Panel Data Econometrics III (PhD)
Static Non-Linear Models

Yonas Alem

Department of Economics
University of Gothenburg

April 24, 2017

Lecture Outline

- Introduction
- Random Effects Probit
- Chamberlain's RE Probit
- Fixed Effects Logit

Reference

- Wooldridge JM. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT Press: Cambridge, MA. Ch. 15

- Consider the panel binary choice model

$$y_{it}^* = \mathbf{x}_{it}'\boldsymbol{\beta} + c_i + u_{it}, \qquad t = 1, 2, ..., T, \tag{1}$$

$$y_{it} = \begin{cases} 1 & \text{if } y_{it}^* > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

- Begin with a linear model and remove the unobserved heterogeneity through a with-in transformation
- We face the same problem as the cross-sectional case (predicted probabilities outside the 0-1 range, the error term is heteroskeastic)
- Less appealing in the presence of c_i

Incidental Parameters Problem

- Assumptions:
  1. $D(y_{it}|x_i, c_i) \equiv D(y_{it}|x_{i1}, x_{i2}, ..., x_{iT}, c_i) = D(y_{it}|x_{it}, c_i),\ t = 1, ..., T$. (strict exogeneity of x_{it} expressed in terms of conditional distributions
  2. $Pr(y_{it} = 1|x_{it}, c_i) = \Phi(\boldsymbol{x_{it}\beta} + c_i)$ a response probability
  3. $y_{i1}, ..., y_{iT}$ are independent conditional on $(\boldsymbol{x_i}, c_i)$
- The density of $(y_{i1}, ..., y_{iT})$ conditional on $(\boldsymbol{x_i}, c_i)$:

$$f(y_1, ..., y_T|\boldsymbol{x}_i, c_i; \boldsymbol{\beta}) = \prod_{t=1}^{T} f(y_t|\boldsymbol{x}_{it}, c_i; \boldsymbol{\beta}), \tag{3}$$

- Where $f(y_t|\boldsymbol{x}_{it}, c_i; \boldsymbol{\beta}) = \Phi(\boldsymbol{x}_t\boldsymbol{\beta} + c)^{y_t}[1 - \Phi(\boldsymbol{x}_t\boldsymbol{\beta} + c)]^{1-y_t}$
- Ideally, we can try estimating the model without restricting the relationship between c_i and $\boldsymbol{x}_{it}$
- Thus, the c_i would be parameters to be estimated along with β

Incidental Parameters Problem Cont.

- Estimation using the log-likelihood function $\sum_{i=1}^{N} \ell_i(c_i, \boldsymbol{\beta})$, where $\ell_i(c_i, \boldsymbol{\beta})$ is the log of (3) would be:
  - Computationally difficult
  - Introduces what is called the **incidental parameters problem**
    - Unlike the linear case, here for a fixed T, as $N \longrightarrow \infty$ estimating c_i for each individual along with $\boldsymbol{\beta}$ leads to inconsistent estimation of $\boldsymbol{\beta}$
    - In maximum likelihood framework the number of regressors is fixed, but here it grows with N
    - The problem persists unless T gets very large
    - Under this case, it is possible to show that $plim\hat{\beta} \longrightarrow 2\beta$ as $N \longrightarrow \infty$

- We need a better approach that integrates out c from the log-likelihood function
- Consider the panel binary choice model in its latent form again

$$y_{it}^* = \boldsymbol{x}_{it}\boldsymbol{\beta} + c_i + u_{it}, \qquad t = 1, 2, ..., T, \tag{4}$$

$$y_{it} = \begin{cases} 1 & \text{if } y_{it}^* > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

Assumptions

- We need a 4^{th} assumption on the relationship between c_i and $\boldsymbol{x}_i$ in addition to [1]-[3]

$$c_i | \boldsymbol{x}_i \sim Normal(0, \sigma_c^2) \tag{6}$$

Thus

$$f(y_1, ..., y_T | \boldsymbol{x}_i; \boldsymbol{\theta}) = \int_{-\infty}^{\infty} [\prod_{t=1}^{T} f(y_t | \boldsymbol{x}_{it}, c; \boldsymbol{\beta})](1/\sigma_c)\phi(c/\sigma_c)dc,$$

Random Effects Probit Estimator Cont.

- Where $f(y_t|\boldsymbol{x}_t, c; \boldsymbol{\beta}) = \Phi(\boldsymbol{x}_t\boldsymbol{\beta} + c)^{y_t}[1 - \Phi(\boldsymbol{x}_t\boldsymbol{\beta} + c)]^{1-y_t}$, and $\boldsymbol{\theta}$ contains $\boldsymbol{\beta}$ and σ_c^2
- One could get the conditional log likelihood for each i by plugging in $y_{it}\ \ \forall t$, and taking the log of the above likelihood function in 7
- $\sqrt{N}$- consistent asymptotically normal estimators can be obtained by maximizing the entire sample of N log-likelihood w.r.t $\boldsymbol{\beta}$ and σ_c^2
- The integral in (7) cannot be solved analytically but can be approximated by a numerical method known as **Gauss-Hermite Quadrature** developed by Butler and Moffitt (1982)
- The correlation in the composite error term $(v_{it} = c_i + u_{it})$ between two periods measured as $\rho = \sigma_c^2/(\sigma_c^2 + 1)$ is a useful indicator of the relative importance of c_i
- Stata command: **xtprobit**

Chamberlain's CRE Probit Model

- Chamberlain's model relaxes assumption (6), i.e.,$(c_i|\boldsymbol{x}_i \sim Normal(0, \sigma_c^2))$ and allows for correlation between c_i and $\boldsymbol{x}_{it}$
- Does so following Mundlak (1978) and assuming a conditional normal distribution with linear expectation and constant variance as:

$$c_i|\boldsymbol{x}_i \sim Normal(\psi + \bar{\boldsymbol{x}}_i\xi, \sigma_a^2) \tag{8}$$

- Where $\bar{\boldsymbol{x}}_i$ is the average of $\boldsymbol{x}_{it}, t = 1, ..., T$ and σ_a^2 is the variance of a_i in the equation $c_i = \psi + \bar{\boldsymbol{x}}_i\xi + a_i$
- Chamberlain allowed more generality by having $\boldsymbol{x}_i$, the vector of all explanatory variables across all time periods, in place of $\bar{\boldsymbol{x}}_i$
- The corresponding latent model would look like the following:

$$y_{it}^* = \psi + \boldsymbol{x}_{it}\boldsymbol{\beta} + \bar{\boldsymbol{x}}_i\xi + a_i + u_{it} \tag{9}$$

Chamberlain's CRE Probit Model

- Where $u_{it} \sim (0, 1)$ and $a_i \sim Normal(0, \sigma_a^2)$
- This model is knows as **Chamberlain's Correlated Random Effects Probit Model**
- The model looks restrictive in the sense that it specifies a certain distribution of c_i over $\boldsymbol{x_i}$ but is appealing since it allows for correlation between the two
- A test $H_0 : \xi = 0$ could be obtained for the RE probit (i.e., tests whether allowing for correlation was useful or not)

The Fixed Effects Logit Model

- Applies a conditional maximum likelihood estimator
- Considers the likelihood function conditional upon as set of statistics n_i that are sufficient for c_i
  - Conditional upon n_i an individual's likelihood contribution no longer depends upon c_i but still depends upon the other parameters β
- In the panel data binary choice model, the existence of a sufficient statistics depends upon the functional form of F, so the approach works only in the logit model
- Let the joint density or probability mass function be given as $y_{i1}, ..., y_{iT} = f(y_{i1}, ..., y_{iT} | c_i, \boldsymbol{\beta})$
- If a sufficient statistics $n_i \equiv \sum_{t=1}^{T} y_{it}$ exists $\Rightarrow$ $f(y_{i1}, ..., y_{iT} | n_i, c_i, \boldsymbol{\beta}) = f(y_{i1}, ..., y_{iT} | n_i, \boldsymbol{\beta})$
- Maximize the conditional likelihood function based up on $f(y_{i1}, ..., y_{iT} | n_i, \boldsymbol{\beta})$ to get a consistent estimator of $\boldsymbol{\beta}$

The Fixed Effects Logit Model cont.

- Consider $T = 2 \implies n_i = \{0, 1, 2\}$
- The conditional distribution of $(y_{i1}, y_{i2})'$ given n_i won't be informative for $\boldsymbol{\beta}$ when $n_i = 0$ or $n_i = 2$ because these values completely determine the outcome of y_i
- But if $n = 1$,

$$P(y_{i2} = 1 | \boldsymbol{x_i}, c_i, n_i = 1) = \frac{P(y_{i2} = 1, n_i = 1 | \boldsymbol{x_i}, c_i)}{P(n_i = 1 | \boldsymbol{x_i}, c_i)} \tag{10}$$

$$= \frac{P(y_{i2} = 1 | \boldsymbol{x_i}, c_i) P(y_{i1} = 0 | \boldsymbol{x_i}, c_i)}{\{P(y_{i1} = 0, y_{12} = 1 | \boldsymbol{x_i}, c_i) + P(y_{i1} = 1, y_{i2} = 0 | \boldsymbol{x_i}, c_i)\}} \tag{11}$$

$$= \frac{\Lambda(\boldsymbol{x}_{i2}\boldsymbol{\beta} + c_i)[1 - \Lambda(\boldsymbol{x}_{i1}\boldsymbol{\beta} + c_i)]}{\{[1 - \Lambda(\boldsymbol{x}_{i1}\boldsymbol{\beta} + c_i)]\Lambda(\boldsymbol{x}_{i2}\beta + c_i) + \Lambda(\boldsymbol{x}_{i1}\boldsymbol{\beta} + c_i)[1 - \Lambda(\boldsymbol{x}_{i2}\boldsymbol{\beta} + c_i)]} \tag{12}$$

$$= \Lambda[(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})\boldsymbol{\beta}]. \tag{13}$$

- Similarly, $P(y_{i1} = 1|\boldsymbol{x}_i, c_i, n_i = 1) = \Lambda[-(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})\boldsymbol{\beta}] = 1 - \Lambda[(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})\boldsymbol{\beta}]$.
- The conditional likelihood for observation i is given as

$$\ell_i = 1[n_i = 1](w_i log \Lambda[(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})\boldsymbol{\beta}] + (1 - w_i) log\{1 - \Lambda[(\boldsymbol{x}_{i2} - x_{i1})\boldsymbol{\beta} \tag{14}$$

- Where $w_i = 1$ if $(y_{i1} = 0, y_{i2} = 1)$ and $w_i = 0$ if $(y_{i1} = 1, y_{i2} = 0)$
- As indicated earlier observations that do not change their status $(i.e., n_i = 0$ or $n_i = 2)$ are excluded and don't contribute to the log-likelihood

The Fixed Effects Logit Model cont.

- The Log-likelihood function stated in Eq (14) is just a standard cross-sectional logit for w_i on $(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})$
- The approach is analogous to first differencing in the linear case
- As the model uses Conditional Maximum Likelihood Estimator on (14), it is also called the **Conditional Logit Estimator**
- Note: this estimator does not arise by treating the c_is as parameters to be estimated along $\boldsymbol{\beta}$ just like we tried in the FE probit model which is significantly biased due to the incidental parameter problem!
- This model is extendable to the case where $T > 2$ in a straightforward fashion.
- Stata Command: **xtlogit, fe**

Other Common Non-linear Static Panel Data Models

- Random Effects Tobit (Stata command: **xttobit**)
- Random Effects Multinomial Logit (Stata command: **Gllamm**)
- Random Effects Ordered Probit (Stata command: **reoprob**)