1/2)\left(\theta y_i + \mathbf{x}_i'\gamma\right)^2)\right]$$

$$\frac{\partial \log L}{\partial \gamma} = \sum_{i=1}^{n}\left[(1\text{-}d_i)\frac{\phi\left(\mathbf{x}_i'\gamma\right)}{\Phi\left(\mathbf{x}_i'\gamma\right)} - d_i e_i\right]\mathbf{x}_i$$

$$\frac{\partial \log L}{\partial \theta} = \sum_{i=1}^{n} d_i\left(\frac{1}{\theta} - e_i y_i\right)$$

*Note on the Uniqueness of the MLE in the Tobit Model," Econometrica, 1978.

Hessian for Tobit Model

$$\log L=\sum_{i=1}^{n}\log\left[(1\text{-}d_i)\log\Phi\left(\mathbf{x}_i'\gamma\right)+d_i(\log\theta+(1/2)\log 2\pi-(1/2)\left(\theta y_i+\mathbf{x}_i'\gamma\right)^2)\right]$$

$$\frac{\partial\log L}{\partial\gamma}=\sum_{i=1}^{n}\left[(1\text{-}d_i)\frac{\phi\left(\mathbf{x}_i'\gamma\right)}{\Phi\left(\mathbf{x}_i'\gamma\right)}-d_i e_i\right]\mathbf{x_i} \qquad \frac{\partial\log L}{\partial\theta}=\sum_{i=1}^{n}d_i\left(\frac{1}{\theta}-e_i y_i\right)$$

$$\frac{\partial^2\log L}{\partial\gamma\partial\gamma'}=\sum_{i=1}^{n}\left[(1\text{-}d_i)\left(-(\mathbf{x}_i'\gamma)\frac{\phi\left(\mathbf{x}_i'\gamma\right)}{\Phi\left(\mathbf{x}_i'\gamma\right)}-\left(\frac{\phi\left(\mathbf{x}_i'\gamma\right)}{\Phi\left(\mathbf{x}_i'\gamma\right)}\right)^2\right)-d_i\right]\mathbf{x_i x_i'}$$

$$\frac{\partial^2\log L}{\partial\gamma\partial\theta}=\sum_{i=1}^{n}\left[-d_i\right]\mathbf{x_i}y_\mathbf{i}$$

$$\frac{\partial^2\log L}{\partial\theta\partial\theta}=\sum_{i=1}^{n}-d_i\left(\frac{1}{\theta^2}\right) \qquad \left[-d_i\right]y_\mathbf{i}y_\mathbf{i}$$

Simplified Hessian

$$\frac{\partial^2 \log L}{\partial\gamma\partial\gamma'} = \sum_{i=1}^{n} \left[(1\text{-}d_i)\left(-(\mathbf{x}_i'\gamma)\frac{\phi(\mathbf{x}_i'\gamma)}{\Phi(\mathbf{x}_i'\gamma)} - \left(\frac{\phi(\mathbf{x}_i'\gamma)}{\Phi(\mathbf{x}_i'\gamma)}\right)^2 \right) - d_i \right]\mathbf{x}_i\mathbf{x}_i'$$

$$\frac{\partial^2 \log L}{\partial\gamma\partial\theta} = \sum_{i=1}^{n} \left[-d_i\right]\mathbf{x}_i y_i$$

$$\frac{\partial^2 \log L}{\partial\theta\partial\theta} = \sum_{i=1}^{n} -d_i\left(\frac{1}{\theta^2}\right) \qquad \left[-d_i\right] y_i y_i$$

$$\frac{\partial^2 \log L}{\partial\begin{pmatrix}\gamma\\ \theta\end{pmatrix}\partial\begin{pmatrix}\gamma' & \theta\end{pmatrix}} = \sum_{i=1}^{n} \begin{bmatrix} ((1-d_i)\delta_i - d_i)\mathbf{x}_i\mathbf{x}_i' & -d_i\mathbf{x}_i y_i \\ -d_i y_i\mathbf{x}_i' & -d_i(1/\theta^2 + y_i^2) \end{bmatrix}$$

$$a_i = \mathbf{x}_i'\gamma,\ \lambda_i = \phi(a_i)/\Phi(a_i),\ \ \delta_i = -\lambda_i(a_i + \lambda_i)$$

Recovering Structural Parameters

$$\begin{pmatrix} \boldsymbol{\beta}(\gamma,\theta) \\ \sigma(\gamma,\theta) \end{pmatrix} = \begin{pmatrix} -\dfrac{1}{\theta}\boldsymbol{\gamma} \\ \dfrac{1}{\theta} \end{pmatrix}$$

Use the delta method to estimate Asy.Var $\begin{pmatrix} \hat{\boldsymbol{\beta}}(\hat{\gamma},\hat{\theta}) \\ \hat{\sigma}(\hat{\gamma},\hat{\theta}) \end{pmatrix}$

$$\frac{\partial \begin{pmatrix} \boldsymbol{\beta}(\gamma,\theta) \\ \sigma(\gamma,\theta) \end{pmatrix}}{\partial \begin{pmatrix} \boldsymbol{\gamma}' & \theta \end{pmatrix}} = \begin{pmatrix} \dfrac{-1}{\theta}\mathbf{I} & \dfrac{1}{\theta^2}\boldsymbol{\gamma} \\ \mathbf{0'} & \dfrac{-1}{\theta^2} \end{pmatrix} = \frac{-1}{\theta^2}\begin{pmatrix} \theta\mathbf{I} & -\boldsymbol{\gamma} \\ \mathbf{0'} & 1 \end{pmatrix} = \mathbf{G}(\gamma,\theta)$$

$$\text{Est.Asy.Var}\begin{pmatrix} \hat{\boldsymbol{\beta}}(\hat{\gamma},\hat{\theta}) \\ \hat{\sigma}(\hat{\gamma},\hat{\theta}) \end{pmatrix} = \mathbf{G}(\hat{\gamma},\hat{\theta}) \times \text{Est.Asy.Var}\begin{pmatrix} \hat{\boldsymbol{\gamma}} \\ \hat{\theta} \end{pmatrix} \times \mathbf{G}(\hat{\gamma},\hat{\theta})'$$