Lecture 13: Estimation Methods for **DYNAMIC** Games of Incomplete Information

Two-step CCP, Sequential NPL, MLE-MPEC, MLE-NFXP

Fedor Iskhakov, Australian National University
Anders Munk-Nielsen, University of Copenhagen
John Rust, Georgetown University
Bertel Schjerning, University of Copenhagen

Empirical IO: Dynamic Structural Models
Department of Economics (NHH)
29 Nov - 3 Dec., 2021

Estimating dynamic discrete-choice games of incomplete information

Egesdahl, Lai and Su (Quantitative Economics, 2014)

Che-Lin Su, 1974 - 2017

Road Map for Lecture 13

We begin with an example: Dynamic exit/entry model similar to the one analyzed in Aguirregabiria and Mira (ECMA, 2007)

Then

- Introduction to dynamic discrete games.
- Estimation methods: MPEC, NFXP, NPL, 2-step methods
- Empirical results from Aguirregabiria and Mira (ECMA, 2007): Dynamic model for entry/entry in oligopoly markets using Chilean data from several retail industries
- Monte Carlo from Egesdal, Lai, Su (QE, 2015)

Econometric Issues

- Indeterminacy problem due to multiple equilibria
- The computational burden in the solution of the game.
- A&M (2007) propose a class of PML estimators that they argue deals with these problems
- A&M (2007) analyze asymptotic and finite sample properties of several estimators in this class.

Extry/Exit Games: An Illustrating Example

- Five firms: $i = 1, ..., N = 5$
- Firm i's decision in period t:

$$a_i^t = 0 : \text{exit (inactive)};\ a_i^t = 1 : \text{enter (active)};$$

- Simultaneous decisions conditional on observing the market size, all firms' decisions in the last period and private shocks

Time	Market Size	Firm 1	Firm 2	Firm 3	Firm 4	Firm 5
0	2	0	0	0	0	0
1	3	0	1	0	0	1
2	4	0	1	0	1	1
3	5	0	1	0	0	1
4	5	1	1	0	0	0
5	5	1	1	0	0	1
6	6	1	1	1	1	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Estimation Methods for Discrete-Choice Games of Incomplete Information

Maximum-Likelihood (ML) estimator

- Efficient estimator, but expensive to compute
- Egesdal, Lai and Su (2015) propose a constrained optimization formulation for the ML estimator to estimate dynamic games
- NFXP relies on full solution methods.
  - Iskhakov, Rust and Schjerning (2016) develop such an algorithm for DDGs (coming lecture)
  - Borkovsky, Doraszelsky and Kryukov (2010): All solution homotopy methods (path following algorithms)

Two-step estimators

- Bajari, Benkard, Levin (2007), Pesendorfer and Schmidt-Dengler (2008), Pakes, Ostrovsky, and Berry (2007)
- Computationally simple, but potentially large finite-sample biases

Nested Pseudo Likelihood (NPL) estimator:

- Aguirregabiria and Mira (2007), Kasahara and Shimotsu (2012)

Road Map for rest of lecture

Plan

1. The dynamic game in AM (2007)
2. State variables
3. Player i's Utility Maximization Problem
4. Equilibrium Concept: Markov Perfect Equilibrium (MPE)
5. Bellman Optimality
6. Bayes-Nash Equilibrium Conditions
7. Solving for the Markov Perfect Equilibrium
8. Estimation methods: MPEC, NFXP, NPL, 2-step methods
9. Monte Carlo results from Egesdal, Lai, Su (QE, 2015)
10. Empirical results from A & M (2007)

The Dynamic Game in AM (2007)

- Discrete time infinite-horizon: $t = 1, 2, ..., \infty$
- N players: $i \in \mathcal{I} = \{1, ..., N\}$
- The market is characterized by size $s^t \in \mathcal{S} = \{s_1, ..., s_L\}$.
  - market size is observed by all players
  - exogenous and stationary market size transition: $f_{\mathcal{S}}(s_{t+1}|s_t)$
- At the beginning of each period t, player i observes (x^t, ε_i^t)
  - x_t: a vector of common-knowledge state variables
  - ε_i^t: private shocks
- Players then simultaneously choose whether to be active in the market in that period
- $a_i^t \in \mathcal{A} = \{0, 1\}$: player i's action in period t
- $a^t = (a_1^t, ..., a_N^t)$: the collection of all players' actions.
- $a_{-i}^t = (a_1^t, ..., a_{i-1}^t, a_{i+1}^t, ..., a_N^t)$: the current actions of all players other than i

State Variables

- Common-knowledge state variables: $\mathbf{x}_t = (s_t, \mathbf{a}^{t-1})$
- Private shocks: $= \varepsilon_i^t = \{\varepsilon_i^t(a_i^t)\}_{a_i^t \in \mathcal{A}}$
- $\varepsilon_i^t(a_i^t)$ has a i.i.d type-I extreme value distribution across actions and players as well as over time
- opposing players know only its probability density function $g(\varepsilon_i^t)$.
- The **conditional independence** assumption on state transition:

$$p\left[\mathbf{x}^{t+1} = (s', \mathbf{a}'), \varepsilon_i^{t+1} | \mathbf{x}^t = (s, \tilde{\mathbf{a}}), \varepsilon_i^t, \mathbf{a}^t\right] = f_{\mathcal{S}}(s'|s)1\{\mathbf{a}' = \mathbf{a}^t\}g(\varepsilon_i^{t+1})$$

Player i's Utility Maximization Problem

- θ : the vector of structural parameters
- $\beta \in (0, 1)$: the discount factor
- player i's per-period payoff function:

$$\tilde{\Pi}_i(a_i^t, \mathbf{a}_{-i}^t, \mathbf{x}^t, \varepsilon_{\mathbf{i}}^{\mathbf{t}}; \theta) = \Pi_i(a_i^t, \mathbf{a}_{-i}^t, \mathbf{x}^t; \theta) + \varepsilon_i^t(a_i^t)$$

- The common-knowledge component of the per-period payoff

$$\begin{aligned}
&\Pi_i(a_i^t, \mathbf{a}_{-i}^t, \mathbf{x}^t; \theta) \\
&= \begin{cases} \theta^{RS} s^t - \theta^{RN} \log\left\{1 + \sum_{j \neq i} a_j^t\right\} - \theta_i^{FC} - \theta^{EC}(1 - a_i^{t-1}), & \text{if } a_i^t = 1 \\ 0, & \text{if } a_i^t = 0 \end{cases}
\end{aligned}$$

- Player i's utility maximization problem:

$$\max_{\{a_i^t, a_i^{t+\mathbf{1}}, a_i^{t+\mathbf{2}}, \ldots.\}} \mathbb{E}\left[\sum_{\tau=t}^{\infty} \tilde{\Pi}_i(a_i^\tau, \mathbf{a}_{-i}^\tau, \mathbf{x}^\tau, \varepsilon_{\mathbf{i}}^\tau; \theta) | (\mathbf{x}^t, \varepsilon_{\mathbf{i}}^{\mathbf{t}})\right]$$

Equilibrium Concept: Markov Perfect Equilibrium

- Equilibrium characterization in terms of the observed states x
- $P_i(a_i|\mathbf{x})$: the conditional choice probability of player i choosing action a_i at state $\mathbf{x}$
- $V_i(\mathbf{x})$: the expected value function for player i at state $\mathbf{x}$
- Define $\mathbf{P} = \{P_i(a_i|\mathbf{x})\}_{i\in\mathcal{I},a_i\in\mathcal{A},\mathbf{x}\in\mathcal{X}}$ and $\mathbf{V} = \{V_i(\mathbf{x})\}_{i\in\mathcal{I},\mathbf{x}\in\mathcal{X}}$
- A Markov perfect equilibrium is a vector $(\mathbf{V}, \mathbf{P})$ that satisfies two systems of nonlinear equations:
    - Bellman equation (for each player i)
    - Bayes-Nash equilibrium conditions

System I: Bellman Optimality

- **Bellman Optimality**. $\forall i \in \mathcal{I}, \mathbf{x} \in \mathcal{X}$

$$V_i(\mathbf{x}) = \sum_{a_i \in \mathcal{A}} P_i(a_i|\mathbf{x}) \left[\pi_i(a_i|\mathbf{x}, \theta) + e_i^{\mathbf{P}}(a_i|\mathbf{x})\right] + \beta \sum_{\mathbf{x} \in \mathcal{X}} V_i(\mathbf{x}') f_{\mathcal{X}}^{\mathbf{P}}(\mathbf{x}'|\mathbf{x})$$

- $\pi_i(a_i|\mathbf{x}, \theta)$: the expected payoff for player i from choosing action a_i at state $\mathbf{x}$ and given beliefs of other players's actions, $P_j(a_j|\mathbf{x})$,

$$\pi_i(a_i|\mathbf{x}, \theta) = \sum_{\mathbf{a}_{-i} \in \mathcal{A}^{N-1}} \left\{ \left[\prod_{a_j \in \mathbf{a}_{-i}} P_j(a_j|\mathbf{x}) \right] \Pi_i(a_i, \mathbf{a}_{-i}, \mathbf{x}; \theta) \right\}$$

- $f_{\mathcal{X}}^{\mathbf{P}}(\mathbf{x}'|\mathbf{x})$: state transition probability of $\mathbf{x}$, given $\mathbf{P}$

$$f_{\mathcal{X}}^{\mathbf{P}}(\mathbf{x}' = (s', \mathbf{a}')|\mathbf{x} = (s, \tilde{\mathbf{a}})) = \left[\prod_{j=1}^{N} P_j(a_j'|\mathbf{x})\right] f_{\mathcal{S}}(s'|s)$$

- Assuming that $\varepsilon_{\mathbf{i}}^{\mathbf{t}}$ is i.i.d. extreme value distributed with scale factor σ, it's conditional expectation is

$$e_i^{\mathbf{P}}(a_i|\mathbf{x}) = \text{Euler's constant} - \sigma \log[P_i(a_i|\mathbf{x})]$$

System II: Bayes-Nash Equilibrium Conditions

- **Bayes-Nash Equilibrium.**

$$P_i(a_i = j|\mathbf{x}) = \frac{\exp[v_i(a_i = j|\mathbf{x})]}{\sum_{k \in \mathcal{A}} \exp[v_i(a_i = k|\mathbf{x})]}, \forall i \in \mathcal{I}, j \in \mathcal{A}, (x) \in \mathcal{X}$$

- $v_i(a_i = j|\mathbf{x})$: choice-specific expected value function

$$v_i(a_i = j|\mathbf{x}) = \pi_i(a_i|\mathbf{x}, \theta) + \beta \sum_{(x) \in (X)} V_i(\mathbf{x}') f_i^{\mathbf{P}}(\mathbf{x}'|\mathbf{x}, a_i)$$

- $f_i^{\mathbf{P}}(\mathbf{x}'|\mathbf{x}, a_i)$: the state transition probability conditional on the current state $\mathbf{x}$, player i's action a_i, and his beliefs $\mathbf{P}$

$$f_i^{\mathbf{P}}(\mathbf{x}' = (s', \mathbf{a}')|\mathbf{x} = (s, \tilde{\mathbf{a}}), a_i)) = f_{\mathcal{S}}(s'|s)\mathbf{1}\{a_i' = a_i\} \prod_{j \in \mathcal{I} \setminus i}^{N} P_j(a_j'|\mathbf{x})$$

Markov Perfect Equilibrium

- **Bellman Optimality**. $\forall i \in \mathcal{I}, x \in \mathcal{X}$

$$V_i(\mathbf{x}) = \sum_{a_i \in \mathcal{A}} P_i(a_i|\mathbf{x}) \left[\pi_i(a_i|\mathbf{x}, \theta) + \mathbf{e}_i^{\mathbf{P}}(a_i|\mathbf{x})\right] + \beta \sum_{\mathbf{x} \in \mathcal{X}} V_i(\mathbf{x}') f_{\mathcal{X}}^{\mathbf{P}}(\mathbf{x}'|\mathbf{x})$$

- **Bayes-Nash Equilibrium**.

$$P_i(a_i = j|\mathbf{x}) = \frac{\exp[v_i(a_i = j|\mathbf{x})]}{\sum_{k \in \mathcal{A}} \exp[v_i(a_i = k|\mathbf{x})]}, \forall i \in \mathcal{I}, j \in \mathcal{A}, (x) \in \mathcal{X}$$

- In compact notation

$$\begin{aligned} \mathbf{V} &= \Psi^{\mathbf{V}}(\mathbf{V}, \mathbf{P}, \theta) \\ \mathbf{P} &= \Psi^{\mathbf{P}}(\mathbf{V}, \mathbf{P}, \theta) \end{aligned}$$

- Set of all Markov Perfect Equilibria

$$SOL(\Psi, \theta) = \left\{ (\mathbf{P}, \mathbf{V}) \,\middle|\, \begin{array}{l} \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V}, \mathbf{P}, \theta) \\ \mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V}, \mathbf{P}, \theta) \end{array} \right\}$$

Data Generating Process

DGP

- θ_0 : the true value of structural parameters in the population
- $(\mathbf{V}^0, \mathbf{P}^0)$: a Markov perfect equilibrium at $\theta_{\mathbf{0}}$
- **Assumption**: If multiple Markov perfect equilibria exist, only one equilibrium is played in the data

Data: $\mathbf{Z} = \{\bar{\mathbf{a}}^{mt}, \bar{\mathbf{x}}^{mt}\}_{m \in \mathcal{M}, t \in \mathcal{T},}$

- Observations from M independent markets over T periods
- In each market m and time period t, researchers observe
  - the common-knowledge state variables $\bar{\mathbf{x}}^{mt}$
  - players' actions $\bar{\mathbf{a}}^{mt} = (\bar{a_1}^{mt}, ..., \bar{a_N}^{mt})$

Maximum-Likelihood Estimation

- For a given θ, let $\left(\mathbf{P}^{\ell}(\theta), \mathbf{V}^{\ell}(\theta)\right) \in SOL(\Psi, \theta)$ the ℓ-the equilibrium
- Given data $\mathbf{Z} = \{\bar{\mathbf{a}}^{mt}, \bar{\mathbf{x}}^{mt}\}_{m \in \mathcal{M}, t \in \mathcal{T}}$ the log-likelihood function is

$$\mathcal{L}(Z, \theta) = \max_{(\mathbf{P}^{\ell}(\theta), \mathbf{V}^{\ell}(\theta) \in SOL(\Psi, \theta)} \frac{1}{M} \sum_{i=1}^{N} \sum_{m=1}^{M} \sum_{t=1}^{T} \log P_i^{\ell}(\bar{a}_i{}^{mt} | \bar{\mathbf{x}}^{mt}; \theta)$$

- The ML estimator is

$$\theta^{ML} = \arg\max_{\theta} \mathcal{L}(Z, \theta)$$

Structural Estimation ML Estimation via Constrained Optimization Approach

- Given data: $\mathbf{Z} = \{\bar{\mathbf{a}}^{mt}, \bar{\mathbf{x}}^{mt}\}_{m \in \mathcal{M}, t \in \mathcal{T},}$ the log of the augmented likelihood function is

$$\mathcal{L}(Z, P) = \frac{1}{M} \sum_{i=1}^{N} \sum_{m=1}^{M} \sum_{t=1}^{T} \log P_i(\bar{a}_i{}^{mt} | \bar{\mathbf{x}}^{mt}; \theta)$$

- The constrained optimization formulation of the ML estimation problem is

$$\begin{aligned} \max_{(\theta, \mathbf{P}, \mathbf{V})} \quad & \mathcal{L}(\mathbf{Z}, \mathbf{P}) \\ \text{subject to} \quad & \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V}, \mathbf{P}, \theta) \\ & \mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V}, \mathbf{P}, \theta) \end{aligned}$$

- Theorem: Both constrained and unconstrained problems have same solutions.

Solving All Equilibria in ML Estimation?

Motivation for MPEC: It has been stated in the literature that researchers using the constrained optimization approach do not need to solve for all the equilibria at each guess of structural parameter vector

MPEC

- Constraints are satisfied (and an equilibrium solved) only at a solution, not at every iteration
- The constrained optimization approach only needs to find those equilibria together with structural parameters that are local solutions and satisfy the corresponding first-order conditions
- These two features eliminate a large set of equilibria together with structural parameters that do not need to be solved
- What could go wrong?

Two-Step Methods: Intuition

- Recall the constrained optimization formulation for the ML estimator is

$$
\begin{aligned}
\max_{(\theta,\mathbf{P},\mathbf{V})} \quad & \mathcal{L}(\mathbf{Z},\mathbf{P}) \\
\text{subject to} \quad & \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V},\mathbf{P},\theta) \\
& \mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V},\mathbf{P},\theta)
\end{aligned}
$$

- Denote the solution by $(\theta, \mathbf{P}^*, \mathbf{V}^*)$
- Suppose we know $\mathbf{P}^*$, how do we recover θ^* (and $\mathbf{V}^*$)?

Two-Step Pseudo Maximum-Likelihood (2S-PML)

- Step 1: nonparametrically estimate the conditional choice probabilities, denoted by $\hat{\mathbf{P}}$ directly from the observed data $\mathbf{Z}$
- Step 2: solve

$$\begin{aligned} \max_{(\theta,\mathbf{P},\mathbf{V})} \quad & \mathcal{L}(\mathbf{Z},\mathbf{P}) \\ \text{subject to} \quad & \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V},\hat{\mathbf{P}},\theta) \\ & \mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V},\hat{\mathbf{P}},\theta) \end{aligned}$$

- or equivalently

$$\begin{aligned} \max_{(\theta,\mathbf{V})} \quad & \mathcal{L}(\mathbf{Z},\Psi^{\mathbf{P}}(\mathbf{V},\hat{\mathbf{P}},\theta)) \\ \text{subject to} \quad & \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V},\hat{\mathbf{P}},\theta) \end{aligned}$$

Reformulation of the Optimization Problem in Step 2

- **Bellman Optimality**. $\forall i \in \mathcal{I}, x \in \mathcal{X}$

$$V_i(\mathbf{x}) = \sum_{a_i \in \mathcal{A}} P_i(a_i|\mathbf{x}) \left[\pi_i(a_i|\mathbf{x},\theta) + \mathbf{e}_i^{\mathbf{P}}(a_i|\mathbf{x})\right] + \beta \sum_{\mathbf{x} \in \mathcal{X}} V_i(\mathbf{x}') f_{\mathcal{X}}^{\mathbf{P}}(\mathbf{x}'|\mathbf{x})$$

- Define:
  $\mathbf{V}_i = [V_i(x)]_{\mathbf{x}\in\mathcal{X}}$, $\hat{\mathbf{P}}_i(a_i) = [\hat{P}_i(a_i)(x)]_{\mathbf{x}\in\mathcal{X}}$, $\mathbf{e}_i^{\mathbf{P}}(a_i) = [\mathbf{e}_i^{\mathbf{P}}(a_i|\mathbf{x})]_{\mathbf{x}\in\mathcal{X}}$,
  $\pi_i(a_i,\theta) = [\pi_i(a_i|\mathbf{x},\theta)]_{\mathbf{x}\in\mathcal{X}}$, $\mathbf{F}_{\mathcal{X}}^{\hat{\mathbf{P}}} = [f_{\mathcal{X}}^{\mathbf{P}}(\mathbf{x}'|\mathbf{x})]_{\mathbf{x},\mathbf{x}'\in\mathcal{X}}$,
- The Bellman equation above can be rewritten as

$$\left[\mathbf{I} - \beta \mathbf{F}_{\mathcal{X}}^{\hat{\mathbf{P}}}\right] \mathbf{V}_i = \sum_{a_i \in \mathcal{A}} \left[\hat{\mathbf{P}}_i(a_i) \times \pi_i(a_i,\theta)\right] + \sum_{a_i \in \mathcal{A}} \left[\hat{\mathbf{P}}_i(a_i) \times \mathbf{e}_i^{\mathbf{P}}(a_i)\right],$$

  or equivalently

$$\mathbf{V}_i = \left[\mathbf{I} - \beta \mathbf{F}_{\mathcal{X}}^{\hat{\mathbf{P}}}\right]^{-1} \left\{\sum_{a_i \in \mathcal{A}} \left[\hat{\mathbf{P}}_i(a_i) \times \pi_i(a_i,\theta)\right] + \sum_{a_i \in \mathcal{A}} \left[\hat{\mathbf{P}}_i(a_i) \times \mathbf{e}_i^{\mathbf{P}}(a_i)\right]\right\},$$

  or in compact notation:

$$\mathbf{V} = \Gamma(\theta, \hat{\mathbf{P}})$$

Reformulation of the Optimization Problem in Step 2

- Replacing the constraint $\mathbf{V} = \Psi(\mathbf{V}, \hat{\mathbf{P}}, \theta)$ by $\mathbf{V} = \Gamma(\theta, \hat{\mathbf{P}})$ through a simple elimination of $\mathbf{V}$, the optimization problem in Step 2 becomes:

$$\max_{\theta} \mathcal{L}(\mathbf{Z}, \Psi^{\mathbf{P}}(\Gamma(\theta, \hat{\mathbf{P}}), \hat{\mathbf{P}}, \theta))$$

- The 2S-PML estimator is defined as

$$\theta^{2S-PML} = \arg\max_{\theta} \mathcal{L}(\mathbf{Z}, \Psi^{\mathbf{P}}(\Gamma(\theta, \hat{\mathbf{P}}), \hat{\mathbf{P}}, \theta))$$

NPL Estimator

- The 2S-PML estimator can have large biases in finite samples
- In an effort to reduce the finite-sample biases associated with the 2S-PML estimator, Aguirregabiria and Mira (2007) propose an NPL estimator
- A NPL fixed point $(\tilde{\theta}, \tilde{\mathbf{P}})$ satisfies the conditions

$$\tilde{\theta} = \arg\max_{\theta} \mathcal{L}(\mathbf{Z}, \Psi^{\mathbf{P}}(\Gamma(\theta, \tilde{\mathbf{P}}), \tilde{\mathbf{P}}, \theta))$$
$$\tilde{\mathbf{P}} = \Psi^{\mathbf{P}}(\Gamma(\theta, \tilde{\mathbf{P}}), \tilde{\mathbf{P}}, \theta)$$

- The NPL algorithm: For $1 \leq K \leq \bar{K}$, iterate over Steps 1 and 2:
  Step 1:
  Given $\tilde{\mathbf{P}}_{K-1}$,
  solve $\tilde{\theta}_K = \arg\max_{\theta} \mathcal{L}(\mathbf{Z}, \Psi^{\mathbf{P}}(\Gamma(\theta, \tilde{\mathbf{P}}_{K-1}), \tilde{\mathbf{P}}_{K-1}, \theta))$
  Step 2:
  Given $\tilde{\theta}_K$, update $\tilde{\mathbf{P}}_K$ by
  $\tilde{\mathbf{P}}_K = \Psi^{\mathbf{P}}(\Gamma(\theta_K, \tilde{\mathbf{P}}_{K-1}), \tilde{\mathbf{P}}_{K-1}, \theta_K)$
  increase K by 1.

A Modified NPL Algorithm: NPL-Λ

- It is now well known that the NPL algorithm may not converge or even if it converges, it may fail to provide consistent estimates
- Kasahara and Shimotsu (Ecta, 2012) propose the $NPL - \lambda$ algorithm that modifies Step 2 of the NPL algorithm to compute the NPL estimator $\tilde{\mathbf{P}}_K = \left(\Psi^{\mathbf{P}}(\Gamma(\theta_K, \tilde{\mathbf{P}}_{K-1}), \tilde{\mathbf{P}}_{K-1}, \theta_K)\right)^{\lambda} \left(\tilde{\mathbf{P}}_{K-1}\right)^{1-\lambda}$ where $\lambda$is chosen to be between 0 and 1.
- The proper value for λ depends on the true parameter values θ_0
- Alternatively, Kasahara and Shimotsu suggest computing the spectral radius (largest eigenvalue) of the mapping

$$\nabla_{\mathbf{P}} \Psi^{\mathbf{P}}(\Gamma(\theta_K, \tilde{\mathbf{P}}_{K-1}), \tilde{\mathbf{P}}_{K-1}, \theta_K)$$

  at every guess of structural parameter vector $\tilde{\mathbf{P}}_K$

Aguirregabiria and Mira (2007) Example

- Discrete time infinite-horizon: $t = 1, 2, ..., \infty$
- $N = 5$ players: $i \in \mathcal{I} = \{1, ..., 5\}$
- The market is characterized by size $s^t \in \mathcal{S} = \{1, ..., 5\}$.
- Total number of grid points in the state space: $|\mathcal{X}| = |\mathcal{S}| \times |\mathcal{A}| = 5 \times 2^5 = 160$
- The discount factor $\beta = 0.95$; the scale parameter of the type I extreme value distribution, $\sigma = 1$
- The common-knowledge component of the per-period payoff

$$
\begin{aligned}
&\Pi_i(a_i^t, \mathbf{a}_{-i}^t, \mathbf{x}^t; \theta) \\
&= \begin{cases} \theta^{RS} \log(s^t) - \theta^{RN} \log\left\{1 + \sum_{j \neq i} a_j^t\right\} - \theta_i^{FC} - \theta^{EC}(1 - a_i^{t-1}), & \text{if } a_i^t = 1 \\ 0, & \text{if } a_i^t = 0 \end{cases}
\end{aligned}
$$

- $\theta = (\theta^{RS}, \theta^{RN}, \theta^{FC}, \theta^{EC})$: the vector of structural parameters with $\theta^{FC} = \{\theta_i^{FC}\}_{i=1}^N$

Descriptive Evidence: Aguirregabiria and Mira (2007)

DESCRIPTIVE STATISTICS: 189 MARKETS; YEARS 1994–1999

Descriptive Statistics	Restaurants	Gas Stations	Bookstores	Shoe Shops	Fish Shops
Number of firms per 10,000 people	14.6	1.0	1.9	0.9	0.7
Markets with					
0 firms	32.2%	58.6%	49.5%	67.1%	74.1%
1 firm	1.3%	15.3%	15.8%	10.8%	9.6%
2 firms	1.2%	7.8%	8.0%	6.7%	5.0%
3 firms	0.5%	5.2%	6.9%	3.8%	3.4%
4 firms	1.2%	4.0%	3.6%	2.7%	2.0%
More than 4 firms	63.5%	9.2%	16.2%	8.9%	5.9%
Herfindahl index (median)	0.169	0.738	0.663	0.702	0.725
Annual revenue per firm (in thousand $)	17.6	67.7	23.3	67.2	124.8
Regression log(1+# firms) on log(market size)[a]	0.383 (0.043)	0.133 (0.019)	0.127 (0.024)	0.073 (0.020)	0.062 (0.018)
Regression log(firm size) on log(market size)[b]	−0.019 (0.034)	0.153 (0.082)	−0.066 (0.050)	0.223 (0.081)	0.097 (0.111)
Entry rate (%)[c]	9.8	14.6	19.7	12.8	21.3
Exit rate (%)[d]	9.9	7.4	13.5	10.4	14.5
Survival rate (hazard rate)					
1 year (%)[e]	86.2 (13.8)	89.5 (10.5)	84.0 (16.0)	86.8 (13.2)	79.7 (20.3)
2 years (%)	69.5 (19.5)	88.5 (1.1)	70.0 (16.6)	71.1 (18.2)	58.1 (27.2)
3 years (%)	60.1 (14.9)	84.6 (4.3)	60.0 (14.3)	52.6 (25.1)	44.6 (23.3)

[a] Market size = population. Regression included time dummies. Standard errors are given in parentheses.
[b] Firm size = revenue per firm. Regression included time dummies. Standard errors are given in parentheses.
[c] Entry rate = entrants / incumbents.

What drives entry/exit decisions

Some observations

- Why so many Restaurants and so few Gas stations and Bookstores?
- Market concentration, smaller in the restaurant industry.
- Turnover rates are very high in all retail industries,
  ...but survival is more likely for gas stations than for other industries.

What explains these facts?

- Economies of scale
  ...smaller fixed cost for restaurants?
- Sunken entry costs
  ...smaller for restaurants, higher for gas stations?
- Strategic Interactions
  ...is product differentiation possible for gas stations?
  ...what about restaurants and bookstores?

Structural Estimates: Aguirregabiria and Mira (2007)

TABLE VIII

NPL ESTIMATION OF ENTRY–EXIT MODEL[a]

Parameters	Restaurants	Gas Stations	Bookstores	Shoe Shops	Fish Shops
Variable profit:					
$\frac{\theta_{RS}}{\sigma_\varepsilon}$	1.743	1.929	2.029	2.030	0.914
	(0.045)	(0.127)	(0.076)	(0.121)	(0.125)
$\frac{\theta_{RN}}{\sigma_\varepsilon}$	1.643	2.818	1.606	2.724	1.395
	(0.176)	(0.325)	(0.201)	(0.316)	(0.234)
Fixed operating cost:					
$\frac{\theta_{\mathrm{FC}}}{\sigma_\varepsilon}$	9.519	12.769	15.997	14.497	6.270
	(0.478)	(1.251)	(0.141)	(1.206)	(1.233)
Entry cost:					
$\frac{\theta_{\mathrm{EC}}}{\sigma_\varepsilon}$	5.756	10.441	5.620	5.839	4.586
	(0.030)	(0.150)	(0.081)	(0.145)	(0.121)
$\frac{\sigma_\omega}{\sigma_\varepsilon}$	1.322	2.028	1.335	2.060	1.880
	(0.471)	(1.247)	(0.133)	(1.197)	(1.231)
Number of observations	945	945	945	945	945
R-squared:					
Entries	0.298	0.196	0.442	0.386	0.363
Exits	0.414	0.218	0.234	0.221	0.298

[a]Standard errors are given in parentheses. These standard errors are computed from the formulae in Section 4, which do not account for the error in the estimation of the parameters in the autoregressive process of market size.

Structural Estimates: Aguirregabiria and Mira (2007)

TABLE IX

NORMALIZED PARAMETERS

Parameters[a]	Restaurants	Gas Stations	Bookstores	Shoe Shops	Fish Shops
$\dfrac{\theta_{\text{FC}}}{\theta_{RS}\ln(S_{\text{Med}})}$	0.590	0.716	0.852	0.772	0.742
$\dfrac{\theta_{\text{EC}}}{\theta_{RS}\ln(S_{\text{Med}})}$	0.357	0.585	0.299	0.311	0.542
$100\dfrac{\theta_{RN}\ln(2)}{\theta_{RS}\ln(S_{\text{Med}})}$	7.1%	10.9%	5.9%	10.1%	11.4%
$\dfrac{\sigma_{\omega}^2}{\theta_{RS}^2\operatorname{var}(\ln(S))+\sigma_{\omega}^2+1}$	0.278	0.436	0.235	0.423	0.642

- $\theta_{FC}/(\theta_{RS} ln(S_{med}))$: Ratio between fixed operating costs and variable profits of a monopolist in a market of median size
- $\theta_{EC}/(\theta_{RS} ln(S_{med}))$: Ratio between sunken entry costs and the variable profit of a monopolist in a market of median size
- $\theta_{RN} ln(2)/(\theta_{RS} ln(S_{med}))$: Pct. reduction in variable profits per firm when we go from monopoly to duopoly in a market of median size.
- $\sigma_w^2/(\theta_{RS}^2 var(ln(S)) + \sigma_w^2 + 1)$: Pct. of cross-market variability in monopoly profits that is explained by the unobserved market type

Structural Estimates: Aguirregabiria and Mira (2007)

TABLE IX

NORMALIZED PARAMETERS

Parameters[a]	Restaurants	Gas Stations	Bookstores	Shoe Shops	Fish Shops
$\dfrac{\theta_{\text{FC}}}{\theta_{RS}\ln(S_{\text{Med}})}$	0.590	0.716	0.852	0.772	0.742
$\dfrac{\theta_{\text{EC}}}{\theta_{RS}\ln(S_{\text{Med}})}$	0.357	0.585	0.299	0.311	0.542
$100\dfrac{\theta_{RN}\ln(2)}{\theta_{RS}\ln(S_{\text{Med}})}$	7.1%	10.9%	5.9%	10.1%	11.4%
$\dfrac{\sigma_\omega^2}{\theta_{RS}^2\operatorname{var}(\ln(S))+\sigma_\omega^2+1}$	0.278	0.436	0.235	0.423	0.642

- Fixed operating costs are a very important component of total profits.
- Entry costs are statistically significant in all five industries. (Gas stations are the retailers with largest sunken costs.)
- The strategic interaction parameter is statistically significant for all five industries.

Summary of findings: Aguirregabiria and Mira (2007)

1. Economies of scale are smaller in the restaurant industry...
   $\rightarrow$ Explains the large number of restaurants.
2. Strategic interactions small among restaurants and bookstores...
   $\rightarrow$ This also contributes to explain the large number of restaurants.
   $\rightarrow$ might be due to more product differentiation in those industries....
3. Bookstore industry: Economies of scale seem very important, but negative strategic interactions are weak
   $\rightarrow$ Relatively large number of bookstores
   (more than gas stations or shoe shops).
4. Sunken entry costs such as industry-specific investments are significant in all five industries
   ... but smaller than annual fixed operating costs.

   Gas stations have largest entry costs
   $\rightarrow$ explains the lower turnover for gas stations.

Monte Carlo from Egesdal, Lai, Su (2015)

We continue with Aguirregabiria and Mira (2007) example

- Market size transition matrix is

$$f_{\mathcal{S}}(s^{t+1}|s^t) = \begin{pmatrix} 0.8 & 0.2 & 0 & .... & 0 & 0 \\ 0.2 & 0.6 & 0.2 & .... & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & ... & 0.2 & 0.6 & 0.2 \\ 0 & 0 & ... & 0 & 0.2 & 0.8 \end{pmatrix}$$

  where market size takes $|\mathcal{S}|$ values, $s^t \in \mathcal{S} = \{1, 2, ..., |\mathcal{S}|\}$
- True values of structural parameters $\theta_0^{FC} = (1.9, 1.8, 1.7, 1.6, 1.5)$ and $\theta_0^{EC} = 1$
- For θ^{RS}, θ^{RN} and $|\mathcal{S}|$ we consider the cases:
  Case 3: $(\theta^{RS}, \theta^{RN}) = (2, 1)$, $|\mathcal{S}| = 5$
  Case 4: $(\theta^{RS}, \theta^{RN}) = (4, 2)$, $|\mathcal{S}| = 5$
  Case 5: $(\theta^{RS}, \theta^{RN}) = (2, 1)$, $|\mathcal{S}| = 10$
  Case 6: $(\theta^{RS}, \theta^{RN}) = (2, 1)$, $|\mathcal{S}| = 10$
- Case 3 is the Experiment 3 in Aguirregabiria and Mira (2007)

MPEC: Number constraints and variables in experiments

- The MPEC implementation in Egesdal, Lai and Su (QE, 2015) impose the following two of state and player specific constraints

$$\mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V}, \mathbf{P}, \theta)$$

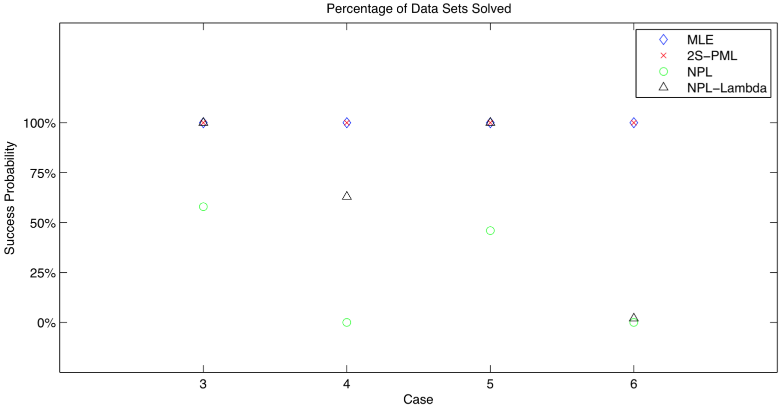
$$\mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V}, \mathbf{P}, \theta)$$

  and constraint that imposes **P** sum to 1 over alternatives.
  $\rightarrow$ 3 constraints for each state and player
- Total number constraints $|\mathcal{X}| \cdot N \cdot 3 = |\mathcal{S}| \cdot 2^N \cdot N \cdot 3$
  case 3: $5 \cdot 2^5 \cdot 3 \cdot 5 = 2,400$
  case 4: $5 \cdot 2^5 \cdot 3 \cdot 5 = 2,400$
  case 5: $10 \cdot 2^5 \cdot 3 \cdot 5 = 4,800$
  case 6: $15 \cdot 2^5 \cdot 3 \cdot 5 = 7,200$
- Total number variables $|\mathcal{X}| \cdot N \cdot 3 + dim(\theta)$
- Essential to utilize sparsity. In case 6 the full Jacobian of constraints has $7200 \cdot 7208 = 51,897,600$ elements, but only 1.08 pct of them are non-zero (561,600 in total).

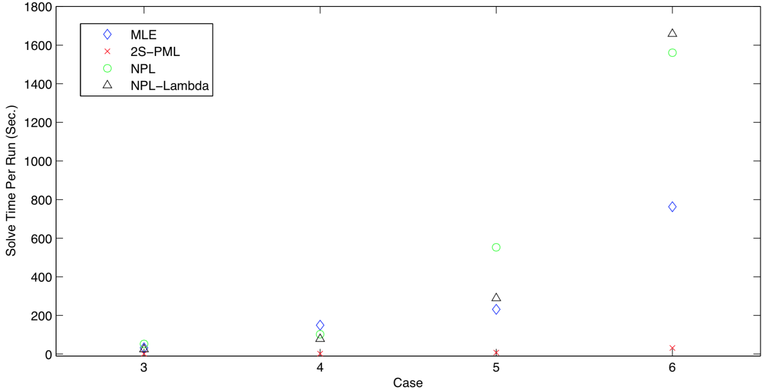
More simulation and implementation details

- In each dataset: $M = 400$ and $T = 10$
- For Case 3 and 4
  - Simulate 100 Monte Carlo data sets for each case
  - MPEC: 10 starting points for each data set
- For Cases 5 and 6
  - Simulate 50 Monte Carlo data sets for each case
  - MPEC: 5 start points for each data sets
- For NPL and NPL-Λ: $\bar{K} = 100$
- For the NPLΛ algorithm: $\lambda = 0.5$

Monte Carlo Results: Percentage of Data Sets Solved



Monte Carlo Results: Avg. Solve Time Per Run



Monte Carlo Results: Estimates for Experiment 1

Figure: Monte Carlo Results

Case	Estimator	Estimates							
		$\theta_{FC,1}$	$\theta_{FC,2}$	$\theta_{FC,3}$	$\theta_{FC,4}$	$\theta_{FC,5}$	θ_{EC}	θ_{RN}	θ_{RS}
	Truth	1.9	1.8	1.7	1.6	1.5	1	2	1
3	MLE	1.895 (0.077)	1.794 (0.078)	1.697 (0.075)	1.597 (0.074)	1.495 (0.073)	0.990 (0.046)	2.048 (0.345)	1.011 (0.095)
3	2S-PML	1.884 (0.066)	1.774 (0.069)	1.662 (0.065)	1.548 (0.062)	1.425 (0.057)	1.040 (0.039)	0.805 (0.251)	0.671 (0.068)
3	NPL	1.894 (0.075)	1.788 (0.077)	1.688 (0.069)	1.581 (0.071)	1.478 (0.073)	1.010 (0.041)	1.812 (0.213)	0.946 (0.061)
3	NPL-Λ	1.896 (0.077)	1.795 (0.079)	1.697 (0.076)	1.597 (0.074)	1.495 (0.073)	0.991 (0.044)	2.039 (0.330)	1.008 (0.091)
	Truth	1.9	1.8	1.7	1.6	1.5	1	4	2
4	MLE	1.897 (0.084)	1.797 (0.084)	1.697 (0.082)	1.594 (0.085)	1.496 (0.095)	0.993 (0.045)	4.015 (0.216)	2.004 (0.086)
4	2S-PML	1.934 (0.090)	1.824 (0.085)	1.703 (0.079)	1.556 (0.079)	1.338 (0.085)	1.123 (0.049)	2.297 (0.330)	1.409 (0.117)
4	NPL	N/A (N/A)	N/A (N/A)	N/A (N/A)	N/A (N/A)	N/A (N/A)	N/A (N/A)	N/A (N/A)	N/A (N/A)
4	NPL-Λ	1.900 (0.079)	1.801 (0.081)	1.700 (0.077)	1.600 (0.080)	1.500 (0.091)	0.991 (0.052)	4.023 (0.255)	2.007 (0.098)

Monte Carlo Results: Estimates for Experiment 2

Figure: Monte Carlo Results

$\|S\|$	Estimator	Estimates							
		$\theta_{FC,1}$	$\theta_{FC,2}$	$\theta_{FC,3}$	$\theta_{FC,4}$	$\theta_{FC,5}$	θ_{EC}	θ_{RN}	θ_{RS}
	Truth	1.9	1.8	1.7	1.6	1.5	1	2	1
10	MLE	1.882 (0.092)	1.780 (0.087)	1.677 (0.079)	1.584 (0.084)	1.472 (0.068)	0.999 (0.046)	2.031 (0.201)	1.004 (0.048)
10	2S-PML	1.884 (0.102)	1.792 (0.088)	1.679 (0.082)	1.583 (0.087)	1.469 (0.068)	1.039 (0.048)	1.065 (0.222)	0.755 (0.053)
10	NPL	1.919 (0.092)	1.810 (0.089)	1.699 (0.068)	1.606 (0.079)	1.485 (0.071)	1.011 (0.050)	1.851 (0.136)	1.966 (0.036)
10	NPL-Λ	1.884 (0.095)	1.781 (0.089)	1.678 (0.081)	1.584 (0.085)	1.472 (0.070)	0.997 (0.049)	2.032 (0.211)	1.005 (0.051)
15	MLE	1.897 (0.098)	1.800 (0.107)	1.694 (0.087)	1.597 (0.093)	1.492 (0.090)	0.983 (0.059)	2.040 (0.311)	1.011 (0.069)
15	2S-PML	1.792 (0.119)	1.705 (0.123)	1.595 (0.119)	1.506 (0.114)	1.394 (0.114)	1.046 (0.059)	0.766 (0.220)	0.664 (0.053)
15	NPL	N/A (N/A)	N/A (N/A)	N/A (N/A)	N/A (N/A)	N/A (N/A)	N/A (N/A)	N/A (N/A)	N/A (N/A)
15	NPL-Λ	1.922 (0.000)	1.821 (0.000)	1.671 (0.000)	1.611 (0.000)	1.531 (0.000)	1.012 (0.000)	1.992 (0.000)	1.007 (0.000)

Conclusion

- Recursive methods (NPL and NPL-Λ algorithms) are not always reliable computational algorithms and should be used with caution.
- The 2S-PML estimator often produces large finite-sample biases
  - Not surprising, see discussion in Pakes, Ostrovsky, and Berry (2007)
  - Can other two-step estimators perform better?
- The constrained optimization approach seems reliable and capable of estimating relevant dynamic game models such as those in Aguirregabiria and Mira (2007)
- Improving the performance of the constrained optimization approach on dynamic games with higher-dimensional state space?
- Is MPEC really as reliable as Nested Fixed point algorithm, when there are a huge multiplicity of equilibria?
  - NFXP requires reliable algorithms that can find all MPE's fast
  - We investigate this next lecture for a specific class of games: Dynamic Directional Games.