Identification of Average Marginal Effects in Fixed Effects Dynamic Discrete Choice Models

Victor Aguirregabiria (University of Toronto)
Jesús Carro (Carlos III, Madrid)

PANEL DATA WORKSHOP

NUFFIELD COLLEGE, OXFORD UNIVERSITY

June 6 & 7, 2022

INTRODUCTION (1/2)

- Consider the **Fixed Effects (FE) Dynamic Binary Logit model** as described by the transition probability:

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

where $p_1(y_{i1}|\alpha_i)$ and $f_\alpha(\alpha_i)$ are unrestricted, i.e., FE model.

- Given panel data $(y_{i1}, y_{i2}, ..., y_{iT})$ with $T \geq 4$, **parameter β is identified** (Chamberlain (1985), Honoré & Kyriazidou (2000)):

$$\beta = \log \mathbb{P}(0,0,1,1) - \log \mathbb{P}(0,1,0,1)$$

where $\mathbb{P}(y_1, y_2, y_3, y_4)$ is the probability of history (y_1, y_2, y_3, y_4).

INTRODUCTION (2/2)

- In this paper, we are interested in the **identification & estimation of Average Marginal Effects (AMEs)**.

- For instance, for the Binary Choice AR(1) model:

$$\begin{aligned} AME &= \mathbb{E}_\alpha \left(\mathbb{E}\left[y_{it} | \alpha_i, y_{i,t-1} = 1\right] - \mathbb{E}\left[y_{it} | \alpha_i, y_{i,t-1} = 0\right] \right) \\ &= \int \left(\frac{\exp\{\alpha_i + \beta\}}{1 + \exp\{\alpha_i + \beta\}} - \frac{\exp\{\alpha_i\}}{1 + \exp\{\alpha_i\}} \right) f_\alpha(\alpha_i)\ d\alpha_i \end{aligned}$$

- **Common wisdom:** these AMEs are not identified in FE models.

  - They depend on the whole distribution $f_\alpha(\alpha_i)$, and this distribution is not identified in FE models.

CONTRIBUTIONS OF THIS PAPER

1. Everything started with the derivation (almost by chance) of a **simple closed-form expression for the AME in BC-AR(1)** in terms of β and probabilities of choice histories.

2. Inspired by this result, we develop a **general method to obtain AMEs** in a broad class of dynamic logit models: binary, multinomial, ordered, with exogenous regressors, with duration dependence.

   - This method is based on the solution of a finite (low-dimension) system of linear equations.

3. Empirical application: dynamic demand of differentiated products with brand switching costs.

RELATED LITERATURE

- Related papers studying the **identification of AMEs in FE discrete choice models** are:

  - Chernozhukov, Fernandez-Val, Hahn, & Newey (ECMA, 2013)
  - Davezies, D'Haultfoeuille, & Laage (WP, 2021)
  - Pakel & Weidner (WP, 2021)

- Some differences between our paper and these papers:

  (1) Point identification (us) versus set identification (them).

  (2) Logit (us) versus more general discrete choice (them).

  (3) Simple closed-form expressions (us) more computationally intensive methods (them).

OUTLINE

1. Identification result for AME in BC-AR(1)

2. General identification method

    - Application of the general identification method

    - Multinomial, Exogenous X, Duration, Ordered Logit.

3. Empirical application

1. IDENTIFICATION OF AME IN BC-AR(1) MODEL

Identification of AME in BC-AR(1) Model (1/3)

- Define the **individual-level transition probabilities**:

$$\begin{aligned}\pi_{01}(\alpha_i) &\equiv P\left(y_{it}=1|\alpha_i, y_{i,t-1}=0\right) = \Lambda\left(\alpha_i\right) \\ \pi_{11}(\alpha_i) &\equiv P\left(y_{it}=1|\alpha_i, y_{i,t-1}=1\right) = \Lambda\left(\alpha_i+\beta\right)\end{aligned}$$

- And the corresponding **average transition probabilities**:

$$\begin{aligned}\Pi_{01} &\equiv \int \pi_{01}(\alpha_i)\ f_\alpha(\alpha_i)\ d\alpha_i \\ \Pi_{11} &\equiv \int \pi_{11}(\alpha_i)\ f_\alpha(\alpha_i)\ d\alpha_i\end{aligned}$$

- Define the **individual-level marginal effect**:

$$\Delta(\alpha_i) \equiv \pi_{11}(\alpha_i) - \pi_{01}(\alpha_i)$$

- And the corresponding **Average Marginal Effect (AME)**:

$$AME \equiv \int \Delta(\alpha_i)\ f_\alpha(\alpha_i)\ d\alpha_i = \Pi_{11} - \Pi_{01}$$

Identification of AME in BC-AR(1) Model (2/3)

- We show the following: **identification results**:

$$\begin{cases} \Pi_{01} = [1 - \exp\{\beta\}] \; \mathbb{P}_{1,0,1} + \mathbb{P}_{1,1} + \mathbb{P}_{0,1} \\ \Pi_{11} = \exp\{\beta\} \; \mathbb{P}_{0,1,0} + \mathbb{P}_{0,1,1} + \mathbb{P}_{1,1} \\ AME = [\exp\{\beta\} - 1] \; [\mathbb{P}_{0,1,0} + \mathbb{P}_{1,0,1}] \end{cases}$$

where:

$\mathbb{P}_{y_1,y_2,y_3}$ = empirical probability of $(y_{i1}, y_{i2}, y_{i3}) = (y_1, y_2, y_3)$

$\mathbb{P}_{y_1,y_2}$ = empirical probability of $(y_{i1}, y_{i2}) = (y_1, y_2)$

Proof of Identification of AME (3/3)

- Key in this proof: following **property** of Logit model. For any α_i:

$$\Delta(\alpha_i) = [\exp\{\beta\} - 1] \;\; \pi_{01}(\alpha_i) \;\; \pi_{10}(\alpha_i) \tag{1}$$

- For any sequence (y_1, y_2, y_3):

$$\mathbb{P}_{y_1,y_2,y_3} = \int p^*(y_1|\alpha_i) \;\; \pi_{y_1,y_2}(\alpha_i) \;\; \pi_{y_2,y_3}(\alpha_i) \;\; f_\alpha(\alpha_i) \;\; d\alpha_i$$

- Applying equation (1) to $\mathbb{P}_{0,1,0}$ and $\mathbb{P}_{1,0,1}$, we have that:

$$\begin{cases} [\exp\{\beta\} - 1] \;\; \mathbb{P}_{0,1,0} & = \displaystyle\int p^*(0|\alpha_i) \;\; \Delta(\alpha_i) \;\; f_\alpha(\alpha_i) \;\; d\alpha_i \\ [\exp\{\beta\} - 1] \;\; \mathbb{P}_{1,0,1} & = \displaystyle\int p^*(1|\alpha_i) \;\; \Delta(\alpha_i) \;\; f_\alpha(\alpha_i) \;\; d\alpha_i \end{cases}$$

- Adding up these two equations:

$$[\exp\{\beta\} - 1] \;\; [\mathbb{P}_{0,1,0} + \mathbb{P}_{1,0,1}] \;\; = \;\; AME$$

Identification of n-periods forward AME

- Using a similar approach, we show the identification of the n-periods forward AME, for any $n \geq 1$:

$$AME^{(n)} \equiv \mathbb{E}_{\alpha}\left(\mathbb{E}\left[y_{i,t+n} | \alpha_i, y_{it} = 1\right] - \mathbb{E}\left[y_{i,t+n} | \alpha_i, y_{it} = 0\right]\right)$$

- We show that, for $T \geq 2n + 1$:

$$AME^{(n)} = \left[\exp\{\beta\} - 1\right]^n \left[\mathbb{P}_{0,\widetilde{\mathbf{10}}^n} + \mathbb{P}_{\widetilde{\mathbf{10}}^n,1}\right]$$

where $\widetilde{\mathbf{10}}^n$ represents the repetition n times of of sequence $1, 0$.

Identification of Average Transition Probability in Multinomial Logit

- A similar procedure shows identification of average transition probability Π_{jj} in a dynamic multinomial logit, for $j = 1, 2, ..., J$:

$$\Pi_{jj} \equiv \int \pi_{jj}(\alpha_i) \; f_\alpha(\boldsymbol{\alpha}_i) \; d\boldsymbol{\alpha}_i$$

  with

$$\pi_{jj}(\alpha_i) \equiv P\left(y_{it} = j | \boldsymbol{\alpha}_i, y_{i,t-1} = j\right)$$

- Logit model implies that for any triple of choice alternatives j, k, ℓ:

$$\exp\left\{\beta_{k\ell} - \beta_{kj} + \beta_{jj} - \beta_{j\ell}\right\} = \frac{\pi_{k\ell}(\boldsymbol{\alpha}_i) \; \pi_{jj}(\boldsymbol{\alpha}_i)}{\pi_{kj}(\boldsymbol{\alpha}_i) \; \pi_{j\ell}(\boldsymbol{\alpha}_i)}$$

- And using this property, we can show that:

$$\Pi_{jj} = \mathbb{P}_{j,j} + \sum_{k \neq j} \left[\mathbb{P}_{k,j,j} + \sum_{\ell \neq j} \exp\left\{\beta_{k\ell} - \beta_{kj} + \beta_{jj} - \beta_{j\ell}\right\} \; \mathbb{P}_{k,j,\ell}\right]$$

2. GENERAL METHOD TO SHOW IDENTIFICATION OF AMEs

General Dynamic Logit Model

- Consider a dynamic logit model that allows for multinomial y, exogenous regressors ($\mathbf{x}$), and duration (d) dependence.

- Let $\mathbf{y}_i \equiv (d_{i1}, y_{i1}, y_{i2}, ..., y_{iT}) \in \mathcal{D} \times \mathcal{Y}^T$ be individual i's choice, and let $\mathbf{x}_i \equiv (\mathbf{x}_{i1}, \mathbf{x}_{i2}, ..., \mathbf{x}_{iT}) \in \mathcal{X}^T$

- Let $\mathbb{P}_{\mathbf{y}|\mathbf{x}}$ represent the probability $P(\mathbf{y}_i = \mathbf{y}|\mathbf{x}_i = \mathbf{x})$.

- According to the model, probability $\mathbb{P}_{\mathbf{y}|\mathbf{x}}$ has the following structure:

$$\mathbb{P}_{\mathbf{y}|\mathbf{x}} = \int G\left(\mathbf{y}^{\{2,T\}}|d_1, y_1, \mathbf{x}, \boldsymbol{\alpha}; \boldsymbol{\theta}\right) \; p^*(d_1, y_1|\boldsymbol{\alpha}, \mathbf{x}) \; f_\alpha(\boldsymbol{\alpha}|\mathbf{x}) \; d\alpha,$$

  where

$$G\left(\mathbf{y}^{\{2,T\}}|y_1, d_1, \mathbf{x}, \boldsymbol{\alpha}; \boldsymbol{\theta}\right) \equiv \prod_{t=2}^{T} \Lambda\left(y_t|y_{t-1}, d_t, \mathbf{x}_t, \boldsymbol{\alpha}; \boldsymbol{\theta}\right)$$

LEMMA 1

- Consider a FE dynamic discrete choice model characterized by the probability function $G(\mathbf{y}^{\{2,T\}}|y_1, d_1, \mathbf{x}, \boldsymbol{\alpha}; \boldsymbol{\theta})$.

- Let $AME(\mathbf{x}) \equiv \int \Delta(\boldsymbol{\alpha}_i, \mathbf{x}, \boldsymbol{\theta}) f_\alpha(\boldsymbol{\alpha}_i|\mathbf{x}) \; d\boldsymbol{\alpha}_i$ be an average marginal effect of interest.

- This *AME* **is point identified if and only if** there is a weighting function $w(\mathbf{y}, \mathbf{x}, \boldsymbol{\theta})$ that satisfies the following equation:

$$\sum_{\mathbf{y}^{\{2,T\}}} w(d_1, y_1, \mathbf{y}^{\{2,T\}}, \mathbf{x}, \boldsymbol{\theta}) \; G\left(\mathbf{y}^{\{2,T\}}|y_1, d_1, \mathbf{x}, \boldsymbol{\alpha}; \boldsymbol{\theta}\right) \;=\; \Delta(\boldsymbol{\alpha}, \mathbf{x}, \boldsymbol{\theta}),$$

  for every value $(d_1, y_1) \in \mathcal{D} \times \mathcal{Y}$ and every $\boldsymbol{\alpha} \in \mathbb{R}^J$.

- Furthermore, this condition implies that:

$$AME(\mathbf{x}) \;=\; \sum_{\mathbf{y}} w(\mathbf{y}, \mathbf{x}, \boldsymbol{\theta}) \; \mathbb{P}_{\mathbf{y}|\mathbf{x}}$$

Particular Structure of FE Dynamic Logit

- Lemma 1 does not impose any restriction on the form of function G.

- In **FE Dynamic Logit model** the probability of a choice history:

$$\log \mathbb{P}\left(\mathbf{y}_i | \mathbf{x}_i, \boldsymbol{\alpha}_i, \boldsymbol{\theta}\right) = \mathbf{s}(\mathbf{y}_i, \mathbf{x}_i)' \, \mathbf{g}(\boldsymbol{\alpha}_i, \mathbf{x}_i, \boldsymbol{\theta}) + \mathbf{c}(\mathbf{y}_i, \mathbf{x}_i)' \, \boldsymbol{\theta}$$

where $\mathbf{s}_i \equiv \mathbf{s}(\mathbf{y}_i, \mathbf{x}_i)$ and $\mathbf{c}_i \equiv \mathbf{c}(\mathbf{y}_i, \mathbf{x}_i)$ are vectors of statistics.

- This equation implies that:

(1) $\mathbf{s}_i$ is a sufficient statistic for $\boldsymbol{\alpha}_i$.

(2) Given $\boldsymbol{\theta}$, the distribution of $\mathbf{s}_i$ contains all the information in the data about the distribution of $\boldsymbol{\alpha}_i$, and therefore, about AMEs.

(3) The form of $\mathbb{P}_{\mathbf{s}|\mathbf{x}}$ is:

$$\mathbb{P}_{\mathbf{s}|\mathbf{x}} = \sum_{\mathbf{y}:\ \mathbf{s}(\mathbf{y},\mathbf{x}) = \mathbf{s}} \left[\int \exp\{ \mathbf{s}(\mathbf{y}, \mathbf{x})' \, \mathbf{g}(\boldsymbol{\alpha}, \mathbf{x}, \boldsymbol{\theta}) + \mathbf{c}(\mathbf{y}, \mathbf{x})' \, \boldsymbol{\theta} \} \, f_\alpha(\boldsymbol{\alpha} | \mathbf{x}) \, d\boldsymbol{\alpha} \right]$$

LEMMA 2

- Consider a FE Dynamic Logit model.
- Let $AME(\mathbf{x}) \equiv \int \Delta(\boldsymbol{\alpha}_i, \mathbf{x}, \boldsymbol{\theta}) f_\alpha(\boldsymbol{\alpha}_i|\mathbf{x})\ d\boldsymbol{\alpha}_i$ be an AME of interest.
- This *AME* **is point identified if and only if** there is a weighting function $m(\mathbf{s}, \mathbf{x}, \boldsymbol{\theta})$ that satisfies the following equation:

$$\sum_{\widetilde{\mathbf{s}} \in \widetilde{\mathbb{S}}} m(d_1, y_1, \widetilde{\mathbf{s}}, \mathbf{x}, \boldsymbol{\theta})\ \exp\{(d_1, y_1, \widetilde{\mathbf{s}})'\ \mathbf{g}(\boldsymbol{\alpha}, \mathbf{x}, \boldsymbol{\theta})\} \ = \ \Delta(\boldsymbol{\alpha}, \mathbf{x}, \boldsymbol{\theta}),$$

  for every value (d_1, y_1) and every $\boldsymbol{\alpha} \in \mathbb{R}^J$.
- Furthermore, this condition implies that:

$$AME(\mathbf{x}) \ = \ \sum_{\mathbf{s} \in \mathcal{S}} \frac{m(\mathbf{s}, \mathbf{x}, \boldsymbol{\theta})}{\sum_{\mathbf{y}:\ \mathbf{s}(\mathbf{y},\mathbf{x})=\mathbf{s}} \exp\{\mathbf{c}(\mathbf{y}, \mathbf{x})'\boldsymbol{\theta}\}} \mathbb{P}_{\mathbf{s}|\mathbf{x}}$$

System with Infinite Restrictions and Finite Unknowns (1/2)

- The identification condition in Lemma 2 defines an infinite system of equations – as many as values of $\boldsymbol{\alpha}_i$.
- The researcher knows functions $\mathbf{g}(\boldsymbol{\alpha}, \mathbf{x}, \boldsymbol{\theta})$ and $\Delta(\boldsymbol{\alpha}, \mathbf{x}, \boldsymbol{\theta})$.
- The unknowns are the weights $m(\mathbf{s}, \mathbf{x}, \boldsymbol{\theta})$.
- Without some structure, this system with infinite restrictions and finite unknowns would not have a solution.

System with Infinite Restrictions and Finite Unknowns (2/2)

- Lemma 3 shows that, in the FE dynamic logit model, the structure of functions $\mathbf{g}(\boldsymbol{\alpha}, \mathbf{x}, \boldsymbol{\theta})$ and $\Delta(\boldsymbol{\alpha}, \mathbf{x}, \boldsymbol{\theta})$ is such that the **identification condition can be represented as a finite order polynomial** in the variables $\exp\{\alpha_i(j)\}$ for $j = 1, 2, ..., J$.

- Since these variables are always strictly positive, there is a solution to the system **if and only if the coefficients multiplying every monomial term in this polynomial are all equal to zero**.

- This property transforms the infinite system of equations into a **finite system with finite unknowns**.

- Furthermore, if a solution exists, this solution implies a closed-form expression for the weights $m(\mathbf{s}, \mathbf{x}, \boldsymbol{\theta})$, and therefore, for AME.

LEMMA 3

- Consider the FE dynamic logit model.
- The identification condition in Lemma 2 can be represented as a finite order polynomial in the variables $\exp\{\alpha_i(j)\}$ for $j = 1, 2, ..., J$.
- This implies a **finite system of linear equations** with unknowns the finite number of weights $m(\mathbf{s}, \mathbf{x}, \boldsymbol{\theta})$ for every $\mathbf{s} \in \mathcal{S}$.

EXAMPLE: AME in BC-AR(1) (1/2)

- $\mathbf{s} = (y_1, y_T, n_1)$ with $n_1 = \sum_{t=2}^{T} y_t$; $\mathbf{c} = \sum_{t=2}^{T} y_{t-1} y_t$, and:

$$\begin{cases} \Delta(\alpha_i) & = \dfrac{e^{\alpha_i}(e^{\beta}-1)}{(1+e^{\alpha_i+\beta})(1+e^{\alpha_i})} \\ e^{\mathbf{s}'\mathbf{g}(\alpha)} & = \left(\dfrac{1}{1+e^{\alpha}}\right)^{T-1} \left(\dfrac{1+e^{\alpha+\beta}}{1+e^{\alpha}}\right)^{y_T-y_1} \left(\dfrac{e^{\alpha}(1+e^{\alpha})}{1+e^{\alpha+\beta}}\right)^{n_1} \end{cases}$$

- Therefore, the identification condition is:

$$\sum_{y_T, n_1} m(y_1, y_T, n_1) \left(\frac{1}{1+e^{\alpha}}\right)^{T-1} \left(\frac{1+e^{\alpha+\beta}}{1+e^{\alpha}}\right)^{y_T-y_1} \left(\frac{e^{\alpha}\,(1+e^{\alpha})}{1+e^{\alpha+\beta}}\right)^{n_1}$$
$$= \frac{e^{\alpha}(e^{\beta}-1)}{(1+e^{\alpha+\beta})(1+e^{\alpha})}$$

EXAMPLE: AME in BC-AR(1) (2/2)

$$\sum_{y_T, n_1} m(y_1, y_T, n_1) \left(\frac{1}{1+e^{\alpha}} \right)^{T-1} \left(\frac{1+e^{\alpha+\beta}}{1+e^{\alpha}} \right)^{y_T - y_1} \left(\frac{e^{\alpha}\left(1+e^{\alpha}\right)}{1+e^{\alpha+\beta}} \right)^{n_1}$$
$$- \frac{e^{\alpha}\left(e^{\beta}-1\right)}{\left(1+e^{\alpha+\beta}\right)\left(1+e^{\alpha}\right)} = 0$$

- Multiplying this equation times $\left(1+e^{\alpha+\beta}\right)\left(1+e^{\alpha}\right)$ to eliminate denominators, we obtain a **polynomial of order $2T-2$ in e^{α}**.
- Since $e^{\alpha} > 0$, this equation holds for every value of α iff the coefficients multiplying each of the $2T-2$ monomials are zero.
- These coefficients are linear in the weights m_{y_1, y_T, n_1}, and this defines a system of $2T-2$ linear equations with $2T-2$ unknowns.

Application of the general identification method

- We apply this general approach to show identifcation of diffetent AMEs in different versions of the FE dynamic logit model.

1. Π_{11}, Π_{00}, and $AME^{(n)}$ in BC-AR(1).
2. Average transition probability Π_{jj} in multinomial and ordered logit.
3. AME of change in duration.
4. All these AMEs in model with exogenous $\mathbf{x}$.

3. EMPIRICAL APPLICATION

PRELIMINARIES

- Demand of differentiated product / state dependence in consumer brand choice.
- Data (Nielsen scanner panel data) from **Erdem, Imai, and Keane, 2003 (EIK)**, and similar model.
- The main goal is to determine the relative contribution of unobserved heterogeneity and state dependence to explain the observed time persistence of consumer brand choices.
- All previous studies estimate Random Effects models.

DATA

- Product: Ketchup.
- Same working sample as EIK. 996 households; 123 weeks.
- For our analysis, a time period is a *household purchase occasion*.
- Number observation purchase occasions $= \sum_{i=1}^{N} T_i = 9,562$

Table 3
Distribution of number of purchase occasions (T_i)

Minimum	*5%*	*25%*	*Median*	*75%*	*95%*	*Maximum*
3	4	5	8	12	21	52

BRAND CHOICE PERSISTENCE

Table 4
Matrix of Transition Probabilities of Brand Choices
(percentage points)

	Brand choice at $t+1$				*Total*
Brand choice at t	*Heinz* ($j=0$)	*Hunts* ($j=1$)	*Del Monte* ($j=2$)	*Store* ($j=3$)	
Heinz ($j=0$)	**78.95**	10.67	6.98	3.40	100.00
Hunts ($j=1$)	45.16	**32.30**	15.76	6.78	100.00
Del Monte ($j=2$)	41.11	18.98	**34.07**	5.83	100.00
Store ($j=3$)	42.32	17.11	13.38	**27.19**	100.00
Market share ($\mathbb{P}_j$)	66.65	15.63	12.19	5.53	100.00
Persistence ($\mathbb{P}_{j\|j} - \mathbb{P}_j$)	12.30	16.67	21.88	21.66	

MODEL

- We consider the following dynamic structural model:

$$y_{it} = \arg \max_{j \in \{0,1,2,3\}} \{ \alpha_i(j) + \varepsilon_{it}(j) + \beta_{jj} 1\{y_{i,t-1} = j\} + v_i(j, y_{i,t-1}) \}$$

- β_{jj} represents habits in the purchase/consumption of brand j: additional utility from buying the same brand as in last purchase.
- To illustrate our method using a short panel, we split the purchasing histories in the original sample into subs-histories of length T, where T is small.
- We present results for $T = 6$ and $T = 8$.

ESTIMATION OF BETA PARAMETERS

Table 5
Conditional Maximum Likelihood Estimates
of Brand Habit (β_{jj}) Parameters

Parameter	$T = 6$ **sub-histories**		$T = 8$ **sub-histories**	
β_{jj}	*Estimate*	*(s.e.)*$^{(1)}$	*Estimate*	*(s.e.)*$^{(1)}$
Heinz	0.00	(.)	0.00	(.)
Hunts	0.2312	(0.0590)	0.2566	(0.0570)
Del Monte	0.1155	(0.0718)	0.1191	(0.0722)
Store	0.3245	(0.1166)	0.4675	(0.1106)
# *histories of length* T	4,764		3,396	

(1) Standard errors ($s.e$) are obtained using a boostrap method. We generate 1,000 resamples (independent, with replacement, and with $N = 996$) from the 996 purchasing histories in the original dataset. Then, we split each history of the bootstrap sample into all the possible sub-histories of length T.

STATE DEPENDENCE vs UH, USING AME

Table 6
FE Estimates of ATPs and AMEs

	$T = 8$ **sub-histories**			
	Pers (s.e.)	*ATP* (*s.e.*)	*ATE* (s.e.)	*UHet* (*s.e.*)
Heinz	0.1230 (0.0033)	0.6708 (0.0062)	0.0043 (0.0067)	0.1187 (0.0069)
Hunts	0.1667 (0.0077)	0.1788 (0.0072)	0.0225 (0.0106)	0.1442 (0.0109)
Del Monte	0.2188 (0.0090)	0.1345 (0.0062)	0.0126 (0.0110)	0.2062 (0.0113)
Store	0.2166 (0.0062)	0.0805 (0.0072)	0.0252 (0.0094)	0.1914 (0.0099)

CONCLUSIONS

- AME are useful parameters to represent causal effects.
- In FE nonlinear panel data models with short panels, the distribution of the UH, and this problem has been associated with the common belief that AMEs are not identified.
- In the context of dynamic logit models, we prove the identification of AMEs associated with changes in lagged dependent variables and in duration variables.
- Our proofs provide simple closed-form expressions for the AMEs in terms of frequencies of choice histories.
- We illustrate our identification results using consumer scanner data in dynamic demand model with state dependence.