Implementation of a Conditional Logit Model with State Dependence Using Standard Preimplemented Procedures

Romain Aeberhardt* Laurent Davezies§

January 30, 2009

Abstract

Disentangling state dependence from unobserved heterogeneity is a common issue in economics. It arises for instance when studying transitions between different states on the labor market. When the outcome variable is binary, one of the usual ways is to use a conditional logit model with an appropriate conditioning suitable for a dynamic framework.

Although static conditional logit procedures are widely available, these procedures can not be used directly with state dependence. That is, it is inappropriate to just use them with a lag dependent variable in the list of regressors. Moreover, reprogramming this kind of procedures in a dynamic framework can prove quite cumbersome because the likelihood can have a very high number of terms when the number of periods increases.

Here, in the case of a binary outcome with one period state dependence, we show how to rearrange the input data in order to use standard procedures originally intended for a framework without state dependence.

Keywords: conditional logit, state dependence.
JEL Classification C23, C25, J62:

*CREST-INSEE, romain.aeberhardt@insee.fr
§CREST-INSEE

1 Introduction

Disentangling state dependence from unobserved heterogeneity is a common issue in economics. It arises for instance when studying transitions between different states on the labor market. See for instance Magnac (2000) for an example concerning subsidized training and youth employment. When the outcome variable is binary, one of the usual ways is to use a conditional logit model with an appropriate conditioning suitable for a dynamic framework as developed in Chamberlain (1985).

The conditional logit is a common procedure to deal with unobserved heterogeneity in panel data econometrics when the outcome is binary and the covariates are exogenous, in particular when there is no state dependence. In that case one needs only to condition on the number of visits in each state. When there is state dependence, for instance when the outcome at date t depends on the outcome at date $t-1$, one needs also to condition on the first and last periods in order to make the individual unobserved heterogeneity terms disappear. In that case, there must be at least four observations by individual for the model to be identified.

Conditional logit procedures are widely available. However, these procedures can not be used directly with state dependence. That is, it is inappropriate to just use them with a lag dependent variable in the list of regressors. Moreover, reprogramming this kind of procedures in a dynamic framework can prove quite cumbersome because the likelihood can have a very high number of terms when the number of periods increases.

Here, in the case of a binary outcome (two states), we provide a simple way to use standard procedures intended for a static framework in a dynamic one, by rearranging the input data.

The second part will give an idea of the method on a simple case, the third part compares the conditional likelihoods in the static and dynamic frameworks, the fourth part gives the algorithm and a proof. The last part concludes.

2 Example with $T = 4$

2.1 Conditional Logit example in a static framework with two independent time periods

The model is characterised by the following equation,

$$\mathbb{P}(z_{it}|x_{it}, \alpha_i) = \frac{e^{z_{it}(\alpha_i + x_{it}\delta)}}{1 + e^{\alpha_i + x_{it}\delta}}$$

Since z_{i1} and z_{i2} are independent,

$$\begin{aligned}
\mathbb{P}(z_{i1} = 1, z_{i2} = 0|\alpha_i, x_i) &= \frac{e^{\alpha_i + \delta x_{i1}}}{1 + e^{\alpha_i + \delta x_{i1}}} \frac{1}{1 + e^{\alpha_i + \delta x_{i2}}} \\
\mathbb{P}(z_{i1} = 0, z_{i2} = 1|\alpha_i, x_i) &= \frac{1}{1 + e^{\alpha_i + \delta x_{i1}}} \frac{e^{\alpha_i + \delta x_{i2}}}{1 + e^{\alpha_i + \delta x_{i2}}} \\
\frac{\mathbb{P}(z_{i1} = 1, z_{i2} = 0|\alpha_i, x_i)}{\mathbb{P}(z_{i1} = 0, z_{i2} = 1|\alpha_i, x_i)} &= e^{\delta(x_{i1} - x_{i2})}
\end{aligned}$$

and hence,

$$\begin{aligned}
\mathbb{P}(z_{i1} = 0, z_{i2} = 1|z_{i1} + z_{i2} = 1, \alpha_i x_i) &= \frac{\mathbb{P}(z_{i1} = 0, z_{i2} = 1|\alpha_i, x_i)}{\mathbb{P}(z_{i1} = 0, z_{i2} = 1|\alpha_i, x_i) + \mathbb{P}(z_{i1} = 1, z_{i2} = 0|\alpha_i, x_i)} \\
&= \frac{1}{1 + \mathbb{P}(z_{i1} = 1, z_{i2} = 0|\alpha_i, x_i)/\mathbb{P}(z_{i1} = 0, z_{i2} = 1|\alpha_i, x_i)} \\
&= \frac{1}{1 + e^{\delta(x_{i1} - x_{i2})}}
\end{aligned}$$

which is independent of α_i and hence,

$$\mathbb{P}(z_{i1} = 0, z_{i2} = 1|z_{i1} + z_{i2} = 1, x_i) = \frac{1}{1 + e^{\delta(x_{i1} - x_{i2})}}$$

and similarly,

$$\mathbb{P}(z_{i1} = 1, z_{i2} = 0|z_{i1} + z_{i2} = 1, x_i) = \frac{e^{\delta(x_{i1} - x_{i2})}}{1 + e^{\delta(x_{i1} - x_{i2})}}$$

Finaly, we consider the following conditional likelihood which does not depend on α_i anymore:

$$\mathbb{P}(z_{i1}, z_{i2}|z_{i1}+z_{i2}, x_i) = \frac{e^{\delta z_{i1}x_{i1}}e^{\delta z_{i2}x_{i2}}}{e^{\delta x_{i1}}+e^{\delta x_{i2}}}$$

This model is identified only on the individuals who change state between the two periods.

2.2 Conditional Logit example in a dynamic framework with four time periods

We consider here that y_{it} depends only on y_{it-1} so that the model is characterized by:

$$\mathbb{P}(y_{it}|y_{it-1}, \alpha_i) = \frac{e^{y_{it}(\alpha_i+y_{it-1}\delta)}}{1+e^{\alpha_i+y_{it-1}\delta}}$$

This time in order to make the α_i vanish, you need to condition on y_{i1}, y_{i4} and $y_{i2}+y_{i3}$. With this particular conditioning, you need to consider the two sets of events: $A = \{y_{i1}, y_{i2} = 0, y_{i3} = 1, y_{i4}\}$ and $B = \{y_{i1}, y_{i2} = 1, y_{i3} = 0, y_{i4}\}$

Then, with the same kind of calculations as in the previous part, the conditional likelihood becomes:

$$\mathbb{P}(y_{i2}, y_{i3}|y_{i1}, y_{i4}, y_{i2}+y_{i3}) = \frac{e^{\delta y_{i2}y_{i1}}e^{\delta y_{i4}y_{i3}}}{e^{\delta y_{i1}}+e^{\delta y_{i4}}}$$

Since the model is identified only on the individuals who change state between periods 2 and 3, we have always $y_{i2} = 0$ or $y_{i3} = 0$ which explains why the term $e^{\delta y_{i3}y_{i2}}$ does not appear at the numerator.

2.3 Link between the static and the dynamic frameworks

In this simple case, if you write $z_{i1} = y_{i2}$, $x_{i1} = y_{i1}$, $z_{i2} = y_{i3}$ and $x_{i2} = y_{i4}$, both likelihoods are equal. That means, that with four time periods, if you feed a standard conditional logit procedure designed for a static framework with a two period dataset modified this way, the software will maximize the "right" conditional likelihood of a dynamic framework.

In the following sections we shall see that this result generalizes to more than four periods.

3 Conditional Likelihoods

Here we consider a binary outcome with T periods in the dynamic framework $(y_{it})_{1 \leq t \leq T}$ and $T-2$ in the static framework $(z_{it})_{1 \leq t \leq T-2}$.

3.1 Covariates but no state dependence

$$\mathbb{P}(z_{i1}, \dots, z_{iT-2} | \sum_{t=1}^{T-2} z_{it}, x_{i1}, \dots, x_{iT-2}) = \frac{\exp\left(\sum_{t=1}^{T-2} z_{it} x_{it} \delta\right)}{\sum_{\tilde{z} \in B'} \exp\left(\sum_{t=1}^{T-2} \tilde{z}_{it} x_{it} \delta\right)}$$

where $B'(T-2) = \left\{(\tilde{z}_{i1}, \dots, \tilde{z}_{iT-2}) | \sum_{t=1}^{T-2} \tilde{z}_{it} = \sum_{t=1}^{T-2} z_{it}\right\}$ is the set of trajectories which are compatible with the conditioning.

3.2 State dependence but no other covariates

$$\mathbb{P}(y_{i1}, ..., y_{iT} | y_{i1}, \sum_{t=2}^{T-1} y_{it}, y_{iT}) = \frac{\exp\left(\sum_{t=2}^{T} y_{it} y_{it-1} \delta\right)}{\sum_{\tilde{y} \in B} \exp\left(\sum_{t=2}^{T} \tilde{y}_{it} \tilde{y}_{it-1} \delta\right)}$$

where $B(T) = \left\{(\tilde{y}_{i1}, \dots, \tilde{y}_{iT}) | \tilde{y}_{i1} = y_{i1}, \tilde{y}_{iT} = y_{iT}, \sum_{t=2}^{T-1} \tilde{y}_{it} = \sum_{t=2}^{T-1} y_{it}\right\}$ is the set of trajectories which are compatible with the conditioning.

3.3 Link between the static and dynamic frameworks

The aim is to appropriately choose a set of $(z_{it})_{1 \leq t \leq T-2}$ and $(x_{it})_{1 \leq t \leq T-2}$, such that the conditional likelihood in a static framework correspond to the likelihood in a dynamic one:

$$\sum_{\tilde{y} \in B} \exp\left(\sum_{t=2}^{T} (\tilde{y}_{it} \tilde{y}_{it-1} - y_{it} y_{it-1}) \delta\right) = \sum_{\tilde{z} \in B'(T-2)} \exp\left(\sum_{t=1}^{T-2} (\tilde{z}_{it} - z_{it}) x_{it} \delta\right)$$

with B and B' defined as before.

4 Algorithm and Proof

4.1 Algorithm

For any $T \geq 4$, the following conditional likelihoods coincide:

$$\mathbb{P}(z_{i1}, \ldots, z_{iT-2} | \sum_{t=1}^{T-2} z_{it}, x_{i1}, \ldots, x_{iT-2}) = \left[\sum_{\tilde{z} \in B'(T-2, \sum_{t=1}^{T-2} z_{it})} \exp\left(\sum_{t=1}^{T-2} (\tilde{z}_{it} - z_{it}) \, x_{it} \delta \right) \right]^{-1}$$

with $B'(T-2, \sum_{t=1}^{T-2} z_{it}) = \left\{ (z_{i1}, \ldots, z_{iT-2}) | \sum_{t=1}^{T-2} \tilde{z}_{it} = \sum_{t=1}^{T} z_{it} \right\}$

and

$$\mathbb{P}(y_{i2}, \ldots, y_{iT-1} | y_{i1}, \sum_{2}^{T-1} y_{it}, y_{iT}) = \left[\sum_{\tilde{y} \in B(T, \sum_{2}^{T-1} y_{it}, y_{i1}, y_{iT})} \exp\left(\sum_{t=2}^{T} (\tilde{y}_{it} \tilde{y}_{it-1} - y_{it} y_{it-1}) \, \delta \right) \right]^{-1}$$

with $B(T, \sum_{2}^{T-1} y_{it}, y_{i1}, y_{iT}) = \left\{ (\tilde{y}_{i1}, \ldots, \tilde{y}_{iT}) | \tilde{y}_{i1} = y_{i1}, \sum_{2}^{T-1} \tilde{y}_{it} = \sum_{2}^{T-1} y_{it}, \tilde{y}_{iT} = y_{iT} \right\}$

if:

- $\forall t \in [1, T-2] \cap \mathbb{N}, \quad z_{it} = y_{it+1}$
- $x_{i1} = y_{i1}$ and $x_{iT-2} = y_{iT}$
- if $T \geq 5$ $\forall t \in [2, T-3] \cap \mathbb{N}, \quad x_{it} = \begin{cases} 0 \text{ if } \sum_{\tau=2}^{t-1} x_{i\tau} = \sum_{\tau=2}^{T-1} y_{i\tau} - 1 \\ 1 \text{ if } \sum_{\tau=2}^{t-1} x_{i\tau} = t + 1 + \sum_{\tau=2}^{T-1} y_{i\tau} - T \\ y_{it} \text{ if } t + 1 + \sum_{\tau=2}^{T-1} y_{i\tau} - T < \sum_{\tau=2}^{t-1} x_{i\tau} < \sum_{\tau=2}^{T-1} y_{i\tau} - 1 \end{cases}$

with the convention: $\sum_{\tau=2}^{1} = 0$

The implicit rule is that $\sum_{t=2}^{T-3} x_{it} = \sum_{t=2}^{T-1} y_{it} - 1$. Therefore, the algorithm corresponds to $x_{it} = y_{it}$ for $t \geq 2$, as long as the number of ones is not already met: $\sum_{\tau=2}^{t-1} x_{i\tau} = \sum_{\tau=2}^{T-1} y_{i\tau} - 1$, or that the number of remaining periods is not exactly what is needed in order to meet the condition: $\sum_{\tau=2}^{t-1} x_{i\tau} + T - t - 2 = \sum_{\tau=2}^{T-1} y_{i\tau} - 1$.

4.2 Proof

4.2.1 Proof for a particular trajectory

Let T and $1 \leq k \leq T-3$, we show the result first for a special trajectory y of length T such that $(y_2 = 1, y_3 = 1, \ldots, y_{k+1} = 1, y_{k+2} = 0, \ldots, y_{T-1} = 0)$.

The inverse of the conditional likelihood for this trajectory y is thus:

$$\left[\mathbb{P}(y_{i2}, \ldots, y_{iT-1} | y_{i1}, \sum_{2}^{T-1} y_{it} = k, y_{iT})\right]^{-1} = e^{-\delta(y_{i1}+(k-1))} \sum_{\tilde{y} \in B(T,k,y_{i1},y_{iT})} \exp\left(\sum_{t=2}^{T} \tilde{y}_{it}\tilde{y}_{it-1}\delta\right)$$

The set $B(T, k, y_{i1}, y_{iT})$ is the disjointed reunion of the following sets:

$$B(T, k, y_{i1}, y_{iT}) = \bigcup_{\substack{j \in \mathbb{N} \\ \kappa \in \{0,1\} \\ \lambda \in \{0,1\}}} A_{k,T}^{\kappa,\lambda}(j) =$$

$$\bigcup_{\substack{j \in \mathbb{N} \\ \kappa \in \{0,1\} \\ \lambda \in \{0,1\}}} \left\{ \tilde{y} | \tilde{y}_1 = y_1, \tilde{y}_2 = \kappa, \tilde{y}_{T-1} = \lambda, \tilde{y}_T = y_T, \sum_{t=2}^{T-1} \tilde{y}_t = k, \sum_{t=3}^{T-1} \tilde{y}_t\tilde{y}_{t-1} = j \right\}$$

j is the number of transitions from state 1 to state 1 and $\forall \tilde{y} \in A_{k,T}^{\kappa,\lambda}(j)$, $\sum_{t=2}^{T} \tilde{y}_t\tilde{y}_{t-1} = \kappa y_1 + \lambda y_T + j$.

In order to evaluate expression $\sum_{\tilde{y} \in B(T,k,y_{i1},y_{iT})} \exp\left(\sum_{t=2}^{T} \tilde{y}_{it}\tilde{y}_{it-1}\delta\right)$, there remains only to know the number of elements in $A_{k,T}^{\kappa,\lambda}(j)$ for all $j \in \mathbb{N}$, $\kappa \in \{0,1\}$, $\lambda \in \{0,1\}$.

To answer this question we shall use the following notations: let $Red(\tilde{y})$ be the "reduced trajectory" of $\tilde{y}$. It is the trajectory which is obtained by concatenation of the successive zeros and ones. For instance, $Red(y_1, 1, 1, 1, 0, 0, 1, 1, 0, y_T) = (y_1, 1, 0, 1, 0, y_T)$. Then, let $(0, 1)_n$ (respectively $(1, 0)_n$) be "n times" sequence $(0, 1)$ (respectively sequence $(1, 0)$). This leads for instance to: $Red(y_1, 1, 1, 1, 0, 0, 1, 1, 0, y_T) = (y_1, (1, 0)_2, y_T)$.

For all $j \in \mathbb{N}$, the following relations hold:

$$A_{k,T}^{0,1}(j) = Red^{-1}(y_1, (0, 1)_{k-j}, y_T)$$

$$A_{k,T}^{1,0}(j) = Red^{-1}(y_1, (1, 0)_{k-j}, y_T)$$

$$A^{1,1}_{k,T}(j) = Red^{-1}(y_1, 1, (0,1)_{k-j-1}, y_T)$$

$$A^{0,0}_{k,T}(j) = Red^{-1}(y_1, (0,1)_{k-j}, 0, y_T)$$

Indeed, considering an element of $A^{0,1}_{k,T}(j)$, its image by function Red is of type $(y_1, (0,1)_n, y_T)$, the difference between k and n is exactly the number of transition from state 1 to state 1. Conversely $Red^{-1}(y_1, (0,1)_{k-j}, y_T)$ belongs to a set of type $A^{0,1}_{k,T}(n)$ with $n + j = k$. The three other relations can be proved the same way.

The number of elements in $Red^{-1}(y_1, (0,1)_{k-j}, y_T)$ is the number of ways to make $k-j$ groups out of k identical elements (the 1's) times the number of ways to make $k-j$ groups out of $T-2-k$ identical elements (the 0's). Using the fact that if you write down the list of identical elements, making n groups of these elements is like setting $n-1$ "barriers" between the elements. The answer is thus $C^{k-j-1}_{k-1} C^{k-j-1}_{T-3-k} = C^{j}_{k-1} C^{k-j-1}_{T-3-k}$.

Similar arguments lead to the following relations (using convention $C^l_m = 0$ if $l < 0$ or if $l > m$):

$$\#A^{0,1}_{k,T}(j) = C^{j}_{k-1} C^{k-j-1}_{T-3-k} \neq 0 \quad \text{for} \quad max(0, 2k+2-T) \leq j \leq k-1$$

$$\#A^{1,0}_{k,T}(j) = C^{j}_{k-1} C^{k-j-1}_{T-3-k} \neq 0 \quad \text{for} \quad max(0, 2k+2-T) \leq j \leq k-1$$

$$\#A^{1,1}_{k,T}(j) = C^{j}_{k-1} C^{k-j-2}_{T-3-k} \neq 0 \quad \text{for} \quad max(0, 2k+1-T) \leq j \leq k-2$$

$$\#A^{0,0}_{k,T}(j) = C^{j}_{k-1} C^{k-j}_{T-3-k} \neq 0 \quad \text{for} \quad max(0, 2k+3-T) \leq j \leq k-1$$

Therefore:

$$\sum_{\tilde{y} \in B(T,k,y_{i1},y_{iT})} \exp\left(\sum_{t=2}^{T} \tilde{y}_{it}\tilde{y}_{it-1}\delta\right) = \sum_{(\kappa,\lambda)\in\{0,1\}^2} \sum_{j=max(0,2k+3-T-\kappa-\lambda)}^{k-1-\kappa\lambda} C^{j}_{k-1} C^{k-j-\kappa-\lambda}_{T-3-k} e^{\delta(\kappa y_1 + \lambda y_T + j)}$$

Now let's go back to the conditional likelihood in the static framework with $(z_{it})_{1\leq t\leq T-2}$ and $(x_{it})_{1\leq t\leq T-2}$ defined as in the algorithm: $(z_1, x_1) = (1, y_1)$, $(z_2, x_2) = (1, 1)$, $\ldots$, $(z_k, x_k) = (1, 1)$, $(z_{k+1}, x_{k+1}) = (0, 0)$, $\ldots$, $(z_{T-3}, x_{T-3}) = (0, 0)$, $(z_{T-2}, x_{T-2}) = (0, y_T)$.

$$\left[\mathbb{P}(z_{i1},\ldots,z_{iT-2}|x,\sum_{1}^{T-2} z_{it}=k)\right]^{-1} = \exp^{-\delta(y_{i1}+(k-1))} \sum_{\tilde{z}\in B'(T-2,k)} \exp\left(\sum_{t=1}^{T-2} x_{it}\tilde{z}_{it}\delta\right)$$

The set $B'(T-2,k)$ is the disjointed reunion of the following sets:

$$B'(T-2,k) = \bigcup_{\substack{j\in\mathbb{N}\\ \kappa\in\{0,1\}\\ \lambda\in\{0,1\}}} A_{k,T}^{'\kappa,\lambda}(j) =$$

$$\bigcup_{\substack{j\in\mathbb{N}\\ \kappa\in\{0,1\}\\ \lambda\in\{0,1\}}} \left\{\tilde{z}\in\{0,1\}^{T-2}|\tilde{z}_1=\kappa,\tilde{z}_{T-2}=\lambda,\sum_{t=1}^{T-2}\tilde{z}_t=k,\sum_{t=2}^{T-3}\tilde{z}_t x_t=j\right\}$$

and $\forall\tilde{z}\in A_{k,T}^{'\kappa,\lambda}(j), \sum_{t=1}^{T-2}\tilde{z}_t x_t = \kappa y_1+\lambda y_T+j$.

$$\#A_{k,T}^{'0,1}(j) = C_{k-1}^{j}C_{T-3-k}^{k-j-1} \neq 0 \quad \text{for} \quad max(0,2k+2-T)\leq j\leq k-1$$

$$\#A_{k,T}^{'1,0}(j) = C_{k-1}^{j}C_{T-3-k}^{k-j-1} \neq 0 \quad \text{for} \quad max(0,2k+2-T)\leq j\leq k-1$$

$$\#A_{k,T}^{'1,1}(j) = C_{k-1}^{j}C_{T-3-k}^{k-j-2} \neq 0 \quad \text{for} \quad max(0,2k+1-T)\leq j\leq k-2$$

$$\#A_{k,T}^{'0,0}(j) = C_{k-1}^{j}C_{T-3-k}^{k-j} \neq 0 \quad \text{for} \quad max(0,2k+3-T)\leq j\leq k-1$$

Therefore, both likelihoods are equal for the trajectories $(y_{it})_{1\leq t\leq T}$ and $(z_{it})_{1\leq t\leq T-2}$ such that:

- $(y_1,1,\ldots,1,0,\ldots,0,y_T)$
- $(z_1,x_1)=(1,y_1)$, $(z_2,x_2)=(1,1)$, $\ldots$, $(z_k,x_k)=(1,1)$, $(z_{k+1},x_{k+1})=(0,0)$, $\ldots$, $(z_{T-3},x_{T-3})=(0,0)$, $(z_{T-2},x_{T-2})=(0,y_T)$.

4.2.2 Extension to any trajectory

Now we treat the case of any trajectory of length T with k visits to state 1 between $t=2$ and $t=T-1$. Let $(z_{it})_{1\leq t\leq T-2}$ and $(x_{it})_{1\leq t\leq T-2}$ be the trajectories as defined in the algorithm.

Because the order of the 1's and 0's in $(x_{it})_{2\leq t\leq T-3}$ does not play any role in the previous part, the following relation holds:

$$\sum_{\tilde{z}\in B'(T-2,k)} \exp\left(\sum_{t=1}^{T-2} x_{it}\tilde{z}_{it}\delta\right) = \sum_{\tilde{z}\in B'(T-2,k)} \exp^{\left(\tilde{z}_{i1}y_{i1}+\sum_{t=2}^{k}\tilde{z}_{it}+\tilde{z}_{iT-2}y_{iT}\right)\delta}$$

Therefore, with the result obtained on trajectory $x = (y_1, 1, \ldots, 1, 0, \ldots, 0, y_T)$, we have: $\sum_{\tilde{z}\in B'(T-2,k)} \exp\left(\sum_{t=1}^{T-2} x_{it}\tilde{z}_{it}\delta\right) = \sum_{\tilde{y}\in B(T,k,y_{i1},y_{iT})} \exp\left(\sum_{t=2}^{T} \tilde{y}_{it}\tilde{y}_{it-1}\delta\right)$.

There remains only to check that $\sum_{t=1}^{T-2} z_{it}x_{it} = \sum_{t=2}^{T} y_{it}y_{it-1}$, which can be shown by distinguishing whether:

- y ends with $(1, 0, \ldots, 0, y_T)$: in that case, let l be the time of the last 1, we have $\forall t < l$, $x_{it} = y_{it}$ and $\forall t \in [l, T-3] \cap \mathbb{N}$, $x_{it} = 0$, $x_{iT-2} = y_{iT}$.

  And thus $\sum_{t=1}^{T-2} z_{it}x_{it} = \sum_{t=1}^{l-1} z_{it}x_{it} = \sum_{t=2}^{l} y_{it}y_{it-1} = \sum_{t=2}^{T} y_{it}y_{it-1}$

- y ends with $(0, 1, \ldots, 1, y_T)$: in that case, let l be the time of the last 0, we have $\forall t < l$, $x_{it} = y_{it}$ and $\forall t \in [l, T-3] \cap \mathbb{N}$, $x_{it} = 1$, $x_{iT-2} = y_{iT}$.

  And thus $\sum_{t=1}^{T-2} z_{it}x_{it} = \sum_{t=1}^{l-2} z_{it}x_{it} + T - l - 1 + y_T = \sum_{t=2}^{l-1} y_{it}y_{it-1} + T - l - 1 + y_T = \sum_{t=2}^{T} y_{it}y_{it-1}$

■

5 Conclusion

In the case of a binary outcome and one period state dependence, we show how a simple rearranging of the input dataset allows to use standard conditional logit procedures in a dynamic framework although these were originally intended for a static one. Further research could probably investigate possibilities of extensions to more than two states and more than one lag of state dependence. Moreover, the proof gives a simple closed form of the likelihood in that particular case with a number of terms which does not increase too much with the number of periods.

References

CHAMBERLAIN, G. (1985): "Heterogeneity, omitted variable bias, and duration dependence," in *Longitudinal Analysis of Labor Market Data*, ed. by J. J. Heckman, and B. Singer. Cambridge University Press.

MAGNAC, T. (2000): "Subsidised training and youth employment: distinguishing unobserved heterogeneity from state dependence in labour market histories," *The Economic Journal*, 110.