STRUCTURAL DYNAMIC DISCRETE CHOICE MODELS WITH FIXED EFFECTS

Econometric Society Summer School
in Dynamic Structural Econometrics

Victor Aguirregabiria (University of Toronto)

August 5, 2024

INTRODUCTION

- Disentangling **true dynamics** (causal effect of past decisions) versus **spurious dynamics** (persistent unobserved heterogeneity [UH] is a fundamental problem in the econometrics of dynamic models.

- Challenges with short panels:

  - **Incidental Parameters Problem (IPP):** Treating UH as fixed parameters implies inconsistent estimation of parameters of interest.

  - **Initial Conditions Problem (ICP):** There is no Nonparametric Identification of the distribution of UH and initial conditions.

- Two alternative approaches to deal with the Nonparametric No-Identification from the ICP:
  - **Random Effects (RE)**.
  - **Fixed Effects (FE)**.

RANDOM EFFECTS (RE) vs. FIXED EFFECTS

- **Random Effects (RE):**
  - We deal with the ICP by imposing parametric & finite support restrictions on the joint distribution of UH and initial conditions.
  - **Pros:** Full identification of structural parameters & distribution of UH.
  - **Cons:** Misspecification of parametric restrictions on UH can introduce substantial biases in the estimates of "true dynamics".
- **Fixed Effects (FE)**:
  - Focus on identification of structural parameters capturing "true dynamics" and not on the identification of the distrribution of UH.
  - **Pros:** NP specification of UH. Robust identification of true dynamics.
  - **Cons:** Distribution of UH is not fully identified. It limits the counterfactuals we can identify.
  - **Cons:** Not all dynamic models have consistent FE estimators.

FIXED EFFECTS IN STRUCTURAL DDC MODELS

- Until recently, all applications of Structural DDC models use RE models to deal with UH.
- The absence of applications using a FE approach was partly because of two common beliefs.

1. Belief that there are not consistent FE estimators in structural models where agents are forward-looking: **problem with continuation values**.
2. Belief that, even if structural parameters are identified, we cannot identify **Average Marginal Effects (AME)** and other Counterfactuals as these depend on the distribution of the UH.

- Recent developments have challenged these beliefs.

THIS LECTURE

- This lecture presents recent identification results and estimation methods of Structural DDC - FE Models.
- It focuses on the following papers.
- **Aguirregabiria, Gu, & Luo (Journal of Econometrics, 2021)**
  - Identification of structural parameters in Structural DDC-FE with lagged decision and duration as state variables.
  - Conditional Max. Like. estimation based on sufficient statistics for UH.
- **Aguirregabiria (Econometrics Journal, 2023)**
  - Application to dynamic demand for differentiated products
- **Aguirregabiria & Carro (Working Paper, 2024)**
  - Point identification of different types of Average Marginal Effects (AMEs) despite the distribution of the UH is not fully identified.

OUTLINE

1. Model – General version

2. Model – Example – Dynamic Demand for Differentiated Product

3. Identification of Structural Parameters.

4. Conditional MLE of Structural Parameters.

5. Identification of Average Marginal Effects.

6. Empirical application – Dynamic Demand for Differentiated Product.

1. MODEL

Model: Decision & State Variables

- **Decision variable**: $y_{it} \in \mathcal{Y} = \{0, 1, ..., J\}$.
- Agent maximizes $\mathbb{E}_t \left[\sum_{s=0}^{\infty} \delta_i^s \; U_{i,t+s} \right]$. U_{it} is the utility function.
- U_{it} depends on current choice, y_{it}, and on:
- **Two types of unobservables** for the researcher, $(\boldsymbol{\alpha}_i, \varepsilon_{it})$;
- **Three types of observable state variables**:

$$\left(\boldsymbol{x}_{it}, \; y_{it-1}, \; d_{it} \right)$$

$\boldsymbol{x}_{it}$ = strictly exogenous with respect to ε_{it};
$y_{i,t-1}$ is the **lagged decision**;
d_{it} is the **duration/experience** in last choice:

Model: Utility/Payoff Function

- The current payoff of choosing alternative j:

$$
\begin{aligned}
U_{it}(j) &= \alpha_i(j) + \varepsilon_{it}(j) + \boldsymbol{x}_{it}\ \boldsymbol{\beta}_x \\
&+ 1\{j \neq y_{it-1}\}\ \beta_y\left(j, y_{it-1}\right) + 1\{j = y_{it-1}\}\ \beta_d\left(j, d_{it}\right)
\end{aligned}
$$

- $\beta_y(j,k)$ **switching cost** from k to j.
- $\beta_d(j,d)$ **return (cost)** of d periods **of experience** in j.
- **Unobservables**:

  - $\varepsilon_{it}(j)$ i.i.d. type I extreme value distributed;

  - **FE model**: $p(\boldsymbol{\alpha}_i, \delta_i \mid y_{i1},\ d_{i1},\ \boldsymbol{x}_i\}$ is unrestricted.

Optimal decision & Conditional Choice Probabilities

- Let $\boldsymbol{s}_{it} \equiv (\boldsymbol{x}_{it}, y_{i,t-1}, d_{it})$. The optimal decision is:

$$y_{it} = \arg\max_{j \in \mathcal{Y}} \left\{ \alpha_i\left(j\right) + \varepsilon_{it}(j) + \beta\left(j, \boldsymbol{s}_{it}\right) + v_i(j, \boldsymbol{s}_{it}) \right\}$$

- $\beta\left(j, \boldsymbol{s}_{it}\right) \equiv \boldsymbol{x}_{it}\boldsymbol{\beta}_x + \beta_y\left(j, y_{it-1}\right) + \beta_d\left(j, d_{it}\right)$
- $v_i(j, \boldsymbol{s}_{it})$ is the **continuation value function**.

- The extreme value type 1 distribution of the unobservables ε, implies the **conditional choice probability (CCP)** function:

$$P_i\left(j | \boldsymbol{s}_{it}\right) = \frac{\exp\left\{\alpha_i(j) + \beta(j, \boldsymbol{s}_{it}) + v_i(j, \boldsymbol{s}_{it})\right\}}{\sum\limits_{k \in \mathcal{Y}} \exp\left\{\alpha_i(k) + \beta(k, \boldsymbol{s}_{it}) + v_i(k, \boldsymbol{s}_{it})\right\}}$$

Examples of dynamic structural models within this class

- *Market entry-exit models.*
- *Machine replacement models.*
- *Occupational choice models.*
- *Dynamic demand of differentiated storable / durable products.*
- *Pricing with menu costs.*
- *Inventory management.*
- ...

2. MODEL: DYNAMIC DEMAND

MODEL: DECISION & STATE VARIABLES

- J **products** indexed by j; consumers by i, calendar time by t.

- **Consumer Decision variable**:
  $y_{it} = 0$ means "no purchase"; $y_{it} = j > 0$ means "purchase j"

- State variables that **depend the consumer's choices**:

  ℓ_{it} = brand choice in last purchase

  $$\ell_{i,t+1} = 1\{y_{it} = 0\}\ell_{it} + 1\{y_{it} > 0\}y_{it}$$

  d_{it} = time duration since last purchase.

  $$d_{i,t+1} = 1 + 1\{y_{it} = 0\}d_{it}$$

- State variables that **DO NOT depend on the consumer's choices**:
  prices, advertising, other time-varying product characteristics.

  $\mathbf{p}_{it} = (p_{it}(j) : j = 1, 2, ..., J)$

MODEL: CONSUMER PREFERENCES

- Consumers maximize expected & discounted intertemporal utility:

$$\mathbb{E}_t \left[\sum_{s=0}^{\infty} \delta_i^s \; U_{i,t+s} \right]$$

δ_i **is unrestricted**: a component of the FE Unobserved Heterogeneity.

- Utility has four components:

$$U_{it} \;=\; b_i \left(y_{it}, \ell_{it}, d_{it} \right) \;+\; m_i \left(y_{it}, \mathbf{p}_{it} \right) \;-\; sc_i \left(y_{it}, \ell_{it} \right) \;+\; \varepsilon_{it}(y_{it})$$

$b_i \left(y_{it}, \ell_{it}, d_{it} \right)$ = utility from consumption of branded product.

$m_i \left(y_{it}, \mathbf{p}_{it} \right)$ = utility from consumption of composite product.

$sc_i \left(y_{it}, \ell_{it} \right)$ = switching cost / habits.

$\varepsilon_{it}(y_{it})$ = i.i.d. Logit / Nested Logit shock.

UTILITY: CONSUMPTION BRANDED PRODUCT

$$b_i\left(y_{it}, \ell_{it}, d_{it}\right) \equiv \begin{cases} \alpha_i(\ell_{it}) + \ln(c_{it}) & \text{if} \quad y_{it} = 0 \\ \\ \alpha_i(j) + \ln(c_{it}) & \text{if} \quad y_{it} = j > 0 \end{cases}$$

- $\alpha_i(j)$ = flow utility for consumer i from consuming brand j.
- We can see $\alpha_i(j)$ as a combination of product & consumer characteristics, observable and unobservable to the researcher.

$$\alpha_i(j) = \mathbf{x}_j' \, \boldsymbol{\beta}_i^x + \boldsymbol{\xi}_j' \, \boldsymbol{\beta}_i^{\xi}$$

- $\boldsymbol{\alpha}_i \equiv (\alpha_i(1), \alpha_i(2), ..., \alpha_i(J))$ are the fixed effects for consumer i.

UTILITY: CONSUMPTION BRANDED PRODUCT [2/2]

- A **fundamental measurement problem** in this literature is that the researcher does not observe (with enough high frequency) a consumer's amounts of consumption c_{it} and inventory i_{it}.

- Here I follow a similar approach as in Erdem, Imai, & Keane (2003) and assume a consumption rule:

$$c_{it} = \begin{cases} \lambda^{dep}(\mathbf{w}_i, \ell_{it}) \; i_{it} & \text{if} \quad y_{it} = 0 \\ i_{it} & \text{if} \quad y_{it} > 0 \end{cases}$$

  where $\lambda^{dep}(\mathbf{w}_i, j) \in (0, 1)$ is an exogenous consumption rate that may vary across products, and across consumers according to observable characteristics $\mathbf{w}_i$.

- Together with the standard transition rule for inventories, we have:

$$\ln(c_{ht}) = constant - \beta^{dep}(\mathbf{w}_i, j) \; d_{it}$$

  with $\beta^{dep}(\mathbf{w}_i, j) = -\ln(1 - \lambda^{dep}(\mathbf{w}_i, j))$.

UTILITY FROM COMPOSITE GOOD

$$m_i\left(y_{it}, \mathbf{p}_{it}\right) \;=\; \gamma(\mathbf{w}_i)\;\left(\mu_i - \sum_{j=1}^{J} p_{it}(j)\; 1\{y_{it} = j\}\right)$$

- μ_i = consumer's disposable income.
- $\gamma(\mathbf{w}_i)$ = marginal utility of the composite good, e.g., $\gamma(\mathbf{w}_i) = \mathbf{w}_i'\,\gamma$
- Identification results extend to the case of **nonlinear in consumption** but linear in parameters utlity from the composite good:

$$\gamma_1\left(\mu_i - \sum_{j=1}^{J} p_{it}(j)\; 1\{y_{it} = j\}\right) + \gamma_2\left(\mu_i - \sum_{j=1}^{J} p_{it}(j)\; 1\{y_{it} = j\}\right)^2$$

UTILITY: SWITCHING COSTS

$$sc_i\left(y_{it}, \ell_{it}\right) = \sum_{k=1}^{J} \sum_{j \neq k} 1\{\ell_{it} = k \;\&\; y_{it} = j\}\; \beta^{sc}(\mathbf{w}_i, k, j)$$

- $\beta^{sc}(\mathbf{w}_i, k, j) =$ cost of switching from brand k to brand j.

UTILITY: LOGIT IDIOSYNCRATIC SHOCKS

- $\varepsilon_{it}(j)$'s are i.i.d. over (i, t, j) type I extreme value distributed.
- I provide identification & estimation results for Nested Logit version.

COMPLETE UTILITY FUNCTION

- Putting together the different components:

$$U_{it} = \begin{cases} \alpha_i(\ell_{it}) - \beta^{dep}(\ell_{it})\ d_{it} + \varepsilon_{it}(0) & \textit{if} \quad y_{it} = 0 \\ \\ \alpha_i(j) + \gamma_i\,(\mu_i - p_{it}(j)) - \beta^{sc}(\ell_{it}, j) + \varepsilon_{it}(j) & \textit{if} \quad y_{it} = j > 0 \end{cases}$$

- We use $\mathbf{x}_{it} = (\ell_{it}, d_{it})$, and:

  $u_{\boldsymbol{\alpha}_i}(y_{it}, \mathbf{x}_{it}, \mathbf{p}_{it}) =$ utility excluding unobservable logit shocks.

MODEL: STOCHASTIC PROCESS FOR PRICES

- $p_{it}(j)$ has two components: persistent, $z_{it}(j)$; and transitory, $e_{it}(j)$.

$$p_{it}(j) \;=\; \rho\left(z_{it}(j),\; e_{it}(j)\right)$$

where $\rho(.)$ is a known function.

- Define $\mathbf{z}_{it} \equiv (z_{it}(j) : j = 1, 2, ..., J)$ and $\mathbf{e}_{it} \equiv (e_{it}(j) : j = 1, 2, ..., J)$.

- **ASSUMPTION 1:**

(i) $\mathbf{z}_{it}$ follows a first order Markov process.

(ii) **Conditional independence of transitory component of prices:** Conditional on $\mathbf{z}_{it}$, $(\mathbf{e}_{i,t+1}, \mathbf{z}_{i,t+1})$ does not depend on $\mathbf{e}_{it}$.

EXAMPLE: HI-LO PRICING

- Many supermarket products: evolution of weekly prices is characterized by the alternation between a regular price and a promotion price. See Hitsch, Hortacsu, & Lin (2019).

- Stochastic process for price:

$$p_{it}(j) \ = \ (1 - e_{it}(j)) \ z_{it}^{reg}(j) \ + \ e_{it}(j) \ z_{it}^{pro}(j)$$

- $z_{it}^{reg}(j)$ = Regular price (follows Markov chain.)

- $z_{it}^{pro}(j)$ = Promotion price (follows Markov chain.)

- $e_{it}(j)$ = Dummy variable for "promotion for product j in market i at period t". Satisfies the Conditional Independence Assumption 1(ii).

STOCHASTIC PROCESS FOR PRICES & IDENTIFICATION

- The stochastic process of prices is not needed for the identification of the parameters β^{sc} and β^{dep}.
- However, it plays a **key role in the identification of the price parameter** γ in a FE forward-looking model.
- Both $\mathbf{z}_{it}$ and $\mathbf{e}_{it}$ affect a consumer's current utility, but expected future utility (the continuation value) depends on $\mathbf{z}_{it}$ but not on $\mathbf{e}_{it}$.
- This exclusion restriction is key in the identification of γ.
- Given data on prices and a specification of the $\rho(.)$ function, it is possible to identify the two components $\mathbf{z}_{it}$ and $\mathbf{e}_{it}$.

CONSUMER DYNAMIC DECISION PROBLEM

- The decision problem of consumer i at period t is:

$$y_{it} = argmax_{j \in \mathcal{Y}} \left\{ u_{\alpha_i}(j, \mathbf{x}_{it}, \mathbf{p}_{it}) + \varepsilon_{it}(j) + v_{\alpha_i}(f_x(j, \mathbf{x}_{it}), \mathbf{z}_{it}) \right\}$$

- $f_x(j, \mathbf{x}_{it})$ = value of $\mathbf{x}_{i,t+1}$ given state $\mathbf{x}_{it}$ and decision $y_{it} = j$.
- $v_{\alpha_i}(f_x(j, \mathbf{x}_{it}), \mathbf{z}_{it})$ = *continuation value function*.
- $P(j|\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}, \alpha_i)$ = **Conditional Choice Probability (CCP)**.
- Model implies:

$$\log P\left(j|\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it}, \alpha_i\right) =$$

$$= u_{\alpha_i}(j, \mathbf{x}_{it}, \mathbf{p}_{it}) + v_{\alpha_i}(f_x(j, \mathbf{x}_{it}), \mathbf{z}_{it}) - \sigma_{\alpha_i}(\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it})$$

where $\sigma_{\alpha_i}(\mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{e}_{it})$ be the log of the denominator in the Logit CCP function (i.e, log of sum of exponentials of utilities).

3. IDENTIFICATION OF STRUCTURAL PARAMETERS

Sufficient Statistics Approach

- Let $\widetilde{\boldsymbol{y}} = \{d_1, y_0, y_1, y_2, ..., y_T\}$ be an individual's observed history

$$\mathbb{P}_i\left(\widetilde{\boldsymbol{y}}\right) = \prod_{t=1}^{T} \frac{\exp\left\{\alpha_i\left(y_t\right) + \beta\left(y_t, \boldsymbol{s}_t\right) + v_i\left(y_t, d_{t+1}(y_t)\right)\right\}}{\sum\limits_{j \in \mathcal{Y}} \exp\left\{\ \alpha_i\left(j\right) + \beta\left(j, \boldsymbol{s}_t\right) + v_i\left(j, d_{t+1}(j)\right)\ \right\}} p_i(d_1, y_0$$

- The log-probability of a choice history has the following form:

$$\ln \mathbb{P}_i\left(\widetilde{\boldsymbol{y}}\right) = S(\widetilde{\boldsymbol{y}})'\ g_\alpha + C(\widetilde{\boldsymbol{y}})'\ \boldsymbol{\beta}$$

- where $S(\widetilde{\boldsymbol{y}})$ and $C(\widetilde{\boldsymbol{y}})$ are vectors of statistics.

Sufficient Statistics Approach (2)

- This structure has several important implications.

$$\ln \mathbb{P}_i\left(\widetilde{\boldsymbol{y}}\right) = S(\widetilde{\boldsymbol{y}})'\ g_{\alpha} + C(\widetilde{\boldsymbol{y}})'\ \boldsymbol{\beta}$$

1. $S(\widetilde{\boldsymbol{y}})$ is a sufficient statistic for $\boldsymbol{\alpha}$.

2. If the elements in the vector $\left[S(\widetilde{\boldsymbol{y}}), C(\widetilde{\boldsymbol{y}})'\right]$ are linearly independent, then CMLE implies the identification of $\boldsymbol{\beta}$.

3. Given $\boldsymbol{\beta}$, the distribution of $S(\widetilde{\boldsymbol{y}})$ contains all the information in the data about the distribution of $\boldsymbol{\alpha}$.

- We consider a **sequential approach**.
  - 1st: identification of $\boldsymbol{\beta}$ from CML.
  - 2nd: identification AMEs given $\boldsymbol{\beta}$ and empirical distribution $S(\widetilde{\boldsymbol{y}})$.

Sufficient Statistics Approach (3)

- We can write the log-likelihood function as the sum of two likelihoods:

$$\ell\left(\boldsymbol{\alpha}, \boldsymbol{\beta}\right) = \ell^{C}\left(\mathbf{C}, \mathbf{S}\ ; \boldsymbol{\beta}\right) + \ell^{S}\left(\mathbf{S}\ ; \boldsymbol{\alpha}, \boldsymbol{\beta}\right)$$

with $\mathbf{C} = \{\mathbf{c}_i : i = 1, 2, ..., N\}$ and $\mathbf{S} = \{\mathbf{s}_i : i = 1, 2, ..., N\}$.

$$\ell^{C}\left(\mathbf{C}, \mathbf{S}\ ; \boldsymbol{\beta}\right) = \sum_{i=1}^{N} \mathbf{c}_i' \boldsymbol{\beta} - \sum_{i=1}^{N} \ln\left[\sum_{\mathbf{y}:\ \mathbf{s}(\mathbf{y})=\mathbf{s}_i} \exp\left\{\mathbf{c}(\mathbf{y})' \boldsymbol{\beta}\right\}\right]$$

and

$$\ell^{S}\left(\mathbf{S}\ ; \boldsymbol{\alpha}, \boldsymbol{\beta}\right) = \sum_{i=1}^{N} \ln\left[\sum_{\mathbf{y}:\mathbf{s}(\mathbf{y})=\mathbf{s}_i} \exp\left\{\mathbf{s}_i'\ \mathbf{g}(\alpha_i) + \mathbf{c}(\mathbf{y})'\ \boldsymbol{\beta}\right\}\right]$$

Sufficient Statistics Approach (4)

- Function $\ell^C(\mathbf{C}, \mathbf{S}\ ; \boldsymbol{\theta})$ is the *conditional log-likelihood function*.

  - It does not depend on the incidental parameters $\boldsymbol{\alpha}$.

  - It is globally concave in $\boldsymbol{\beta}$.

- Function $\ell^S(\mathbf{S}\ ; \boldsymbol{\alpha}, \boldsymbol{\beta})$ is the likelihood for the sufficient statistic $\mathbf{s}_i$.

  - All the information about the incidental parameters appears in this function.

  - It depends on the data only through the sufficient statistics $\mathbf{s}_i$.

- Therefore, given $\boldsymbol{\beta}$, **all the information in the data about the AMEs appears in the empirical distribution of the sufficient statistics**.

SUFFICIENT STATISTICS APPROACH

- I follow Aguirregabiria, Gu, & Luo (2021) who consider a sufficient statistic - conditional likelihood approach in the spirit of Cox (1958), Rasch (1960).

- Let $\boldsymbol{y}_i = \{\ell_1, d_1, y_1, y_2, ..., y_T\}$ be an individual's observed history; and let $\widetilde{\mathbf{z}}_i \equiv (\mathbf{z}_{i1}, \mathbf{z}_{i2}, ..., \mathbf{z}_{iT})$ and $\widetilde{\mathbf{e}}_i \equiv (\mathbf{e}_{i1}, \mathbf{e}_{i2}, ..., \mathbf{e}_{iT})$.

- $\boldsymbol{\theta}$ is the vector of structural parameters: $\beta^{sc}(.)$, $\beta^{dep}(.)$, and γ

- Probability of $\mathbf{y}_i$ conditional on history of prices $\widetilde{\mathbf{z}}_i$, $\widetilde{\mathbf{e}}_i$ and $\boldsymbol{\alpha}_i$ is:

$$\mathbb{P}\left(\mathbf{y}_i | \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i, \boldsymbol{\alpha}_i, \boldsymbol{\theta}\right) =$$

$$\mathrm{p}^*(\ell_{i1}, d_{i1} | \boldsymbol{\alpha}_i) \prod_{t=2}^{T} \frac{\exp\{u_{\boldsymbol{\alpha}_i}(y_{it}, \mathbf{x}_{it}, \mathbf{p}_{it}) + v_{\boldsymbol{\alpha}_i}(y_{it}, \mathbf{x}_{it}, \mathbf{z}_{it})\}}{\sum_{j=0}^{J} \exp\{u_{\boldsymbol{\alpha}_i}(j, \mathbf{x}_{it}, \mathbf{p}_{it}) + v_{\boldsymbol{\alpha}_i}(j, \mathbf{x}_{it}, \mathbf{z}_{it})\}}$$

SUFFICIENT STATISTICS APPROACH (2/2)

- This log-probability has the following structure:

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

- This structure has several important implications.

1. $\mathbf{s}(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)$ is a sufficient statistic for $\boldsymbol{\alpha}$.

2. If the elements in the vector $\left[\mathbf{s}(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i))', \mathbf{c}(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\right]$ are linearly independent, then CMLE implies the identification of $\boldsymbol{\theta}$.

A MORE INTUITIVE DESCRIPTION

- For every parameter in the vector $\boldsymbol{\theta}$, say θ_k, there exist two choice histories, say A and B, such that $\mathbf{s}(A) = \mathbf{s}(B)$ and $\mathbf{c}(A) - \mathbf{c}(B)$ is a vector where all the elements are zero except element k that is one.

- Under these conditions, we have that:

$$\theta_k \;=\; \log \mathbb{P}(A) - \log \mathbb{P}(B),$$

- Parameter θ_k is identified from the log odds ratio of histories A and B.

IDENTIFICATION OF β^{sc} AND γ

- For $k, j \geq 1$ with $k \neq j$, and any two natural numbers n_1 and n_2, consider the following choice histories ($\mathbf{0}_n$ = vector of n zeros):

$$A = (k,\ \mathbf{0}_{n_1},\ j,\ \mathbf{0}_{n_2},\ k,\ \mathbf{0}_{n_2},\ j)\,;\, B = (k,\ \mathbf{0}_{n_1},\ k,\ \mathbf{0}_{n_2},\ j,\ \mathbf{0}_{n_2},\ j)$$

- And the following condition on the history of prices:

$\mathbf{z}_{it}$ is constant from period $n_1 + 2$ to $n_1 + 2n_2 + 4$

- Under these conditions, we have that:

$$\log \mathbb{P}(A) - \log \mathbb{P}(B) =$$

$$- \widetilde{\beta}^{sc}(k, j) - \gamma\ \left(e_{n_1+2}(j) - e_{n_1+3}(j) - e_{n_1+2}(k) + e_{n_1+3}(k)\right)$$

IDENTIFICATION OF β^{sc} AND γ (2/2)

$$\log \mathbb{P}(A) - \log \mathbb{P}(B) =$$

$$-\ \widetilde{\beta}^{sc}(k,j) - \gamma\ \left(e_{n_1+2}(j) - e_{n_1+3}(j) - e_{n_1+2}(k) + e_{n_1+3}(k)\right)$$

- This equation shows that:
  1. A change between periods n_1+2 and n_1+3 in the transitory component of the price of product j or k identifies parameter γ.
  2. The switching cost parameter $\widetilde{\beta}^{sc}(k,j)$ is identified from histories where this transitory component is constant.

IDENTIFICATION OF β^{dep}

- **ASSUMPTION 2.** *For any product j, there is a value of duration d_j^* – which can vary across products – such that $\beta^{dep}(j,n) = \beta^{dep}(j,d_j^*)$ for any duration $n \geq d_j^*$.* ■

- **PROPOSITION.** *For any product j and any duration n, define the pair of histories:*

$$A_{j,n} = (j,\ \mathbf{0}_{n-1},\ j,\ \mathbf{0}_{n+1}) \qquad \text{and} \qquad B_{j,n} = (j,\ \mathbf{0}_n,\ j,\ \mathbf{0}_n).$$

If $d_j^* \leq (T-1)/2$, then d_j^* is identified from the following expression:

$$d_j^* \ = \ max\{n \ : \ \log \mathbb{P}(A_{j,n}) - \log \mathbb{P}(B_{j,n}) \neq 0\} \qquad ■$$

IDENTIFICATION OF β^{dep} (2/2)

- Then, for $n = d_j^* - 1$, we have that:

$$\log \mathbb{P}(A_{j,n}) - \log \mathbb{P}(B_{j,n}) \;=\; -\beta^{dep}(j, d_j^*) + \beta^{dep}(j, d_j^* - 1)$$

- The (local) depreciation rate $\beta^{dep}(j, d_j^*) - \beta^{dep}(j, d_j^* - 1)$ is identified.

- If $\beta^{dep}(j, d)$ is a linear function, i.e., $\beta^{dep}(j, d) = \overline{\beta}_j^{dep}\, d$, then the product-specific depreciation rate $\overline{\beta}_j^{dep}$ is identified.

4. ESTIMATION OF STRUCTURAL PARAMETERS

CML ESTIMATION

- Let represent a sufficient statistic for $\boldsymbol{\alpha}_i$ as a binary indicator that combines the condition $\mathbf{y}_i \in \{A \cup B\}$, and restrictions on prices, that we represent as $r(\widetilde{\mathbf{z}}_\mathbf{i}, \widetilde{\mathbf{e}}_\mathbf{i}) = \mathbf{0}$. That is:

$$s_i = 1\{\mathbf{y}_i \in A \cup B \text{ and } r(\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i) = \mathbf{0}\}$$

- There are many of these binary sufficient statistics. Let index them by $m \in \{1, 2, ..., M\}$. Then, the *conditional log-likelihood function* is:

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{m=1}^{M} \sum_{i=1}^{N} 1\{\mathbf{y}_i \in A^m \cup B^m\} \ 1\{r^m(\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i) = \mathbf{0}\}$$

$$\log \left(\frac{\exp\{c^m(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\}}{\exp\{c^m(A^m, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\} + \exp\{c^m(B^m, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\}} \right)$$

CML ESTIMATION (2/2)

- Imposing exactly the restrictions on prices typically implies loosing a substantial amount of observations.

- To deal with this issue, we follow the Kernel weighting in Honore & Kyriazidou (2000)

- The Kernel Weighted conditional log-likelihood function is:

$$\mathcal{L}^{KW}(\boldsymbol{\theta}) = \sum_{m=1}^{M} \sum_{i=1}^{N} 1\{\mathbf{y}_i \in A^m \cup B^m\} \; K\left(\frac{r^m(\widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)}{b_N}\right)$$

$$\log\left(\frac{\exp\{c^m(\mathbf{y}_i, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\}}{\exp\{c^m(A^m, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\} + \exp\{c^m(B^m, \widetilde{\mathbf{z}}_i, \widetilde{\mathbf{e}}_i)'\boldsymbol{\theta}\}}\right)$$

5. IDENTIFICATION OF AVERAGE MARGINAL EFFECTS

INTRODUCTION

- 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.

OUTLINE

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

b. General identification method

- Application of the general identification method
- Multinomial, Exogenous X, Duration, Ordered Logit.

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:

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

- 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 $T \geq 2n + 1$:

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

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:

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

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)

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

- Multiplying this equation times $(1+e^{\alpha+\beta})(1+e^{\alpha})$ 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 $\mathbf{x}$.

6. EMPIRICAL APPLICATION

Dynamic Demand for Differentiated Product

Laundry Detergent

DATA

- NIELSEN scanner data from Chicago-Kilts center.
- Period 2006-2019. Current estimates using only years 2017-2018.
- More than 40k participating households all over US.
- Rich demographics ($\boldsymbol{w}_i$): ZIP code, income, age, education, occupation, race, family size, family composition, type of residence,
- Data on every shopping trip.
- Product: Laundry detergent

ESTIMATION OF DEMAND PARAMETERS

Fixed Effects provide precise enough estimates ($N = 19,776$).

Estimates of Structural Parameters

Parameter	FE Kernel W. CML Estimate	FE Kernel W. CML (s.e.)	RE (2 types) + $\mathbf{w}_i'\alpha(j)$ Estimate	RE (2 types) + $\mathbf{w}_i'\alpha(j)$ (s.e.)
γ *Price*	1.7392	(**0.3018**)	1.155	(**0.1221**)
β^{sc}(*habits*) *Brand 1*	0.3804	(**0.0290**)	0.7551	(**0.0101**)
β^{sc}(*habits*) *Brand 2*	0.2556	(**0.0573**)	0.6695	(**0.0110**)
β^{sc}(*habits*) *Brand 3*	0.2388	(**0.0591**)	0.7360	(**0.0162**)
β^{dep}(*linear*) *Brand 1*	0.0597	(**0.0112**)	-0.0089	(**0.0040**)
β^{dep}(*linear*) *Brand 2*	0.0611	(**0.0118**)	-0.0161	(**0.0046**)
β^{dep}(*linear*) *Brand 3*	0.0692	(**0.0172**)	-0.0208	(**0.0072**)
Hausman test (p-value)	0.0000			

ESTIMATION OF DEMAND PARAMETERS

Hausman test clearly rejects the Random Effects model.

Estimates of Structural Parameters				
	FE Kernel W. CML		**RE (2 types) + $\mathbf{w}_i'\alpha(j)$**	
Parameter	*Estimate*	*(s.e.)*	*Estimate*	*(s.e.)*
γ *Price*	1.7392	(0.3018)	1.155	(0.1221)
$\beta^{sc}(habits)$ *Brand 1*	0.3804	(0.0290)	0.7551	(0.0101)
$\beta^{sc}(habits)$ *Brand 2*	0.2556	(0.0573)	0.6695	(0.0110)
$\beta^{sc}(habits)$ *Brand 3*	0.2388	(0.0591)	0.7360	(0.0162)
$\beta^{dep}(linear)$ *Brand 1*	0.0597	(0.0112)	-0.0089	(0.0040)
$\beta^{dep}(linear)$ *Brand 2*	0.0611	(0.0118)	-0.0161	(0.0046)
$\beta^{dep}(linear)$ *Brand 3*	0.0692	(0.0172)	-0.0208	(0.0072)
Hausman test (p-val)	**0.0000**			

ESTIMATION OF STRUCTURAL PARAMETERS

Random Effects model over-estimates habits parameters.

Estimates of Structural Parameters

	FE Kernel W. CML		**RE (2 types) + $\mathbf{w}_i'\alpha(j)$**	
Parameter	*Estimate*	*(s.e.)*	*Estimate*	*(s.e.)*
γ *Price*	1.7392	(0.3018)	1.155	(0.1221)
β^{sc}(*habits*) *Brand 1*	**0.3804**	(0.0290)	**0.7551**	(0.0101)
β^{sc}(*habits*) *Brand 2*	**0.2556**	(0.0573)	**0.6695**	(0.0110)
β^{sc}(*habits*) *Brand 3*	**0.2388**	(0.0591)	**0.7360**	(0.0162)
β^{dep}(*linear*) *Brand 1*	0.0597	(0.0112)	-0.0089	(0.0040)
β^{dep}(*linear*) *Brand 2*	0.0611	(0.0118)	-0.0161	(0.0046)
β^{dep}(*linear*) *Brand 3*	0.0692	(0.0172)	-0.0208	(0.0072)
Hausman test (p-value)	0.0000			

ESTIMATION OF STRUCTURAL PARAMETERS

Random Effects model provides wrong sign for duration dependence.

Estimates of Structural Parameters				
	FE Kernel W. CML		**RE (2 types) + $\mathbf{w}_i'\alpha(j)$**	
Parameter	*Estimate*	*(s.e.)*	*Estimate*	*(s.e.)*
γ *Price*	1.7392	(0.3018)	1.155	(0.1221)
$\beta^{sc}(habits)$ *Brand 1*	0.3804	(0.0290)	0.7551	(0.0101)
$\beta^{sc}(habits)$ *Brand 2*	0.2556	(0.0573)	0.6695	(0.0110)
$\beta^{sc}(habits)$ *Brand 3*	0.2388	(0.0591)	0.7360	(0.0162)
$\beta^{dep}(linear)$ *Brand 1*	**0.0597**	(0.0112)	**-0.0089**	(0.0040)
$\beta^{dep}(linear)$ *Brand 2*	**0.0611**	(0.0118)	**-0.0161**	(0.0046)
$\beta^{dep}(linear)$ *Brand 3*	**0.0692**	(0.0172)	**-0.0208**	(0.0072)
Hausman test (p-value)	0.0000			

ESTIMATION OF DEMAND PARAMETERS

Random Effects model under-estimates price-sensitivity of demand.

Estimates of Structural Parameters

	FE Kernel W. CML		**RE (2 types) + $\mathbf{w}_i'\alpha(j)$**	
Parameter	*Estimate*	*(s.e.)*	*Estimate*	*(s.e.)*
γ *Price*	**1.7392**	(0.3018)	**1.155**	(0.1221)
β^{sc}(*habits*) *Brand 1*	0.3804	(0.0290)	0.7551	(0.0101)
β^{sc}(*habits*) *Brand 2*	0.2556	(0.0573)	0.6695	(0.0110)
β^{sc}(*habits*) *Brand 3*	0.2388	(0.0591)	0.7360	(0.0162)
β^{dep}(*linear*) *Brand 1*	0.0597	(0.0112)	−0.0089	(0.0040)
β^{dep}(*linear*) *Brand 2*	0.0611	(0.0118)	−0.0161	(0.0046)
β^{dep}(*linear*) *Brand 3*	0.0692	(0.0172)	−0.0208	(0.0072)
Hausman test (p-value)	0.0000			

CONCLUSIONS / EXTENSIONS

- This paper presents a Fixed Effects dynamic panel data model of demand for different products where consumers are forward looking.

- Some relevant extensions:

1. Identification of aggregate price elasticities (i.e., AME) following recent results.

2. Consumer purchases of multiple units (for inventory).

3. Dynamics from state variables other than ℓ_{it} and d_{it}.

4. Combining this dynamic demand model with dynamic model of price competition.