Censored/Truncated Outcome Panel Data

Example Questions and Solutions

230347: Advanced Microeconometrics

Question 1

Tobit Random Effects Model

Consider the model:

$$y_{it}^{\star} = \alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it} \qquad i = 1, \ldots, N \quad t = 1, \ldots, T$$

where $\alpha_i \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\alpha^2\right)$ and $\varepsilon_{it} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$. A researcher observes $\boldsymbol{x}_{it}$ and y_{it} for each i, t, where:

$$y_{it} = \begin{cases} y_{it}^{\star} & \text{if } y_{it}^{\star} > 0 \\ 0 & \text{if } y_{it}^{\star} \leq 0 \end{cases}$$

Note: You may write the pdf and cdf of the standard normal distribution as $\phi(\cdot)$ and $\Phi(\cdot)$ respectively. If $x \sim \mathcal{N}\left(\mu, \sigma^2\right)$, then the pdf of x is $\frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$.

(a) What is individual i's contribution to the likelihood at time t, $f\left(y_{it} | \boldsymbol{x}_{it}, \alpha_i, \boldsymbol{\beta}, \sigma_\varepsilon^2\right)$?

(b) What is $f\left(\boldsymbol{y}_i \,\middle|\, \boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, \sigma_\varepsilon^2\right)$, where $\boldsymbol{X}_i = (\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{iT})$, i.e. what is individual i's contribution to the likelihood?

(c) Write down individual i's contribution to the likelihood after the random effect α_i has been integrated out, $f\left(\boldsymbol{y}_i \,\middle|\, \boldsymbol{X}_i, \boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\alpha^2\right)$. Write your answer in the form:

$$f\left(\boldsymbol{y}_i \,\middle|\, \boldsymbol{X}_i, \boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\alpha^2\right) = \int_{\infty}^{\infty} f\left(\boldsymbol{y}_i \,\middle|\, \boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, \sigma_\varepsilon^2\right) f_\alpha\left(\alpha_i, \sigma_\alpha^2\right) d\alpha_i$$

(d) Perform a change in variables to be able to write $f\left(\boldsymbol{y}_i \,\middle|\, \boldsymbol{X}_i, \boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\alpha^2\right)$ in the form:

$$f\left(\boldsymbol{y}_i \,\middle|\, \boldsymbol{X}_i, \boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\alpha^2\right) = \int_{-\infty}^{\infty} g\left(r_i, \boldsymbol{y}_i, \boldsymbol{X}_i, \boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\alpha^2\right) \exp\left(-r_i^2\right) dr_i$$

(e) Using (d), describe how you would estimate the model's parameters, $\left(\boldsymbol{\beta}, \sigma_\alpha^2\right)$, in practice.

Solution

(a) When we observe $y_{it} = 0$, it means that $y_{it}^\star \leq 0$ which occurs with probability:

$$\begin{aligned}
\Pr\left(y_{it}^\star \leq 0 \,\middle|\, \boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, \sigma_\varepsilon^2\right) &= \Pr\left(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it} \leq 0 \,\middle|\, \boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, \sigma_\varepsilon^2\right) \\
&= \Pr\left(\varepsilon_{it} \leq -\alpha_i - \boldsymbol{x}_{it}'\boldsymbol{\beta} \,\middle|\, \boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, \sigma_\varepsilon^2\right) \\
&= \Phi\left(\frac{-\alpha_i - \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right) \\
&= 1 - \Phi\left(\frac{\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)
\end{aligned}$$

When we observe $y_{it} > 0$, $y_{it} = y_{it}^\star$, where $y_{it}^\star | \boldsymbol{x}_{it}, \alpha_i, \boldsymbol{\beta}, \sigma_\varepsilon^2 \sim \mathcal{N}\left(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta}, \sigma_\varepsilon^2\right)$. Then density of $y_{it}^\star | \boldsymbol{x}_{it}, \alpha_i, \boldsymbol{\beta}, \sigma_\varepsilon^2$ is then $\frac{1}{\sigma_\varepsilon}\phi\left(\frac{y_{it} - \alpha_i - \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)$. Let $d_{it} = \mathbb{1}\{y_{it} > 0\}$. Then:

$$f\left(y_{it} \,\middle|\, \boldsymbol{x}_{it}, \alpha_i, \boldsymbol{\beta}, \sigma_\varepsilon^2\right) = \left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{y_{it} - \alpha_i - \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{d_{it}} \left[1 - \Phi\left(\frac{\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{1-d_{it}}$$

(b)

$$f\left(\boldsymbol{y}_i \,\middle|\, \boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, \sigma_\varepsilon^2\right) = \prod_{t=1}^{T} \left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{y_{it} - \alpha_i - \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{d_{it}} \left[1 - \Phi\left(\frac{\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{1-d_{it}}$$

(c)

$$\begin{aligned}
f\left(\boldsymbol{y}_i \,\middle|\, \boldsymbol{X}_i, \boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\alpha^2\right) &= \int_{-\infty}^{\infty} \prod_{t=1}^{T} \left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{y_{it} - \alpha_i - \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{d_{it}} \left[1 - \Phi\left(\frac{\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{1-d_{it}} f_\alpha\left(\alpha_i, \sigma_\alpha^2\right) d\alpha_i \\
&= \int_{-\infty}^{\infty} \prod_{t=1}^{T} \left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{y_{it} - \alpha_i - \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{d_{it}} \left[1 - \Phi\left(\frac{\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{1-d_{it}} \frac{\exp\left(-\frac{\alpha_i^2}{2\sigma_\alpha^2}\right)}{\sqrt{2\pi\sigma_\alpha^2}} d\alpha_i
\end{aligned}$$

(d) Let $r_i = \frac{\alpha_i}{\sqrt{2\sigma_\alpha^2}}$ so $\alpha_i = \sqrt{2\sigma_\alpha^2}r_i$. Then $d\alpha_i = \sqrt{2\sigma_\alpha^2}dr_i$. Using these:

$$\begin{aligned}
&f\left(\boldsymbol{y}_i \,\middle|\, \boldsymbol{X}_i, \boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\alpha^2\right) \\
&= \int_{-\infty}^{\infty} \prod_{t=1}^{T} \left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{y_{it} - \sqrt{2\sigma_\alpha^2}r_i - \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{d_{it}} \left[1 - \Phi\left(\frac{\sqrt{2\sigma_\alpha^2}r_i + \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{1-d_{it}} \frac{\exp\left(-\frac{2\sigma_\alpha^2 r_i^2}{2\sigma_\alpha^2}\right)}{\sqrt{2\pi\sigma_\alpha^2}} \sqrt{2\sigma_\alpha^2}dr_i \\
&= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \prod_{t=1}^{T} \left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{y_{it} - \sqrt{2\sigma_\alpha^2}r_i - \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{d_{it}} \left[1 - \Phi\left(\frac{\sqrt{2\sigma_\alpha^2}r_i + \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{1-d_{it}} \exp\left(-r_i^2\right) dr_i
\end{aligned}$$

(e) This integral for the likelihood can be approximated using Gauss-Hermite quadrature. For H evaluation points (e.g. 12), tables or software will give you evaluation nodes z_h and weights w_h. The approximation for individual i's contribution to the likelihood is then:

$$\begin{aligned}
f\left(\boldsymbol{y}_i \,\middle|\, \boldsymbol{X}_i, \boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\alpha^2\right) \approx& \widetilde{f}\left(\boldsymbol{y}_i \,\middle|\, \boldsymbol{X}_i, \boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\alpha^2\right) \\
=& \frac{1}{\sqrt{\pi}} \sum_{h=1}^{H} w_h \prod_{t=1}^{T} \left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{y_{it} - \sqrt{2\sigma_\alpha^2}z_h - \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{d_{it}} \left[1 - \Phi\left(\frac{\sqrt{2\sigma_\alpha^2}z_h + \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)\right]^{1-d_{it}}
\end{aligned}$$

You would then estimate $\left(\boldsymbol{\beta}, \sigma_\alpha^2, \sigma_\varepsilon^2\right)$ by maximizing the (approximated) log likelihood:

$$\left(\widehat{\boldsymbol{\beta}}, \widehat{\sigma}_\alpha^2, \widehat{\sigma}_\varepsilon^2\right) = \underset{(\boldsymbol{\beta}, \sigma_\alpha^2, \sigma_\varepsilon^2)}{\arg\max} \sum_{i=1}^{N} \log\left(\widetilde{f}\left(\boldsymbol{y}_i \,|\boldsymbol{X}_i, \boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\alpha^2\right)\right)$$

Question 2

Honoré (1992) Truncated Least Squares

Consider the following data generating process:

$$y_{it}^{\star} = \alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it}, \qquad i = 1, \dots, N \quad t = 1, 2$$

where $\varepsilon_{it} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$.

A researcher observes $y_{it} = y_{it}^{\star}$ and $\boldsymbol{x}_{it}$ when $y_{it}^{\star} > 0$ and does not observe data when $y_{it} \leq 0$.

Note: If $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$, then $\mathbb{E}\left[X | X > a\right] = \mu + \sigma \frac{\phi(b)}{1 - \Phi(b)}$, where $b = \frac{a - \mu}{\sigma}$.

(a) Find $\mathbb{E}\left[y_{it} | \alpha_i, \boldsymbol{x}_{it}, \boldsymbol{\beta}, y_{it} > 0\right]$

(b) Show that $\mathbb{E}\left[y_{i2} | \alpha_i, \boldsymbol{x}_{i2}, \boldsymbol{\beta}, y_{i2} > 0\right] - \mathbb{E}\left[y_{i1} | \alpha_i, \boldsymbol{x}_{i1}, \boldsymbol{\beta}, y_{i1} > 0\right]$ in general depends on α_i.

(c) Show that if the researcher knew $\boldsymbol{\beta}$ and restricted their sample according to:

$$y_{i1} \geq -\left(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1}\right)' \boldsymbol{\beta} \qquad \text{and} \qquad y_{i2} \geq \left(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1}\right)' \boldsymbol{\beta}$$

the difference:

$$\mathbb{E}\left[y_{i2} | \boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, y_{i2} > \left(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1}\right)' \boldsymbol{\beta}\right] - \mathbb{E}\left[y_{i1} | \boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, y_{i1} > -\left(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1}\right)' \boldsymbol{\beta}\right]$$

is a constant function of α_i (does not depend on α_i), where $\boldsymbol{X}_i = \left(\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}\right)$. You should consider each of the three cases:

(a) $\left(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1}\right)' \boldsymbol{\beta} < 0$

(b) $\left(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1}\right)' \boldsymbol{\beta} = 0$

(c) $\left(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1}\right)' \boldsymbol{\beta} > 0$

Solution

(a)

$$\begin{aligned}
\mathbb{E}\left[y_{it} | \alpha_i, \boldsymbol{x}_{it}, \boldsymbol{\beta}, y_{it} > 0\right] &= \mathbb{E}\left[\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it} | \alpha_i, \boldsymbol{x}_{it}, \boldsymbol{\beta}, y_{it} > 0\right] \\
&= \alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \mathbb{E}\left[\varepsilon_{it} | \alpha_i, \boldsymbol{x}_{it}, \boldsymbol{\beta}, \varepsilon_{it} > -\alpha_i - \boldsymbol{x}_{it}'\boldsymbol{\beta}\right] \\
&= \alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \frac{\sigma_\varepsilon \phi\left(\frac{-\alpha_i - \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)}{1 - \Phi\left(\frac{-\alpha_i - \boldsymbol{x}_{it}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)}
\end{aligned}$$

(b)

$$\begin{aligned}
&\mathbb{E}\left[y_{i2}|\alpha_i, \boldsymbol{x}_{i2}, \boldsymbol{\beta}, y_{i2} > 0\right] - \mathbb{E}\left[y_{i1}|\alpha_i, \boldsymbol{x}_{i1}, \boldsymbol{\beta}, y_{i1} > 0\right] \\
&= (\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta} + \frac{\sigma_\varepsilon \phi\left(\frac{-\alpha_i - \boldsymbol{x}_{i2}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)}{1 - \Phi\left(\frac{-\alpha_i - \boldsymbol{x}_{i2}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)} - \frac{\sigma_\varepsilon \phi\left(\frac{-\alpha_i - \boldsymbol{x}_{i1}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)}{1 - \Phi\left(\frac{-\alpha_i - \boldsymbol{x}_{i1}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)}
\end{aligned}$$

Unless $\boldsymbol{x}_{i1} = \boldsymbol{x}_{i2}$, the terms containing α_i do not cancel.

(c)

- *Case 1:* $(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta} < 0$.
  The truncation does not affect y_{i2} but does affect y_{i1}:

$$\begin{aligned}
&\mathbb{E}\left[y_{i2}|\boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, y_{i2} > (\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta}\right] - \mathbb{E}\left[y_{i1}|\boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, y_{i1} > -(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta}\right] \\
&= \mathbb{E}\left[y_{i2}|\boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, y_{i2} > 0\right] - \mathbb{E}\left[y_{i1}|\boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, \alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta} + \varepsilon_{i1} > -(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta}\right] \\
&= \mathbb{E}\left[y_{i2}|\boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, \varepsilon_{i2} > -\alpha_i - \boldsymbol{x}_{i2}'\boldsymbol{\beta}\right] - \mathbb{E}\left[y_{i1}|\boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, \varepsilon_{i1} > -\alpha_i - \boldsymbol{x}_{i2}'\boldsymbol{\beta}\right] \\
&= (\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta} + \frac{\sigma_\varepsilon \phi\left(\frac{-\alpha_i - \boldsymbol{x}_{i2}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)}{1 - \Phi\left(\frac{-\alpha_i - \boldsymbol{x}_{i2}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)} - \frac{\sigma_\varepsilon \phi\left(\frac{-\alpha_i - \boldsymbol{x}_{i2}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)}{1 - \Phi\left(\frac{-\alpha_i - \boldsymbol{x}_{i2}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)} \\
&= (\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta}
\end{aligned}$$

- *Case 2:* $(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta} = 0$.
  Neither y_{i1} nor y_{i2} are affected by the additional truncation. From (b) we can see that when $\boldsymbol{x}_{i1}'\boldsymbol{\beta} = \boldsymbol{x}_{i2}'\boldsymbol{\beta}$, the two inverse Mills ratio terms will cancel and the difference is zero in expectation.
- *Case 3:* $(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta} > 0$.
  The truncation does not affect y_{i1} but does affect y_{i2}:

$$\begin{aligned}
&\mathbb{E}\left[y_{i2}|\boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, y_{i2} > (\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta}\right] - \mathbb{E}\left[y_{i1}|\boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, y_{i1} > -(\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta}\right] \\
&= \mathbb{E}\left[y_{i2}|\boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, \alpha_i + \boldsymbol{x}_{i2}'\boldsymbol{\beta} + \varepsilon_{i2} > (\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta}\right] - \mathbb{E}\left[y_{i1}|\boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, y_{i1} > 0\right] \\
&= \mathbb{E}\left[y_{i2}|\boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, \varepsilon_{i2} > -\alpha_i - \boldsymbol{x}_{i1}'\boldsymbol{\beta}\right] - \mathbb{E}\left[y_{i1}|\boldsymbol{X}_i, \alpha_i, \boldsymbol{\beta}, \varepsilon_{i1} > -\alpha_i - \boldsymbol{x}_{i1}'\boldsymbol{\beta}\right] \\
&= (\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta} + \frac{\sigma_\varepsilon \phi\left(\frac{-\alpha_i - \boldsymbol{x}_{i1}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)}{1 - \Phi\left(\frac{-\alpha_i - \boldsymbol{x}_{i1}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)} - \frac{\sigma_\varepsilon \phi\left(\frac{-\alpha_i - \boldsymbol{x}_{i1}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)}{1 - \Phi\left(\frac{-\alpha_i - \boldsymbol{x}_{i1}'\boldsymbol{\beta}}{\sigma_\varepsilon}\right)} \\
&= (\boldsymbol{x}_{i2} - \boldsymbol{x}_{i1})' \boldsymbol{\beta}
\end{aligned}$$

In all three cases, the difference is a constant function of α_i.

Question 3

Honoré and Weidner (2020)

Suppose the outcome variable y_{it} is binary. We assume y_{it} is generated according to a dynamic logit model where the probability that $y_{it} = 1$ is a function of its one-period lag y_{it-1}, a vector of strictly exogenous regressors, $\boldsymbol{x}_{it}$ and an unobserved individual fixed effect $\alpha_i \in \mathbb{R}$. We assume the initial period y_{i1} is generated according to an unspecified function $p_1(\boldsymbol{x}_{i1}, \alpha_i)$. Therefore:

$$\Pr(y_{i1} = 1 | \boldsymbol{x}_{it}, \alpha_i, \rho, \boldsymbol{\beta}) = p_1(\boldsymbol{x}_{i1}, \alpha_i)$$
$$\Pr(y_{it} = 1 | y_{it-1}, \boldsymbol{x}_{it}, \alpha_i, \rho, \boldsymbol{\beta}) = \frac{\exp(\rho y_{it-1} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \alpha_i)}{1 + \exp(\rho y_{it-1} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \alpha_i)} \qquad \text{for } t > 1$$

We observe N individuals at time periods $t = 1, 2, \ldots, T$. In what follows we will restrict ourselves to $T = 4$.

We can use a set of moments to estimate $(\rho, \boldsymbol{\beta})$. We look for moments in such a way that they are independent of the α_i:

$$\mathbb{E}[\boldsymbol{m}(\boldsymbol{y}_i, \boldsymbol{x}_i, \rho, \boldsymbol{\beta}) | y_{i1}, \boldsymbol{x}_i, \alpha_i] = \boldsymbol{0} \qquad \forall \alpha_i \in \mathbb{R}$$

where $\boldsymbol{x}_i = (\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{iT})$. One moment that satisfies this is:

$$m_0^a(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho) = \begin{cases} \exp\left((\boldsymbol{x}_{i2} - \boldsymbol{x}_{i3})'\boldsymbol{\beta}\right) & \text{if } (y_{i1}, y_{i2}, y_{i3}, y_{i4}) = (0,0,1,0) \\ \exp\left((\boldsymbol{x}_{i2} - \boldsymbol{x}_{i4})'\boldsymbol{\beta} - \rho\right) & \text{if } (y_{i1}, y_{i2}, y_{i3}, y_{i4}) = (0,0,1,1) \\ -1 & \text{if } (y_{i1}, y_{i2}, y_{i3}) = (0,1,0) \\ \exp\left((\boldsymbol{x}_{i4} - \boldsymbol{x}_{i3})'\boldsymbol{\beta} - 1\right) & \text{if } (y_{i1}, y_{i2}, y_{i3}, y_{i4}) = (0,1,1,0) \\ 0 & \text{otherwise} \end{cases}$$

Alternatively expressed, $m_0^a(\cdot)$ for every combination $(y_{i2}, y_{i3}, y_{i4}) \in \{0,1\}^3$ is shown in the table below:

y_{i1}	(y_{i2}, y_{i3}, y_{i4})	$m_0^a(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho)$
0	$(0,0,0)$	0
0	$(0,0,1)$	0
0	$(0,1,0)$	$\exp\left((\boldsymbol{x}_{i2} - \boldsymbol{x}_{i3})'\boldsymbol{\beta}\right)$
0	$(0,1,1)$	$\exp\left((\boldsymbol{x}_{i2} - \boldsymbol{x}_{i4})'\boldsymbol{\beta} - \rho\right)$
0	$(1,0,0)$	-1
0	$(1,0,1)$	-1
0	$(1,1,0)$	$\exp\left((\boldsymbol{x}_{i4} - \boldsymbol{x}_{i3})'\boldsymbol{\beta}\right) - 1$
0	$(1,1,1)$	0

(i) Show that this moment restriction is valid. That is, show that:

$$\mathbb{E}[m_0^a(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho) | y_{i1} = 0, \boldsymbol{x}_i, \alpha_i] = 0$$

where

$$\mathbb{E}[m_0^a(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho) | y_{i1} = 0, \boldsymbol{x}_i, \alpha_i] = \sum_{(y_{i2}, y_{i3}, y_{i4}) \in \{0,1\}^3} \prod_{t=2}^{4} \Pr(y_{it} | y_{it-1}, \boldsymbol{x}_{it}, \alpha_i, \rho, \boldsymbol{\beta}) \, m_0^a(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho)$$

(ii) Given this moment condition is valid, assume now that you have 3 other valid moment conditions: $m_0^b(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho)$, $m_1^a(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho)$ and $m_1^b(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho)$. These are shown in the table below but you don't need them to answer this question.

In the lectures, we saw that for this particular model if we condition on the events $y_{i2} + y_{i3} = 1$ and

$\boldsymbol{x}_{i3} = \boldsymbol{x}_{i4}$, the conditional likelihood was no longer a function of α_i and we could estimate $(\rho, \boldsymbol{\beta})$ with maximum likelihood. Given we have these valid moment conditions that do not depend on α_i, what do these imply about the estimation approach we discussed in the lectures? Which approach would be better to use?

A verbal discussion is all that is necessary to answer this question. No derivations are required.

y_{i1}	(y_{i2}, y_{i3}, y_{i4})	$m_0^a(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho)$	$m_0^b(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho)$
0	$(0,0,0)$	0	0
0	$(0,0,1)$	0	$\exp\left((\boldsymbol{x}_{i3} - \boldsymbol{x}_{i4})' \boldsymbol{\beta}\right) - 1$
0	$(0,1,0)$	$\exp\left((\boldsymbol{x}_{i2} - \boldsymbol{x}_{i3})' \boldsymbol{\beta}\right)$	-1
0	$(0,1,1)$	$\exp\left((\boldsymbol{x}_{i2} - \boldsymbol{x}_{i4})' \boldsymbol{\beta} - \rho\right)$	-1
0	$(1,0,0)$	-1	$\exp\left((\boldsymbol{x}_{i4} - \boldsymbol{x}_{i2})' \boldsymbol{\beta}\right)$
0	$(1,0,1)$	-1	$\exp\left((\boldsymbol{x}_{i3} - \boldsymbol{x}_{i2})' \boldsymbol{\beta} + \rho\right)$
0	$(1,1,0)$	$\exp\left((\boldsymbol{x}_{i4} - \boldsymbol{x}_{i3})' \boldsymbol{\beta}\right) - 1$	0
0	$(1,1,1)$	0	0
y_{i1}	(y_{i2}, y_{i3}, y_{i4})	$m_1^a(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho)$	$m_1^b(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho)$
1	$(0,0,0)$	0	0
1	$(0,0,1)$	0	$\exp\left((\boldsymbol{x}_{i3} - \boldsymbol{x}_{i4})' \boldsymbol{\beta}\right) - 1$
1	$(0,1,0)$	$\exp\left((\boldsymbol{x}_{i2} - \boldsymbol{x}_{i3})' \boldsymbol{\beta} + \rho\right)$	-1
1	$(0,1,1)$	$\exp\left((\boldsymbol{x}_{i2} - \boldsymbol{x}_{i4})' \boldsymbol{\beta}\right)$	-1
1	$(1,0,0)$	-1	$\exp\left((\boldsymbol{x}_{i4} - \boldsymbol{x}_{i2})' \boldsymbol{\beta} - \rho\right)$
1	$(1,0,1)$	-1	$\exp\left((\boldsymbol{x}_{i3} - \boldsymbol{x}_{i2})' \boldsymbol{\beta}\right)$
1	$(1,1,0)$	$\exp\left((\boldsymbol{x}_{i4} - \boldsymbol{x}_{i3})' \boldsymbol{\beta}\right) - 1$	0
1	$(1,1,1)$	0	0

Solution

(i) We want to show that:

$$\begin{aligned}
\mathbb{E}\left[m_0^a(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho) \,|\, y_{i1} = 0, \boldsymbol{x}_i, \alpha_i\right] &= \sum_{(y_{i2}, y_{i3}, y_{i4}) \in \{0,1\}^3} \prod_{t=2}^{4} \Pr\left(y_{it} | y_{it-1}, \boldsymbol{x}_{it}, \alpha_i, \rho, \boldsymbol{\beta}\right) m_0^a(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{\beta}, \rho) \\
&= 0
\end{aligned}$$

There are 8 terms in the sum (every combination of $(y_{i2}, y_{i3}, y_{i4}) \in \{0,1\}^3$), but the combinations $(0,0,0)$, $(0,0,1)$ and $(1,1,1)$ have $m_0^a = 0$ so we can remove those from the sum. This leaves us with 5 terms:

$$(0,1,0) \qquad \frac{1}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} \times \frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}}{e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}} +$$

$$(0,1,1) \qquad \frac{1}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\alpha_i}} \frac{e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} \times \frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}}{e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}} +$$

$$(1,0,0) \qquad \frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\alpha_i}} \times (-1) +$$

$$(1,0,1) \qquad \frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} \frac{e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\alpha_i}} \times (-1) +$$

$$(1,1,0) \qquad \frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} \times \left(\frac{e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}}{e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}} - 1\right)$$

Note that for the 3rd and 4th term, the expressions are identical expect for the 3rd fraction in the product. However, these fractions sum to 1 so we can combine them to $-\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}$. Doing this, cancelling some terms and rearranging:

$$\begin{aligned}
&\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} + \\
&\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}} \frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} + \\
&\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} \times (-1) + \\
&\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} \times \left(\frac{e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}}{e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}} - 1\right)
\end{aligned}$$

Now the first two terms are the same except for the 2nd fraction. But these two fractions sum to 1 so we can combine them to $\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}$. Doing this and expanding the last term (with some rearranging):

$$\begin{aligned}
&\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} + \\
&\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} \times (-1) + \\
&\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} \frac{e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} + \\
&\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}} \frac{e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} \frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}} \times (-1)
\end{aligned}$$

Since $\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}$ is common to all parts of the sum, we can write this as:

$$\begin{aligned}
&\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}\left[\frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}\left[1-\frac{e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}\right]+\frac{1}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}\left[\frac{e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}-1\right]\right]\\
=&\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}\left[\frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}\left[\frac{1}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}\right]+\frac{1}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}\left[-\frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}\right]\right]\\
=&\frac{e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}{1+e^{\boldsymbol{x}'_{i2}\boldsymbol{\beta}}e^{\alpha_i}}\left[\frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}\frac{1}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}-\frac{1}{1+e^{\boldsymbol{x}'_{i3}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}\frac{1}{1+e^{\boldsymbol{x}'_{i4}\boldsymbol{\beta}}e^{\rho}e^{\alpha_i}}\right]\\
=&0
\end{aligned}$$

Therefore this moment restriction is valid.

(ii) Using these moment restrictions we can estimate the model without restricting ourselves to the subsample where $\boldsymbol{x}_{i3} = \boldsymbol{x}_{i4}$. This is especially useful in situations where we have continuous variables or many covariates where this is unlikely to hold. Furthermore, observations where $y_{i2} + y_{i3} = 0$ ($y_{i2} + y_{i3} = 2$) are still used to identify $\boldsymbol{\beta}$, provided $y_{i4} = 1$ ($y_{i4} = 0$).