Estimating Dynamic Discrete-Choice Games of Incomplete Information

Che-Lin Su

The University of Chicago
Booth School of Business

joint work with Michael Egesdal and Zhenyu Lai (Harvard University)

2014 Workshop on Optimization for Modern Computation
BICMR
September 2–4, 2014

Roadmap of the Talk

- Introduction / Literature Review
- The Model
- Estimation
- Monte Carlo Experiments / Results
- Conclusion

Part I

Introduction

Discrete-Choice Games

- An active research topic in applied econometrics, empirical IO and marketing
- Classical application: entry/exit decisions
    - Bresnahan and Reiss (1987, 1991), Berry (1992)
    - Determining the sources of firms profitability
    - Understanding how firms react to competition
- Other applications:
    - Location choices: Seim (2006), Orhun (2012)
    - Pricing strategy (EDLP vs. Promotion): Ellickson and Misra (2008), Ellickson, Misra and Nair (2012)
    - Technology innovation: Igami (2012)
- Identification: Sweeting (2009), de Paula and Tang (2012)

Extry/Exit Games: An Illustrating Example

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

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

- 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	?	?	?	?	?
2						
3						
4						
5						
6						
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Extry/Exit Games: An Illustrating Example

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

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

- 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						
3						
4						
5						
6						
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Extry/Exit Games: An Illustrating Example

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

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

- 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	?	?	?	?	?
3						
4						
5						
6						
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Extry/Exit Games: An Illustrating Example

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

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

- 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						
4						
5						
6						
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Extry/Exit Games: An Illustrating Example

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

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

- 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	?	?	?	?	?
4						
5						
6						
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Extry/Exit Games: An Illustrating Example

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

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

- 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						
6						
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Extry/Exit Games: An Illustrating Example

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

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

- 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	?	?	?	?	?
5						
6						
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Extry/Exit Games: An Illustrating Example

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

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

- 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						
6						
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Extry/Exit Games: An Illustrating Example

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

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

- 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	0
6						
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Extry/Exit Games: An Illustrating Example

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

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

- 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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Estimation Methods for Discrete-Choice Games of Incomplete Information

- Maximum-Likelihood (ML) estimator
  - Efficient estimator in large-sample theory
  - Expensive to compute
- Two-step estimators: Bajari, Benkard, Levin (2007), Pesendorfer and Schmidt-Dengler (2008), Pakes, Ostrovsky, and Berry (2007)
  - Computationally simple
  - Potentially large finite-sample biases
- Nested Pseudo Likelihood (NPL) estimator: Aguirregabiria and Mira (2007), Kasahara and Shimotsu (2012)
- Moment inequality estimator: Pakes, Porter, Ho, and Ishii (2011)
  - does not require the assumption that only one equilibrium is played in the data
- Constrained optimization approach: Su and Judd (2012), Dubé, Fox and Su (2012)

What We Do in This Paper

- Based on Su and Judd (2012), propose a constrained optimization formulation for the ML estimator to estimate dynamic games
- Conduct Monte Carlo experiments to compare performance of different estimators
  - Two-step pseudo maximum likelihood (2S-PML) estimator
  - NPL estimator implemented by NPL algorithm and NPL-Λ algorithm
  - ML estimator via the constrained optimization approach

Part II

The Model

The Dynamic Game Model 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 $s^t \in \mathcal{S} = \{s_1, \ldots, s_L\}$.
  - 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 $(\boldsymbol{x}^t, \boldsymbol{\varepsilon}_i^t)$
  - $\boldsymbol{x}^t$: a vector of common-knowledge state variables
  - $\boldsymbol{\varepsilon}_i^t$: 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
  - $\boldsymbol{a}^t = (a_1^t, \ldots, a_N^t)$: the collection of all players' actions.
  - $\boldsymbol{a}_{-i}^t = (a_1^t, \ldots, a_{i-1}^t, a_{i+1}^t, \ldots, a_N^t)$: the current actions of all players other than i

State Variables

- Common-knowledge state variables: $\boldsymbol{x}^t = \left(s^t, \boldsymbol{a}^{t-1}\right)$
- Private shocks: $\boldsymbol{\varepsilon}_i^t = \left\{\varepsilon_i^t\left(a_i^t\right)\right\}_{a_i^t \in \mathcal{A}}$
  - $\varepsilon_i^t\left(a_i^t\right)$ 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 $g(\boldsymbol{\varepsilon}_i^t)$.
- The **conditional independence** assumption on state transition:

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

Player i's Utility Maximization Problem

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

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

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

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

- Player i's utility maximization problem:

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

Equilibrium Concept: Markov Perfect Equilibrium

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

System I: Bellman Optimality

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

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

- $\pi_i(a_i|\boldsymbol{x}, \boldsymbol{\theta})$: the expected payoff of $\Pi_i(a_i, \boldsymbol{a}_{-i}, \boldsymbol{x}; \boldsymbol{\theta})$ for player i from choosing action a_i at state $\boldsymbol{x}$ and given $P_j(a_j|\boldsymbol{x})$,

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

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

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

-

$$e_i^{\boldsymbol{P}}(a_i, \boldsymbol{x}) = \text{Euler's Constant} - \sigma \log\left[P_i(a_i|\boldsymbol{x})\right]$$

System II: Bayes-Nash Equilibrium Conditions

- **Bayes-Nash Equilibrium**.

$$P_i\left(a_i = j|\boldsymbol{x}\right) = \frac{\exp\left[v_i\left(a_i = j|\boldsymbol{x}\right)\right]}{\sum\limits_{k\in\mathcal{A}}\exp\left[v_i\left(a_i = k|\boldsymbol{x}\right)\right]}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{A}, \boldsymbol{x} \in \mathcal{X},$$

- $v_i\left(a_i|\boldsymbol{x}\right)$: choice-specific expected value function

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

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

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

Markov Perfect Equilibrium

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

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

- **Bayes-Nash Equilibrium**.

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

- In compact notation

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

- Set of all Markov Perfect Equilibria

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

Part III

Estimation

Data Generating Process

- $\boldsymbol{\theta}^0$: the true value of structural parameters in the population
- $(\boldsymbol{V}^0, \boldsymbol{P}^0)$: a Markov perfect equilibrium at $\boldsymbol{\theta}^0$
- **Assumption:** If multiple Markov perfect equilibria exist, only one equilibrium is played in the data
- Data: $\boldsymbol{Z} = \left\{\bar{\boldsymbol{a}}^{mt}, \bar{\boldsymbol{x}}^{mt}\right\}_{m\in\mathcal{M},t\in\mathcal{T}}$
  - observations from M independent markets over T periods
  - In each market m and time period t, researchers observe
    - the common-knowledge state variables $\bar{\boldsymbol{x}}^{mt}$
    - players' actions $\bar{\boldsymbol{a}}^{mt} = (\bar{a}_1^{mt}, \ldots, \bar{a}_N^{mt})$

Maximum-Likelihood Estimation

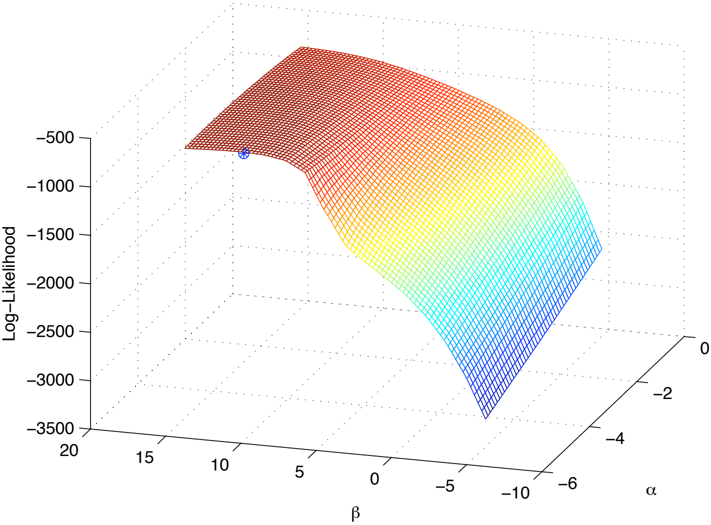
- For a given $\boldsymbol{\theta}$, let $\left(\boldsymbol{P}^{\ell}(\boldsymbol{\theta}), \boldsymbol{V}^{\ell}(\boldsymbol{\theta})\right) \in SOL(\Psi, \boldsymbol{\theta})$ be the ℓ-th equilibrium
- Given data $\boldsymbol{Z} = \left\{\bar{\boldsymbol{a}}^{mt}, \bar{\boldsymbol{x}}^{mt}\right\}_{m\in\mathcal{M},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

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

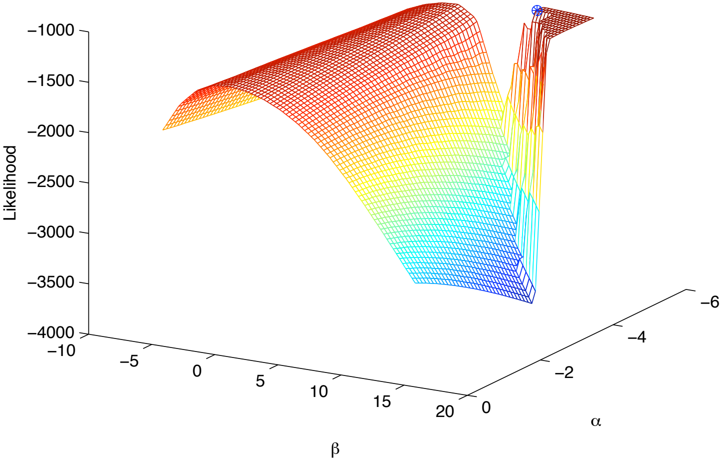
- The ML estimator is

$$\boldsymbol{\theta}^{ML} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\ L(Z, \boldsymbol{\theta}) \tag{1}$$

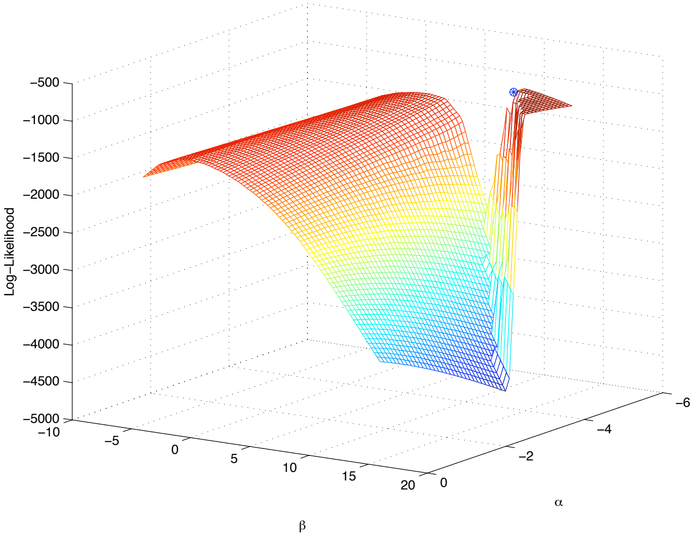
NFXP's Likelihood as a Function of (α, β) – Eq 1



NFXP's Likelihood as a Function of (α, β) – Eq 2



NFXP's Likelihood as a Function of (α, β) – Eq 3



ML Estimation via Constrained Optimization Approach

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

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

- The constrained optimization formulation of the ML estimation problem is

$$\begin{array}{ll} \max\limits_{(\boldsymbol{\theta}, \boldsymbol{P}, \boldsymbol{V})} & \mathcal{L}\left(\boldsymbol{Z}, \boldsymbol{P}\right) \\ \text{subject to} & \boldsymbol{V} = \Psi^{\boldsymbol{V}}\left(\boldsymbol{V}, \boldsymbol{P}, \boldsymbol{\theta}\right) \\ & \boldsymbol{P} = \Psi^{\boldsymbol{P}}\left(\boldsymbol{V}, \boldsymbol{P}, \boldsymbol{\theta}\right) \end{array} \qquad (2)$$

- **Proposition 1.** Problem (1) and (2) have the same solution.

Asymptotic Properties of ML Estimator via Constrained Optimization Approach

- **Theorem.** The constrained maximum likelihood estimator is consistent and asymptotic normal.

  See Appendix.

  Aitchison and Silvey (1958) and Section 10.3 in Gourieroux and Monfort (1995).

Two-Step Methods: Intuition

- Recall the constrained optimization formulation for the ML estimator is

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

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

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

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

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

or, equivalently,

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

Reformulation of the Optimization Problem in Step 2

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

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

Reformulation of the Optimization Problem in Step 2

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

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

- Define $\boldsymbol{V}_i = [V_i(x)]_{\boldsymbol{x} \in \mathcal{X}}$, $\widehat{\boldsymbol{P}}_i(a_i) = [\widehat{P}_i(a_i|\boldsymbol{x})]_{\boldsymbol{x}}$, $\boldsymbol{e}_i^{\widehat{\boldsymbol{P}}}(a_i) = [e_i^{\widehat{\boldsymbol{P}}}(a_i, \boldsymbol{x})]_{\boldsymbol{x}}$, $\boldsymbol{\pi}_i(a_i, \boldsymbol{\theta}) = [\pi_i(a_i|\boldsymbol{x}, \boldsymbol{\theta})]_{\boldsymbol{x}}$, and $\boldsymbol{F}_{\mathcal{X}}^{\widehat{\boldsymbol{P}}} = \left[f_{\mathcal{X}}^{\widehat{\boldsymbol{P}}}(\boldsymbol{x}'|\boldsymbol{x})\right]_{\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}}$

Reformulation of the Optimization Problem in Step 2

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

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

- Define $\boldsymbol{V}_i = [V_i(x)]_{\boldsymbol{x} \in \mathcal{X}}$, $\widehat{\boldsymbol{P}}_i(a_i) = [\widehat{P}_i(a_i|\boldsymbol{x})]_{\boldsymbol{x}}$, $\boldsymbol{e}_i^{\widehat{\boldsymbol{P}}}(a_i) = [e_i^{\widehat{\boldsymbol{P}}}(a_i, \boldsymbol{x})]_{\boldsymbol{x}}$, $\boldsymbol{\pi}_i(a_i, \boldsymbol{\theta}) = [\pi_i(a_i|\boldsymbol{x}, \boldsymbol{\theta})]_{\boldsymbol{x}}$, and $\boldsymbol{F}_{\mathcal{X}}^{\widehat{\boldsymbol{P}}} = \left[f_{\mathcal{X}}^{\widehat{\boldsymbol{P}}}(\boldsymbol{x}'|\boldsymbol{x})\right]_{\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}}$
- The Bellman equation above can be rewritten as

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

  or equivalently

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

  or in a compact notation

$$\boldsymbol{V} = \boldsymbol{\Gamma}(\boldsymbol{\theta}, \widehat{\boldsymbol{P}}).$$

Reformulation of the Optimization Problem in Step 2

- Replacing the constraint $\boldsymbol{V} = \Psi^{\boldsymbol{V}}\left(\boldsymbol{V}, \widehat{\boldsymbol{P}}, \boldsymbol{\theta}\right)$ by $\boldsymbol{V} = \boldsymbol{\Gamma}(\boldsymbol{\theta}, \widehat{\boldsymbol{P}})$ through a simple elimination of variables $\boldsymbol{V}$, the optimization problem in Step 2 becomes

$$\max_{\boldsymbol{\theta}} \mathcal{L}\left(\boldsymbol{Z}, \Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\boldsymbol{\theta}, \widehat{\boldsymbol{P}}), \widehat{\boldsymbol{P}}, \boldsymbol{\theta}\right)\right).$$

- The 2S-PML estimator is defined as

$$\boldsymbol{\theta}^{2S-PML} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \mathcal{L}\left(\boldsymbol{Z}, \Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\boldsymbol{\theta}, \widehat{\boldsymbol{P}}), \widehat{\boldsymbol{P}}, \boldsymbol{\theta}\right)\right).$$

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
- An NPL fixed point $(\tilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{P}})$ satisfies the conditions:

$$\begin{aligned} \tilde{\boldsymbol{\theta}} &= \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\ \mathcal{L}\left(\boldsymbol{Z}, \Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\boldsymbol{\theta}, \widetilde{\boldsymbol{P}}), \widetilde{\boldsymbol{P}}, \boldsymbol{\theta}\right)\right) \\ \widetilde{\boldsymbol{P}} &= \Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\tilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{P}}), \widetilde{\boldsymbol{P}}, \tilde{\boldsymbol{\theta}}\right) \end{aligned} \tag{3}$$

NPL Algorithm

- The NPL algorithm: For $1 \leq K \leq \bar{K}$, iterate over Steps 1 and 2 below until convergence:

Step 1. Given $\widetilde{\boldsymbol{P}}_{K-1}$,
solve $\tilde{\boldsymbol{\theta}}_K = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \, \mathcal{L}\left(\boldsymbol{Z}, \Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\boldsymbol{\theta}, \widetilde{\boldsymbol{P}}_{K-1}), \widetilde{\boldsymbol{P}}_{K-1}, \boldsymbol{\theta}\right)\right).$

Step 2. Given $\tilde{\boldsymbol{\theta}}_K$, update $\widetilde{\boldsymbol{P}}_K$ by
$\widetilde{\boldsymbol{P}}_K = \Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\tilde{\boldsymbol{\theta}}_K, \widetilde{\boldsymbol{P}}_{K-1}), \widetilde{\boldsymbol{P}}_{K-1}, \tilde{\boldsymbol{\theta}}_K\right)$; increase K by 1

- Convergence criterion:

$$\left\|(\tilde{\boldsymbol{\theta}}_K, \widetilde{\boldsymbol{P}}_K) - (\tilde{\boldsymbol{\theta}}_{K-1}, \widetilde{\boldsymbol{P}}_{K-1})\right\| \leq \texttt{tol}_{\text{NPL}}$$

$\texttt{tol}_{\text{NPL}}$: the convergence tolerance, for example, `1.0e-6`

- If the NPL algorithm converges, $(\tilde{\boldsymbol{\theta}}_K, \widetilde{\boldsymbol{P}}_{K-1})$ approximately satisfies the NPL fixed-point conditions (3):

$$\|\widetilde{\boldsymbol{P}}_{K-1} - \Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\tilde{\boldsymbol{\theta}}_K, \widetilde{\boldsymbol{P}}_{K-1}), \widetilde{\boldsymbol{P}}_{K-1}, \tilde{\boldsymbol{\theta}}_K\right)\| \leq \texttt{tol}_{\text{NPL}}$$

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; Pesendorfer and Schmidt-Dengler (2010)
- Kasahara and Shimotsu (2012) propose the NPL-Λ algorithm that modifies Step 2 of the NPL algorithm to compute the NPL estimator

$$\widetilde{\boldsymbol{P}}_K = \left(\Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\tilde{\boldsymbol{\theta}}_K, \widetilde{\boldsymbol{P}}_{K-1}), \widetilde{\boldsymbol{P}}_{K-1}, \tilde{\boldsymbol{\theta}}_K\right)\right)^{\lambda} \left(\widetilde{\boldsymbol{P}}_{K-1}\right)^{1-\lambda}$$

  where λ is chosen to be between 0 and 1
  - $\lambda = 0$: two-step PML estimator
  - $\lambda = 1$: NPL algorithm
- The proper value for λ depends on the true parameter values $\boldsymbol{\theta}^0$
- Alternatively, Kasahara and Shimotsu suggest using a small number for the spectral radius

Convergence Criteria for the NPL-Λ Algorithm

- The NPL-Λ algorithm: For $1 \leq K \leq \bar{K}$, iterate over Steps 1 and 2 below until convergence:

Step 1. Given $\widetilde{\boldsymbol{P}}_{K-1}$,
solve $\tilde{\boldsymbol{\theta}}_K = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \, \mathcal{L}\left(\boldsymbol{Z}, \Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\boldsymbol{\theta}, \widetilde{\boldsymbol{P}}_{K-1}), \widetilde{\boldsymbol{P}}_{K-1}, \boldsymbol{\theta}\right)\right)$.

Step 2. Given $\tilde{\boldsymbol{\theta}}_K$, update $\widetilde{\boldsymbol{P}}_K$ by

$$\widetilde{\boldsymbol{P}}_K = \left(\Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\tilde{\boldsymbol{\theta}}_K, \widetilde{\boldsymbol{P}}_{K-1}), \widetilde{\boldsymbol{P}}_{K-1}, \tilde{\boldsymbol{\theta}}_K\right)\right)^{\lambda} \left(\widetilde{\boldsymbol{P}}_{K-1}\right)^{1-\lambda};$$

increase K by 1

- Convergence criterion used in Kasahara and Shimotsu (2012):

$$\left\|(\tilde{\boldsymbol{\theta}}_K, \widetilde{\boldsymbol{P}}_K) - (\tilde{\boldsymbol{\theta}}_{K-1}, \widetilde{\boldsymbol{P}}_{K-1})\right\| \leq \mathtt{tol}_{\mathrm{NPL}}$$

- If the NPL-Λ algorithm converges, does $(\tilde{\boldsymbol{\theta}}_K, \widetilde{\boldsymbol{P}}_{K-1})$ approximately satisfy the NPL fixed-point conditions (3)?

$$\|\widetilde{\boldsymbol{P}}_{K-1} - \Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\tilde{\boldsymbol{\theta}}_K, \widetilde{\boldsymbol{P}}_{K-1}), \widetilde{\boldsymbol{P}}_{K-1}, \tilde{\boldsymbol{\theta}}_K\right)\| \leq \mathtt{tol}_{\mathrm{NPL}}??$$

Convergence Criteria for the NPL-Λ Algorithm

- Using the previous convergence criterion, if the NPL-Λ algorithm converges,

$$\|\widetilde{\boldsymbol{P}}_{K-1} - \Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\tilde{\boldsymbol{\theta}}_K, \widetilde{\boldsymbol{P}}_{K-1}), \widetilde{\boldsymbol{P}}_{K-1}, \tilde{\boldsymbol{\theta}}_K\right)\| \leq \frac{\texttt{tol}_{\text{NPL}}}{\lambda}$$

  - If one uses a very small value for λ, e.g., $\lambda = \texttt{1.0e-5}$, and $\texttt{tol}_{\text{NPL}} = \texttt{1.0e-6}$, then $\frac{\texttt{tol}_{\text{NPL}}}{\lambda} = 0.1$

- Appropriate convergence criterion:

$$\left\| \begin{array}{c} (\tilde{\boldsymbol{\theta}}_K, \widetilde{\boldsymbol{P}}_K) - (\tilde{\boldsymbol{\theta}}_{K-1}, \widetilde{\boldsymbol{P}}_{K-1}) \\ \widetilde{\boldsymbol{P}}_{K-1} - \Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\tilde{\boldsymbol{\theta}}_K, \widetilde{\boldsymbol{P}}_{K-1}), \widetilde{\boldsymbol{P}}_{K-1}, \tilde{\boldsymbol{\theta}}_K\right) \end{array} \right\| \leq \texttt{tol}_{\text{NPL}}.$$

Part IV

Monte Carlo

Monte Carlo

Experiment Design

- Three experiment specifications with two cases in each experiment
- Experiment 1: Kasahara and Shimotsu (2012) example
- Experiment 2: Aguirregabiria and Mira (2007) example
- Experiment 3: Examples with increasing $|\mathcal{S}|$, the number of market size values
- Market size transition matrix is

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

Experiment 2: Aguirregabiria and Mira (2007) Example

- $N = 5$ players
- $\mathcal{S} = \{1, 2, \ldots, 5\}$
- Total number of grid points in the state space: $|\mathcal{X}| = |\mathcal{S}| \times |\mathcal{A}|^N = 5 \times 2^5 = 160$
- 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

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

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

Experiment 2: Cases 3 and 4

- True values of structural parameters $\boldsymbol{\theta}_{FC}^0 = (1.9, 1.8, 1.7, 1.6, 1.5)$ and $\theta_{EC}^0 = 1$
- Consider two sets of true parameter values for θ_{RS} and θ_{RN}

$$\text{Case 3:} \quad (\theta_{RN}^0, \theta_{RS}^0) = (2, 1);$$
$$\text{Case 4:} \quad (\theta_{RN}^0, \theta_{RS}^0) = (4, 2).$$

- Case 3 is Experiment 3 in Aguirregabiria and Mira (2007)
- The ML estimator solves the constrained optimization problem with 2,400 constraints and 2,408 variables.

Experiment 3: Cases 5 and 6

- Consider two sets of market size values:

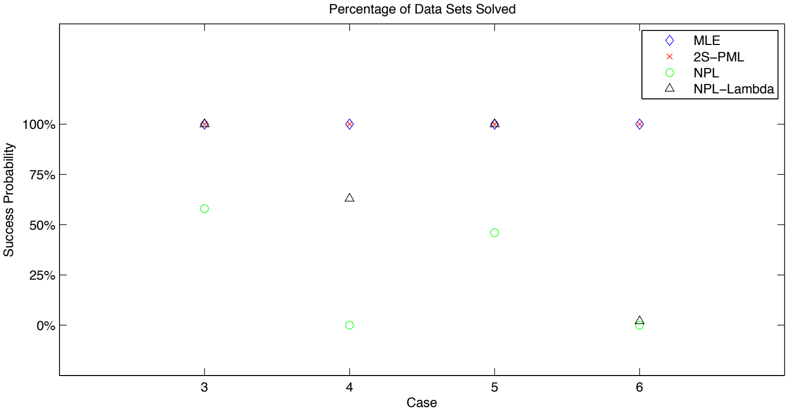
$$\text{Case 5:} \quad |\mathcal{S}| = 10 \text{ with } \mathcal{S} = \{1, 2, \ldots, 10\};$$
$$\text{Case 6:} \quad |\mathcal{S}| = 15 \text{ with } \mathcal{S} = \{1, 2, \ldots, 15\}.$$

- All other specifications remain the same as those in Case 3 in Experiment 2
- Case 5: The ML estimator solves the constrained optimization problem with 4,800 constraints and 4,808 variables.
- Case 6: The ML estimator solves the constrained optimization problem with 7,200 constraints and 7,208 variables.

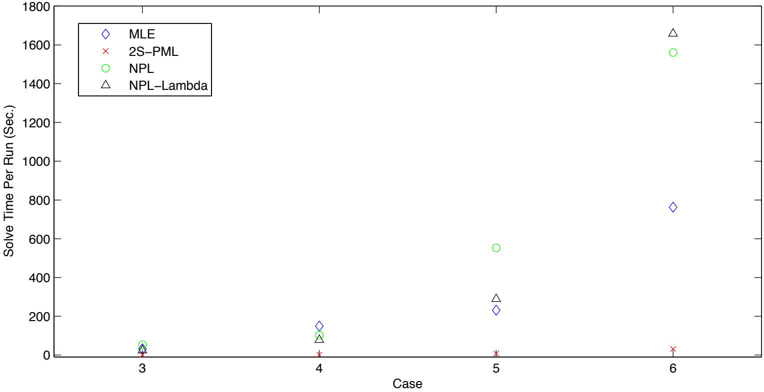
Data Simulation and Algorithm Implementation

- Data simulation: MATLAB
- Optimization: AMPL (programming language) / KNITRO (NLP solver), providing first-order / second-order analytical derivatives
- In each data set: $M = 400$ and $T = 10$
- For Case 3 and 4 in Experiments 2
  - Construct 100 data sets for each case
  - 10 starting points for each data set
- For Cases 5 and 6 in Experiments 3
  - Construct 50 data sets for each case
  - 5 start points for each data sets
- For NPL and NPL-Λ: $\bar{K} = 100$
- For the NPL-Λ 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 2

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 3

$\lvert S \rvert$	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)

Implementation Improvements and Robustness Checks

- ML estimator
  Can we improve the performance (reduce computational time) of the constrained optimization approach for the ML estimator?
  - Use 2S-PML estimates as starting values for the constrained optimization problem for the ML estimator
- NPL-Λ algorithm
  Can we improve the convergence results of the NPL-Λ algorithm by using different values for λ or $\bar{K}$?
  - Use $\lambda \in \{0.1, 0.3, 0, 5, 0.7, 0.9\}$

ML Estimator/Constr. Opt. using 2S-PML Estimates as Starting Values for Cases 3 and 4

Case 3

	$\theta_{FC,1}$	$\theta_{FC,2}$	$\theta_{FC,3}$	$\theta_{FC,4}$	$\theta_{FC,5}$	θ_{EC}	θ_{RN}	θ_{RS}	Data Sets	CPU Time
	1.9	1.8	1.7	1.6	1.5	1	2	1		
T	Estimates								Conv.	(sec.)
1	1.949 (0.254)	1.849 (0.236)	1.764 (0.241)	1.651 (0.247)	1.563 (0.250)	0.983 (0.150)	2.257 (1.086)	1.086 (0.310)	99	42.35
10	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)	100	25.05
20	1.903 (0.056)	1.801 (0.050)	1.701 (0.050)	1.600 (0.049)	1.502 (0.050)	0.996 (0.028)	2.020 (0.241)	1.005 (0.067)	100	23.61

Case 4

	$\theta_{FC,1}$	$\theta_{FC,2}$	$\theta_{FC,3}$	$\theta_{FC,4}$	$\theta_{FC,5}$	θ_{EC}	θ_{RN}	θ_{RS}	Data Sets	CPU Time
	1.9	1.8	1.7	1.6	1.5	1	4	2		
T	Estimates								Conv.	(sec.)
1	1.947 (0.310)	1.845 (0.291)	1.741 (0.282)	1.632 (0.287)	1.538 (0.316)	1.006 (0.181)	3.989 (0.906)	2.011 (0.343)	100	42.19
10	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)	100	29.19
20	1.908 (0.057)	1.806 (0.056)	1.707 (0.053)	1.607 (0.055)	1.514 (0.059)	0.991 (0.031)	4.046 (0.137)	2.017 (0.054)	100	27.43

ML Estimator/Constr. Opt. using 2S-PML Estimates as Starting Values for Cases 5 and 6

	$\theta_{FC,1}$	$\theta_{FC,2}$	$\theta_{FC,3}$	$\theta_{FC,4}$	$\theta_{FC,5}$	θ_{EC}	θ_{RN}	θ_{RS}	Data Sets Conv.	CPU Time (sec.)
Truth	1.9	1.8	1.7	1.6	1.5	1	2	1		
$\lvert S\rvert$	Estimates									
10	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)	50	91.41
15	1.899 (0.098)	1.803 (0.106)	1.697 (0.085)	1.600 (0.093)	1.494 (0.090)	0.983 (0.059)	2.034 (0.304)	1.010 (0.067)	49	449.06

NPL-Λ Algorithm using Different λ Values for Case 4

		$\theta_{FC,1}$	$\theta_{FC,2}$	$\theta_{FC,3}$	$\theta_{FC,4}$	$\theta_{FC,5}$	θ_{EC}	θ_{RN}	θ_{RS}	Data Sets Conv.	CPU Time (sec.)
		1.9	1.8	1.7	1.6	1.5	1	4	2		
T	λ	Estimates									
1	0.9	2.009 (0.266)	1.869 (0.282)	1.743 (0.285)	1.571 (0.311)	1.339 (0.275)	1.301 (0.119)	2.234 (0.222)	1.414 (0.107)	8	78.38
1	0.7	1.970 (0.238)	1.873 (0.241)	1.741 (0.210)	1.612 (0.201)	1.460 (0.170)	1.111 (0.129)	3.349 (0.584)	1.790 (0.185)	54	61.89
1	0.3	2.006 (0.277)	1.916 (0.298)	1.797 (0.279)	1.619 (0.287)	1.409 (0.265)	1.167 (0.151)	2.819 (0.507)	1.621 (0.192)	25	84.27
1	0.1	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)	0	87.83
10	0.9	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)	0	88.53
10	0.7	1.879 (0.081)	1.782 (0.081)	1.678 (0.077)	1.571 (0.073)	1.454 (0.076)	1.016 (0.047)	3.876 (0.216)	1.949 (0.083)	33	76.30
10	0.3	1.873 (0.110)	1.786 (0.098)	1.683 (0.107)	1.560 (0.102)	1.407 (0.102)	1.058 (0.049)	3.581 (0.181)	1.845 (0.085)	11	83.84
10	0.1	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)	0	88.03
20	0.9	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)	0	92.59
20	0.7	1.896 (0.084)	1.787 (0.084)	1.697 (0.082)	1.591 (0.085)	1.485 (0.095)	1.016 (0.045)	3.935 (0.216)	1.972 (0.086)	22	84.54
20	0.3	1.932 (0.068)	1.834 (0.066)	1.731 (0.068)	1.623 (0.065)	1.513 (0.069)	1.016 (0.026)	3.884 (0.133)	1.969 (0.053)	15	85.49
20	0.1	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)	0	92.67

NPL-Λ Algorithm using Different λ Values for Case 6

		$\theta_{FC,1}$	$\theta_{FC,2}$	$\theta_{FC,3}$	$\theta_{FC,4}$	$\theta_{FC,5}$	θ_{EC}	θ_{RN}	θ_{RS}	Data Sets Conv.	CPU Time (sec.)
		1.9	1.8	1.7	1.6	1.5	1	4	2		
T	λ	Estimates									
10	0.9	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)	0	1706.26
10	0.7	1.922 (N/A)	1.821 (N/A)	1.671 (N/A)	1.611 (N/A)	1.531 (N/A)	1.012 (N/A)	1.992 (N/A)	1.007 (N/A)	1	1679.52
10	0.3	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)	0	1766.75
10	0.1	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)	0	1764.13

Final Comment

- Lyapunov-Stable Equilibria
  - Aguirregabiria and Nevo (2012) have argued that with multiple equilibria, it is reasonable to assume that only Lyapunov-stable (or best-response stable) equilibria will be played in the data, in which case the NPL algorithm should converge
  - Lyapunov-stable (or best-response stable) equilibria:

$$\rho\left(\nabla_{\boldsymbol{P}}\Psi^{\boldsymbol{P}}\left(\boldsymbol{\Gamma}(\boldsymbol{\theta}^0, \boldsymbol{P}^0), \boldsymbol{P}^0, \boldsymbol{\theta}^0\right)\right) < 1$$

  - The spectral radius of the mapping above depends not only on $\boldsymbol{\theta}^0$ but also on the grid of the market size values, market size transition, etc
- Ongoing work:
  - Robustness check for NPL-Λ algorithm with different choices of λ value
  - Performance of other two-step estimators?
  - Improving the performance of the constrained optimization approach on dynamic games with higher-dimensional state space?