Limited Dependent Variables and Panel Data

Tibor Hanappi – 30.06.2010

Limited Dependent Variables

Discrete: Variables that can take only a countable number of values

Censored/Truncated: Data points in some specific range cannot be observed

➡ Focus on **discrete dependent variables!**

Outlook

1. Review: Cross-sectional models
2. Panel data: Fixed effects
3. Panel data: Random effects
4. Discussion: FE and RE
5. Implementation in STATA

Discrete Dependent Variables

➡ Estimating probabilities of observing a particular outcome (e.g. binary outcomes)

Applications:

- **Labour market outcomes**: Employment, self selection, labour supply
- **Consumer demand**: Purchase decisions, investments
- **Programme participation**: Health, education, insurance schemes
- **Other**: Retirement decisions, transportation mode choice

Discrete Dependent Variables

- Binary outcomes:

$$y_{it} = \{0,1\}$$
$$y_{it} = 1 \text{ if } y_{it}^{*} = x_{it}'\beta + u_{it} > 0$$
$$y_{it} = 0 \text{ otherwise}$$

- Outcome for individual *i* at time point *t*
- Cross-sectional analysis: drop subscript *t*

Cross-Sectional Analysis

- Interested in: $\Pr(y_i = 1 | x_i)$

$$
\begin{aligned}
\Pr(y_i = 1 | x_i) &= \Pr(y_i^* > 0 | x_i) \\
&= \Pr(x_i'\beta + u_i > 0) \\
&= \Pr(u_i > -x_i'\beta) \\
&= 1 - \Pr(u_i \le -x_i'\beta)
\end{aligned}
$$

- Probability that the outcome is positive for individual *i*

Cross-Sectional Analysis

In the binary case it follows that:

$$E(y_i | x_i) = 0 \cdot \Pr(y_i = 0 | x_i) + 1 \cdot \Pr(y_i = 1 | x_i) = \Pr(y_i = 1 | x_i)$$

→ Possible approach: Linear probability model (LPM):

$$y_i = E(y_i | x_i) = \Pr(y_i = 1 | x_i) + e_i = x_i'\beta + e_i$$

- Usual panel data methods could be applied
- But: Estimated probabilities are not restricted to the unit interval

Cross-Sectional Analysis

Non-linear models:

$$\Pr(y_i = 1 | x_i) = \Pr(u_i > -x_i'\beta) = 1 - \Pr(u_i \le -x_i'\beta)$$
$$= \Pr(u_i \le x_i'\beta) = F(x_i'\beta)$$

- The last line holds only as long as the distribution of u_i is symmetric around zero
- Function *F*(.) restricts the outcomes to be within the unit interval

Cross-Sectional Analysis

→ Intuitive illustration:

$$\Pr(u_i > -x_i'\beta) = \int_{-\infty}^{+\infty} I\left(u_i > -x_i'\beta\right) f(u_i)\,du_i$$

$$= \int_{u_i = -x_i'\beta}^{+\infty} f(u_i)\,du_i$$

- Integral over an indicator function showing whether the outcome is positive given the values of the error term
- Other way to see this: Integrate over all those values of u_i for which the outcome is positive
- Gives the probability that the error term is such that a positive outcome occurs (given what is known about $x_i'\beta$)

Cross-Sectional Analysis

Specifying a distribution for u_i:

- Binary Logit: $$F(c) = \Lambda(c) = \frac{\exp(c)}{1+\exp(c)}$$

- Binary Probit: $$F(c) = \Phi(c) = \int_{-\infty}^{c} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right) dz$$

→ Logit leads to closed form outcome probabilities

→ Probit is computationally more intense; but offers a more general treatment

Cross-Sectional Analysis

Maximum Likelihood:

- Outcome probabilities are independent of each other
- Likelihood function takes the following form

$$L(\beta) = \prod_{i=1}^{N} \Pr(y_i = 1 | X) \Pr(y_i = 0 | X)$$

- Outcome probabilities are $F(x_i'\beta)$ and $F(-x_i'\beta)$ respectively
- Where *F*(.) is either $\Lambda(.)$ or $\Phi(.)$
- Maximization of log $L(\beta)$ with respect to β gives MLE
- Probit: Outcome probabilities have to be approximated numerically (e.g. through simulation)

Panel Data: Fixed Effects

- For panel data $u_{it} = \mu_i + \nu_{it}$
- Outcome probabilities in the fixed effects model

$$\Pr(y_{it} = 1 | x_{it}) = \Pr(y_{it}^* > 0) = \Pr(\nu_{it} > -x_{it}'\beta - \mu_i)$$
$$= F(x_{it}'\beta + \mu_i)$$

- Logit: View μ_i and β as unknown parameters to be estimated by maximizing log $L(\beta,\mu_i)$
- However: As $N \to \infty$, for a fixed *T*, μ_i increase with *N*
- μ_i cannot be consistently estimated for a fixed *T*

➡ Incidental parameter problem!

Panel Data: Fixed Effects

- Cannot get rid of μ_i through within-transformation (as in linear models)
- Possible solution: Find a minimum *sufficient statistic* for μ_i
- For the logit model $\sum_{t=1}^{T} y_{it}$ is a minimum sufficient statistic for μ_i
- By definition, conditioning $L(\beta)$ on the minimum sufficient statistic gives a conditional likelihood function that does not depend on μ_i (Chamberlain, 1980)

$$L_C(\beta) = \prod_{i=1}^{N} \Pr\left(y_{i1}, \ldots, y_{iT} \middle| \sum_{t=1}^{T} y_{it}, X\right)$$

Panel Data: Fixed Effects

- **Logit model with FE**: conditioning on the minimum sufficient statistic for μ_i yields outcome probabilities that are independent of μ_i
- Maximizing the *conditional* likelihood gives consistent β estimates (while retaining closed-form outcome probabilities)
- However: Only observations for individuals who switched status can be used in estimation
- Dependent variable takes the value 1 if y_{it} switches from 0 to 1 and 0 if y_{it} switches from 1 to 0
- In this case the differences $x'_{it} - x'_{it-1}$ are used as independent variables

Panel Data: Random Effects

- In the **Probit model with RE** $u_{it} = \mu_i + \nu_{it}$ with $\mu_i \sim IIN(0, \sigma_\mu^2)$ and $\nu_{it} \sim IIN(0, \sigma_\nu^2)$ independent of each other and the x_{it}
- Now: $E(u_{it}u_{is}) = \sigma_\mu^2$ for $t \neq s$ and the joint likelihood of $(y_{i1}, \ldots, y_{iT})$ involves a T-dimensional integral

$$L(\beta, \sigma_\mu) = \prod_{i=1}^{N} \Pr\left(y_{i1}, \ldots, y_{iT} \middle| X\right) = \int \ldots \int_{-\infty}^{+\infty} f(u_{i1}, u_{i2}, \ldots, u_{iT}) du_{i1} du_{i2} \ldots du_{iT}$$

- Maximization with respect to β and σ_μ gets to be infeasible if *T* is large
- Numerical approximation involves simulation as well as non-simulation procedures

Panel Data: Preliminary Summary

→ Possible combinations of models with effects specifications:

	Logit	Probit
Fixed Effects	Yes	No
Random Effects	Yes	Yes

- Fixed effects specification not possible in the Probit framework (conditional likelihood approach does not lead to simplifications)
- Logit model with random effects is feasible, but (maybe) not very useful
- Potential advantage of Logit: Closed-form outcome probabilities (lost in case of RE!)
- Disadvantage of Logit: Only outcome switches can be used for estimation

Panel Data: Discussion

Possible interpretations of **fixed individual effects** in models with discrete dependent variables:

- Individual-specific unobserved effects on outcome that are not picked up by x_{it}
- Labour market context: E.g. ability
- Other contexts: E.g. time-invariant individual preferences

- Logit: To test for FE use a variant of the Hausman-test based on the difference between the conditional MLE and the usual MLE without FE

Panel Data: Discussion

Possible interpretations of **random individual effects:**

- Idea: *N* individuals are randomly drawn from a large population; N is large; FE would imply large losses in degrees of freedom
- E.g. household panel studies with representative samples
- Individual effects viewed as random and estimates valid for the population from which the sample was drawn
- However: Population does not consist of an infinity of individuals
- Alternative view: Population as an infinity of decisions (Haavelmo, 1944) → "behavioural" interpretation

- Probit: To test for RE use LR-principle to evaluate the likelihood for the pooled regression and for the RE estimator

Implementation in STATA

`xtlogit`

- Uses RE as default, FE optional
- Automatically omits groups of observations without within-group variation
- However: Conditional likelihood approach has to be implemented through data transformation (switches, differences)

`xtprobit`

- Only RE possible
- Uses mean and variance adaptive Gauss-Hermite quadrature as integration method
- Other variants of Gauss-Hermite quadrature can be defined
- Apparently: Other simulation methods not available (Accept-Reject, smoothed Accept-Reject, GHK simulator)
- Estimation of RE Probit models sometimes not possible because maximum of the likelihood function difficult to find

Definition

Sufficient statistic: If T(X) is a sufficient statistic for S, then any inference about S should depend on the sample X only through the value of T(X). That is if x and y are two sample points such that T(x)=T(y) then inference about S should be the same whether X=x or X=y is observed.

(Casella and Berger, *Statistical Inference*, 2002, p.272)

Literature

- Baltagi, Badi, 2001, *Econometric Analysis of Panel Data*, 2nd Edition, John Wiley & Sons.
- Casella, George and Roger L. Berger, 2002, *Statistical Inference*, Duxbury.
- Chamberlain, G., 1980, "Analysis of Covariance with Qualitative Data", Review of Economic Studies, Vol.47, *pp*. 225-238.
- Greene, William, 2008, "Discrete Choice Modelling", in *The Handbook of Econometrics: Vol. 2, Applied Econometrics, Part 4.2.,* ed. T. Mills and K. Patterson, Palgrave, London.
- Haavelmo, Trygve, 1944, "The Probability Approach in Econometrics", *Econometrica*, Vol.12 Supplement, *pp*. iii-115.
- Train, Kenneth, 2009, *Discrete Choice Methods with Simulation*, 2nd Edition, Cambridge University Press.