Identification and Estimation of Demand Models with Endogenous Product Entry & Exit

Victor Aguirregabiria (University of Toronto)
Alessandro Iaria (University of Bristol)
Senay Sokullu (University of Bristol)

DSE CONFERENCE ON POLICY EVALUATION & HETEROGENEITY MEASUREMENT

Madison, August 7, 2024

Context – Selection Problem in Demand Estimation

- Demand estimation is usually based on data from multiple geographic regions and/or time periods (markets).
- Often, **not all products are offered in all markets**.
  - Airlines; Retail Markets; Radios; Computers; ...
- This can create an **Endogenous Selection Problem** because:
  - A product is offered when expected demand is larger.
  - Firms' information about expected demand may include variables that are unobserved to the researcher.
- However, most applications **assume product availability is exogenous**.

Context – Non-standard Selection Problem

- The reason why most applications have not dealt with this selection problem is that it is non-standard, and in fact quite challenging.

- A key feature creating this **non-standard selection problem** is that the selection equation depends nonlinearly on multiple unobservables:

$$a_j = 1 \Longleftrightarrow \pi_j(\boldsymbol{x}, \xi_1, \xi_2, ..., \xi_J) \geq 0$$

where ξ_1, ξ_2, ..., ξ_J are unobservables affecting demand.

- This structure implies that the **propensity score** (entry probability) **cannot control** for the selection bias in the demand equation:

$$P(a_j = 1|\boldsymbol{x}) \text{ is not sufficient to control for } \mathbb{E}\left(\xi_j \mid \boldsymbol{x}, a_j = 1\right)$$

$\Longrightarrow$ **Lack of identification of demand parameters** using well-known **two-step**, "Heckman-like," semiparametric control-function methods.

Context – Full Solution Estimation Methods

- This issue has motivated the development of recent **full-solution methods**: to deal with endogenous product selection:
  - Ciliberto, Murry, & Tamer (JPE, 2021).
  - Li, Mazur, Roberts, & Sweeting (RAND, 2022).

- Despite the great merit of these full-solution methods, they have limitations in terms of computational cost and robustness:
  - Nested fixed point algorithms are computationally demanding, especially given multiple equilibria.
  - Strong parametric assumptions for all the structural functions and the distribution of unobservables.

THIS PAPER

1. We establish **identification of demand using a two-step approach**.

2. We propose a **simple two-step estimator** that builds on and extends traditional methods to address endogenous selection.

3. Our approach emphasizes **robustness, flexibility, and computational simpicity**:

   - Nonparametric specification of the unobservables and expected profit.
   - Flexible information structure including complete information, private information, and multiple equilibria unobservables as particular cases.
   - The method and its computational advantages apply both to static and dynamic games of market entry & exit.

4. We illustrate the proposed method with an application.

Outline – Rest of this Presentation

1. **Model**
2. **Identification & Estimation**
3. **Empirical Application**

1. MODEL

DEMAND: BLP with Selection

- J products indexed by $j \in \mathcal{J} = \{1, 2, ..., J\}$ can be offered in market t.
- $a_{jt} \in \{0, 1\}$: indicator that product j is available in market t.
- Market shares:

$$s_{jt} = d_j(\boldsymbol{\delta}_t, \boldsymbol{\sigma}) = \int \frac{a_{jt} \ \exp\left\{\delta_{jt} + v(p_{jt}, \boldsymbol{x}_{jt}, v)\right\}}{1 + \sum_{i=1}^{J} a_{it} \ \exp\left\{\delta_{it} + v(p_{it}, \boldsymbol{x}_{it}, v)\right\}} dF_v(v|\boldsymbol{\sigma}).$$

LEMMA. *If the outside option $j = 0$ is available, then Berry (1994)'s demand invertibility applies to the sub-system of available products:*

$$\delta_{jt} = d_j^{-1}(\boldsymbol{s}_t, \boldsymbol{\sigma}) = \alpha \ p_{jt} + \boldsymbol{x}_{jt}' \ \boldsymbol{\beta} + \xi_{jt} \ \text{ if and only if } a_{jt} = 1.$$

DEMAND: BLP with Selection

- This Lemma implies that the selection/censoring condition in the regression equation for product j depends on whether this product is offered but on the offering of the other products, a_k for $k \neq j$:

$$\delta_{jt} = \begin{cases} \alpha \ p_{jt} + \boldsymbol{x}'_{jt} \ \boldsymbol{\beta} + \xi_{jt} & \text{if } a_{jt} = 1 \\ \text{unobserved} & \text{if } a_{jt} = 0 \end{cases}$$

- Therefore, the selection bias term in this regression equation is:

$$\mathbb{E}\left[\xi_{jt} \mid \boldsymbol{x}_t, a_{jt} = 1\right]$$

- In contrast, in *Almost Ideal Demand Systems* (Deaton and Muellbauer, 1980) product j's selection term would depend on the availability profile of the other products:

$$\mathbb{E}\left[\xi_{jt} \mid \boldsymbol{x}_t, a_{jt} = 1, \boldsymbol{a}_{-jt} = \boldsymbol{a}_{-j}\right]$$

SUPPLY: Dynamic Entry-Exit & Static Price Competition

- Standard model in IO: BLP + Ericson-Pakes.

- **Exogenous state variables:** Product characteristics affecting demand and marginal costs, $\boldsymbol{x}_t \equiv (\boldsymbol{x}_{1t}, ..., \boldsymbol{x}_{Jt})$, $\boldsymbol{\xi}_t \equiv (\xi_{1t}, ..., \xi_{Jt})$, $\boldsymbol{\omega}_t \equiv (\omega_{1t}, ..., \omega_{Jt})$ follow Markov processes.

- **Endogenous state variables:** $\boldsymbol{a}_{t-1} \equiv (\boldsymbol{a}_{1,t-1}, ..., \boldsymbol{a}_{J,t-1})$, as $a_{j,t-1}$ determines whether the firm needs to pay an entry cost or not.

- **Product entry / exit decisions:** Every period t, firms decide which products to offer in the market: $\boldsymbol{a}_t \equiv (a_{1t}, ..., a_{Jt})$ to maximize their intertemporal profit in a dynamic game of product entry/exit.

- **Price competition:** Given the products offered at period t, firms compete in prices: p_{jt} and s_{jt} for products with $a_{jt} = 1$ are determined in a static Bertrand-Nash equilibrium.

MODEL: Entry/Exit Game & Information Structure

- When making product entry/exit decision at time t, firm j's **information set** is:

$$\mathcal{I}_{jt} = \{\boldsymbol{a}_{t-1},\ \boldsymbol{x}_t,\ \boldsymbol{\kappa}_t,\ \eta_{jt}\}$$

(i) $\boldsymbol{a}_{t-1}$ and $\boldsymbol{x}_t$ are common knowledge to firms and observable to the researcher.

(ii) $\boldsymbol{\kappa}_t$ is common knowledge to firms & unobservable to the researcher.

- We do not restrict what is included in $\boldsymbol{\kappa}_t$.
- A case included in our model is: $\boldsymbol{\kappa}_t = (\boldsymbol{\xi}_t, \boldsymbol{\omega}_t)$.
- But firms might have uncertainty about $(\boldsymbol{\xi}_t, \boldsymbol{\omega}_t)$ at the moment of product entry decision.

(iii) η_{jt} is private information shock in firm j's entry cost, independent of $(\boldsymbol{\kappa}_t, \boldsymbol{x}_t)$, and i.i.d. over firms with CDF F_η.

MODEL: MARKOV PERFECT EQUILIBRIUM

- A **Markov Perfect Equilibrium (MPE)** of the product entry/exit game is a J-tuple of probability functions (CCPs):

$$P_j\left(\boldsymbol{a}_{t-1}, \boldsymbol{x}_t, \boldsymbol{\kappa}_t\right) = Prob\left(a_{jt} = 1 \mid \boldsymbol{a}_{t-1}, \boldsymbol{x}_t, \boldsymbol{\kappa}_t\right)$$

- Equilibrium CCPs are based on the best-reply conditions:

$$a_{jt} = 1 \iff V_j^P\left(\boldsymbol{a}_{t-1}, \boldsymbol{x}_t, \boldsymbol{\kappa}_t\right) - \eta_{jt} \;\geq\; 0$$

- Implying the equilibrium conditions:

$$P_j\left(\boldsymbol{a}_{t-1}, \boldsymbol{x}_t, \boldsymbol{\kappa}_t\right) \;=\; F_\eta\left(\; V_j^P\left(\boldsymbol{a}_{t-1}, \boldsymbol{x}_t, \boldsymbol{\kappa}_t\right)\;\right)$$

2. IDENTIFICATION & ESTIMATION

STRUCTURE OF SELECTION BIAS IN DEMAND

- For notational simplicity, here I use $\boldsymbol{x}_t$ to represent $(\boldsymbol{a}_{t-1}, \boldsymbol{x}_t)$
- If the researcher could observe $\boldsymbol{\kappa}_t$, the selection bias term in the demand equation would have a standard structure:

$$
\begin{aligned}
\mathbb{E}\left[\xi_{jt} \mid \boldsymbol{x}_t, \boldsymbol{\kappa}_t, a_{jt} = 1\right] &= \mathbb{E}\left[\xi_{jt} \mid \boldsymbol{x}_t, \boldsymbol{\kappa}_t, \eta_{jt} \leq V_j^P\left(\boldsymbol{x}_t, \boldsymbol{\kappa}_t\right)\right] \\
&= \mathbb{E}\left[\xi_{jt} \mid \boldsymbol{x}_t, \boldsymbol{\kappa}_t, \eta_{jt} \leq F_\eta^{-1}\left(P_j\left(\boldsymbol{x}_t, \boldsymbol{\kappa}_t\right) \right)\right] \\
&= \psi_j\left(P_j\left(\boldsymbol{x}_t, \boldsymbol{\kappa}_t\right), \boldsymbol{\kappa}_t\right)
\end{aligned}
$$

- Regression equation for demand is:

$$
d_j^{-1}(\mathbf{s}_t, \boldsymbol{\sigma}) = \alpha\ p_{jt} + \boldsymbol{x}_{jt}'\ \boldsymbol{\beta} + \psi_j\left(P_j\left(\boldsymbol{x}_t, \boldsymbol{\kappa}_t\right), \boldsymbol{\kappa}_t\right) + \widetilde{\xi}_{jt},
$$

that provides identification of demand parameters.

STRUCTURE OF SELECTION BIAS IN DEMAND (2)

- Since $\boldsymbol{\kappa}_t$ is unobservable, the selection bias term is:

$$\begin{aligned}\mathbb{E}\left[\xi_{jt} \mid \boldsymbol{x}_t, a_{jt}=1\right] &= \int \psi_j\left(P_j\left(\boldsymbol{x}_t, \boldsymbol{\kappa}_t\right), \boldsymbol{\kappa}_t\right) \; f(\boldsymbol{\kappa}_t | \boldsymbol{x}_t, a_{jt}=1) \; d\boldsymbol{\kappa}_t \\ &= \int \psi_j\left(P_j\left(\boldsymbol{x}_t, \boldsymbol{\kappa}_t\right), \boldsymbol{\kappa}_t\right) \frac{P_j\left(\boldsymbol{x}_t, \kappa_t\right)}{\bar{P}_j\left(\boldsymbol{x}_t\right)} f_\kappa(\kappa_t) \; d\boldsymbol{\kappa}_t\end{aligned}$$

  where $\bar{P}_j\left(\boldsymbol{x}_t\right)$ is the Propensity Score.

- It seems that, without further restrictions we cannot identify this selection term / control function.

A USEFUL REPRESENTATION RESULT

- Kargas & Sidiropoulos (IEEE, 2019) establish this convenient nonparametric finite mixture representation.

LEMMA *For any* $(\boldsymbol{a}, \boldsymbol{x}) \in \{0,1\}^J \times \mathcal{X}$ *with* $J \geq 3$, *any arbitrary probability mass function* $\Pr(\boldsymbol{a}_t = \boldsymbol{a} \mid \boldsymbol{x}_t = \boldsymbol{x})$ *admits the nonparametric finite mixture representation:*

$$\Pr(\boldsymbol{a}_t = \boldsymbol{a} \mid \boldsymbol{x}_t = \boldsymbol{x}) =$$

$$\sum_{\kappa^* \in \mathcal{K}(\boldsymbol{x})} f_{\kappa^*}(\kappa^* \mid \boldsymbol{x}) \left[\prod_{j=1}^{J} \left[P_j(\boldsymbol{x}, \kappa^*) \right]^{a_j} \left[1 - P_j(\boldsymbol{x}, \kappa^*) \right]^{1-a_j} \right]$$ *with*

$\mathcal{K}(\boldsymbol{x})$ *a discrete and finite collection of latent classes with at most* $|\mathcal{K}(\boldsymbol{x})| \leq 2^{J-1}$ *components.* ■

Identification and Estimation: Sequential Approach

- **Step 1:** Given $\Pr(\boldsymbol{a}_t \mid \boldsymbol{x}_t)$, non-parametric identification of $\boldsymbol{f}_t \equiv (f_\kappa(\boldsymbol{\kappa} \mid \boldsymbol{x}_t) : \boldsymbol{\kappa} \in \mathcal{K})$ and $\boldsymbol{P}_{jt} \equiv (P_j(\boldsymbol{x}_t, \boldsymbol{\kappa}) : \boldsymbol{\kappa} \in \mathcal{K})$.

- **Step 2:** Given $(\boldsymbol{f}_t, \boldsymbol{P}_{jt})$, identification of $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}, \boldsymbol{\sigma})$ from partially linear model:

$$d_j^{-1}(\mathbf{s}_t, \boldsymbol{\sigma}) = \alpha \; p_{jt} + \boldsymbol{x}'_{jt} \; \boldsymbol{\beta} + \boldsymbol{f}'_t \; \boldsymbol{\psi}_j(\boldsymbol{P}_{jt}) + \widetilde{\xi}_{jt},$$

where $\boldsymbol{\psi}_j(\boldsymbol{P}_{jt}) \equiv \left(\psi_j\left(P_j\left(\boldsymbol{x}_t, \boldsymbol{\kappa}\right), \boldsymbol{\kappa}\right) : \boldsymbol{\kappa} \in \mathcal{K}\right)$.

Identification and Estimation: Step 1

- Step 1 identifies the non-parametric finite mixture model:

$$\Pr\left(\boldsymbol{a}_t = \boldsymbol{a} | \boldsymbol{x}_t = \boldsymbol{x}\right) = \sum_{\boldsymbol{\kappa} \in \mathcal{K}} f_\kappa(\boldsymbol{\kappa} | \boldsymbol{x}) \left[\prod_{j=1}^{J} \left[P_j(\boldsymbol{x}, \boldsymbol{\kappa})\right]^{a_j} \left[1 - P_j(\boldsymbol{x}, \boldsymbol{\kappa})\right]^{1-a_j}\right].$$

- We follow Bonhome et al. (2016) and Aguirregabiria and Mira (2019).
  - Identification based on independence among firms' entry decisions for given $(\boldsymbol{x}, \boldsymbol{\kappa})$.
  - Number of equations 2^J "large enough" relative to number of components $|\mathcal{K}|$.
  - $f_\kappa(\boldsymbol{\kappa} \mid \boldsymbol{x}) > 0$ and linear independence of $P_j(\boldsymbol{x}, \boldsymbol{\kappa})$ across $\boldsymbol{\kappa} \in \mathcal{K}$.

- Extend existing procedures to inclusion of continuous regressors in $\boldsymbol{x}_t$, preserving $\sqrt{T}$ consistence and asymptotic normality.

Identification and Estimation: Step 2

- Given $(\boldsymbol{f}_t, \boldsymbol{P}_{jt})$, we can difference out $\boldsymbol{f}_t' \, \boldsymbol{\psi}_j(\boldsymbol{P}_{jt})$:

$$\delta_{jt} - \mathbb{E}\left[\delta_{jt} \mid \boldsymbol{f}_t, \boldsymbol{P}_{jt}\right] \;=\; \alpha \left(p_{jt} - \mathbb{E}\left[p_{jt} \mid \boldsymbol{f}_t, \boldsymbol{P}_{jt}\right]\right) + \left(\boldsymbol{x}_{jt}' - \mathbb{E}\left[\boldsymbol{x}_{jt}' \mid \boldsymbol{f}_t, \boldsymbol{P}_{jt}\right]\right)\boldsymbol{\beta} + \widetilde{\xi}_{jt},$$

- The remaining problem is the endogeneity of $\left(p_{jt} - \mathbb{E}\left[p_{jt} \mid \boldsymbol{f}_t, \boldsymbol{P}_{jt}\right]\right)$.
- Given an instrument for price, we can then identify $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta})$ by IV.
  - We require instruments $\boldsymbol{z}_{jt}$ such that $\mathbb{E}\left[\widetilde{\xi}_{jt}(\boldsymbol{\theta}) \; \boldsymbol{z}_{jt}\boldsymbol{\theta}'\right]$ has rank $\dim(\boldsymbol{\theta})$.
- We follow the literature on estimation of partially linear models.
  - Either **pairwise differencing** as in Aradillas-Lopez et al. (2007).
  - Or direct estimation of $\boldsymbol{\psi}_j(\cdot)$ by **sieves** as in Newey (2009).

3. EMPIRICAL APPLICATION

Empirical Application: US Airline Markets

- US airlines. Quarterly data from 2012-Q1 to 2013-Q4.
- 7 players: American Airlines (AA), Delta Airlines (DL), United Airlines (UA), US Airways (US), Southwest Airlines (WN), Other Low Cost (LCC), and Rest of carriers.
- Demand: Directional routes between the airports at the 100 largest Metropolitan Statistical Areas (MSA) in US.
- For entry: Market defined as non-directional airport pair.
- We include in our sample only markets with at least 1 incumbent airline in 1 year over 40 years covered in the USDOT database. This accounts for $2,230$ non-directional markets.

Distribution of Markets by Number of Entrants

Number of airlines	Frequency Markets-quarters	Market size million people	Market distance in miles
0 airlines	5,583 (31.8%)	2.88	737
1 airline	8,204 (46.7%)	3.42	916
2 airlines	2,614 (14.9%)	4.44	955
3 airlines	844 (4.8%)	5.36	1,112
4 airlines	221 (1.2%)	5.44	1,136
5 airlines	68 (0.4%)	8.61	1,185
$\geq$ 6 airlines	7 (0.1%)	6.95	314
Total	17,541 (100.0%)	3.54	881

Entry Frequencies/Probabilities by Airline

Airline	Frequency # markets-quarters (%)	Avg. market size in million people	Avg. market distance in miles
WN	4,602 (26.23%)	3.63	981
DL	3,257 (18.56%)	4.07	876
UA	3,221 (18.36%)	4.50	968
LCC	2,382 (13.57%)	4.61	1,171
US	1,933 (11.02%)	3.98	879
AA	1,815 (10.34%)	5.32	962

Estimation of the model for market entry

- Sieve finite mixture Logit.
- Vector of explanatory variables x_t includes:
  - Market size (`msize`), as measured by the sum of populations in the MSAs of the two airports.
  - Market distance (`mktdistance`), as the geodesic distance between the two airports.
  - Airline's own hub-size in the market (`ownhub-size`), as measured by the sum of the airline's hub-size in the two airports.
  - Average hub-size of the other airlines (`comphub-size`).
  - Airline$\times$Time dummies.

Estimation Market Entry Model — Goodness-of-Fit

Statistics	**Logit 1 type**	**Mixture Logit 2 types**	**Mixture Logit 3 types**	**Mixture Logit 4 types**
# Obs.	$17,155$	$17,155$	$17,155$	$17,155$
Parameters	72	145	218	287
Log-like.	$-20,378$	$-18,985$	$-18,022$	$-17,621$
AIC	$40,900$	$38,261$	$36,481$	$35,817$
BIC	$41,458$	$39,385$	$38,170$	$38,041$

Estimation of Demand Parameters

- Nested logit demand as in Ciliberto et al. (JPE, 2021) but controlling for selection bias using control function method.

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \alpha\ p_{jt} + \boldsymbol{x}'_{jt}\boldsymbol{\beta} + \sigma\ \ln\left(\frac{s_{jt}}{1-s_{0t}}\right) + \boldsymbol{h}'_{jt}\ \gamma_j^{\psi} + \widetilde{\xi}_{jt}.$$

- We present different estimators according to the specification of the control function:

    1. *Parametric Nested Logit*.
    2. *Semiparametric Nested Logit without* κ_t.
    3. *Semiparametric mixture Nested Logit*.

- To deal with endogenous prices, we use as IVs: number of competitors in the market and average *hub-size* of the rest of the airlines.

Estimation of Demand Parameters

	Not control. for sel.		*Controlling for endogenous selection*			
	OLS	**2SLS**	**2SLS Heckman**	**2SLS Semi-P.**	**2SLS Fin-Mix** $\lvert\mathcal{K}\rvert = 2$	**2SLS Fin-Mix** $\lvert\mathcal{K}\rvert = 3$
Price (100$) ($\alpha$)	-0.643	-2.180	-2.193	-2.261	-2.574	-2.708
	(0.0105)	(0.1378)	(0.1348)	(0.1298)	(0.1549)	(0.1662)
Within Share (σ)	0.371	0.409	0.413	0.431	0.547	0.570
	(0.0058)	(0.0351)	(0.0389)	(0.0372)	(0.0509)	(0.0565)
Distance (1000mi)	0.729	2.130	2.196	2.264	2.497	2.472
	(0.0306)	(0.1372)	(0.1365)	(0.1310)	(0.1524)	(0.1572)
Distance^2	-0.216	-0.424	-0.453	-0.462	-0.496	-0.511
	(0.0112)	(0.0244)	(0.0252)	(0.0250)	(0.0276)	(0.0289)
hub-size orig. (100s)	1.637	2.272	1.999	1.320	1.593	1.383
	(0.0263)	(0.0382)	(0.0593)	(0.0625)	(0.0869)	(0.0989)
hub-size dest. (100s)	1.613	2.242	1.995	1.310	1.587	1.377
	(0.0267)	(0.0385)	(0.0595)	(0.0633)	(0.0872)	(0.0994)
Airline$\times$Quarter FE	Y	Y	Y	Y	Y	Y
# control var. entry	0	0	6	18	36	54
Observations	35,763	35,763	35,763	35,763	35,763	35,763

Average Own-Price Elasticities and Lerner Indexes

	Not control. for sel.		*Controllin for endogenous selection*			
	OLS	**2SLS**	**2SLS Heckman**	**2SLS Semi-P.**	**2SLS Fin-Mix** $\lvert\mathcal{K}\rvert = 2$	**2SLS Fin-Mix** $\lvert\mathcal{K}\rvert = 3$
Own-Price Elasticity	-1.596	-5.549	-5.601	-5.849	-7.406	-8.000
AA	-1.722	-6.013	-6.071	-6.363	-8.169	-8.857
DL	-1.761	-6.082	-6.133	-6.382	-7.871	-8.450
UA	-1.887	-6.573	-6.636	-6.936	-8.847	-9.573
US	-1.665	-5.801	-5.856	-6.122	-7.809	-8.450
WN	-1.354	-4.680	-4.719	-4.913	-6.068	-6.517
LCC	-1.370	-4.808	-4.857	-5.095	-6.674	-7.265
Others	-1.332	-4.705	-4.757	-5.006	-6.706	-7.337
Lerner Index	68.8%	19.9%	19.7%	18.9%	15.4%	14.4%
AA	62.7%	18.0%	17.9%	17.1%	13.8%	12.8%
DL	60.4%	17.5%	17.3%	16.7%	13.7%	12.8%
UA	56.9%	16.4%	16.2%	15.6%	12.6%	11.7%
US	65.9%	19.0%	18.9%	18.1%	14.8%	13.8%
WN	78.4%	22.8%	22.6%	21.8%	18.2%	17.1%
LCC	82.1%	23.5%	23.3%	22.2%	17.5%	16.3%
Others	79.2%	22.5%	22.3%	21.3%	16.4%	15.2%
Observations	35,763	35,763	35,763	35,763	35,763	35,763

Empirical Distribution of Estimated Elasticities