Binary Outcome Panel Data

Example Questions and Solutions

230347: Advanced Microeconometrics

Question 1

Fixed Effects Logit Model with $T = 2$

Consider the model:

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

$$y_{it} = \begin{cases} 1 & \text{if } y_{it}^{\star} > 0 \\ 0 & \text{otherwise} \end{cases}$$

ε_{it} are iid logistic so:

$$\Pr(y_{it} = 1 | \boldsymbol{x}_{it}, \boldsymbol{\beta}, \alpha_i) = \frac{\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta} + \alpha_i)}{1 + \exp(\boldsymbol{x}_{it}'\boldsymbol{\beta} + \alpha_i)}$$

(i) Show that the likelihood of $\Pr(y_{i1}, y_{i2} | \boldsymbol{x}_i, \boldsymbol{\beta}, \alpha_i)$, where $\boldsymbol{x}_i = (\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2})$ can be written as:

$$\Pr(y_{i1}, y_{i2} | \boldsymbol{x}_i, \boldsymbol{\beta}, \alpha_i) = \frac{\exp\left(\sum_{t=1}^{2} y_{it}(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta})\right)}{[1 + \exp(\alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta})][1 + \exp(\alpha_i + \boldsymbol{x}_{i2}'\boldsymbol{\beta})]}$$

(ii) What is the probability that $y_{i1} + y_{i2} = 1$?

(iii) What is the probability that $y_{i1} = 1$ and $y_{i2} = 0$ conditional on $y_{i1} + y_{i2} = 1$?

(iv) Show that the individual effects α_i cancel in the conditional likelihood in (iii).

Solution

(i) Since ε_{it} is iid logistic, the likelihood of (y_{i1}, y_{i2}) is:

$$\begin{aligned}
f(y_{i1}, y_{i2} | \boldsymbol{x}_i, \boldsymbol{\beta}, \alpha_i) &= \prod_{t=1}^{2} \left(\frac{\exp(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta})}{1 + \exp(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta})}\right)^{y_{it}} \left(\frac{1}{1 + \exp(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta})}\right)^{1-y_{it}} \\
&= \prod_{t=1}^{2} \left(\frac{\exp(y_{it}(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta}))}{1 + \exp(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta})}\right) \\
&= \frac{\exp(y_{i1}(\alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta}))}{1 + \exp(\alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta})} \times \frac{\exp(y_{i2}(\alpha_i + \boldsymbol{x}_{i2}'\boldsymbol{\beta}))}{1 + \exp(\alpha_i + \boldsymbol{x}_{i2}'\boldsymbol{\beta})} \\
&= \frac{\exp\left(\sum_{t=1}^{2} y_{it}(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta})\right)}{[1 + \exp(\alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta})][1 + \exp(\alpha_i + \boldsymbol{x}_{i2}'\boldsymbol{\beta})]}
\end{aligned}$$

(ii) If $y_{i1} + y_{i2} = 1$, then we have either $(y_{i1}, y_{i2}) = (1, 0)$ or $(y_{i1}, y_{i2}) = (0, 1)$. Therefore:

$$\Pr(y_{i1} + y_{i2} = 1) = \Pr((1,0)) + \Pr((0,1))$$

Using the answer in part (i):

$$\Pr((1,0)) = \frac{\exp(\alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta})}{[1+\exp(\alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta})][1+\exp(\alpha_i + \boldsymbol{x}_{i2}'\boldsymbol{\beta})]} \tag{1}$$

$$\Pr((0,1)) = \frac{\exp(\alpha_i + \boldsymbol{x}_{i2}'\boldsymbol{\beta})}{[1+\exp(\alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta})][1+\exp(\alpha_i + \boldsymbol{x}_{i2}'\boldsymbol{\beta})]}$$

So:

$$\Pr(y_{i1} + y_{i2} = 1) = \frac{\exp(\alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta}) + \exp(\alpha_i + \boldsymbol{x}_{i2}'\boldsymbol{\beta})}{[1+\exp(\alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta})][1+\exp(\alpha_i + \boldsymbol{x}_{i2}'\boldsymbol{\beta})]} \tag{2}$$

(iii) The conditional probability is:

$$\Pr((1,0)|y_{i1} + y_{i2} = 1) = \frac{\Pr(1,0)}{\Pr(y_{i1} + y_{i2} = 1)}$$

This is simply dividing equations (1) and (2) above. Since the denominators are identical, this is:

$$\Pr((1,0)|y_{i1} + y_{i2} = 1) = \frac{\exp(\alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta})}{\exp(\alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta}) + \exp(\alpha_i + \boldsymbol{x}_{i2}'\boldsymbol{\beta})}$$

(iv)

$$\begin{aligned}
\Pr((1,0)|y_{i1} + y_{i2} = 1) &= \frac{\exp(\alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta})}{\exp(\alpha_i + \boldsymbol{x}_{i1}'\boldsymbol{\beta}) + \exp(\alpha_i + \boldsymbol{x}_{i2}'\boldsymbol{\beta})} \\
&= \frac{e^{\alpha_i}\exp(\boldsymbol{x}_{i1}'\boldsymbol{\beta})}{e^{\alpha_i}\exp(\boldsymbol{x}_{i1}'\boldsymbol{\beta}) + e^{\alpha_i}\exp(\boldsymbol{x}_{i2}'\boldsymbol{\beta})} \\
&= \frac{\exp(\boldsymbol{x}_{i1}'\boldsymbol{\beta})}{\exp(\boldsymbol{x}_{i1}'\boldsymbol{\beta}) + \exp(\boldsymbol{x}_{i2}'\boldsymbol{\beta})}
\end{aligned}$$

Question 2

Dynamic Fixed Effects Logit Without Regressors

Suppose we have the model $y_{it} = \mathbb{1}\{\alpha_i + \rho y_{it-1} + \varepsilon_{it} > 0\}$, for $t = 2, \ldots, T$, where ε_{it} is distributed logit. Then:

$$\Pr(y_{it} = 1 | y_{it-1}, \alpha_i, \rho) = \frac{\exp(\alpha_i + \rho y_{it-1})}{1 + \exp(\alpha_i + \rho y_{i,t-1})}$$

Assume $\Pr(y_{i1} = 1|\alpha_i) = p_1(\alpha_i)$ and let $\widetilde{p}_1(\alpha_i, y_{i1}) = [p_1(\alpha_i)]^{y_{i1}}[1 - p_1(\alpha_i)]^{1-y_{i1}}$. Individual i's contribution to the likelihood is:

$$f(\boldsymbol{y}_i|\boldsymbol{y}_{-i}, \alpha_i, \rho) = \widetilde{p}_1(\alpha_i, y_{i1}) \frac{\exp\left(\alpha_i \sum_{t=2}^T y_{it}\right)\exp\left(\rho \sum_{t=2}^T y_{it-1}y_{it}\right)}{\prod_{t=2}^T [1 + \exp(\alpha_i + \rho y_{it-1})]}$$

The goal of this exercise is to show that the likelihood conditional on y_{i1}, y_{iT}, and $\sum_{t=2}^T y_{it}$ is independent of α_i.

(i) Show that we can rewrite the denominator of the likelihood as:

$$\prod_{t=2}^{T}\left[1+\exp\left(\alpha_i+\rho y_{it-1}\right)\right]=\left[1+\exp\left(\alpha_i\right)\right]^{T-1-y_{i1}+y_{iT}-\sum_{t=2}^{T}y_{it}}\left[1+\exp\left(\alpha_i+\rho\right)\right]^{y_{i1}-y_{iT}+\sum_{t=2}^{T}y_{it}}$$

Hint: You should use that $\sum_{t=2}^{T}y_{it-1}=y_{i1}-y_{iT}+\sum_{t=2}^{T}y_{it}$.

(ii) Define the set $\mathcal{B}_i=\left\{\boldsymbol{d}_i : d_{i1}=y_{i1}, d_{iT}=y_{iT}, \sum_{t=2}^{T}d_{it}=\sum_{t=2}^{T}y_{it}\right\}$. This is the set of possible vectors of length T where the first element is y_{i1}, the last element is y_{iT} and the sum of elements 2 to T is $\sum_{t=2}^{T}y_{it}$. Using your answer to (i), write down:

$$\Pr\left(d_{i1}=y_{i1}, d_{iT}=y_{iT}, \sum_{t=2}^{T}d_{it}=\sum_{t=2}^{T}y_{it}\right)=\sum_{\boldsymbol{d}_i\in\mathcal{B}_i}\Pr\left(\boldsymbol{d}_i\right)$$

explicity.

(iii) Use the answers in (i) and (ii) to show that the likelihood conditional on y_{i1}, y_{iT} and $\sum_{t=1}^{T}y_{it}$ is equal to:

$$f\left(\boldsymbol{y}_i\middle| y_{i1}=d_{i1}, y_{iT}=d_{iT}, \sum_{t=2}^{T}y_{it}=\sum_{t=2}^{T}d_{it}\right)=\frac{\exp\left(\rho\sum_{t=2}^{T}y_{it-1}y_{it}\right)}{\sum_{\boldsymbol{d}_i\in\mathcal{B}_i}\exp\left(\rho\sum_{t=2}^{T}d_{it-1}d_{it}\right)}$$

(iv) Is it possible to identify ρ in this model with $T=3$? Explain why or why not.

Solution

(i)

$$\begin{aligned}\prod_{t=2}^{T}\left[1+\exp\left(\alpha_i+\rho y_{it-1}\right)\right] &= \prod_{t=2}^{T}\left[1+\exp\left(\alpha_i\right)\right]^{1-y_{it-1}}\left[1+\exp(\alpha_i+\rho)\right]^{y_{it-1}}\\ &= \left[1+\exp\left(\alpha_i\right)\right]^{\sum_{t=2}^{T}1-y_{it-1}}\left[1+\exp(\alpha_i+\rho)\right]^{\sum_{t=2}^{T}y_{it-1}}\\ &= \left[1+\exp\left(\alpha_i\right)\right]^{T-1-y_{i1}+y_{iT}-\sum_{t=2}^{T}y_{it}}\left[1+\exp(\alpha_i+\rho)\right]^{y_{i1}-y_{iT}+\sum_{t=2}^{T}y_{it}}\end{aligned}$$

(ii)

$$\begin{aligned}&\Pr\left(d_{i1}=y_{i1}, d_{iT}=y_{iT}, \sum_{t=2}^{T}d_{it}=\sum_{t=2}^{T}y_{it}\right)\\ &=\sum_{\boldsymbol{d}_i\in\mathcal{B}_i}\Pr\left(\boldsymbol{d}_i\right)\\ &=\sum_{\boldsymbol{d}_i\in\mathcal{B}_i}\widetilde{p}_1\left(\alpha_i, d_{i1}\right)\frac{\exp\left(\alpha_i\sum_{t=2}^{T}d_{it}\right)\exp\left(\rho\sum_{t=2}^{T}d_{it-1}d_{it}\right)}{\prod_{t=2}^{T}\left[1+\exp\left(\alpha_i+\rho d_{it-1}\right)\right]}\\ &=\sum_{\boldsymbol{d}\in\mathcal{B}}\widetilde{p}_1\left(\alpha_i, d_{i1}\right)\frac{\exp\left(\alpha_i\sum_{t=2}^{T}d_{it}\right)\exp\left(\rho\sum_{t=2}^{T}d_{it-1}d_{it}\right)}{\left[1+\exp\left(\alpha_i\right)\right]^{T-1-d_{i1}+d_{iT}-\sum_{t=2}^{T}d_{it}}\left[1+\exp(\alpha_i+\rho)\right]^{d_{i1}-d_{iT}+\sum_{t=2}^{T}d_{it}}}\end{aligned}$$

(iii)

$$f\left(\boldsymbol{y}_i \middle| d_{i1} = y_{i1}, d_{iT} = y_{iT}, \sum_{t=2}^{T} d_{it} = \sum_{t=2}^{T} y_{it}\right) = \frac{\Pr\left(\boldsymbol{y_i}, d_{i1} = y_{i1}, d_{iT} = y_{iT}, \sum_{t=2}^{T} d_{it} = \sum_{t=2}^{T} y_{it}\right)}{\Pr\left(d_{i1} = y_{i1}, d_{iT} = y_{iT}, \sum_{t=2}^{T} d_{it} = \sum_{t=2}^{T} y_{it}\right)}$$
$$= \frac{\Pr\left(\boldsymbol{y_i}\right)}{\Pr\left(d_{i1} = y_{i1}, d_{iT} = y_{iT}, \sum_{t=2}^{T} d_{it} = \sum_{t=2}^{T} y_{it}\right)}$$

as knowing $\sum_{t=2}^{T} y_{it}$ does not add to the knowledge of $\boldsymbol{y}_i$. The numerator of this resulting expression is:

$$\widetilde{p}_1\left(\alpha_i, y_{i1}\right) \frac{\exp\left(\alpha_i \sum_{t=2}^{T} y_{it}\right) \exp\left(\rho \sum_{t=2}^{T} y_{it-1} y_{it}\right)}{\left[1 + \exp\left(\alpha_i\right)\right]^{T-1-y_{i1}+y_{iT}-\sum_{t=2}^{T} y_{it}} \left[1 + \exp(\alpha_i + \rho)\right]^{y_{i1}-y_{iT}+\sum_{t=2}^{T} y_{it}}}$$

The denominator is the answer to (ii):

$$\sum_{\boldsymbol{d}_i \in \mathcal{B}_i} \widetilde{p}_1\left(\alpha_i, d_{i1}\right) \frac{\exp\left(\alpha_i \sum_{t=2}^{T} d_{it}\right) \exp\left(\rho \sum_{t=2}^{T} d_{it-1} d_{it}\right)}{\left[1 + \exp\left(\alpha_i\right)\right]^{T-1-d_{i1}+d_{iT}-\sum_{t=2}^{T} d_{it}} \left[1 + \exp(\alpha_i + \rho)\right]^{d_{i1}-d_{iT}+\sum_{t=2}^{T} d_{it}}}$$

Since $d_{i1} = y_{i1}$, $d_{iT} = y_{iT}$ and $\sum_{t=2}^{T} d_{it} = \sum_{t=2}^{T} y_{it}$, the denominators of these expressions cancel, as well as the $\widetilde{p}_1\left(\alpha_i, y_{i1}\right)$. This leaves us with:

$$\frac{\exp\left(\alpha_i \sum_{t=2}^{T} y_{it}\right) \exp\left(\rho \sum_{t=2}^{T} y_{it-1} y_{it}\right)}{\sum_{\boldsymbol{d}_i \in \mathcal{B}_i} \exp\left(\alpha_i \sum_{t=2}^{T} d_{it}\right) \exp\left(\rho \sum_{t=2}^{T} d_{it-1} d_{it}\right)}$$

Finally, since $\sum_{t=2}^{T} d_{it} = \sum_{t=2}^{T} y_{it}$, we get:

$$\frac{\exp\left(\rho \sum_{t=2}^{T} y_{it-1} y_{it}\right)}{\sum_{\boldsymbol{d}_i \in \mathcal{B}_i} \exp\left(\rho \sum_{t=2}^{T} d_{it-1} d_{it}\right)}$$

(iv) No. If we condition on $y_{i1} = d_1$ and $y_{i3} = d_3$, then there is only one possible sequence of y_{it}s such that $y_{i2} + y_{i3} = d_2 + d_3$ since d_3 is fixed. Since $\mathcal{B}$ is always a singleton set when $T = 3$, the conditional likelihood is 1 for any ρ.

Question 3

In this question we will consider inference in the static pooled logit model:

$$y_{it} = \mathbb{1}\left\{\boldsymbol{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it} > 0\right\}$$

where $\boldsymbol{x}_{it}$ has K elements (including a constant) and ε_{it} is distributed logistic. Individual i's contribution to the likelihood is:

$$f\left(y_{it} | \boldsymbol{x}_{it}, \boldsymbol{\beta}\right) = \frac{\exp\left(y_{it}\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1 + \exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}$$

The log-likelihood function is:

$$\ell\ell\left(\boldsymbol{\beta}\right)=\sum_{i=1}^{N}\sum_{t=1}^{T}\log\left(f\left(y_{it}|\boldsymbol{x}_{it},\boldsymbol{\beta}\right)\right)$$

The maximum likelihood estimator of $\boldsymbol{\beta}$ is:

$$\widehat{\boldsymbol{\beta}}=\arg\max_{\boldsymbol{\beta}}\ \ell\ell\left(\boldsymbol{\beta}\right)$$

(i) Show that the jth element of the gradient of the log likelihood is equal to:

$$\frac{\partial\ell\ell\left(\boldsymbol{\beta}\right)}{\partial\beta_j}=\sum_{i=1}^{N}\sum_{t=1}^{T}x_{itj}\left(y_{it}-\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}\right)$$

where β_j and x_{itj} are the jth elements of $\boldsymbol{\beta}$ and $\boldsymbol{x}_{it}$, respectively.

(ii) Show that the jth row and kth column of the Hessian of the log likelihood is:

$$\frac{\partial^2\ell\ell\left(\boldsymbol{\beta}\right)}{\partial\beta_j\partial\beta_k}=-\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)x_{itj}x_{itk}}{\left[1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)\right]^2}$$

(iii) This part begins with describing the variance-covariance estimator for $\widehat{\boldsymbol{\beta}}$ (and more generally any maximum likelihood estimator). Taking a Taylor series approximation of the likelihood around the true $\boldsymbol{\beta}$:

$$\ell\ell\left(\widetilde{\boldsymbol{\beta}}\right)=\ell\ell\left(\boldsymbol{\beta}\right)+\ell\ell'\left(\boldsymbol{\beta}\right)\left(\widetilde{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)+\frac{1}{2}\left(\widetilde{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\ell\ell''\left(\boldsymbol{\beta}\right)\left(\widetilde{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)$$

Taking first-order conditions of this, i.e. $\ell\ell'\left(\widetilde{\boldsymbol{\beta}}\right)=\mathbf{0}$, gives our maximum likelihood estimator $\widehat{\boldsymbol{\beta}}$:

$$\ell\ell'\left(\boldsymbol{\beta}\right)+\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\ell\ell''\left(\boldsymbol{\beta}\right)=\mathbf{0}$$

$\ell\ell'\left(\boldsymbol{\beta}\right)$ is $1\times K$ and $\ell\ell''\left(\boldsymbol{\beta}\right)$ is $K\times K$. Transposing and rearranging:

$$\ell\ell''\left(\boldsymbol{\beta}\right)\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)=\left[\ell\ell'\left(\boldsymbol{\beta}\right)\right]'$$

Pre-multiplying by $\left[-\ell\ell''\left(\boldsymbol{\beta}\right)\right]^{-1}$ gives:

$$\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}=\left[-\ell\ell''\left(\boldsymbol{\beta}\right)\right]^{-1}\left[\ell\ell'\left(\boldsymbol{\beta}\right)\right]'$$

The variance-covariance matrix of the estimator $\widehat{\boldsymbol{\beta}}$ is then the expectation of the outer product of the above expression:

$$\operatorname{Var}\left(\widehat{\boldsymbol{\beta}}\middle|\boldsymbol{X}\right)=\mathbb{E}\left[\left[-\ell\ell''\left(\boldsymbol{\beta}\right)\right]^{-1}\left[\ell\ell'\left(\boldsymbol{\beta}\right)\right]'\left[\ell\ell'\left(\boldsymbol{\beta}\right)\right]\left[-\ell\ell''\left(\boldsymbol{\beta}\right)\right]^{-1}\middle|\boldsymbol{X},\boldsymbol{\beta}\right]$$

where $\boldsymbol{X}$ is the matrix of all NT observations and K regressors. If the model is correctly specified (in particular, each observation (i,t) being independent), then $\mathbb{E}\left[-\ell\ell''\left(\boldsymbol{\beta}\right)\right]^{-1}=\mathbb{E}\left[\left[\ell\ell'\left(\boldsymbol{\beta}\right)\right]'\left[\ell\ell'\left(\boldsymbol{\beta}\right)\right]\right]$ and the formula reduces to:

$$\operatorname{Var}\left(\widehat{\boldsymbol{\beta}}\middle|\boldsymbol{X}\right)=\mathbb{E}\left[\left[-\ell\ell''\left(\boldsymbol{\beta}\right)\right]^{-1}\middle|\boldsymbol{X},\boldsymbol{\beta}\right]$$

In this part, you will show that $\mathbb{E}\left[-\ell\ell''(\boldsymbol{\beta})\right]^{-1} = \mathbb{E}\left[\left[\ell\ell'(\boldsymbol{\beta})\right]'\left[\ell\ell'(\boldsymbol{\beta})\right]\right]$ for the case of the static pooled logit model. Using the expressions in parts (i) and (ii) above, show that if each observation (i,t) is independent, the jth row and kth column of the outer product of the gradient vector equals the negative of the corresponding element of the Hessian matrix in expectation:

$$\mathbb{E}\left[\frac{\partial \ell\ell(\boldsymbol{\beta})}{\partial \beta_j}\frac{\partial \ell\ell(\boldsymbol{\beta})}{\partial \beta_k}\middle| \boldsymbol{X},\boldsymbol{\beta}\right] = -\mathbb{E}\left[\frac{\partial^2 \ell\ell(\boldsymbol{\beta})}{\partial \beta_j \partial \beta_k}\middle| \boldsymbol{X},\boldsymbol{\beta}\right] = \sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\, x_{itj}x_{itk}}{\left[1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\right]^2}$$

Solution

(i)

$$\begin{aligned}
\frac{\partial \ell\ell(\boldsymbol{\beta})}{\partial \beta_j} &= \sum_{i=1}^{N}\sum_{t=1}^{T}\left(\frac{1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})}{\exp(y_{it}\boldsymbol{x}_{it}'\boldsymbol{\beta})}\right)\frac{\left[1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\right]\exp(y_{it}\boldsymbol{x}_{it}'\boldsymbol{\beta})\, y_{it}x_{itj} - \exp(y_{it}\boldsymbol{x}_{it}'\boldsymbol{\beta})\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\, x_{itj}}{\left[1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\right]^2} \\
&= \sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\left[1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\right] y_{it}x_{itj} - \exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\, x_{itj}}{\left[1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\right]} \\
&= \sum_{i=1}^{N}\sum_{t=1}^{T} x_{itj}\left[y_{it}\left(\frac{1}{1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})}\right) + (1-y_{it})\frac{-\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})}{\left[1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\right]}\right] \\
&= \sum_{i=1}^{N}\sum_{t=1}^{T} x_{itj}\left[y_{it}\left(\frac{1}{1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})}\right) + \frac{-\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})}{\left[1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\right]} + y_{it}\frac{\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})}{\left[1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\right]}\right] \\
&= \sum_{i=1}^{N}\sum_{t=1}^{T} x_{itj}\left(y_{it} - \frac{\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})}\right)
\end{aligned}$$

(ii)

$$\begin{aligned}
\frac{\partial^2 \ell\ell(\boldsymbol{\beta})}{\partial \beta_j \beta_k} &= -\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\left[1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\right]\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\, x_{itj}x_{itk} - \exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\, x_{itj}x_{itk}}{\left[1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\right]^2} \\
&= -\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\, x_{itj}x_{itk}}{\left[1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})\right]^2}
\end{aligned}$$

(iii)

$$\begin{aligned}
&\mathbb{E}\left[\frac{\partial \ell\ell(\boldsymbol{\beta})}{\partial \beta_j}\frac{\partial \ell\ell(\boldsymbol{\beta})}{\partial \beta_k}\middle| \boldsymbol{X},\boldsymbol{\beta}\right] = \\
&\mathbb{E}\left[\sum_{i=1}^{N}\sum_{t=1}^{T} x_{itj}\left(y_{it} - \frac{\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}_{it}'\boldsymbol{\beta})}\right)\sum_{j=1}^{N}\sum_{s=1}^{T} x_{jsk}\left(y_{js} - \frac{\exp(\boldsymbol{x}_{js}'\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}_{js}'\boldsymbol{\beta})}\right)\middle| \boldsymbol{X},\boldsymbol{\beta}\right]
\end{aligned}$$

This is a sum of $(N \times T)^2$ terms. If $i \neq j$ or $t \neq s$ (or both), then one element of the sum is:

$$\begin{aligned}
&\mathbb{E}\left[x_{itj}\left(y_{it}-\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}\right)x_{jsk}\left(y_{js}-\frac{\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)}\right)\middle|\boldsymbol{X},\boldsymbol{\beta}\right] \\
&=\mathbb{E}\Big[x_{itj}x_{jsk}y_{it}y_{js}-x_{itj}x_{jsk}y_{it}\frac{\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)}-x_{itj}x_{jsk}y_{js}\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)} \\
&\qquad+x_{itj}x_{jsk}\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}\frac{\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)}\Big|\boldsymbol{X},\boldsymbol{\beta}\Big]
\end{aligned}$$

Each observation is independent. Since $\mathbb{E}\left[y_{it}|\boldsymbol{x}_{it},\boldsymbol{\beta}\right]=\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}$, all these terms cancel in expectation. To see this:

$$\begin{aligned}
\mathbb{E}\left[x_{itj}x_{jsk}y_{it}y_{js}|\boldsymbol{X},\boldsymbol{\beta}\right]&=x_{itj}x_{jsk}\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}\frac{\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)} \\
\mathbb{E}\left[x_{itj}x_{jsk}y_{it}\frac{\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)}-\middle|\boldsymbol{X},\boldsymbol{\beta}\right]&=x_{itj}x_{jsk}\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}\frac{\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)} \\
\mathbb{E}\left[x_{itj}x_{jsk}y_{js}\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}-\middle|\boldsymbol{X},\boldsymbol{\beta}\right]&=x_{itj}x_{jsk}\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}\frac{\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{js}'\boldsymbol{\beta}\right)}
\end{aligned}$$

So now we are left with:

$$\mathbb{E}\left[\frac{\partial\ell\ell\left(\boldsymbol{\beta}\right)}{\partial\beta_j}\frac{\partial\ell\ell\left(\boldsymbol{\beta}\right)}{\partial\beta_k}\middle|\boldsymbol{X},\boldsymbol{\beta}\right]=\mathbb{E}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}x_{itj}x_{itk}\left(y_{it}-\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}\right)^2\middle|\boldsymbol{X},\boldsymbol{\beta}\right]$$

Consider one element of the sum:

$$\begin{aligned}
\mathbb{E}\left[x_{itj}x_{itk}\left(y_{it}-\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}\right)^2\middle|\boldsymbol{x}_{it},\boldsymbol{\beta}\right]=&x_{itj}x_{itk}\left(0-\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}\right)^2\frac{1}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)} \\
&+x_{itj}x_{itk}\left(1-\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}\right)^2\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}
\end{aligned}$$

This is:

$$\begin{aligned}
&x_{itj}x_{itk}\frac{\left[\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)\right]^2}{\left[1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)\right]^2}\frac{1}{\left[1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)\right]}+x_{itj}x_{itk}\frac{1}{\left[1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)\right]^2}\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)} \\
&=x_{itj}x_{itk}\left[\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)\right]\frac{\left[1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)\right]}{\left[1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)\right]^3} \\
&=x_{itj}x_{itk}\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{\left[1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)\right]^2}
\end{aligned}$$

Taking all terms together:

$$\mathbb{E}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}x_{itj}x_{itk}\left(y_{it}-\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}\right)^2\middle|\boldsymbol{X},\boldsymbol{\beta}\right]=\sum_{i=1}^{N}\sum_{t=1}^{T}x_{itj}x_{itk}\frac{\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)}{\left[1+\exp\left(\boldsymbol{x}_{it}'\boldsymbol{\beta}\right)\right]^2}$$