5. ***Binary Response Models – Dynamic Models***

$$y_{it}^{*} = \gamma\, y_{it-1} + \mathbf{x}_{it}\boldsymbol{\beta} + c_i + u_{it} \qquad\qquad y_{it} = 1[y_{it}^{*} > 0] \qquad (1)$$

- Regressor is y_{t-1} rather than y^{*}_{t-1}.
- Can have more than one lag of y.
- Can have time-invariant regressors.
- Have state-dependence. i.e. Prob of being in the current state depends on the previous state y_{t-1}. If $\gamma > 0$ **+ve State dependence**

- c_i is **unobserved heterogeneity.**

- **Spurious state dep issue**: reason for the observed correlation?

- Assume st.exog of **x.**

- $P(y_{i1}=1|\mathbf{x_{it}},c_i,y_{it-1}) = G(\gamma\, y_{it-1} + \mathbf{x_{it}}\boldsymbol{\beta} + c_i)$

5.1 Estimation

MLE: Observe $(y_1, y_2,\ldots, y_T)$. Assuming the y_ts are indep cond on c,

we have

$$P(y_{iT},.., y_{i1}|\mathbf{x}_i,c_i) = \prod_{t=2}^{T} f(y_{it} \mid y_{it-1}, \mathbf{x_{it}}, c_i)\, g(y_{i1} \mid c_i, \mathbf{x_{i1}}) \quad (2)$$

- f and g need not be the same type function….
- g is an unconditional specification! i.e. Marginal model for y_{i1}.
- y_{i1} – initial condition of the process…period 1 is the start of the sample period – need not coincide with the start of the stochastic process.
- y_{i1} is generally stochastic and correlated with c_i – **initial conditions problem**

- If y_{i1} is non-stochastic or independent of c_i, specification of g becomes easier.
- Generally, have to model the initial conditions explicitly.
- As $T \to \infty$ the initial conditions problem diminishes.

5.2 Modelling initial conditions

5.2.1 Heckman's method

We need $g(y_{i1}, c_i \mid \mathbf{x}_{i1})$ and a distribution for c.

If *f* is assumed to be Φ, then Heckman specified a distribution for

$y_{i1}| c_i$ and other covariates as

$$P(y_{i1}=1|\mathbf{z}_i, c_i) \cong \Phi(\mathbf{z}_i\delta + \theta c_i) \qquad \text{[note } \Phi\text{]}$$

i.e. $y_{i1}^* = \mathbf{z}_i\delta + \theta c_i + u_{i1}$ is the eq for the 1st period. (3)

- Where **z** consists of variables from $\mathbf{x}_{i1},\ldots,\mathbf{x}_{iT}$ and possibly others too.

- θ picks up correlation between y_{i1} and $c_{i.}$ $\theta=0$ is a test of no initial conditions problem.

- This gives the joint probability of the observed sample conditional on the c:

$$P(y_{iT},., y_{i1}|\mathbf{x}_i,z_i,c_i) = \prod_{t=2}^{T} \Phi(y_{it} \mid y_{it-1}, \mathbf{x_{it}}, c_i)\ \Phi(y_{i1} \mid c_i, z_i) \qquad (4)$$

- Before we can proceed with the estimation, we have to assume a distribution for c in order to integrate it out of the likelihood function.

- Can assume a discrete dist or some other cont distribution.

- Need to write your own routine to estimate this model!

5.2.2 Orme's method (2-step)

Start from $y_{it}^* = \gamma\, y_{it-1} + \mathbf{x_{it}}\beta + c_i + u_{it}$ $\qquad y_{it} = 1[y_{it}^* > 0]$ (1)

Problem here is the $\text{corr}(y_{it-1}, c_i) \neq 0$.

So replace c_i with something that is uncorrelated with the y_{it-1}.

Only need to worry about the first observation.

Write $c_i = \delta\, \eta_i + e_i$ (5)

By construction, η_i and e_i are orthogonal.

Substituting this in (1) gives you:

$$y_{it}^{*} = \gamma\, y_{it-1} + \mathbf{x_{it}\beta} + \delta\, \eta_i + e_i + u_{it} \qquad (6)$$

NOTE:

- e_i is orthogonal to the regressors by construction
- η_i is unobserved. Need to replace it with something.
- Let $y_{i1}^* = \mathbf{z}_i\delta + \text{error}$ (similar to (3)) (7)
- Assuming bivariate normality of (η_i and c_i) implies

$$E(\eta_i \mid y_{i1}) = \frac{(2y_{i1} - 1)\varphi(\mathbf{z}_i\delta)}{\Phi\left[\left\{2y_{i1} - 1\right\}\mathbf{z}_i\delta\right]} \tag{8}$$

which is the generalised error in the probit – inverse Mill's ratio!

- The two-steps are therefore:

  Step 1: estimate a probit of (7) and construct (8)

  Step2: use the estimated (8) in place of η in (6) and note it is just a RE probit eq.!

- **Problem**: unfortunately, e_i in (6) is heteroskedastic – MLE inconsistent. But simulation shows it is ok even for correlation (c,δ) of 0.5.

5.2.3 Wooldridge's method

Instead of working with $P(y_{iT},.., y_{i1}| \mathbf{x}_i,c_i)$ and specifying the $P(y_{i1}| \mathbf{x}_i,c_i)$,

here we work with $P(y_{iT},.., y_{i2}| y_{i1}, \mathbf{x}_i,c_i) \text{ x } P(c_i| y_{i1}, \mathbf{x_i})$

i.e we need to specify a distrib for $c_i|y_{i1}$.

$$c_i = \psi + \zeta_1 y_{i1} + \mathbf{z_i} \zeta + a_i \qquad (9)$$

Subst into the eq (1):

$$y_{it}^* = \gamma y_{it-1} + \mathbf{x_{it}}\beta + \psi + \zeta_1 y_{i1} + \mathbf{z_i} \zeta + a_i + u_{it} \qquad t=2,..,T \qquad (10)$$

Assuming $a_i \sim \text{Normal}(0,\sigma^2_a)$ and indep of y_{i1}, (10) is like a RE probit eq.

- Easy to estimate.

- APEs can be easily calculated.

5.2.4 Conditional logit

Can use CMLE assuming a logit model.

- Advantage - dist for c not required.

- Disadvantage: (i) Can't estimate APE since no c.

  (ii) sufficient stat exists for model without any **x.**

- See Narendranathan & Elias (1993) Oxford Bulletin for an application.

  Need T >=4 for the 1st order model.

5.2.5 Other methods

There are other methods that require different types of assumptions and have their own advantages and disadvantages. See Hsiao for a list.