Binary Response Models for Panel Data

(GB: Chapters 13.8, 15.8)

Isabel Casas

Syddansk Universitet

Summary today's class

1. Motivation
2. Without unknown individual effects
   - The linear probability model and panel data
   - Pooled Probit and Logit models for panel data
3. Unobserved Effects Probit (with strict exogeneity)
   - Fixed Effects estimation
   - Random Effects estimation
   - Alternatives (pooled, etc)
4. Unobserved Effects Logit (with strict exogeneity)
   - Random Effects estimation
   - Fixed Effects estimation
5. Dynamic Unobserved Effects Models (without strict exogeneity)

Motivation

- We have previously considered binary response models:
  - LPM
  - Probit models
  - Logit models
- But without a time dimension
- Individuals were observed once
- If observed more than once $\Rightarrow$ panel data
- Panel methods can be applied.

Panel data set

j	t	y_{jt}	x^1_{jt}	x^2_{jt}	$\ldots$	x^k_{jt}
1	1	y_{11}	x^1_{11}	x^2_{11}	$\ldots$	x^k_{11}
1	2	y_{12}	x^1_{12}	x^2_{12}	$\ldots$	x^k_{12}
1	3	y_{13}	x^1_{13}	x^2_{13}	$\ldots$	x^k_{13}
2	1	y_{21}	x^1_{21}	x^2_{21}	$\ldots$	x^k_{21}
2	2	y_{22}	x^1_{22}	x^2_{22}	$\ldots$	x^k_{22}
2	3	y_{23}	x^1_{23}	x^2_{23}	$\ldots$	x^k_{23}
$\vdots$					$\vdots$	
n	1	y_{n1}	x^1_{n1}	x^2_{n1}	$\ldots$	x^k_{31}
n	2	y_{n2}	x^1_{n2}	x^2_{n2}	$\ldots$	x^k_{32}
n	3	y_{n3}	x^1_{n3}	x^2_{n3}	$\ldots$	x^k_{33}

y_{jj} = 1 or 0

Binary models **without** unknown effects

- The linear probability model with panel data
  - Same problem than in the cross-sectional case
- To obtain estimates in [0,1]
  - Pooled Probit
  - Pooled Logit
- If ϵ serially correlated or X has lags
  - Need to use robust standard errors
- If dynamic completeness
  - Both models work fine
  - Same R functions than for the cross-section (glm)

Probability linear model for panel data

Examples of binary responses:

- Choice of transportation mode on day t
- Choice of marital status in year t
- Choice of owning a house in year t
- Being unemployed in week t
- Being in prison in year t

It is common that the same individual has repeated observations over time

- How to deal with that (what are the extra problems)?
- Can we somehow exploit that (besides having more observations)?

SYDDANSK UNIVERSITET

Probability linear model for panel data

The simplest panel model is the LPM (without unknown effects):

$$P(y_{jt} = 1|\mathbf{x}_{jt}) = \mathbf{x}_{jt}\boldsymbol{\beta}$$

- y_{jt} is the choice of individual j in period t
- $\mathbf{x}_{jt}$ are individual characteristics that may vary over time, e.g. age, income, and education

- We can estimate this as we saw before with the POLS.
- Problems? Estimates might be out of [0,1]
- Assumptions?

Probability linear model for panel data

LPM with unknown individual effects:

$$P(y_{jt} = 1|\mathbf{x}_{jt}) = \mathbf{x}_{jt}\boldsymbol{\beta} + c_j$$

- POLS: consistent estimates if c_j is uncorrelated with X but inefficient
  - Solution: RE
- POLS: inconsistent estimates if c_j is correlated with X
  - Solution: FE or FD

However, we will have the same problems than in the cross-sectional case:

Probability linear model for panel data

LPM with unknown individual effects:

$$P(y_{jt} = 1|\mathbf{x}_{jt}) = \mathbf{x}_{jt}\boldsymbol{\beta} + c_j$$

- POLS: consistent estimates if c_j is uncorrelated with X but inefficient
  - Solution: RE
- POLS: inconsistent estimates if c_j is correlated with X
  - Solution: FE or FD

However, we will have the same problems than in the cross-sectional case:
Estimates of this probability might be out of [0,1]
Any ideas about how to solve this?

Pooled probit and logit

The model **without** individual unknown effects but with panel data:

$$P(y_{jt} = 1|\mathbf{x}_{jt}) = G(\mathbf{x}_{jt}\boldsymbol{\beta})$$

- $G = \Phi$ – cdf of the normal dist (Probit) or
- $G = \Lambda$ – cdf of the logistic dist (Logit)
- $\mathbf{x}_{jt}$ can contain:
  - Time dummies (e.g. dummies for each year)
  - Time-varying variables (e.g. income for person j)
  - Time-invariant variables (e.g. gender of person j)
  - Cross product of time dummies and time-invariant variables
  - Lagged dependent variables (e.g. labour market status last year) $\Rightarrow Y$'s are correlated over time

Pooled probit and logit

The model:

$$P(y_{jt} = 1|\mathbf{x}_{jt}) = G(\mathbf{x}_{jt}\boldsymbol{\beta})$$

Estimation problem:

- We don't know the full distribution of $(y_{j1}, \ldots, y_{jT})$ given $(\mathbf{x}_{j1}, \ldots, \mathbf{x}_{jT})$
- It could be complicated (multidimensional), as e.g. y_{j1} and y_{j2} could be correlated
- This is a problem when we use maximum likelihood!
- Instead, we apply partial maximum likelihood techniques
- Not in general the same as full maximum likelihood
- But in some cases it is.

Pooled probit and logit

Full conditional maximum likelihood:

- Use full distribution of $(y_{j1}, \ldots, y_{jT})$ given $(\mathbf{x}_{j1}, \ldots, \mathbf{x}_{jT})$

Partial maximum likelihood (for each individual):

- Use distribution of y_t given $\mathbf{x}_t$: $f(y_t|\mathbf{x}_t, \boldsymbol{\beta})$
- The individual likelihood contribution is:

$$\ell_j(\boldsymbol{\beta}) = \sum_{t=1}^{T} \log f(y_t|\mathbf{x}_t, \boldsymbol{\beta})$$

- This is not necessarily the joint density of y_j, unless the y_{jt} are independent
- Correlated y_{jt} would result in more complicated joint densities

Pooled probit and logit

Example:

$$
\begin{aligned}
y^*_{jt} =& \mathbf{x}_{jt}\boldsymbol{\beta} + \epsilon_{jt} \\
y_{jt} =& \mathbf{1}[y^*_{jt} > 0]
\end{aligned}
$$

The partial log-likelihood for observation j is then:

$$
\ell_j(\boldsymbol{\beta}) = \sum_{t=1}^{T} \{y_{jt} \log G(\mathbf{x}_{jt}\boldsymbol{\beta}) + (1 - y_{jt}) \log(1 - G(\mathbf{x}_{jt}\boldsymbol{\beta}))\}
$$

And the total partial log-likelihood for the sample:

$$
\ell(\boldsymbol{\beta}) = \sum_{j=1}^{n} \ell_j(\boldsymbol{\beta}) = \sum_{j=1}^{n} \sum_{t=1}^{T} \{y_{jt} \log G(\mathbf{x}_{jt}\boldsymbol{\beta}) + (1 - y_{jt}) \log(1 - G(\mathbf{x}_{jt}\boldsymbol{\beta})
$$

Which has the same shape as in cross sections $\Rightarrow$ we can use the same command in R for logit/probit (glm)

Pooled probit and logit

If **strict exogeneity** is satisfied:

ϵ_{jt} is independent of $\mathbf{x}_{j1}, \ldots, \mathbf{x}_{jT}$

$$\Downarrow$$

ϵ_{jt} are serially independent

$$\Downarrow$$

y_{jt} are independent conditional on X.

$$\Downarrow$$

Joint distribution of the y_{jt}s is given by the product of the marginal (individual) distributions

$$\Downarrow$$

Partial ML = full conditional ML

Exercise

We use the data keane.dat before for the probit and logit models:

- employ = 1 if working
- black = 1 if black
- exper: years of experience
- educ: years of education
- each individual recorded from 1981 until 1987.
- Dummies y81, y82, y83, y84,y85,y86, y87

Exercise (10 minutes)

Using the glm function run the pooled probit model:

1. Exogenous variables are $black, exper, educ$
2. As above plus year dummies

Pooled probit and logit

But partial ML works also with weaker assumptions.

Case 1 : ϵ_{jt} may be serially correlated

- Example: when $\mathbf{x}_{jt}$ contains lagged dependent variables (e.g. choice last period) $\Rightarrow$ strict exogeneity is violated

Then pooled probit/logit still provides:

- Consistent estimators
- Asymptotically normal estimators

But the usual standard errors are incorrect due to potential serial correlation

Pooled probit and logit

Case 2 : The model is dynamically complete, i.e. all relevant lags of $\mathbf{y}$ and $\mathbf{X}$ are contained in $\mathbf{x}_{jt}$

Then pooled probit/logit still provides:

- Consistent estimators
- Asymptotically normal estimators

And we can use usual standard errors in this case!

Should we always include lagged variables of exogenous and dependent variables in our model?

- Only if it makes economic sense
- Explaining behaviour today by behaviour yesterday

Pooled probit and logit

A simple test of dynamic completeness:

1. Run probit and save residuals $\hat{\epsilon}_{jt} = y_{jt} - \Phi(\mathbf{x}_{jt}\hat{\boldsymbol{\beta}})$
2. Run $P(y_{jt} = 1|\mathbf{x}_{jt}, \hat{\epsilon}_{j,t-1}) = \Phi(\mathbf{x}_{jt}\boldsymbol{\beta} + \lambda\hat{\epsilon}_{j,t-1})$, $t = 2, \ldots, T$
3. If $\hat{\lambda}$ is significantly zero $\Rightarrow$ then there is dynamic completeness

SYDDANSK UNIVERSITET

Binary models with unknown effects

- The previous models are inconsistent
- FE probit model
  - Assumptions (A1)-(A3)
  - Need assumption regarding the dist. of $c_j|\mathbf{x}_j$ to work
- RE probit model
  - Assumptions (A1)-(A4)
- Alternatives to RE probit
  - Pooled probit (A1)+ (A2)+ (A4) $\Rightarrow$ good APE
  - Chamberlain's RE (A1)+ (A2)+ (A3)
- RE logit model (similar to RE probit)
- FE logit model
  - Assumptions (A1)-(A3)
  - Consistent estimates

Probit with unobserved effects

We have seen that if we assumed no unobserved individual effects, c_j $\Rightarrow$ we could use pooled probit/logit

However unobserved effects are likely to be important

Examples:

- Individual taste affects choice of transportation, choice of marriage, etc
- Individual ability affects risk of unemployment
- Unobservable characteristics affect risk of being in prison

Thus, in many cases we should allow for such unobservable effects

Probit with unobserved effects

Now the model is:

$$P(y_{jt} = 1 | \mathbf{x}_{jt}, c_j) = \Phi(\mathbf{x}_{jt}\boldsymbol{\beta} + c_j) \quad t = 1, \ldots, T$$

Assumptions:

(A1) c_j is an unobserved (individual) effect

(A2) Strict exogeneity of $\mathbf{x}_{jt}$ conditional on c_j

- No lagged variables in X
- $E(\mathbf{x}_{js}\epsilon_{jt}) = 0$ for $s, t = 1, \ldots, T$

(A3) $y_{j1}, \ldots, y_{jT}$ are independent conditional on $\mathbf{x}_j$ and c_j

Probit with unobserved effects

Under (A1)-(A3), the joint density of $y_{j1}, \ldots, y_{jT}$ is the product of the marginal densities:

$$f(y_1, \ldots, y_T | \mathbf{x}_j, c_j; \boldsymbol{\beta}) = \prod_{t=1}^{T} f(y_t | \mathbf{x}_{jt}, c_j; \boldsymbol{\beta})$$

Where:

$$f(y_t | \mathbf{x}_t, c; \boldsymbol{\beta}) = [\Phi(\mathbf{x}_t \boldsymbol{\beta} + c)]^{y_t} [1 - \Phi(\mathbf{x}_t \boldsymbol{\beta} + c)]^{1-y_t}$$

How to estimate this model?

Probit with unobserved effects

$$f(y_1, \ldots, y_T | \mathbf{x}_j, c_j; \boldsymbol{\beta}) = \prod_{t=1}^{T} f(y_t | \mathbf{x}_{jt}, c_j; \boldsymbol{\beta})$$

1) **FE probit estimation**:

- We can't get rid of the c_j by differencing as in linear models
- Then, we have to estimate c_j along with $\boldsymbol{\beta}$
- An incidental parameters problem: we get inconsistent estimates of $\boldsymbol{\beta}$ if we have to estimate too many c_j
- No restrictions on relationship between c_j and $\mathbf{x}_{jt}$ (one would be needed)
- It tends to give very biased results

Probit with unobserved effects

$$f(y_1, \ldots, y_T | \mathbf{x}_j, c_j; \boldsymbol{\beta}) = \prod_{t=1}^{T} f(y_t | \mathbf{x}_{jt}, c_j; \boldsymbol{\beta})$$

2) **RE probit estimation**:

- We need to make an additional assumption about c_j and $\mathbf{x}_j$:

  (A4)
  - Independence of c_j and $\mathbf{x}_j$ and
  - c_j is random
  - $c_j | \mathbf{x}_j \sim N(0, \sigma_c^2)$
- These are very strong assumptions

Probit with unobserved effects

$$f(y_1, \ldots, y_T | \mathbf{x}_j, c_j; \boldsymbol{\beta}) = \prod_{t=1}^{T} f(y_t | \mathbf{x}_{jt}, c_j; \boldsymbol{\beta})$$

2) **RE probit estimation**:

- Under (A1)-(A4), we can integrate c_j's out of the log-likelihood function I.e. find joint density that does not depend on c_j:

$$\begin{aligned} \ell_j(\boldsymbol{\beta}, \sigma_c) &= \log(f(y_{j1}, \ldots, y_{jT} | \mathbf{x}_{j1}, \ldots, \mathbf{x}_{jT}; \boldsymbol{\beta}, \sigma_c^2)) \\ &= \int \left[\prod_{t=1}^{T} f(y_{jt} | \mathbf{x}_{j1}, \ldots, \mathbf{x}_{jT}, c; \boldsymbol{\beta}) \right] \underbrace{\frac{1}{\sigma_c} \phi\left(\frac{c}{\sigma_c}\right)}_{density of c_j} dc \end{aligned}$$

- $\ell(\boldsymbol{\beta}, \sigma_c) = \sum_{j=1}^{n} \ell_j(\boldsymbol{\beta}, \sigma_c)$
- Log-lik is maximised wrt $\boldsymbol{\beta}$ and σ_c

Probit with unobserved effects

Alternatives to RE estimation (relaxing some of the assumptions):
A) Pooled probit, (A1)+(A2)+(A4)

$$P(y_{jt} = 1|\mathbf{x}_{jt}) = \Phi(\mathbf{x}_{jt}\boldsymbol{\beta})$$

- Error term is then: $c_j + \epsilon_{jt}$ with variance: $1 + \sigma_c^2$
- We estimate $\beta_c = \frac{\boldsymbol{\beta}}{\sqrt{1+\sigma_c^2}}$
- β_c is enough to compute APE for $\mathbf{x}_{tj}$:

$$\frac{\beta_j}{\sqrt{1+\sigma_c^2}}\phi\left(\frac{\mathbf{x}_t\boldsymbol{\beta}}{\sqrt{1+\sigma_c^2}}\right)$$

NB serial correlation due to $c_j \Rightarrow$ use eq. (13.53) to compute standard errors

Probit with unobserved effects

Alternatives to RE estimation (relaxing some of the assumptions):
B) Assume particular correlation structure for the ϵ_{jt}'s, (A1)+(A2)+(A4)

- Hard to estimate in practice.

C) Chamberlain's random effects model (A1)+(A2)+(A3)

- Relaxes assumption that c_j is independent of $\mathbf{x}_j$
- (A4).2 Specify distribution of c_j given $\mathbf{x}_j$ (some dependence):

$$c_j = \Psi + \bar{\mathbf{x}}_j \xi + a_j \quad a_j | \mathbf{x}_j \sim N(0, \sigma_a^2)$$

- Hence the latent variable is: $y_{jt}^* = \Psi + \mathbf{x}_{jt}\boldsymbol{\beta} + \bar{\mathbf{x}}_j \xi + a_j + \epsilon_{jt}$
- This is just a RE probit $P(y_{jt} = | \mathbf{x}_{jt}, c_j) = \Phi(\Psi + \mathbf{x}_{jt}\boldsymbol{\beta} + \bar{\mathbf{x}}_j \xi)$

Logit with unobserved effects

The model is similar to the probit case:

$$P(y_{jt} = 1 | \mathbf{x}_{jt}, c_j) = \Lambda(\mathbf{x}_{jt}\boldsymbol{\beta} + c_j)$$

Assumptions:

(A1) c_j is an unobserved (individual) effect

(A2) Strict exogeneity of $\mathbf{x}_{jt}$ conditional on c_j

(A3) $y_{j1}, \ldots, y_{jT}$ are independent conditional on $\mathbf{x}_j$ and c_j

(A4) $c_j | \mathbf{x}_j \sim N(0, \sigma_c^2)$

Logit with unobserved effects

Under A1-A4:

- The RE logit model
- Similar in principle to RE probit
- But much tougher to estimate $\Rightarrow$ less used

Under A1-A3:

- The FE logit model
- Similar in principle to FE probit
- But (as opposed to probit), it yields consistent estimates

Logit with unobserved effects

The FE logit model - the idea:

- We can derive joint distribution of $y_{j1}, \ldots, y_{jT}$ that does not depend on c_j's.
- But it depends on $n_j = \sum_{t=1}^{T} y_{jt}$ (# successes for person j)

Example:

- Assume $T = 2$ (two periods)
- Hence, n_j is either 0, 1 or 2
- We must derive conditional joint distribution of y_{j1} and y_{j2}

Logit with unobserved effects

Example:

- If $n_j = 0$, i.e. $y_{j1} = 0$ and $y_{j2} = 0$
- Person j yields no information about $\boldsymbol{\beta}$
- With fixed effects, we need variation within a person

- If $n_j = 2$, i.e. $y_{j1} = 1$ and $y_{j2} = 1$
- Same, person j yields no information about $\boldsymbol{\beta}$
- Only when $n_j = 1$, is person j informative about parameters in $\boldsymbol{\beta}$

Logit with unobserved effects

Example:

If $n_j = 1$, there are two possibilities:

1. $y_{j1} = 0$ and $y_{j2} = 1$ or
2. $y_{j1} = 1$ and $y_{j2} = 0$

Consider case 1:

$$
\begin{aligned}
P(y_{j1} = 0, y_{j2} = 1 | \mathbf{x}_j, c_j, n_j = 1) =& \frac{P(y_{j1} = 0 | \mathbf{x}_j, c_j) P(y_{j2} = 1 | \mathbf{x}_j, c_j)}{P(n_j = 1 | \mathbf{x}_j, c_j)} \\
=& \frac{P(y_{j1} = 0 | \mathbf{x}_j, c_j) P(y_{j2} = 1 | \mathbf{x}_j, c_j)}{P(y_{j1} = 0, y_{j2} = 1 | \mathbf{x}_j, c_j) + P(y_{j1} = 1, y_{j2} = 0 | \mathbf{x}_j, c_j)} \\
=& \Lambda((\mathbf{x}_{j2} - \mathbf{x}_{j1}) \boldsymbol{\beta})
\end{aligned}
$$

Logit with unobserved effects

It can be derived that:

$$P(y_{j1} = 0, y_{j2} = 1 | \mathbf{x}_j, c_j, n_j = 1) = \Lambda((\mathbf{x}_{j2} - \mathbf{x}_{j1})\boldsymbol{\beta})$$

And hence:

$$P(y_{j1} = 1, y_{j2} = 0 | \mathbf{x}_j, c_j, n_j = 1) = 1 - \Lambda((\mathbf{x}_{j2} - \mathbf{x}_{j1})\boldsymbol{\beta})$$

And these do not depend on c_j!

Logit with unobserved effects

The log-lik then becomes:

$$\ell_j(\boldsymbol{\beta}) = \mathbf{1}[n_j = 1]\,\{w_j \log \Lambda((\mathbf{x}_{j2} - \mathbf{x}_{j1})\boldsymbol{\beta}) + (1 - w_j)\log(1 - \Lambda((\mathbf{x}_{j2} - \mathbf{x}_{j1})\boldsymbol{\beta})$$

Note that:

- $w_j = 1$ if $y_{j1} = 0$ and $y_{j2} = 1$
- $w_j = 0$ if $y_{j1} = 1$ and $y_{j2} = 0$
- Observations with $n_j = 0$ or $n_j = 2$ do not contribute to log-lik
- Log-lik does not depend on c_j, only n_j (which is observed)
- We use change in X to explain change in behaviour

Logit with unobserved effects

- For $T > 2$, the log-lik is more complicated to derive, but the principle is the same
- c_j can be eliminated from joint distribution (and from the log-lik), see (15.73)
- Then we can maximise the log-lik without having to estimate the c_j's
- I.e. much like differencing in the linear panel data model.

NB!
We don't estimate the c_j's, but c_j's are needed to calculate partial effects and average partial effects! $\Rightarrow$ we must calculate these for some choice of c_j!

Dynamic binary models with unknown effects

- Condition (A2) ruled out lagged variables
- Dynamic unobserved effect models allow them
- RE probit with additional explanatory variables

Dynamic Unobserved Effects Models

The Probit and Logit models with unobserved effects require strict exogeneity of $\mathbf{x}_{jt}$ conditional on c_j

- I.e. errors uncorrelated with X's from all other periods
- This rules out lagged endogenous variables,
  - Example: If $\mathbf{x}_{jt} = y_{j,t-1}$ then $\mathbf{x}_{jt}$ is correlated with $\epsilon_{j,t-1}$
- But sometimes, we want to include lagged endogenous on the RHS
- We could do that in the absence of unobserved effects, c_j
- But what happens when c_j is present?

Dynamic Unobserved Effects Models

Example:

$$P(y_{jt} = 1 | y_{j,t-1}, y_{j,t-2}, \ldots, y_{j0}, Z_{jt}, c_j) = G(Z_{jt}\delta + \rho y_{j,t-1} + c_j)$$

- Probability depends on event last period
- Probability of being in prison in year $t-1$
- Probability of smoking in week $t-1$
- Probability of watching Paradise Hotel on day $t-1$

These events all exhibit *State Dependence*

- What you do today depends on what you did yesterday
- And in these cases state dependence makes economic sense

Dynamic Unobserved Effects Models

$$P(y_{jt} = 1 | y_{j,t-1}, y_{j,t-2}, \ldots, y_{j0}, Z_{jt}, c_j) = G(Z_{jt}\delta + \rho y_{j,t-1} + c_j)$$

Note:

- G can be probit or logit
- Z_{jt} satisfies strict exogeneity assumption conditional on c_j
- Lagged endogenous variable $\Rightarrow$ violates strict exogeneity (error in $t-1$ is correlated explanatory variables in t)
- $H_A : \rho \neq 0$, i.e. state dependence after controlling for c_j?
- How do we estimate δ and ρ?

Dynamic Unobserved Effects Models

$$P(y_{jt} = 1 | y_{j,t-1}, y_{j,t-2}, \ldots, y_{j0}, Z_{jt}, c_j) = G(Z_{jt}\delta + \rho y_{j,t-1} + c_j)$$

The joint density:

$$f(y_{jt}, \ldots, y_{jT} | y_{j0}, Z_{jt}, c_j) =$$

$$\prod_{t=1}^{T} G(Z_{jt}\delta + \rho y_{j,t-1} + c_j)^{y_{jt}} [1 - G(Z_{jt}\delta + \rho y_{j,t-1} + c_j)]^{1-y_{jt}}$$

- To estimate this, we need to integrate out the c_j's (as in the RE models)
- But first we must figure out what to do about the y_{j0} (the initial conditions problem)?

Dynamic Unobserved Effects Models

Several solutions available:

1. Treat y_{j0} as a non-stochastic starting value for each person, and assume a distribution of c_j given Z_j to integrate out the c_j's
   - we implictly assume that c_j and y_{j0} are independent
2. Specify a density for y_{j0} given Z_j and c_j, and multiply it on the joint density from previous slide
   - Distribution of $y_{j0}, y_{j1}, \ldots, y_{jT}$ given Z_j and c_j
   - But where does this density come from?
   - For example Heckman (1981) assumes a probit for y_{j0} given Z_j and c_j
   - Then he assumes c_j given Z_j is normal $\Rightarrow$ we can integrate out the c_j's

Dynamic Unobserved Effects Models

Several solutions available:

3. Condition on $y_{j0} \Rightarrow$ we need distribution of c_j given y_{j0} and Z_j
   - Example:

$$c_j = \Psi + \xi_0 \; y_{j0} + \xi Z_j + a_j \quad a_j \sim N(0, \sigma_a^2)$$

   - Where a_j is independent of y_{j0} and $Z_j \Rightarrow$
   - Resulting model is:

$$y_{jt} = \mathbf{1} \left[\Psi + Z_{jt}\delta + \rho y_{j,t-1} + \xi_0 \; y_{j0} + \xi Z_j + a_j + \epsilon_{jt} > 0 \right]$$

Dynamic Unobserved Effects Models

$$y_{jt} = \mathbf{1}\left[\Psi + Z_{jt}\delta + \rho y_{j,t-1} + \xi_0 \, y_{j0} + \xi Z_j + a_j + \epsilon_{jt} > 0\right]$$

This model can be estimated as a RE probit:

- $Z_{jt}, y_{j,t-1}, y_{j0}$ and Z_j are the explanatory variables
- a_j is the random effect (normally distributed and independent of the explanatory variables)
- This is simply a RE probit with additional explanatory variables.
- how would you do it in R?

Logit RE in R

- package: lme4
- function: glmer
- syntaxis:
  glmer (y∼1 + X + (1|subject), data=data,
  family=binomial("logit"))

Summary

- Panel data on binary choices
  - If no unobserved effect $\Rightarrow$ pooled probit and logit can be used
  - Also with serial correlation
- If unobserved effect and strict exogeneity
  - FE: allows for correlation with X, but only works in logit case
  - RE: does not allow for correlation (more restrictive)
- If unobserved effect $+$ lagged endogenous variable
  - Dynamic unobserved effects models
  - This is basically a RE estimation with additional explanatory variables