INCAE PhD SUMMER ACADEMY
DYNAMIC GAMES IN EMPIRICAL IO

Lecture 4:
Structural estimation of dynamic games

Victor Aguirregabiria (University of Toronto)

June 23, 2022

Lecture 4: Identification & Estimation of Dynamic Games: Outline

1. **Datasets in applications**

2. **Full Solution Methods**
   [2.1.] Nested Fixed Point algorithm (NFXP)
   [2.2.] Nested Pseudo Likelihood (NPL)

3. **Two-step CCP methods**

1. Datasets in Applications

Type of Data in most Empirical Applications

- Panel data of M geographic markets, over T periods, and N firms.

$$Data = \{\boldsymbol{a}_{mt},\ \mathbf{x}_{mt} : m = 1, 2, ..., M;\ t = 1, 2, ..., T\}$$

- **Example 1**: Major airlines in US ($N = 10$), in the markets/routes defined by all the pairs of top-50 US airports ($M = 1,275$), over $T = 20$ quarters (5 years).

- **Example 2**: Supermarket chains in Ontario ($N = 6$), in the geographic markets defined by census tracts ($M > 1k$), over $T = 24$ months.

Type of Data in most Empirical Applications [2]

- This data structure applies to industries characterized by **many geographic markets**, where a separate (dynamic) game is played in each market: e.g., retail industries, services, airline markets, procurement auctions, ...

- However, there are many manufacturing industries where **competition is more global**: a single national or even international market: e.g., microchips.

- For these "global" industries, applications rely on sample variability that comes from a **combination of modest N, M, and T**.

- Some other industries are characterized by a **large number of heterogeneous firms** (large N), e.g., NYC taxis.

2. Full Solution Estimation Methods

ESTIMATION METHODS

- The primitives of the model, $\{\pi_i, \beta_i, F_x, G_\varepsilon : i \in \mathcal{I}\}$, can be described in terms of a vector of structural parameters $\boldsymbol{\theta}$ that is unknown to the researcher.

- We study methods for the estimation of $\boldsymbol{\theta}$.

- It is convenient to distinguish three components in the vector of structural parameters: $\boldsymbol{\theta} = (\boldsymbol{\theta}_\pi, \boldsymbol{\theta}_f, \boldsymbol{\beta})$.

- **Full Solution Methods** impose the equilibrium restrictions in the estimated structural parameters $(\widehat{\boldsymbol{\theta}})$ and CCPs $(\widehat{\mathbf{P}})$:

$$\widehat{\mathbf{P}} \ = \ \Psi\left(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{P}}\right)$$

6. Full Solution Methods

MLE-NFXP with equilibrium uniqueness

- Rust (1987) NFXP algorithm is a gradient method to obtain MLE.

- Originally proposed for single-agent models, it has been applied to the estimation of games with unique equilibrium for every $\boldsymbol{\theta}$.

- Let $\{P_i(a_i|\mathbf{x},\boldsymbol{\theta}) : i \in \mathcal{I}\}$ be the equilibrium CCPs associated with $\boldsymbol{\theta}$. The **full log-likelihood function** is: $\ell(\boldsymbol{\theta}) = \sum_{m=1}^{M} \ell_m(\boldsymbol{\theta})$, where $\ell_m(\boldsymbol{\theta})$ is the contribution of market m:

$$\ell_m(\boldsymbol{\theta}) = \sum_{i=1}^{N}\sum_{t=1}^{T} \log P_i(a_{imt}|\mathbf{x}_{mt},\boldsymbol{\theta}) + \log f_x(\mathbf{x}_{m,t+1}|\boldsymbol{a}_{mt},\mathbf{x}_{mt},\boldsymbol{\theta}_f)$$

MLE-NFXP with equilibrium uniqueness [2]

- NFXP combines BHHH iterations (**outer algorithm**) with equilibrium solution algorithm (**inner algorithm**) for each trial value $\boldsymbol{\theta}$.

- A BHHH iteration is:

$$\widehat{\boldsymbol{\theta}}_{k+1} = \widehat{\boldsymbol{\theta}}_k + \left(\sum_{m=1}^{M} \frac{\partial \ell_m(\widehat{\boldsymbol{\theta}}_k)}{\partial \boldsymbol{\theta}} \frac{\partial \ell_m(\widehat{\boldsymbol{\theta}}_k)}{\partial \boldsymbol{\theta}'} \right)^{-1} \left(\sum_{m=1}^{M} \frac{\partial \ell_m(\widehat{\boldsymbol{\theta}}_k)}{\partial \boldsymbol{\theta}} \right)$$

- The score vector $\partial \ell_m(\widehat{\boldsymbol{\theta}}_k) / \partial \boldsymbol{\theta}$ depends on $\partial \log P_i(a_{imt} | \mathbf{x}_{mt}, \widehat{\boldsymbol{\theta}}_k) / \partial \boldsymbol{\theta}$. To obtain these derivatives, the inner algorithm of NFXP solves for the equilibrium CCPs given $\widehat{\boldsymbol{\theta}}_k$.

MLE-NFXP with multiple equilibria

- With Multiple Equilibria, $\ell_m(\boldsymbol{\theta})$ is not a function but a correspondence.

- To define the MLE in a model with multiple equilibria, it is convenient to define an *extended* or **Pseudo Likelihood function**.

- For arbitrary values of $\boldsymbol{\theta}$ and firms' CCPs $\mathbf{P}$, define:

$$Q(\boldsymbol{\theta}, \mathbf{P}) = \sum_{m=1}^{M} \sum_{i=1}^{N} \sum_{t=1}^{T} \log \Psi_i(a_{imt} \mid \mathbf{x}_{mt}, \boldsymbol{\theta}, \mathbf{P})$$

where Ψ_i is the *best response probability function*.

MLE-NFXP with multiple equilibria [2]

- A modified version of NFXP can be applied to obtain the MLE in games with multiple equilibria.

- The MLE is the pair $(\widehat{\boldsymbol{\theta}}_{MLE}, \widehat{\mathbf{P}}_{MLE})$ that maximizes the Q subject to the constraint that CCPs are equilibrium strategies associated:

$$(\widehat{\boldsymbol{\theta}}_{MLE}, \widehat{\mathbf{P}}_{MLE}, \widehat{\boldsymbol{\lambda}}_{MLE}) = \arg \max_{(\boldsymbol{\theta},\mathbf{P},\boldsymbol{\lambda})} \; Q(\boldsymbol{\theta}, \mathbf{P}) + \boldsymbol{\lambda}' \left[\mathbf{P} - \Psi(\boldsymbol{\theta}, \mathbf{P}) \right]$$

- The F.O.C. are the Lagrangian equations:

$$\left\{ \begin{array}{rcl} \widehat{\mathbf{P}}_{MLE} - \Psi(\widehat{\boldsymbol{\theta}}_{MLE}, \widehat{\mathbf{P}}_{MLE}) & = & \mathbf{0} \\ \nabla_{\boldsymbol{\theta}} Q(\widehat{\boldsymbol{\theta}}_{MLE}, \widehat{\mathbf{P}}_{MLE}) - \widehat{\boldsymbol{\lambda}}'_{MLE} \; \nabla_{\boldsymbol{\theta}} \Psi(\widehat{\boldsymbol{\theta}}_{MLE}, \widehat{\mathbf{P}}_{MLE}) & = & \mathbf{0} \\ \nabla_{\mathbf{P}} Q(\widehat{\boldsymbol{\theta}}_{MLE}, \widehat{\mathbf{P}}_{MLE}) - \widehat{\boldsymbol{\lambda}}'_{MLE} \; \nabla_{\mathbf{P}} \Psi(\widehat{\boldsymbol{\theta}}_{MLE}, \widehat{\mathbf{P}}_{MLE}) & = & \mathbf{0} \end{array} \right.$$

MLE-NFXP with multiple equilibria [3]

- A Newton method can be used to obtain a root of this system of Lagrangian equations.
- A key computational problem is the very high dimensionality of this system of equations.
- The most costly part of this algorithm is the calculation of the Jacobian matrix $\nabla_{\mathbf{P}} \Psi(\widehat{\theta}, \widehat{\mathbf{P}})$. In dynamic games, in general, this is not a sparse matrix, and can contain billions or trillions of elements.
- The evaluation of the best response mapping $\Psi(\theta, \mathbf{P})$ for a new value of $\mathbf{P}$ requires solving for a valuation operator and solving a system of equations with the same dimension as $\mathbf{P}$.
- Due to serious computational issues, there are no empirical applications of dynamic games with multiple equilibria that compute the MLE, with either the NFXP or MPEC algorithms.

Nested Pseudo Likelihood (NPL)

- Imposes equilibrium restrictions but does NOT require:

  - Repeatedly solving for MPE for each trial value of $\boldsymbol{\theta}$ (as NFXP)
  - Computing $\nabla_{\mathbf{P}} \Psi(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{P}})$ (as NFXP and MPEC)

- A NPL $(\widehat{\boldsymbol{\theta}}_{NPL}, \widehat{\mathbf{P}}_{NPL})$, that satisfy two conditions:

  (1) given $\widehat{\mathbf{P}}_{NPL}$, $\widehat{\boldsymbol{\theta}}_{NPL} = \arg\max_{\boldsymbol{\theta}} \; Q(\boldsymbol{\theta}, \widehat{\mathbf{P}}_{NPL})$;

  (2) given $\widehat{\boldsymbol{\theta}}_{NPL}$, $\widehat{\mathbf{P}}_{NPL} = \Psi(\widehat{\boldsymbol{\theta}}_{NPL}, \widehat{\mathbf{P}}_{NPL})$.

- The NPL estimator is consistent and asymptotically normal under the same regularity conditions as the MLE. For dynamic games, the NPL estimator has larger asymptotic variance than the MLE.

Nested Pseudo Likelihood (NPL) [2]

- An algorithm to compute the NPL is the **NPL fixed point algorithm**.
- Starting with an initial $\widehat{\mathbf{P}}_0$, at iteration $k \geq 1$:

  (Step 1) given $\widehat{\mathbf{P}}_{k-1}$, $\widehat{\boldsymbol{\theta}}_k = \arg\max_{\boldsymbol{\theta}} \; Q(\boldsymbol{\theta}, \widehat{\mathbf{P}}_{k-1})$;

  (Step 2) given $\widehat{\boldsymbol{\theta}}_k$, $\widehat{\mathbf{P}}_k = \Psi(\widehat{\boldsymbol{\theta}}_k, \widehat{\mathbf{P}}_{k-1})$.
- Step 1 is very simple in most applications, as it is equivalent to obtaining the MLE in a static single-agent discrete choice model.
- Step 2 is equivalent to solving once a system of linear equations with the same dimension as $\mathbf{P}$.
- A limitation of this fixed point algorithm is that **convergence is not guaranteed**. An alternative algorithm that has been used to compute NPL is a **Spectral Residual algorithm**.

3. Two-step CCP Methods

Hotz-Miller CCP Method

- To avoid the computational cost of full-solution methods, simpler two-step methods have been proposed.
- Hotz & Miller (1993) was a seminal contribution on this class of methods. They show that the conditional choice values are known functions of CCPs, transition probabilities, and $\boldsymbol{\theta}$.
- When $\pi_i(\boldsymbol{a}_t, \mathbf{x}_t) = h(\boldsymbol{a}_t, \mathbf{x}_t)\ \boldsymbol{\theta}_{\pi,i}$:

$$v_i(a_{it}, \mathbf{x}_t) \ = \ \widetilde{h}_i^{\mathbf{P}}(a_{it}, \mathbf{x}_t)\ \boldsymbol{\theta}_{\pi,i} \ + \ \widetilde{e}_i^{\mathbf{P}}(a_{it}, \mathbf{x}_t)$$

with:

$$\widetilde{h}_i^{\mathbf{P}}(a_{it}, \mathbf{x}_t) \ = \ \mathbb{E}\left(\sum_{j=0}^{\infty} \beta_i^j \ h(\boldsymbol{a}_{t+j}, \mathbf{x}_{t+j}) \mid a_{it}, \mathbf{x}_t\right)$$

$$\widetilde{e}_i^{\mathbf{P}}(a_{it}, \mathbf{x}_t) \ = \ \mathbb{E}\left(\sum_{j=0}^{\infty} \beta_i^j \ [\gamma - \ln P_i(a_{i,t+j}|\mathbf{x}_{t+j})] \mid a_{it}, \mathbf{x}_t\right)$$

Hotz-Miller CCP Method [2]

- Given this representation of conditional choice values, the pseudo likelihood function $Q(\boldsymbol{\theta}, \mathbf{P})$ has practically the same structure as in a static or reduced form discrete choice model.

- Best response probabilities that enter in $Q(\boldsymbol{\theta}, \mathbf{P})$ can be seen as the choice probabilities in a standard random utility model:

$$\Psi_i(a_{imt} | \mathbf{x}_{mt}, \boldsymbol{\theta}, \mathbf{P}) \;=$$

$$\Pr\left(a_{imt} = \arg\max_j \left\{ \widetilde{h}_i^{\mathbf{P}}(j, \mathbf{x}_{mt})\ \boldsymbol{\theta}_i + \widetilde{e}_i^{\mathbf{P}}(j, \mathbf{x}_{mt}) + \varepsilon_{it}(j) \right\}\right).$$

- Given $\widetilde{h}_i^{\mathbf{P}}(., \mathbf{x}_{mt})$ and $\widetilde{e}_i^{\mathbf{P}}(., \mathbf{x}_{mt})$ and a parametric specification for the distribution of ε (e.g., logit, probit), the vector of parameters $\boldsymbol{\theta}_i$ can be estimated as in a standard logit or probit model.

Hotz-Miller CCP Method [3]

- The method proceeds in two steps.
- Let $\widehat{\mathbf{P}}^0$ be a consistent nonparametric estimator of true $\mathbf{P}^0$. The two-step estimator of $\boldsymbol{\theta}$ is defined as:

$$\widehat{\boldsymbol{\theta}}_{2S} \;=\; \arg\max_{\boldsymbol{\theta}} \; Q(\boldsymbol{\theta}, \widehat{\mathbf{P}}^0)$$

- Under standard regularity conditions, this two-step estimator is root-M consistent and asymptotically normal.
- It can be extended to incorporate market unobserved heterogeneity (e.g., Aguirregabiria & Mira (2007); Arcidiacono & Miller (2011)).
- Monte Carlo Simulation can be used to compute present values: Bajari, Benkard, & Levin (2007).
- Limitation: Finite sample bias due to imprecise estimates of CCPs in the first step.