Online Appendix of the Article: Identification of Average Marginal Effects in Fixed Effects Dynamic Discrete Choice Models

Victor Aguirregabiria*
University of Toronto, CEPR

Jesús M. Carro†
Universidad Carlos III de Madrid

September 17, 2024

A.1 Proof of Lemma 1

For notational simplicity, define $\alpha_{i,\mathbf{x}} \equiv \alpha_i + \mathbf{x}'\boldsymbol{\gamma}$. Using the definition of $\Delta(\alpha_i, \mathbf{x})$, we have:

$$\begin{aligned}\Delta(\alpha_i, \mathbf{x}) &= \frac{e^{\alpha_{i,\mathbf{x}}+\beta}}{1+e^{\alpha_{i,\mathbf{x}}+\beta}} - \frac{e^{\alpha_{i,\mathbf{x}}}}{1+e^{\alpha_{i,\mathbf{x}}}} = \frac{e^{\alpha_{i,\mathbf{x}}}\ \left[e^{\beta}-1\right]}{\left[1+e^{\alpha_{i,\mathbf{x}}}\right]\ \left[1+e^{\alpha_{i,\mathbf{x}}+\beta}\right]} \\ &= \left[e^{\beta}-1\right]\ \pi_{01}(\alpha_i, \mathbf{x})\ \pi_{10}(\alpha_i, \mathbf{x}).\end{aligned} \tag{A.1}$$

that gives us equation (10) in Lemma 1. We also have that:

$$\begin{aligned}\frac{\pi_{11}(\alpha_i, \mathbf{x})}{\pi_{10}(\alpha_i, \mathbf{x})}\frac{\pi_{00}(\alpha_i, \mathbf{x})}{\pi_{01}(\alpha_i, \mathbf{x})} &= \frac{e^{\alpha_{i,\mathbf{x}}+\beta}/[1+e^{\alpha_{i,\mathbf{x}}+\beta}]}{1/[1+e^{\alpha_{i,\mathbf{x}}+\beta}]}\frac{1/[1+e^{\alpha_{i,\mathbf{x}}}}{e^{\alpha_{i,\mathbf{x}}}/[1+e^{\alpha_{i,\mathbf{x}}}]} \\ &= \frac{e^{\alpha_{i,\mathbf{x}}+\beta}}{e^{\alpha_{i,\mathbf{x}}}} = e^{\beta}.\end{aligned} \tag{A.2}$$

that corresponds to equation (11) in Lemma 1. ■

*Department of Economics, University of Toronto. 150 St. George Street, Toronto, ON, M5S 3G7, Canada, victor.aguirregabiria@utoronto.ca.

†Department of Economics, Universidad Carlos III de Madrid. C./ Madrid, 126, 28903 Getafe (Madrid), Spain, jcarro@eco.uc3m.es.

A.2 Proof of Proposition 1

W.l.o.g., we consider $T = 3$.[1] For any choice sequence (y_1, y_2, y_3) and any sequence of covariates $\mathbf{x}^{\{1,3\}}$ with $\mathbf{x}_2 = \mathbf{x}_3 = \mathbf{x}$:, we have:

$$\mathbb{P}_{y_1 y_2 y_3 \mid \mathbf{x}^{\{1,3\}}} = \int p^*(y_1|\alpha_i, \mathbf{x}^{\{1,3\}}) \; \pi_{y_1 y_2}(\alpha_i, \mathbf{x}) \; \pi_{y_2 y_3}(\alpha_i, \mathbf{x}) \; f_{\alpha|\mathbf{x}^{\{1,T\}}}(\alpha_i|\mathbf{x}^{\{1,3\}}) \; d\alpha_i \qquad \text{(A.3)}$$

Applying equation (10) in Lemma 1 to equation (A.3) for $\mathbb{P}_{010 \mid \mathbf{x}^{\{1,3\}}}$ and $\mathbb{P}_{101 \mid \mathbf{x}^{\{1,3\}}}$:

$$\left\{ \begin{array}{lcl} \mathbb{P}_{010 \mid \mathbf{x}^{\{1,3\}}} & = & \dfrac{1}{e^{\beta} - 1} \displaystyle\int p^*(0|\alpha_i, \mathbf{x}^{\{1,3\}}) \; \Delta(\alpha_i, \mathbf{x}) \; f_{\alpha|\mathbf{x}^{\{1,T\}}}(\alpha_i|\mathbf{x}^{\{1,3\}}) \; d\alpha_i \\ \\ \mathbb{P}_{101 \mid \mathbf{x}^{\{1,3\}}} & = & \dfrac{1}{e^{\beta} - 1} \displaystyle\int p^*(1|\alpha_i, \mathbf{x}^{\{1,3\}}) \; \Delta(\alpha_i, \mathbf{x}) \; f_{\alpha|\mathbf{x}^{\{1,T\}}}(\alpha_i|\mathbf{x}^{\{1,3\}}) \; d\alpha_i \end{array} \right. \qquad \text{(A.4)}$$

Adding up these two equations, multiplying the resulting equation times $e^{\beta} - 1$, and taking into account that $p^*(0|\alpha_i, \mathbf{x}^{\{1,3\}}) + p^*(1|\alpha_i, \mathbf{x}^{\{1,3\}}) = 1$, we have that $AME(\mathbf{x}) = [e^{\beta} - 1]$ $[\mathbb{P}_{010 \mid (\mathbf{x}_1, \mathbf{x}, \mathbf{x})} + \mathbb{P}_{101 \mid (\mathbf{x}_1, \mathbf{x}, \mathbf{x})}]$ such that $AME(\mathbf{x}, \mathbf{x}^{\{1,3\}})$ is identified. ■

A.3 Proof of Proposition 2

In this proof, for notational simplicity but w.l.o.g., we omit $\mathbf{x}^{\{1,T\}}$ and $\boldsymbol{\theta}$ as arguments in all the functions. Remember equation (17) in Proposition 2:

$$\sum_{\mathbf{y}^{\{2,T\}} \in \mathcal{Y}^{T-1}} w_{y_1, \mathbf{y}^{\{2,T\}}} \; G\left(\mathbf{y}^{\{2,T\}}|y_1, \alpha_i\right) \; = \; \Delta(\alpha_i). \qquad \text{(A.5)}$$

(A) Sufficient condition. Multiplying (A.5) times $p^*(y_1|\alpha) \; f_\alpha(\alpha)$, integrating over α, and taking into account that, as defined in (15), $\int G(\mathbf{y}^{\{2,T\}}|y_1, \alpha) \; p^*(y_1|\alpha) \; f_\alpha(\alpha) \; d\alpha$ is equal to

[1] Given identification with $T = 3$, it is obvious that there is also identification for any value of T greater than 3, as we can take sub-histories with three periods.

$P_{\mathbf{y}^{\{1,T\}}}$, we obtain:

$$\sum_{\mathbf{y}^{\{2,T\}} \in \mathcal{Y}^{T-1}} w_{y_1, \mathbf{y}^{\{2,T\}}} \ P_{\mathbf{y}^{\{1,T\}}} \ = \ \int \Delta(\alpha) \ p^*(y_1|\alpha) \ f_\alpha(\alpha) \ d\alpha. \tag{A.6}$$

We can sum equation (A.6) over all the possible values of y_1 in $\mathcal{Y}$. Given that the sum of $p^*(y_1|\alpha)$ over all values of y_1 in $\mathcal{Y}$ is equal to 1, the right-hand-side becomes $\int \Delta(\alpha) \ f_\alpha(\alpha)$ $d\alpha$, which is the definition of AME. Furthermore, the sum of the equation (A.6) over all the possible values of y_1 implies the following equation:

$$\sum_{\mathbf{y} \in \mathcal{D} \times \mathcal{Y}^T} w_{\mathbf{y}} \ P_{\mathbf{y}^{\{1,T\}}} \ = \ AME, \tag{A.7}$$

which is equation (18) in Proposition 2.

(B) Necessary condition. The proof has two parts. In part (i), we prove that function $h(P_{\mathcal{Y}|\mathcal{X}})$ should be linear in $P_{\mathcal{Y}|\mathcal{X}}$. Then, in part (ii), we show that the linearity of the function $h(P_{\mathcal{Y}|\mathcal{X}})$ implies that equation (A.5) should hold.

Necessary – Part (i). Equality $h(P_{\mathcal{Y}|\mathcal{X}}) = AME$ should hold for every distribution f_α. Consequently, it should hold for the following three specific cases: (Case 1) a degenerate distribution where $\alpha_i = c$ with probability one, where c is constant; (Case 2) a degenerate distribution where $\alpha_i = c'$ with probability one, where c' is a constant different to c; and (Case 3) a distribution with c and c' as the only two points of support, with $q \equiv f_\alpha(c)$. For each of these three cases, AME has the following form: $AME = \Delta(c)$ in Case 1; $AME = \Delta(c')$ in Case 2; and $AME = q \ \Delta(c) + (1-q) \ \Delta(c')$ in Case 3. Therefore, function $h(P_{\mathcal{Y}|\mathcal{X}})$ should satisfy the following three restrictions:

$$\begin{cases} \text{Case 1} & : \quad h\left(\boldsymbol{P}_{\mathcal{Y}|\mathcal{X}}\right) = \Delta(\mathbf{c}) \\ \text{Case 2} & : \quad h\left(\boldsymbol{P}^{(2)}_{\mathcal{Y}|\mathcal{X}}\right) = \Delta(\mathbf{c}') \\ \text{Case 3} & : \quad h\left(\boldsymbol{P}^{(3)}_{\mathcal{Y}|\mathcal{X}}\right) = q \ \Delta(\mathbf{c}) + \ (1-q) \ text\Delta(\mathbf{c}') \end{cases} \tag{A.8}$$

Where $P_{\mathcal{Y}|\mathcal{X}}$, $P^{(2)}_{\mathcal{Y}|\mathcal{X}}$, and $P^{(3)}_{\mathcal{Y}|\mathcal{X}}$ represent the distributions of $\mathbf{y}^{\{1,T\}}$ conditional on $\mathbf{x}^{\{1,T\}}$ under the DGPs of cases 1, 2, and 3, respectively. Note that, by construction:

$$P^{(3)}_{\mathcal{Y}|\mathcal{X}} = q \ P_{\mathcal{Y}|\mathcal{X}} + (1-q) \ \ P^{(2)}_{\mathcal{Y}|\mathcal{X}} \tag{A.9}$$

This condition should hold for any arbitrary values of the constants c, c', and $q \in [0,1]$. Multiplying equation (A.8)(Case 1) times q, multiplying equation (A.8)(Case 2) times $(1-q)$, adding up these two results, and then subtracting equation (A.8)(Case 3), we get that function $h\left(\boldsymbol{P}_{\mathcal{Y}|\mathcal{X}}\right)$ should satisfy the following equation:

$$q \ h\left(\boldsymbol{P}_{\mathcal{Y}|\mathcal{X}}\right) + (1-q) \ \ h\left(\boldsymbol{P}^{(2)}_{\mathcal{Y}|\mathcal{X}}\right) = h\left(q \ \boldsymbol{P}_{\mathcal{Y}|\mathcal{X}} + (1-q) \ \ \boldsymbol{P}^{(2)}_{\mathcal{Y}|\mathcal{X}}\right). \tag{A.10}$$

The only possibility that equation (A.10) holds for any arbitrary value of c, c', and $q \in [0,1]$ is that function $h\left(\boldsymbol{P}_{\mathcal{Y}|\mathcal{X}}\right)$ is linear in $P_{\mathcal{Y}|\mathcal{X}}$, such that $h\left(\boldsymbol{P}_{\mathcal{Y}|\mathcal{X}}\right) = \sum_{\mathbf{y}^{\{1,T\}}} w_{\mathbf{y}^{\{1,T\}}} \ P_{\mathbf{y}^{\{1,T\}}}$.

Necessary – Part (ii).We need to prove that, if equation $\sum_{\mathbf{y}^{\{1,T\}}} w_{\mathbf{y}^{\{1,T\}}} \ P_{\mathbf{y}^{\{1,T\}}} = AME$ holds, then equation (A.5) should hold for every value α. The proof is by contradiction. Suppose that: (a) equation $\sum_{\mathbf{y}^{\{1,T\}}} w_{\mathbf{y}^{\{1,T\}}} \ P_{\mathbf{y}^{\{1,T\}}} = AME$ holds for any distribution f_α in the DGP; and (b) there is a value $\alpha = c$ and a value y_1) of the initial condition such that equation (A.5) does not hold: $\sum_{\mathbf{y}^{\{2,T\}}} w_{y_1, y^{\{2,T\}}} \ G\left(\mathbf{y}^{\{2,T\}}|y_1, c\right) \ \neq \ \Delta(c)$. We show below that condition (b) implies that there is a density function f_α (in fact, a continuum of density functions) such that condition (a) does not hold.

W.l.o.g., consider distributions of α with only two points support, c and c' with $f_\alpha(c) = q$. Define the following function $d(y_1, \alpha)$ that measures the extent in which equation (A.5) is not satisfied:

$$d(y_1, \alpha) \equiv \sum_{\mathbf{y}^{\{2,T\}}} w_{y_1, \mathbf{y}^{\{2,T\}}} \ G\left(\mathbf{y}^{\{2,T\}}|y_1, \alpha\right) \ - \ \Delta(\alpha) \tag{A.11}$$

Condition (b) implies that $d(y_1, c) \neq 0$. For notational simplicity but w.l.o.g., consider that the initial condition y_1) has binary support $\{0, 1\}$. Applying the same operations as in the proof of the sufficient condition, we get:

$$\sum_{\mathbf{y}^{\{1,T\}}} w_{\mathbf{y}^{\{1,T\}}} \; \mathbb{P}_{\mathbf{y}^{\{1,T\}}} - AME = \tag{A.12}$$
$$q \;\; [p^*(0|c) \; d(0,c) + p^*(1|c) \; d(1,c)] \;\; + \;\; (1-q) \;\; [p^*(0|c') \; d(0,c') + p^*(1|c') \; d(1,c')]$$

By definition, each value $d(y_1, \alpha)$ is for a particular value of α, and therefore, it does not depend on the distribution f_α. More specifically, $d(y_1, \alpha)$ does not depend on the value of q. Therefore, there always exists (a continuum of) values of q such that the right-hand side of (A.12) is different from zero, and condition (a) does not hold. ■

A.4 Proof of Proposition 3

A.4.1 Part A: Polynomial in e^α

In this proof, for notational simplicity but w.l.o.g., we omit $\mathbf{x}$ and $\boldsymbol{\theta}$ as arguments in all the functions. Using the structure of the function G in the binary choice model, as presented in equation (16), and the expression for $\Delta(\alpha)$ in Lemma 1, we can rewrite equation (17) as follows:

$$\sum_{\mathbf{y}^{\{2,T\}}} w_{y_1, \mathbf{y}^{\{2,T\}}} \prod_{t=2}^{T} \frac{e^{y_t[\alpha + \beta y_{t-1}]}}{1 + e^{\alpha + \beta y_{t-1}}} = \left(e^\beta - 1\right) \frac{e^\alpha}{1 + e^\alpha} \; \frac{1}{1 + e^{\alpha+\beta}} \tag{A.13}$$

Multiplying this equation times $[1 + e^\alpha]^{T-1} [1 + e^{\alpha+\beta}]^{T-1}$ to eliminate the denominators, and using the Binomial Theorem to expand the terms $[1 + e^\alpha]^n$ as $\sum_{k=0}^{n} \binom{n}{k} [e^\alpha]^k$, and the terms $\left[1 + e^{\alpha+\beta}\right]^n$ as $\sum_{k=0}^{n} \binom{n}{k} [e^\beta]^k \; [e^\alpha]^k$, we can represent this equation as a polynomial in e^α. Therefore, this system of equations holds for every value $\alpha \in \mathbb{R}$ if and only if the coefficients multiplying each monomial term in the polynomial are equal to zero. This result defines a finite system of equations. More specifically, the coefficients multiplying each monomial term are linear functions of the weights $w_{y_1, y^{\{2,T\}}}$. The finite system of equations that makes the monomial coefficients equal to zero is linear in the weights $w_{y_1, y^{\{2,T\}}}$. ■

A.4.2 Part B: Identification – Model with covariates

W.l.o.g. we consider $T = 3$ and $t = 3$. In this case, equation (17) takes the following form:

$$\sum_{y_2,y_3} w_{y_1,y_2,y_3|\mathbf{x}^{\{1,3\}}} \frac{e^{y_2[\alpha+\beta y_1+\mathbf{x}_2'\boldsymbol{\gamma}]}}{1+e^{\alpha+\beta y_1+\mathbf{x}_2'\boldsymbol{\gamma}}} \frac{e^{y_3[\alpha+\beta y_2+\mathbf{x}_3'\boldsymbol{\gamma}]}}{1+e^{\alpha+\beta y_2+\mathbf{x}_3'\boldsymbol{\gamma}}} = \frac{\left(e^{\beta}-1\right)\ e^{\alpha+\mathbf{x}_3'\boldsymbol{\gamma}}}{\left(1+e^{\alpha+\mathbf{x}_3'\boldsymbol{\gamma}}\right)\left(1+e^{\alpha+\beta+\mathbf{x}_3'\boldsymbol{\gamma}}\right)} \tag{A.14}$$

Multiplying this equation by the factor that eliminates the denominators, we get the following equality of polynomials in e^{α}:

$$\begin{aligned}
& w_{000} + w_{000}\ e^{\mathbf{x}_3'\boldsymbol{\gamma}} e^{\beta}\ e^{\alpha} + w_{001}\ e^{\mathbf{x}_3'\boldsymbol{\gamma}}\ e^{\alpha} + w_{001}\ (e^{\mathbf{x}_3'\boldsymbol{\gamma}})^2 e^{\beta}\ (e^{\alpha})^2 \\
+ \quad & w_{010}\ e^{\mathbf{x}_2'\boldsymbol{\gamma}}\ e^{\alpha} + w_{010}\ e^{\mathbf{x}_2'\boldsymbol{\gamma}} e^{\mathbf{x}_3'\boldsymbol{\gamma}}\ (e^{\alpha})^2 \\
+ \quad & w_{011}\ e^{\mathbf{x}_2'\boldsymbol{\gamma}} e^{\mathbf{x}_3'\boldsymbol{\gamma}} e^{\beta}\ (e^{\alpha})^2 + w_{011}\ e^{\mathbf{x}_2'\boldsymbol{\gamma}} (e^{\mathbf{x}_3'\boldsymbol{\gamma}})^2 e^{\beta}\ (e^{\alpha})^3 \\
= \quad & e^{\mathbf{x}_3'\boldsymbol{\gamma}}(e^{\beta}-1)\ e^{\alpha} + e^{\mathbf{x}_2'\boldsymbol{\gamma}} e^{\mathbf{x}_3'\boldsymbol{\gamma}}(e^{\beta}-1)\ (e^{\alpha})^2
\end{aligned} \tag{A.15}$$

To ensure that a solution to this condition holds for every α, the coefficients of each monomial term must be equal on both sides of the equation. Therefore, the solution implies the following four equations:

$$\begin{aligned}
w_{000} &= 0 \\
w_{000}\ e^{\mathbf{x}_3'\boldsymbol{\gamma}} e^{\beta} + w_{001}\ e^{\mathbf{x}_3'\boldsymbol{\gamma}} + w_{010}\ e^{\mathbf{x}_2'\boldsymbol{\gamma}} &= e^{\mathbf{x}_3'\boldsymbol{\gamma}}(e^{\beta}-1) \\
w_{001}\ (e^{\mathbf{x}_3'\boldsymbol{\gamma}})^2 e^{\beta} + w_{010}\ e^{\mathbf{x}_2'\boldsymbol{\gamma}} e^{\mathbf{x}_3'\boldsymbol{\gamma}} + w_{011}\ e^{\mathbf{x}_2'\boldsymbol{\gamma}} e^{\mathbf{x}_3'\boldsymbol{\gamma}} e^{\beta} &= e^{\mathbf{x}_2'\boldsymbol{\gamma}} e^{\mathbf{x}_3'\boldsymbol{\gamma}}(e^{\beta}-1) \\
w_{011}\ e^{\mathbf{x}_2'\boldsymbol{\gamma}} (e^{\mathbf{x}_3'\boldsymbol{\gamma}})^2 e^{\beta} &= 0
\end{aligned} \tag{A.16}$$

The unique solution to this system is $w_{000} = w_{011} = 0$, $w_{001} = e^{[\mathbf{x}_2-\mathbf{x}_3]'\boldsymbol{\gamma}} - 1$, and $w_{010} = e^{[\mathbf{x}_3-\mathbf{x}_2]'\boldsymbol{\gamma}} e^{\beta} - 1$.

We can proceed in the same way for the case of $y_1 = 1$ to obtain the following unique solution for the weights: $w_{100} = w_{111} = 0$, $w_{101} = e^{[\mathbf{x}_2-\mathbf{x}_3]'\boldsymbol{\gamma}} e^{\beta} - 1$, and $w_{110} = e^{[\mathbf{x}_3-\mathbf{x}_2]'\boldsymbol{\gamma}} - 1$.

Putting these pieces together, we have that:

$$AME(\mathbf{x}_3, \mathbf{x}^{\{1,3\}}) = \begin{array}{ll} & w_{0,0,1,\mathbf{x}^{\{1,3\}}}\ \mathbb{P}_{0,0,1|\mathbf{x}^{\{1,3\}}} + w_{0,1,0,\mathbf{x}^{\{1,3\}}}\ \mathbb{P}_{0,1,0|\mathbf{x}^{\{1,3\}}} \\ + & w_{1,0,1,\mathbf{x}^{\{1,3\}}}\ \mathbb{P}_{1,0,1|\mathbf{x}^{\{1,3\}}} + w_{1,1,0,\mathbf{x}^{\{1,3\}}}\ \mathbb{P}_{1,1,0|\mathbf{x}^{\{1,3\}}} \end{array} \tag{A.17}$$

with

$$\begin{array}{l} w_{0,0,1,\mathbf{x}^{\{1,3\}}} = -1 + e^{[\mathbf{x}_2 - \mathbf{x}_3]'\boldsymbol{\gamma}}\ ;\ w_{0,1,0,\mathbf{x}^{\{1,3\}}} = -1 + e^{\beta + [\mathbf{x}_3 - \mathbf{x}_2]'\boldsymbol{\gamma}}\ ; \\ \\ w_{1,0,1,\mathbf{x}^{\{1,3\}}} = -1 + e^{\beta + [\mathbf{x}_2 - \mathbf{x}_3]'\boldsymbol{\gamma}}\ ;\ w_{1,1,0,\mathbf{x}^{\{1,3\}}} = -1 + e^{[\mathbf{x}_3 - \mathbf{x}_2]'\boldsymbol{\gamma}}. \end{array} \tag{A.18}$$

In the model without covariates, the vector of parameters γ equals zero. Consequently, the weights w_{001} and w_{110} are also zero, and we get $AME = \left[e^{\beta} - 1\right]\ [\mathbb{P}_{0,1,0} + \mathbb{P}_{1,0,1}]$. ■

A.4.3 Over-identification when T is greater than three

The identification result using only 3 periods proves identification for any $T \geq 3$ because, with more than 3 periods, we can always take 3 periods. Nonetheless, it is possible to obtain close form expression for higher values of T using the same procedure based on Proposition 2. This expression will use all T periods without having to combine several 3-periods estimates.

For $T > 3$, as said, there is overidentification, so more than one combination of the probability of histories exists. One way to choose one of them is to focus on the probabilities of the sufficient statistics used to identify β in the CMLE. Specifically, for this model and other logit models, the log-probability of a choice history has the following structure:

$$\ln \mathbb{P}\left(\mathbf{y}_i | \alpha_i, \beta\right) = \mathbf{s}(\mathbf{y}_i)'\ \mathbf{g}(\alpha_i) + \mathbf{c}(\mathbf{y}_i)'\ \beta \tag{A.19}$$

where $s(y_i)$ and $c(y_i)$ are vectors of statistics (functions of y_i), and $g(\alpha_i)$ is a vector of functions α_i. $s(y)$ is a sufficient statistic for α_i because $P(y_i|\alpha_i, \beta, s_i) = P(y_i|\beta, s_i)$.[2] Let S_T be the set of possible values of $s(y)$, let $P_{\mathbf{s}}$ be the probability of a value s of $s(y)$, and let

[2]See Aguirregabiria, Gu, and Luo (2021) for further details on this decomposition of the probability choice and sufficient statistics for discrete choice logit models.

$P_{\mathbf{s}} \equiv \{P_{\mathbf{s}} : s \in S_T\}$ be the probability distribution of this statistic. Given θ, the empirical distribution $P_{\mathbf{s}}$ contains all the information in the data about the distribution of α_i, and therefore, about AMEs. Taking into account the structure of the probability of a choice history in the equation (A.19), the model implies:

$$\mathbb{P}_{\mathbf{s}} = \sum_{\mathbf{y}:\, \mathbf{s}(\mathbf{y})=\mathbf{s}} \left[\int e^{\mathbf{s}'\, \mathbf{g}(\alpha_i) + \mathbf{c}(\mathbf{y})'\, \theta} \; f_\alpha(\alpha_i)\; d\alpha_i \right] \tag{A.20}$$

If two sequences, say k and l, have the same $s(y_j)$, the ratio of the probabilities of these two sequences is equal to $e^{c(\mathbf{y}_k)'\,\beta - c(\mathbf{y}_l)'\,\beta}$, which is not a function of α_i. This includes the case in which $P(y_j\ |\beta, \alpha_i)$ is the same for both sequences. Therefore, the set of sequences with the same or proportional $P(y_j\ |\beta, \alpha_i)$ is the set of sequences with the same value of the sufficient statistic $s(y_j)$. This result leads to an infinite number of combinations of these sequences, with the only restriction being that all the combinations have to sum up to the same number (overall weight). We choose the combination in which all these sequences have the same weight w, and, therefore, look for combinations of $P_{\mathbf{s}}$ instead of $P_{\mathbf{y}}$.

In the BC-AR1 model the sufficient statistics $s(y_i)$ is the vector $\left(y_{i1}, y_{iT}, \sum_{t=2}^{T} y_{it}\right)'$ –see Aguirregabiria, Gu, and Luo (2021)– and it can take $4T-4$ different values, $2T-2$ values with $y_{i1}=0$ and $2T-2$ values with $y_{i1}=1$. The conditions from Proposition 2 are:

$$\left.\begin{array}{rcl} \displaystyle\sum_{j=1}^{2T-2} w_j\ \mathbb{P}\left(\mathbf{s_j} \mid y_{j1}=0, \beta, \alpha_i\right) & = & \Delta(\alpha_i) \\ \displaystyle\sum_{j=2T-2+1}^{4T-4} w_j\ \mathbb{P}\left(\mathbf{s_j} \mid y_{j1}=1, \beta, \alpha_i\right) & = & \Delta(\alpha_i) \end{array}\right\} \text{ for every } \alpha_i \in \mathbb{R}, \tag{A.21}$$

where $P\left(\mathbf{s_j} \mid y_{j1}=0, \beta, \alpha_i\right) = \sum_{\mathbf{y}:\mathbf{s}(\mathbf{y})=\mathbf{s}_j} P\left(\mathbf{y}^{\{2,T\}} | y_1 = 0, \beta, \alpha_i\right)$

Proceeding similarly, we obtain the weights for AME in the binary choice AR(1) model for different values of T. These are in Tables A.1 and A.2.

Table A.1: **Weights for histories with** $y_1 = 0$

$(y_1, y_T, \sum_{t=2}^{T} y_t)$	$T=4$	$T=5$	$T=6$	$T=7$
$(0,0,0)$	0	0	0	0
$(0,0,1)$	$\frac{e^\beta-1}{2}$	$\frac{e^\beta-1}{3}$	$\frac{e^\beta-1}{4}$	$\frac{e^\beta-1}{5}$
$(0,1,1)$	0	0	0	0
$(0,0,2)$	0	$\frac{e^\beta-1}{1+2e^\beta}$	$\frac{2(e^\beta-1)}{3+3e^\beta}$	$\frac{3(e^\beta-1)}{6+4e^\beta}$
$(0,1,2)$	$\frac{e^\beta-1}{1+e^\beta}$	$\frac{e^\beta-1}{2+e^\beta}$	$\frac{e^\beta-1}{3+e^\beta}$	$\frac{e^\beta-1}{4+e^\beta}$
$(0,0,3)$	Not possible	0	$\frac{e^\beta-1}{2+2e^\beta}$	$\frac{(e^\beta-1)(1+2e^\beta)}{1+6e^\beta+3e^{2\beta}}$
$(0,1,3)$	0	$\frac{e^\beta-1}{2+e^\beta}$	$\frac{(e^\beta-1)(1+e^\beta)}{1+4e^\beta+e^{2\beta}}$	$\frac{(e^\beta-1)(2+e^\beta)}{3+6e^\beta+e^{2\beta}}$
$(0,0,4)$	Not possible	Not possible	0	$\frac{e^\beta-1}{3+2e^\beta}$
$(0,1,4)$	Not possible	0	$\frac{e^\beta-1}{3+e^\beta}$	$\frac{(e^\beta-1)(2+e^\beta)}{3+6e^\beta+e^{2\beta}}$
$(0,0,5)$	Not possible	Not possible	Not possible	0
$(0,1,5)$	Not possible	Not possible	0	$\frac{e^\beta-1}{4+e^\beta}$
$(0,1,6)$	Not possible	Not possible	Not possible	0

Table A.2: **Weights for histories with** $y_1 = 1$

$(y_1, y_T, \sum_{t=2}^{T} y_t)$	$T=4$	$T=5$	$T=6$	$T=7$
$(1,0,0)$	0	0	0	0
$(1,0,1)$	$\frac{e^\beta-1}{1+e^\beta}$	$\frac{e^\beta-1}{2+e^\beta}$	$\frac{e^\beta-1}{3+e^\beta}$	$\frac{e^\beta-1}{4+e^\beta}$
$(1,1,1)$	0	0	0	0
$(1,0,2)$	0	$\frac{e^\beta-1}{2+e^\beta}$	$\frac{(e^\beta-1)(1+e^\beta)}{1+4e^\beta+e^{2\beta}}$	$\frac{(e^\beta-1)(2+e^\beta)}{3+6e^\beta+e^{2\beta}}$
$(1,1,2)$	$\frac{e^\beta-1}{2}$	$\frac{e^\beta-1}{1+2e^\beta}$	$\frac{e^\beta-1}{2+2e^\beta}$	$\frac{e^\beta-1}{3+2e^\beta}$
$(1,0,3)$	Not possible	0	$\frac{e^\beta-1}{3+e^\beta}$	$\frac{(e^\beta-1)(2+e^\beta)}{3+6e^\beta+e^{2\beta}}$
$(1,1,3)$	0	$\frac{e^\beta-1}{3}$	$\frac{2(e^\beta-1)}{3+e^\beta}$	$\frac{(e^\beta-1)(1+2e^\beta)}{1+6e^\beta+3e^{2\beta}}$
$(1,0,4)$	Not possible	Not possible	0	$\frac{e^\beta-1}{4+e^\beta}$
$(1,1,4)$	Not possible	0	$\frac{e^\beta-1}{4}$	$\frac{3(e^\beta-1)}{6+4e^\beta}$
$(1,0,5)$	Not possible	Not possible	Not possible	0
$(1,1,5)$	Not possible	Not possible	0	$\frac{e^\beta-1}{5}$
$(1,1,6)$	Not possible	Not possible	Not possible	0

A.5 Proof of Proposition 4

We omit $\mathbf{x}$ as an argument throughout this proof for notational simplicity. However, one should understand that the probability of the initial conditions p^*, the density function of $\boldsymbol{\alpha}_i$, the empirical probabilities of choice histories and the average transition probabilities are all conditional on $\mathbf{x}_i^{\{1,3\}} = [\mathbf{x}_1, \mathbf{x}, \mathbf{x}]$.

By definition of Π_{00}, and taking into account that $p^*(0|\alpha_i) + p^*(1|\alpha_i) = 1$, we have that:

$$\Pi_{00} = \int [p^*(0|\alpha_i) + p^*(1|\alpha_i)] \;\; \pi_{00}(\alpha_i) \; f_\alpha(\alpha_i) \; d\alpha_i \tag{A.22}$$

This expression includes the term $\int p^*(0|\alpha_i) \; \pi_{00}(\alpha_i) \; f_\alpha(\alpha_i) \; d\alpha_i$, which is equal to the choice history probability $\mathbb{P}_{00}$. However, it also includes the "counterfactual" $\int p^*(1|\alpha_i) \; \pi_{00}(\alpha_i)$ $f_\alpha(\alpha_i) \; d\alpha_i$. Denote this counterfactual as δ_{100}. Given that $\pi_{10}(\alpha_i) + \pi_{11}(\alpha_i) = 1$, we can represent this counterfactual as:

$$\delta_{100} = \int p^*(1|\alpha_i) \;\; [\pi_{10}(\alpha_i) + \pi_{11}(\alpha_i)] \;\; \pi_{00}(\alpha_i) \; f_\alpha(\alpha_i) \tag{A.23}$$

This equation shows that δ_{100} is the sum of two terms. The first term is $\int p^*(1|\alpha_i) \; \pi_{10}(\alpha_i)$ $\pi_{00}(\alpha_i) \; f_\alpha(\alpha_i)$, which is equal to the choice history probability $\mathbb{P}_{100}$. The second term is $\int p^*(1|\alpha_i) \; \pi_{11}(\alpha_i) \; \pi_{00}(\alpha_i) \; f_\alpha(\alpha_i)$, which in principle is a counterfactual. However, by Lemma 1, we have that $\pi_{11}(\alpha_i)\pi_{00}(\alpha_i)$ is equal to $e^\beta \; \pi_{10}(\alpha_i)\pi_{01}(\alpha_i)$. Therefore, the counterfactual $\int p^*(1|\alpha_i) \; \pi_{11}(\alpha_i) \; \pi_{00}(\alpha_i) \; f_\alpha(\alpha_i)$ is equal to $e^\beta \int p^*(1|\alpha_i) \; \pi_{10}(\alpha_i) \; \pi_{01}(\alpha_i) \; f_\alpha(\alpha_i)$, and in turn this is equal to $e^\beta \; \mathbb{P}_{101}$.

Putting all the pieces together, we have that:

$$\Pi_{00} = \mathbb{P}_{00} + \mathbb{P}_{100} + e^\beta \; \mathbb{P}_{101} \tag{A.24}$$

Using the same procedure, we can show that $\Pi_{11} = \mathbb{P}_{11} + \mathbb{P}_{011} + e^\beta \; \mathbb{P}_{010}$. ■

A.6 Proof of Lemma 2

For clarity in notation, we refrain from explicitly including $\mathbf{x}$ as an argument throughout this proof. It is important to note, however, that all probabilities and expectations in this proof are conditioned on $\mathbf{x}_i^{\{1,T\}} = [\mathbf{x}_1, \mathbf{x}, ..., \mathbf{x}]$, where $\mathbf{x}_1$ is free, and $\mathbf{x}_2 = ... = \mathbf{x}_T = \mathbf{x}$.

Using the Markov structure of the model and the chain rule, we have that:

$$\begin{aligned}\mathbb{E}(y_{i,t+n} \mid \alpha_i, y_{it}) &= \mathbb{P}\left(y_{i,t+n-1} = 0 \mid \alpha_i, y_{it}\right)\ \pi_{01}(\alpha_i) + \mathbb{P}\left(y_{i,t+n-1} = 1 \mid \alpha_i, y_{it}\right)\ \pi_{11}(\alpha_i) \\ &= \pi_{01}(\alpha_i) + \mathbb{E}\left(y_{i,t+n-1} \mid \alpha_i, y_{it}\right)\ \left[\pi_{11}(\alpha_i) - \pi_{01}(\alpha_i)\right]\end{aligned} \tag{A.25}$$

Given the definition of $\Delta^{(n)}(\alpha_i)$ as $\mathbb{E}(y_{i,t+n} \mid \alpha_i, y_{it} = 1) - \mathbb{E}(y_{i,t+n}|\alpha_i, y_{it} = 0)$, and applying equation (A.25), we have that:

$$\begin{aligned}\Delta^{(n)}(\alpha_i) &= \left[\mathbb{E}\left(y_{i,t+n-1} \mid \alpha_i, y_{it} = 1\right) - \mathbb{E}\left(y_{i,t+n-1} \mid \alpha_i, y_{it} = 0\right)\right]\ \left[\pi_{11}(\alpha_i) - \pi_{01}(\alpha_i)\right] \\ &= \Delta^{(n-1)}(\alpha_i)\ \left[\pi_{11}(\alpha_i) - \pi_{01}(\alpha_i)\right]\end{aligned} \tag{A.26}$$

Applying this expression recursively, we obtain that $\Delta^{(n)}(\alpha_i) = [\pi_{11}(\alpha_i) - \pi_{01}(\alpha_i)]^n = [\Delta(\alpha_i)]^n$. Finally, as established in Lemma 1, $\Delta(\alpha_i) = \left[e^{\beta} - 1\right]\ \pi_{10}(\alpha_i)\ \pi_{01}(\alpha_i)$. Thus, we have that $\Delta^{(n)}(\alpha_i) = \left[e^{\beta} - 1\right]^n\ [\pi_{10}(\alpha_i)]^n\ [\pi_{01}(\alpha_i)]^n$. ■

A.7 Proof of Proposition 5

Similarly as in the proof of Lemma 2 above, we omit $\mathbf{x}$ as an argument, but one should understand that all the probabilities in this proof are conditioned on $\mathbf{x}_i^{\{1,T\}} = [\mathbf{x}, \mathbf{x}, ..., \mathbf{x}]$. W.l.o.g., we consider that $T = 2n + 1$. Given the definition of histories $(0, \widetilde{\mathbf{10}}^n)$ and $(\widetilde{\mathbf{10}}^n, 1)$, it is straightforward to see that:

$$\begin{cases} \mathbb{P}_{0,\widetilde{\mathbf{10}}^n} &= \displaystyle\int p^*(0|\alpha_i) \ \ [\pi_{10}(\alpha_i)]^n \ \ [\pi_{01}(\alpha_i)]^n \ \ f_\alpha(\alpha_i) \ d\alpha_i \\ \\ \mathbb{P}_{\widetilde{\mathbf{10}}^n,1} &= \displaystyle\int p^*(1|\alpha_i) \ \ [\pi_{10}(\alpha_i)]^n \ \ [\pi_{01}(\alpha_i)]^n \ \ f_\alpha(\alpha_i) \ d\alpha_i \end{cases} \tag{A.27}$$

Applying equation (29) from Lemma 2, we have that:

$$\begin{cases} \mathbb{P}_{0,\widetilde{\mathbf{10}}^n} &= \dfrac{1}{[e^\beta - 1]^n} \displaystyle\int p^*(0|\alpha_i) \ \Delta^{(n)}(\alpha_i) \ f_\alpha(\alpha_i) \ d\alpha_i \\ \\ \mathbb{P}_{\widetilde{\mathbf{10}}^n,1} &= \dfrac{1}{[e^\beta - 1]^n} \displaystyle\int p^*(1|\alpha_i) \ \Delta^{(n)}(\alpha_i) \ f_\alpha(\alpha_i) \ d\alpha_i \end{cases} \tag{A.28}$$

Adding up these two equations, multiplying the resulting equation times $\left[e^\beta - 1\right]^n$, and taking into account that $p^*(0|\alpha_i) + p^*(1|\alpha_i) = 1$, we have that $AME^{(n)} = [e^\beta - 1]^n \ [\mathbb{P}_{0,\widetilde{\mathbf{10}}^n} + \mathbb{P}_{\widetilde{\mathbf{10}}^n,1}]$ such that $AME^{(n)}$ is identified. ■

A.8 Proof of Lemma 3

Given the expression for the choice probabilities in the logit model, it is simple to verify that:

$$\frac{\pi_{k\ell}(\boldsymbol{\alpha}_i, \mathbf{x})}{\pi_{kj}(\boldsymbol{\alpha}_i, \mathbf{x})} = \exp\{\alpha_i(\ell) - \alpha_i(j) + \mathbf{x}'\left(\gamma_\ell - \gamma_j\right) - \beta_j \mathbb{1}\{k = j\}\} \tag{A.29}$$

and

$$\frac{\pi_{jj}(\boldsymbol{\alpha}_i, \mathbf{x})}{\pi_{j\ell}(\boldsymbol{\alpha}_i, \mathbf{x})} = \exp\{\alpha_i(j) - \alpha_i(\ell) + \mathbf{x}'\left(\gamma_j - \gamma_\ell\right) + \beta_j - \beta_j \mathbb{1}\{\ell = j\}\} \tag{A.30}$$

The product of these two expressions is equation (39). ■

A.9 Proof of Proposition 6: Model with duration

Here, we prove the identification of $AME_d(1)$. Using a similar argument, we can establish the identification of $AME_d(d)$ for any $d \geq 1$. First, we write the expression of the probabilities of choice histories conditional on α_i, $P_{y_1,y_2,y_3,y_4 \mid \alpha_i}$, implied by the model. Second, for each of these probabilities, we multiply the equation times the weights w_{y_1,y_2,y_3,y_4} in the enunciate of Proposition 6. For the probabilities with non-zero weights, we have:

$$\begin{aligned}
\frac{e^{\beta+\delta}-1}{2}\left[\mathbb{P}_{0,0,1,0 \mid \alpha_i}+\mathbb{P}_{0,1,0,0 \mid \alpha_i}\right] &= p^*(0|\alpha_i)\,\frac{\left(e^{\beta+\delta}-1\right)e^{\alpha_i}}{\left(1+e^{\alpha_i+\beta+\delta}\right)\left(1+e^{\alpha_i}\right)^2} \\
\frac{e^{\beta+\delta}-1}{e^{\beta+\delta}}\,\mathbb{P}_{0,0,1,1 \mid \alpha_i} &= p^*(0|\alpha_i)\,\frac{\left(e^{\beta+\delta}-1\right)e^{\alpha_i}e^{\alpha_i}}{\left(1+e^{\alpha_i+\beta+\delta}\right)\left(1+e^{\alpha_i}\right)^2} \qquad \text{(A.31)} \\
\left(e^{\beta+\delta}-1\right)\left[\mathbb{P}_{1,0,1,0 \mid \alpha_i}+\mathbb{P}_{1,0,1,1 \mid \alpha_i}\right] &= p^*(1|\alpha_i)\,\frac{\left(e^{\beta+\delta}-1\right)e^{\alpha_i}\left(e^{\alpha_i+\beta+\delta}+1\right)}{\left(1+e^{\alpha_i+\beta+\delta}\right)^2\left(1+e^{\alpha_i}\right)}
\end{aligned}$$

Third, we sum up these three equations. Simplifying factors and taking into account that $p^*(0|\alpha_i)+ p^*(1|\alpha_i)=1$, we get:

$$\begin{aligned}
&\frac{e^{\beta+\delta}-1}{2}\left[\mathbb{P}_{0,0,1,0 \mid \alpha_i}+\mathbb{P}_{0,1,0,0 \mid \alpha_i}\right]+\frac{e^{\beta+\delta}-1}{e^{\beta+\delta}}\mathbb{P}_{0,0,1,1 \mid \alpha_i}+\left(e^{\beta+\delta}-1\right)\left[\mathbb{P}_{1,0,1,0 \mid \alpha_i}+\mathbb{P}_{1,0,1,1 \mid \alpha_i}\right] \\
&= \frac{\left(e^{\beta+\delta}-1\right)e^{\alpha_i}}{\left(1+e^{\alpha_i+\beta+\delta}\right)\left(1+e^{\alpha_i}\right)} = \frac{e^{\alpha_i+\beta+\delta}}{\left(1+e^{\alpha_i+\beta+\delta}\right)} - \frac{e^{\alpha_i}}{\left(1+e^{\alpha_i}\right)} = \Delta_d(\alpha_i,1)
\end{aligned} \qquad \text{(A.32)}$$

Finally, we integrate the two sides of this equation over the distribution of α_i to obtain

$$\frac{e^{\beta+\delta}-1}{2}\left[\mathbb{P}_{0,0,1,0}+\mathbb{P}_{0,1,0,0}\right]+\frac{e^{\beta+\delta}-1}{e^{\beta+\delta}}\mathbb{P}_{0,0,1,1}+\left(e^{\beta+\delta}-1\right)\left[\mathbb{P}_{1,0,1,0}+\mathbb{P}_{1,0,1,1}\right]=AME_d(1) \qquad \text{(A.33)}$$

such that $AME_d(1)$ is identified.

We can proceed similarly to prove the identification of $AME_d(d)$ for any value $d \geq 1$.

For instance, we can prove that:

$$\begin{aligned} AME_d(2) &= \frac{e^{\beta+2\delta}-1}{2}\left[\mathbb{P}_{0,0,1,0}+\mathbb{P}_{0,1,0,0}\right]+\frac{e^{\beta+2\delta}-1}{e^{\beta+\delta}}\ \mathbb{P}_{0,0,1,1} \\ &+ \left(\frac{e^{\beta+2\delta}\left(1-e^{\beta+2\delta}\right)}{e^{\beta+\delta}}+e^{\beta+2\delta}-1\right)\mathbb{P}_{0,1,1,0} \\ &+ \left(e^{\beta+\delta}-\frac{e^{\beta+\delta}}{e^{\beta+2\delta}}\right)\left[\mathbb{P}_{1,0,1,0}+\mathbb{P}_{1,0,1,1}\right]+\left(\frac{e^{\beta+2\delta}-1}{e^{\beta+\delta}}-1+\frac{1}{e^{\beta+2\delta}}\right)\mathbb{P}_{1,1,0,0} \qquad \blacksquare \end{aligned} \tag{A.34}$$

A.10 Proof of Proposition 7

We omit $\mathbf{x}$ as an argument throughout this proof for notational simplicity. However, one should understand that the probability of the initial conditions p^*, the density function of $\boldsymbol{\alpha}_i$, the empirical probabilities of choice histories and the average transition probabilities are all conditional on $\mathbf{x}_i^{\{1,3\}}=[\mathbf{x},\mathbf{x},\mathbf{x}]$. We can write Π_{jj} as:

$$\Pi_{jj}=\int\left[p^*(0|\alpha_i)+p^*(1|\alpha_i)...+p^*(J|\alpha_i)\right]\ \pi_{jj}(\alpha_i)\ f_\alpha(\alpha_i)\ d\alpha_i \tag{A.35}$$

This expression includes the term $\int p^*(j|\alpha_i)\ \pi_{jj}(\alpha_i)\ f_\alpha(\alpha_i)\ d\alpha_i$ that is equal to the choice history probability $\mathbb{P}_{j,j}$. However, it also includes the "counterfactuals" $\delta_{k,j,j}\equiv\int p^*(k|\alpha_i)$ $\pi_{jj}(\alpha_i)\ f_\alpha(\alpha_i)\ d\alpha_i$ for $k\neq j$. We can represent each of these counterfactuals as:

$$\delta_{k,j,j}=\int p^*(k|\alpha_i)\ \left[\pi_{k0}(\alpha_i)+\pi_{k1}(\alpha_i)+...+\pi_{kJ}(\alpha_i)\right]\ \pi_{jj}(\alpha_i)\ f_\alpha(\alpha_i) \tag{A.36}$$

That is, we have that $\delta_{k,j,j}=\sum_{\ell=0}^{J}\delta^{(2)}_{k,\ell,j,j}$, with $\delta^{(2)}_{k,\ell,j,j}\equiv\int p^*(k|\alpha_i)\ \pi_{k\ell}(\alpha_i)\ \pi_{jj}(\alpha_i)\ f_\alpha(\alpha_i)$. For $\ell=j$, we have that $\delta^{(2)}_{k,j,j,j}$ corresponds to the choice history probability $\mathbb{P}_{k,j,j}$. For the rest of the terms $\delta^{(2)}_{k,\ell,j,j}$, we apply Lemma 3. According to Lemma 3, we have that $\pi_{k\ell}(\alpha_i)$ $\pi_{jj}(\alpha_i)=\exp\{\beta_{k\ell}-\beta_{kj}-\beta_{j\ell}\}\ \pi_{kj}(\alpha_i)\ \pi_{j\ell}(\alpha_i)$. Finally, note that $\int p^*(k|\alpha_i)\ \pi_{kj}(\alpha_i)\ \pi_{j\ell}(\alpha_i)$

$f_\alpha(\alpha_i)$ is the choice history probability $\mathbb{P}_{k,j,\ell}$. Putting all the pieces together, we have the expression in equation (40). $\blacksquare$

A.11 Proof of Proposition 8

We omit $\mathbf{x}$ as an argument throughout this proof for notational simplicity. Equation (17), from Proposition 2, provides the necessary and sufficient condition for identifying an AME. Applying this condition to the model defined by equation (35) and Assumption 1-MNL, with $J+1=3$, $T=3$, and $y_1=0$, we get:

$$\begin{aligned} & w_{0,0,0}\ \mathbb{P}_{0,0,0\ |\ (y_1=0,\alpha_i)} + w_{0,0,1}\ \mathbb{P}_{0,0,1\ |\ (y_1=0,\alpha_i)} + w_{0,0,2}\ \mathbb{P}_{0,0,2\ |\ (y_1=0,\alpha_i)} \\ + \quad & w_{0,1,0}\ \mathbb{P}_{0,1,0\ |\ (y_1=0,\alpha_i)} + w_{0,1,1}\ \mathbb{P}_{0,1,1\ |\ (y_1=0,\alpha_i)} + w_{0,1,2}\ \mathbb{P}_{0,1,2\ |\ (y_1=0,\alpha_i)} \\ + \quad & w_{0,2,0}\ \mathbb{P}_{0,2,0\ |\ (y_1=0,\alpha_i)} + w_{0,2,1}\ \mathbb{P}_{0,2,1\ |\ (y_1=0,\alpha_i)} + w_{0,2,2}\ \mathbb{P}_{0,2,2\ |\ (y_1=0,\alpha_i)} \quad = \pi_{10}(\alpha_i) \end{aligned} \tag{A.37}$$

Let's denote $d_j \equiv 1 + e^{\beta_j\ 1\{j=1\}+\alpha_i(1)} + e^{\beta_j\ 1\{j=2\}+\alpha_i(2)}$, for $j=0,1,2$. Replacing the probabilities by their expression based on the logistic CDF:

$$\begin{aligned} & w_{0,0,0}\frac{1}{d_0^2} + w_{0,0,1}\frac{e^{\alpha_i(1)}}{d_0^2} + w_{0,0,2}\frac{e^{\alpha_i(2)}}{d_0^2} \\ + \quad & w_{0,1,0}\frac{e^{\alpha_i(1)}}{d_0 d_1} + w_{0,1,1}\frac{e^{\alpha_i(1)}\ e^{\beta_1+\alpha_i(1)}}{d_0 d_1} + w_{0,1,2}\frac{e^{\alpha_i(1)}\ e^{\alpha_i(2)}}{d_0 d_1} \\ + \quad & w_{0,2,0}\frac{e^{\alpha_i(2)}}{d_0 d_2} + w_{0,2,1}\frac{e^{\alpha_i(1)}\ e^{\alpha_i(2)}}{d_0 d_2} + w_{0,2,2}\frac{e^{\alpha_i(2)}\ e^{\beta_2+\alpha_i(2)}}{d_0 d_2} \quad = \frac{1}{d_1} \end{aligned} \tag{A.38}$$

After some algebra to undo the fractions on both sides:

$$\begin{aligned}
& w_{0,0,0}\ d_1 d_2 + w_{0,0,1}\ e^{\alpha_i(1)} d_1 d_2 + w_{0,0,2}\ e^{\alpha_i(2)} d_1 d_2 \\
+ \quad & w_{0,1,0}\ e^{\alpha_i(1)} d_0 d_2 + w_{0,1,1}\ e^{\alpha_i(1)} e^{\beta_1+\alpha_i(1)} d_0 d_2 + w_{0,1,2}\ e^{\alpha_i(1)} e^{\alpha_i(2)} d_0 d_2 \\
+ \quad & w_{0,2,0}\ e^{\alpha_i(2)} d_0 d_1 + w_{0,2,1}\ e^{\alpha_i(1)} e^{\alpha_i(2)} d_0 d_1 + w_{0,2,2}\ e^{\beta_2+\alpha_i(2)} e^{\alpha_i(2)} d_0 d_1 \quad = d_0^2 d_2
\end{aligned} \tag{A.39}$$

Expanding this equation by doing the products of d_j and of the exponential, we obtain on both sides of the equality a polynomial in $(e^{\alpha_i(1)})^h\ (e^{\alpha_i(2)})^\ell$, where the minimum value of h and ℓ is 0, and the maximum value is 4. Following Lemma 3, equating the coefficient of each monomial in both sides of the equality, we get a system of linear equation whose unknowns are the weights $w_{0,0,0}, ..., w_{0,2,2}$. This condition on the monomials of $(e^{\alpha_i(1)})^2$ and $(e^{\alpha_i(1)})^3$ imply, respectively:

$$\begin{aligned}
w_{0,0,1} + w_{0,1,0} + w_{0,0,1}\ e^{\beta_1} + w_{0,1,0} &= 1 \\
w_{0,0,1}\ e^{\beta_1} + w_{0,1,0} &= 0
\end{aligned} \tag{A.40}$$

which leads to

$$w_{0,0,1} + w_{0,1,0} = 1. \tag{A.41}$$

At the same time, the condition on the monomial of $(e^{\alpha_i(1)})^1$ implies:

$$w_{0,0,1} + w_{0,1,0} = 2, \tag{A.42}$$

which is incompatible with the condition $w_{0,0,1} + w_{0,1,0} = 1$. Therefore, no weights can solve the system of equations in (A.37). By Proposition 2, this implies that Π_{10} is not point identified. ■

A.12 Proof of Proposition 9: Ordered Logit

The weights in the statement of Proposition 9 were obtained using the general procedure established in Proposition 2. Here, we present the proof for Π_{20}, but it proceeds the same for other Π_{kj}.

By definition, $\Pi_{20} = \displaystyle\int \pi_{20}(\alpha_i)\ f_{\alpha}(\alpha_i)\ d\alpha_i$, and according to equation (42), $\pi_{20}(\alpha_i) = \left[1 + e^{\beta_2 - \lambda_0 + \alpha_i}\right]^{-1}$. We start with the probabilities of choice histories conditional on α_i, that is, $P_{(y_1, y_2, y_3)\ |\ \alpha_i}$. We write the expression for these model probabilities as functions of parameters $\boldsymbol{\beta}$, $\boldsymbol{\lambda}$, and α_i. For each of these probabilities, we multiply the equation times the weights w_{y_1, y_2, y_3} that appear in the statement of Proposition 9. For the probabilities with non-zero weights for Π_{20}, we have:

$$
\begin{aligned}
\mathbb{P}_{0,0,0\ |\ \alpha_i} + \mathbb{P}_{0,0,1\ |\ \alpha_i} &= p^*(0|\alpha_i)\ \frac{1}{\left(1 + e^{\beta_0 - \lambda_1 + \alpha_i}\right)\left(1 + e^{\beta_0 - \lambda_0 + \alpha_i}\right)} \\
e^{\lambda_1 - \lambda_0}\ \mathbb{P}_{0,0,2\ |\ \alpha_i} &= p^*(0|\alpha_i)\ \frac{e^{\beta_0 - \lambda_0 + \alpha_i}}{\left(1 + e^{\beta_0 - \lambda_1 + \alpha_i}\right)\left(1 + e^{\beta_0 - \lambda_0 + \alpha_i}\right)} \\
\left(1 - \frac{e^{\beta_2 - \lambda_0}}{e^{\beta_0 - \lambda_1}}\right)\mathbb{P}_{0,2,0\ |\ \alpha_i} &= p^*(0|\alpha_i)\ \frac{e^{\beta_0 - \lambda_1 + \alpha_i} - e^{\beta_2 - \lambda_0 + \alpha_i}}{\left(1 + e^{\beta_0 - \lambda_1 + \alpha_i}\right)\left(1 + e^{\beta_2 - \lambda_0 + \alpha_i}\right)} \\
\sum_{k=0}^{l}\sum_{\ell=0}^{J-1}\mathbb{P}_{1,k,\ell\ |\ \alpha_i} &= p^*(1|\alpha_i)\left(\frac{1}{1 + e^{\beta_2 - \lambda_0 + \alpha_i}} - \frac{1 - e^{\beta_2 - \lambda_0 + \alpha_i}}{\left(1 + e^{\beta_1 - \lambda_1 + \alpha_i}\right)\left(1 + e^{\beta_2 - \lambda_0 + \alpha_i}\right)}\right) \\
\left(1 - \frac{e^{\beta_2 - \lambda_0}}{e^{\beta_1 - \lambda_1}}\right)\mathbb{P}_{1,2,0\ |\ \alpha_i} &= p^*(1|\alpha_i)\ \frac{1 - e^{\beta_2 - \lambda_0 + \alpha_i}}{\left(1 + e^{\beta_1 - \lambda_1 + \alpha_i}\right)\left(1 + e^{\beta_2 - \lambda_0 + \alpha_i}\right)} \\
\sum_{k=0}^{J-1}\mathbb{P}_{2,0,k\ |\ \alpha_i} &= p^*(2|\alpha_i)\ \frac{1}{1 + e^{\beta_2 - \lambda_0 + \alpha_i}}
\end{aligned}
\tag{A.43}
$$

Summing up these equations, simplifying factors, and taking into account that $p^*(0|\alpha_i) +$

$p^*(1|\alpha_i) + p^*(2|\alpha_i) = 1$, we get:

$$\begin{aligned}
&\mathbb{P}_{0,0,0 \mid \alpha_i} + \mathbb{P}_{0,0,1 \mid \alpha_i} + e^{\lambda_1-\lambda_0}\mathbb{P}_{0,0,2 \mid \alpha_i} + \left(1 - \frac{e^{\beta_2-\lambda_0}}{e^{\beta_0-\lambda_1}}\right)\mathbb{P}_{0,2,0 \mid \alpha_i} \\
&\quad + \sum_{k=0}^{l}\sum_{\ell=0}^{J}\mathbb{P}_{1,k,\ell \mid \alpha_i} + \left(1 - \frac{e^{\beta_2-\lambda_0}}{e^{\beta_1-\lambda_1}}\right)\mathbb{P}_{1,2,0 \mid \alpha_i} + \sum_{k=0}^{J}\mathbb{P}_{2,0,k \mid \alpha_i} \quad = \quad \pi_{20}(\alpha_i)
\end{aligned} \tag{A.44}$$

Finally, we integrate the two sides of this equation over the distribution of α_i to obtain the expression for Π_{20} in equation (43). ■

References

Aguirregabiria, V., J. Gu, and Y. Luo (2021): "Sufficient statistics for unobserved heterogeneity in dynamic structural logit models," *Journal of Econometrics*, 223(2), 280–311.