Solving and Estimating Games of Incomplete Information

Zurich Initiative on Computational Economics (ZICE) 2017

Fedor Iskhakov, Australian National University
John Rust, Georgetown University
Bertel Schjerning, University of Copenhagen

Wednesday, February 1st, 2017

Road Map for Lecture 9

PART I: Structural Estimation of Static Games of Incomplete Information

- Methods: NFXP, MPEC, CCP and NPL
- Example: Simple static entry model

PART II: Structural Estimation of Dynamic Games of Incomplete Information

- Methods: NFXP, MPEC, CCP and NPL
- Example: Dynamic exit/entry model

PART I (a)

Structural Estimation of Static Games of Incomplete Information

Bertel Schjenring

Estimating Discrete-Choice Games of Incomplete Information

Estimating Discrete-Choice Games of Incomplete Information

- Aguirregabiria and Mira (2007): NPL (Recursive 2-Step)
- Bajari, Benkard and Levin (2007): 2-Step Minimum Distance Estimator
- Pakes, Ostrovsky and Berry (2007): Various 2-Step (PML, MoM, min χ^2)
- Pesendorfer and Schmidt-Dengler (2008): 2-Step Least Squares
- Pesendorfer and Schmidt-Dengler (2010): comments on AM (2007)
- Kasahara and Shimotsu (2012): Modified NPL
- Su (2013), Egesdal, Lai and Su (2015): Constrained Optimization

Example: Static Game Entry of Incomplete Information

- Two firms: a and b
- Actions: each firm has two possible actions:

$$d_a = \begin{cases} 1, & \text{if firm } a \text{ choose to enter the market} \\ 0, & \text{if firm } a \text{ choose not to enter the market} \end{cases} \tag{1}$$

$$d_b = \begin{cases} 1, & \text{if firm } b \text{ choose to enter the market} \\ 0, & \text{if firm } b \text{ choose not to enter the market} \end{cases} \tag{2}$$

Example: Static Entry Game of Incomplete Information

Utility: Ex-post payoff to firms

$$u_a(d_a, d_b, x_a, \epsilon_a) = \begin{cases} [\alpha + d_b * (\beta - \alpha)]x_a + \epsilon_{a1}, & \text{if } d_a = 1 \\ 0 + \epsilon_{a0}, & \text{if } d_a = 0 \end{cases}$$

$$u_b(d_a, d_b, x_a, \epsilon_b) = \begin{cases} [\alpha + d_a * (\beta - \alpha)]x_b + \epsilon_{b1}, & \text{if } d_b = 1 \\ 0 + \epsilon_{b0}, & \text{if } d_b = 0 \end{cases}$$

- (α, β): structural parameters to be estimated
- (x_a, x_b): firms' observed types; **common knowledge**
- $\epsilon_a = (\epsilon_{a0}, \epsilon_{a1}), \epsilon_b = (\epsilon_{b0}, \epsilon_{b1})$: firms' unobserved types, **private information**
- (ϵ_a, ϵ_b) are observed only by each firm, but not by their opponent firm nor by the econometrician

Example: Static Entry Game of Incomplete Information

- Assume the error terms (ϵ_a, ϵ_b) have a standardized type III extreme value distribution
- A Bayesian Nash equilibrium (p_a, p_b) satisfies

$$
\begin{aligned}
p_a &= \frac{\exp[p_b \beta x_a + (1 - p_b)\alpha x_a]}{1 + \exp[p_b \beta x_a + (1 - p_b)\alpha x_a]} \\
&= \frac{1}{1 + \exp[-\alpha x_a + p_b x_a(\alpha - \beta)]} \\
&\equiv \Psi_a(p_b, x_a; \alpha, \beta)
\end{aligned}
$$

$$
\begin{aligned}
p_b &= \frac{1}{1 + \exp[-\alpha x_b + p_a x_b(\alpha - \beta)]} \\
&\equiv \Psi_b(p_a, x_b; \alpha, \beta)
\end{aligned}
$$

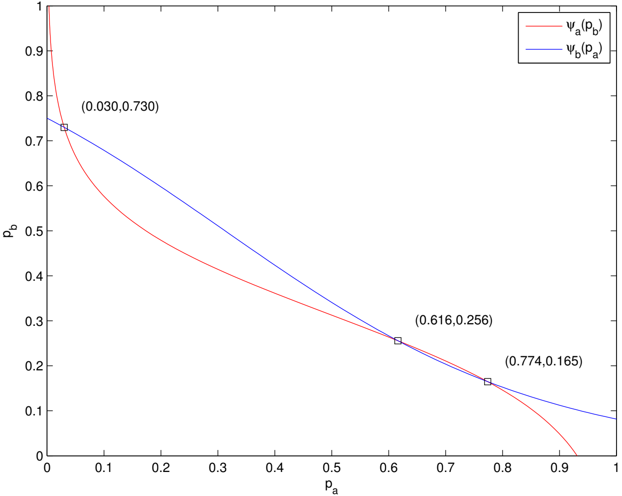
Static Game Example: Parameters

We consider a very contestable game throughout

- Monopoly profits: $\alpha * x_j = 5 * x_j$
- Duopoly profits: $\beta * x_j = -11 * x_j$
- Firm types: $(x_a, x_b) = (0.52, 0.22)$

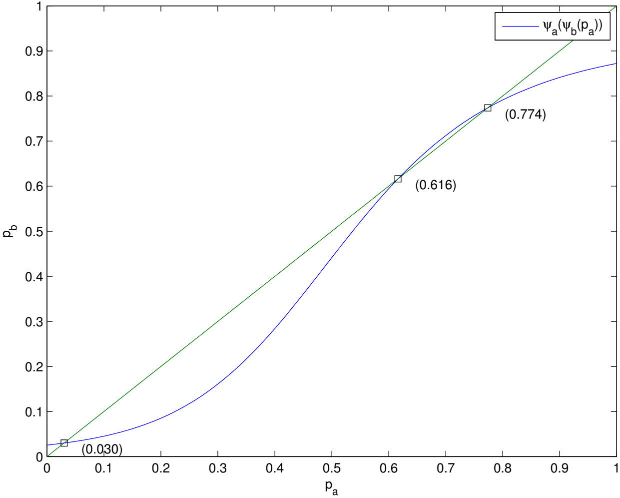
Static Game Example: Three Bayesian Nash Equilibria

Figure: Equilibria at intersections of best response functions



Static Game Example: Solving for Equilibria

Figure: Fixed points on second order best response function



Static Game Example: Solving for Equilibria

Solution method: Combination of succesive approximations and bisection algorithm

Succesive approximations (SA)

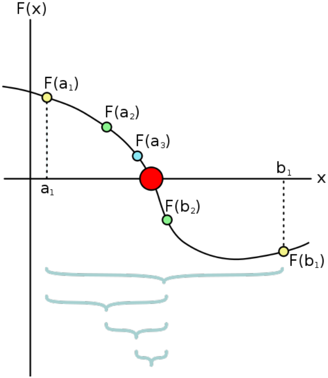
- Converge to nearest stable equilibrium.
- Start SA at $p_a = 0$ and $p_a = 1$.
- Unique equilibrium (K=1): SA will converge to it from anywhere.
- Three equilibria (K=3): Two will be stable, and one will be unstable.
- More equilibria (K>3): Not in this model.

Bisection method

- Use this to find the unstable equilibrium (if K=3).
- The bisection method that repeatedly bisects an interval and then selects a subinterval in which the fixed point (or root) must lie.
- The two stable equilibria, defines the initial interval to search over.
- The bisection method is a very simple and robust method, but it is also relatively slow.

Static Game Example: Solving for Equilibria

Figure: Bisection method



A few steps of the bisection method applied over the starting range [a1;b1]. The bigger red dot is the root of the function.

Static Game Example: Data Generation and Identification

- Data Generating Process (DGP): the data are generated by a single equilibrium
- The two players use the same equilibrium to play 1000 times
- Data: $X = \{(d_a^i, d_b^i)_{i=1}^{1000}, (x_a, x_b) = (0.52, 0.22)\}$
- Given data X, we want to recover structural parameters α and β

Static Game Example: Maximum Likelihood Estimation

- Maximize the likelihood function

$$
\begin{aligned}
\max_{\alpha,\beta} \quad \log \quad & \mathcal{L}(p_a(\alpha,\beta); X) \\
= \quad & \sum_{i=1}^{N} \left(d_a^i * \log(p_a(\alpha,\beta)) + (1 - d_a^i *) \log(1 - p_a(\alpha,\beta)) \right) \\
+ \quad & \sum_{i=1}^{N} \left(d_b^i * \log(p_b(\alpha,\beta)) + (1 - d_b^i *) \log(1 - p_b(\alpha,\beta)) \right)
\end{aligned}
$$

- $p_a(\alpha,\beta)$ and $p_a(\alpha,\beta)$ are the solutions of the Bayesian-Nash Equilibrium equations

$$
\begin{aligned}
p_a &= \frac{1}{1+\exp[-\alpha x_a + p_b x_a(\alpha-\beta)]} \equiv \Psi_a(p_b, x_a; \alpha, \beta) \\
p_b &= \frac{1}{1+\exp[-\alpha x_b + p_a x_b(\alpha-\beta)]} \equiv \Psi_b(p_a, x_b; \alpha, \beta)
\end{aligned}
$$

Static Game Example: MLE via NFXP

- Outer Loop
  - Choose (α, β) to maximize the likelihood function

$$\log \mathcal{L}(p_a(\alpha, \beta), p_b(\alpha, \beta); X)$$

- Inner loop:
  - For a given (α, β), solve the BNE equations for **ALL** equilibria: $(p_a^k(\alpha, \beta), p_b^k(\alpha, \beta)), k = 1, ..., K$
  - Choose the equilibrium that gives the highest likelihood value:

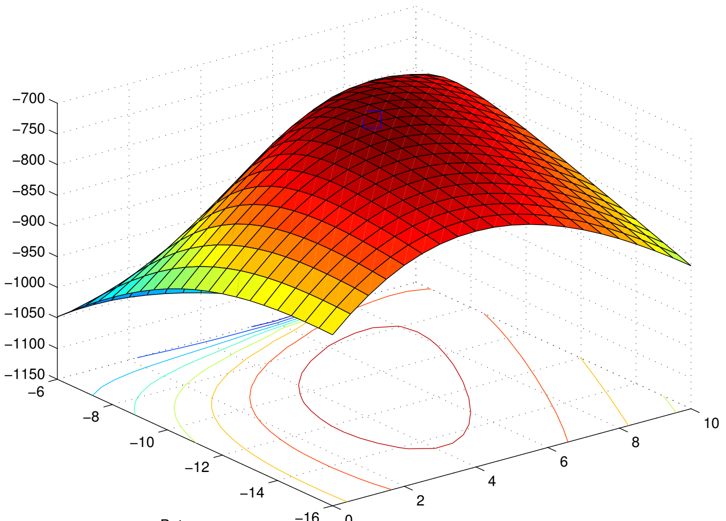
$$k^* = \arg \max_{k=1,\dots,K} \log \mathcal{L}(p_a^k(\alpha, \beta), p_b^k(\alpha, \beta); X)$$

  such that

$$(p_a(\alpha, \beta), p_b(\alpha, \beta)) = (p_a^{k*}(\alpha, \beta), p_b^{k*}(\alpha, \beta))$$

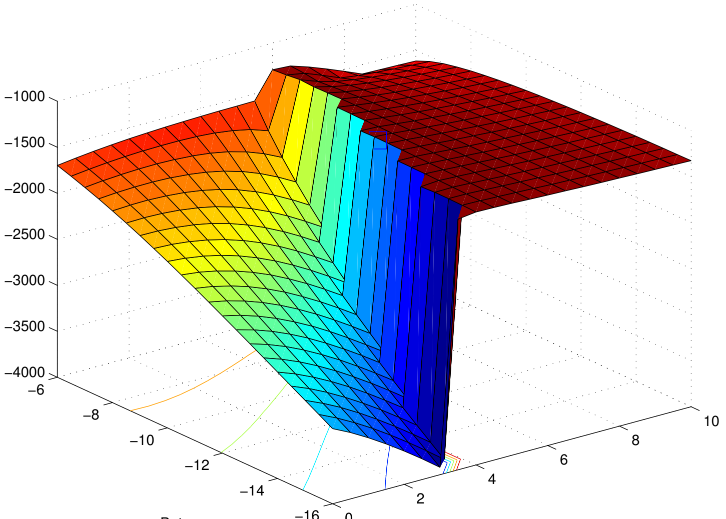
NFXP's Likelihood as a Function of (α, β) - Eq 1

Figure: Data generated from equilibrium 1



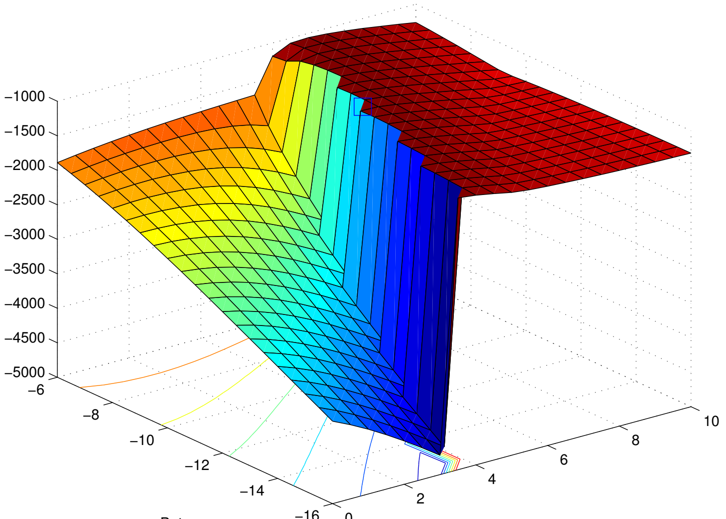
NFXP's Likelihood as a Function of (α, β) - Eq 2

Figure: Data generated from equilibrium 2



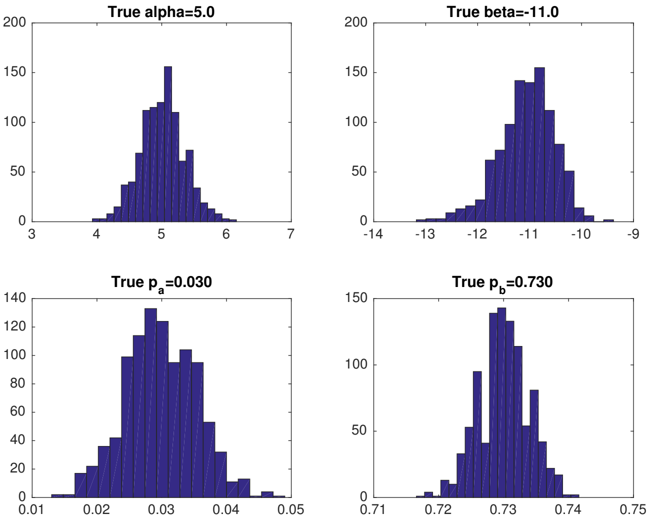
NFXP's Likelihood as a Function of (α, β) - Eq 3

Figure: Data generated from equilibrium 3



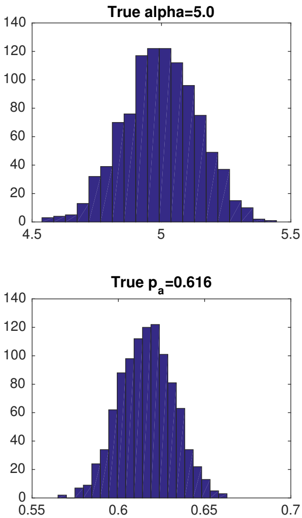
Monte Carlo Results: NFXP with Eq1

Figure: Data generated from equilibrium 1



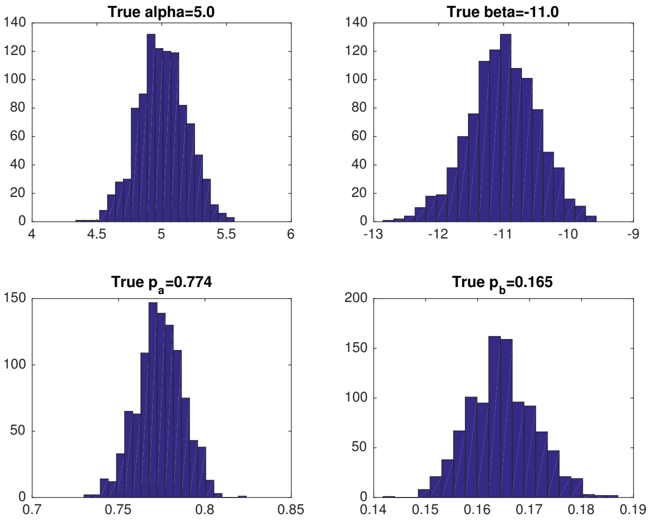
Monte Carlo Results: NFXP with Eq2

Figure: Data generated from equilibrium 2



Monte Carlo Results: NFXP with Eq3

Figure: Data generated from equilibrium 3



Constrained Optimization Formulation for Maximum Likelihood Estimation

- Maximize the likelihood function

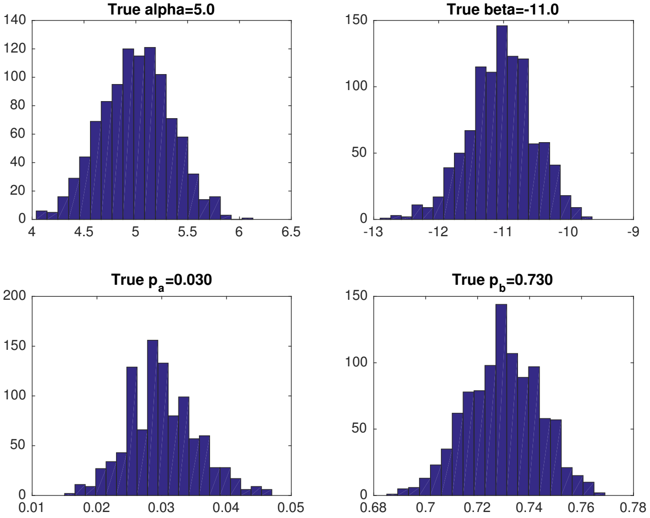
$$\begin{aligned} \max_{\alpha,\beta,p_a,p_b} \quad & \log \quad \mathcal{L}(p_a; X) \\ & = \sum_{i=1}^{N} \left(d_a^i * \log(p_a) + (1 - d_a^i *) \log(1 - p_a) \right) \\ & + \sum_{i=1}^{N} \left(d_b^i * \log(p_b) + (1 - d_b^i *) \log(1 - p_b) \right) \end{aligned}$$

- Subject to p_a and p_a are the solutions of the Bayesian-Nash Equilibrium equations

$$\begin{aligned} p_a &= \frac{1}{1 + \exp[-\alpha x_a + p_b x_a(\alpha - \beta)]} \\ p_b &= \frac{1}{1 + \exp[-\alpha x_b + p_a x_b(\alpha - \beta)]} \\ 0 &\leq p_a, p_b \leq 1 \end{aligned}$$

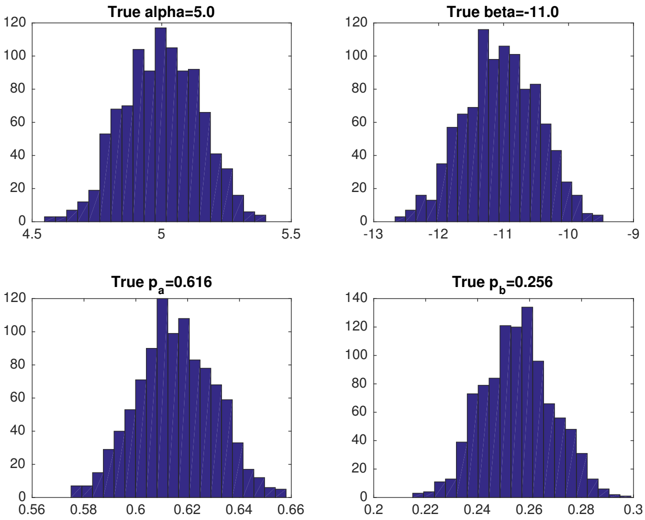
Monte Carlo Results: MPEC with Eq1

Figure: Data generated from equilibrium 1



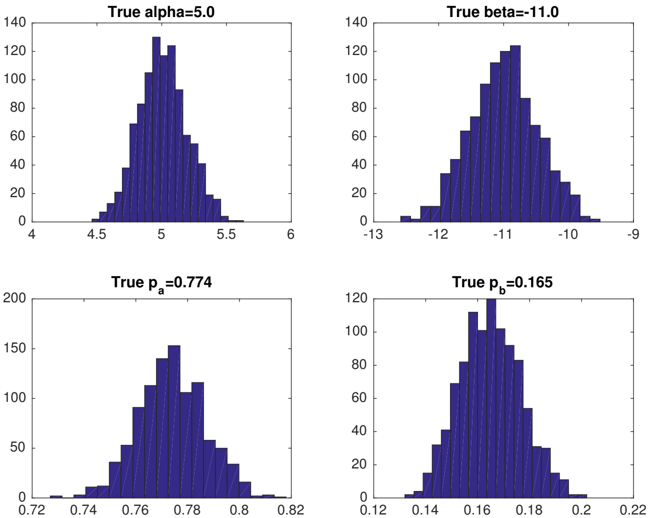
Monte Carlo Results: MPEC with Eq2

Figure: Data generated from equilibrium 2



Monte Carlo Results: MPEC with Eq3

Figure: Data generated from equilibrium 3



Static Game Example: Maximum Likelihood Estimation

- Maximize the likelihood function

$$
\begin{aligned}
\max_{\alpha,\beta} \quad \log \quad & \mathcal{L}(p_a(\alpha,\beta);X) \\
= \quad & \sum_{i=1}^{N} \left(d_a^i * \log(p_a(\alpha,\beta)) + (1-d_a^i *) \log(1-p_a(\alpha,\beta)) \right) \\
+ \quad & \sum_{i=1}^{N} \left(d_b^i * \log(p_b(\alpha,\beta)) + (1-d_b^i *) \log(1-p_b(\alpha,\beta)) \right)
\end{aligned}
$$

- $p_a(\alpha,\beta)$ and $p_a(\alpha,\beta)$ are the solutions of the Bayesian-Nash Equilibrium equations

$$
\begin{aligned}
p_a &= \frac{1}{1+\exp[-\alpha x_a + p_b x_a(\alpha-\beta)]} \equiv \Psi_a(p_b, x_a; \alpha, \beta) \\
p_b &= \frac{1}{1+\exp[-\alpha x_b + p_a x_b(\alpha-\beta)]} \equiv \Psi_b(p_a, x_b; \alpha, \beta)
\end{aligned}
$$

Discussion

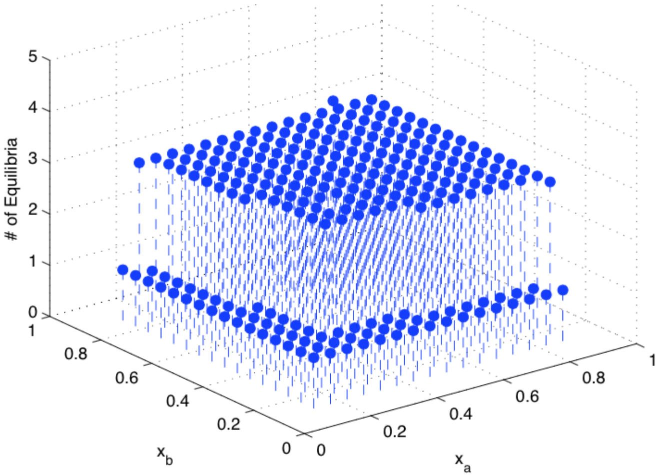
- Q: Is the likelihood function smooth in α and β for NFXP? What about MPEC - is objective function and constraints smooth in parameters, $\theta = (\alpha, \beta, p_a, p_b)$?
- Q: Sensitivity to starting values?
- Q: Can we identify what equilibrium is played in the data, i.e. the equilibrium selection rule?
- Q: Can we use standard theorems for inference? Is true value in interior of parameter space? Is it differentiable? Is objective function continuous?
- Q: This problem is extremely simple. p_a and p_b are scalars. How would you solve for p_b and p_b when they are solutions to players Bellman equations?
- Can we be sure to find all equilibria by iterating on player's Bellman equations? Why/why not?

Estimation with Multiple Markets

- There 25 different markets, i.e., 25 pairs of observed types $(x_a^m, x_b^m), m = 1, ..., 25$
- The grid on x_a has 5 points equally distributed between the interval $[0.12, 0.87]$, and similarly for x_b
- Use the same true parameter values: (α_0, β_0)
- For each market with (x_a^m, x_b^m), solve BNE conditions for (p_a^m, p_b^m).
- There are multiple equilibria in most of 25 markets
- For each market, we (randomly) choose an equilibrium to generate 1000 data points for that market
- The equilibrium used to generate data can be different in different markets - we flip a coin at each market.

of Equilibria with Different (x_a^m, x_b^m)

Figure: Number of equilibria



NFXP - Estimation with Multiple Markets

Inner loop:

$$\max_{\alpha,\beta} \quad \log \mathcal{L}(p_a^m(\alpha,\beta), p_b^m(\alpha,\beta); X)$$

Outer loop: For a given values of (α, β) solve BNE equations for ALL equilibria, $k = 1, .., K$ at each market, $m = 1, .., M$: That is, $p_a^{m,k}(\alpha, \beta)$ and $p_a^{m,k}(\alpha, \beta)$ are the solutions to

$$\begin{aligned} p_a^m &= \Psi_a(p_b^m, x_a^m; \alpha, \beta) \\ p_b^m &= \Psi_b(p_a^m, x_b^m; \alpha, \beta) \\ m &= 1, ..., M \end{aligned}$$

where we again choose the equilibrium, that gives the highest likelihood value at each market m

$$k^* = \arg \max_{k=1,\dots,K} \log \mathcal{L}(p_a^{m,k}(\alpha,\beta), p_b^{m,k}(\alpha,\beta); X)$$

such that

$$(p_a^m(\alpha,\beta), p_b^m(\alpha,\beta)) = (p_a^{m,k*}(\alpha,\beta), p_b^{m,k*}(\alpha,\beta))$$

Estimation with Multiple Markets - MPEC

Constrained optimization formulation

$$\max_{\alpha, \beta, p_a^m, p_b^m} \quad \log \mathcal{L}(p_a^m, p_b^m; X)$$

subject to

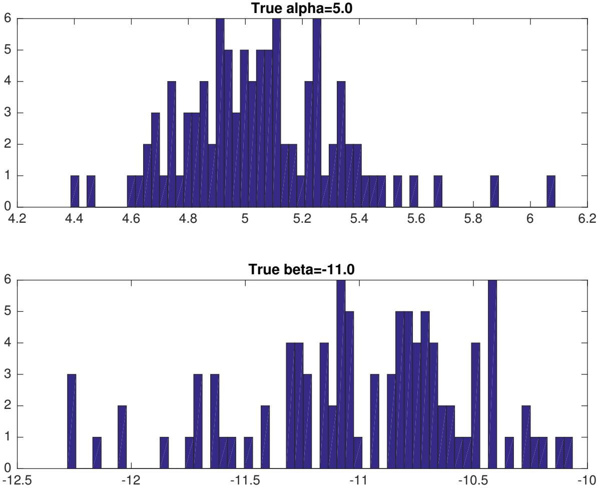
$$\begin{aligned} p_a^m &= \Psi_a(p_b^m, x_a^m; \alpha, \beta) \\ p_b^m &= \Psi_b(p_a^m, x_b^m; \alpha, \beta) \\ 0 &\leq p_a^m, p_b^m \leq 1, m = 1, ..., M \end{aligned}$$

- MPEC does not explicitly solve the BNE equations to find ALL equilibria at each market - for every trial value of parameters.
- But the number of parameters is much larger.
- Both MPEC and NFXP are based on Full Information Maximum Likelihood (FIML) estimators.

NFXP: Monte Carlo - Multiple Markets (M=25, T=50)

Starting values $\alpha_0 = \alpha$, $\beta_0 = \beta$

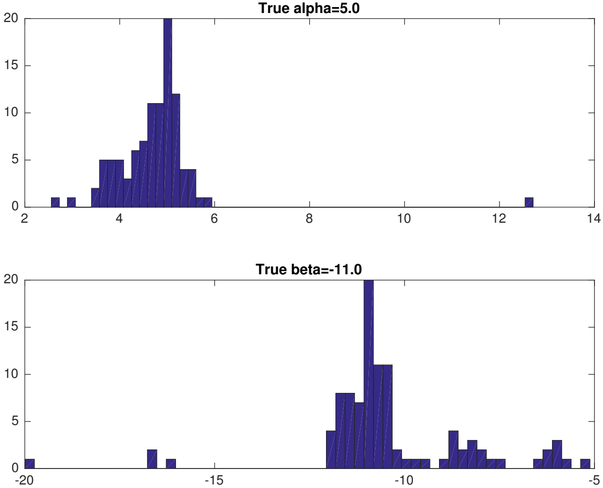
Random equilibrium selection in different markets



MPEC: Monte Carlo - Multiple Markets (M=25, T=50)

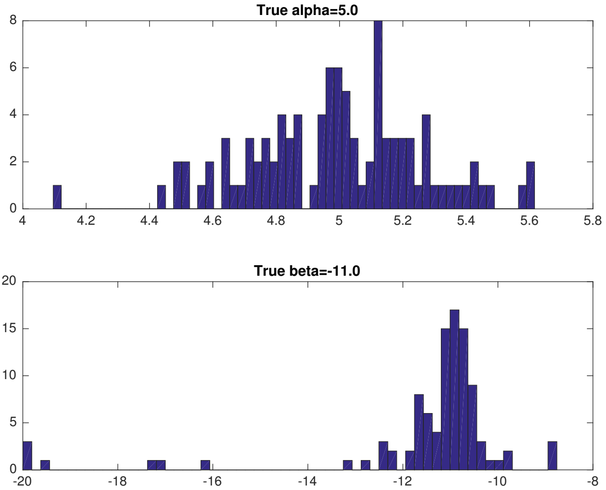
Starting values $\alpha_0 = \alpha$, $\beta_0 = \beta$

Random equilibrium selection in different markets



MPEC: Monte Carlo - Multiple Markets (M=2, T=625)

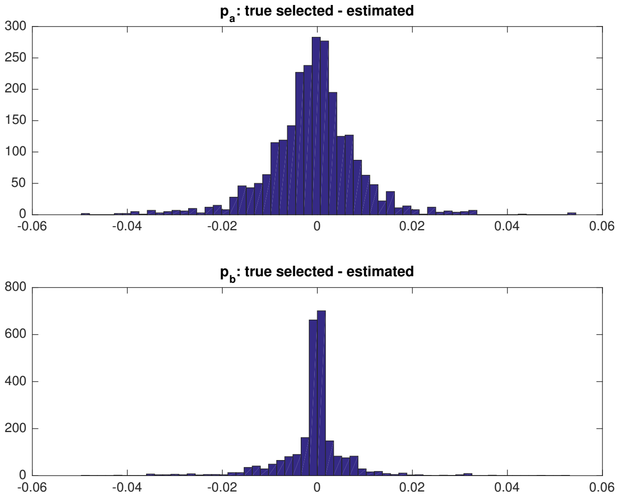
Figure: Random equilibrium selection in different markets



NFXP: Monte Carlo - Multiple Markets (M=25, T=50)

Starting values $\alpha_0 = \alpha$, $\beta_0 = \beta$

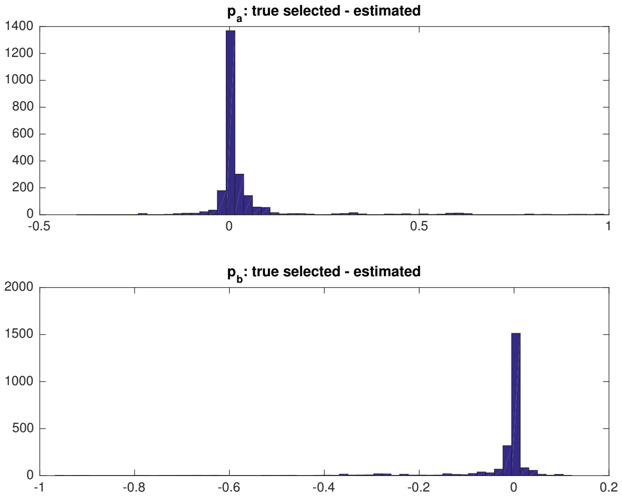
Random equilibrium selection in different markets



MPEC: Monte Carlo - Multiple Markets (M=25, T=50)

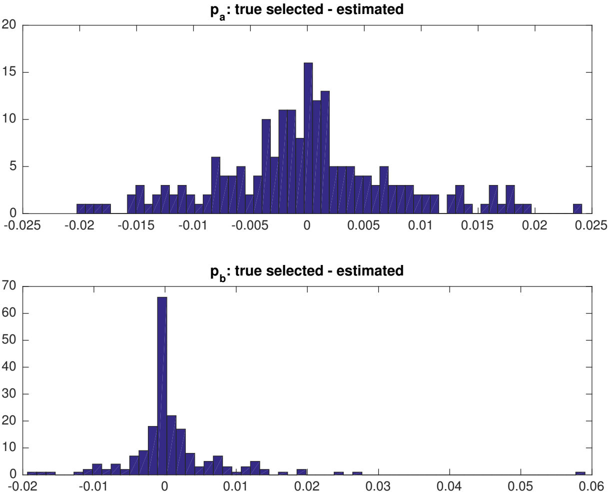
Starting values $\alpha_0 = \alpha$, $\beta_0 = \beta$

Random equilibrium selection in different markets



MPEC: Monte Carlo - Multiple Markets (M=2, T=625)

Figure: Random equilibrium selection in different markets



MPEC and NFXP: multiple markets

NFXP:

- 2 parameters in optimization problem
- we can estimate the equilibrium played in the data, $p_a^{m,k*}$ and $p_b^{m,k*}$ (but in models with observationally equivalent equilibria it may not be possible to obtain joint identification of structural parameters and equilibrium probabilities)
- Needs to find ALL equilibria at each market (very hard in more complex problems)
- Good full solution methods required

MPEC:

- $2 + 2M$ parameters in optimization problem
- Does not always converge towards the equilibrium played in the data, although NFXP indicates that $p_a^{m,k*}$ and $p_b^{m,k*}$ are actually identifiable
- Local minima with many markets.
- Disclaimer: Quick and dirty implementation of MPEC. Use AMPL/Knitro

2-Step Methods

Recall the constrained optimization formulation for FIML is

$$\max_{\alpha,\beta,p_a^m,p_b^m} \quad \log \mathcal{L}(p_a^m, p_b^m; X)$$

subject to

$$\begin{aligned} p_a^m &= \Psi_a(p_b^m, x_a^m; \alpha, \beta) \\ p_b^m &= \Psi_b(p_a^m, x_b^m; \alpha, \beta) \\ 0 &\leq p_a^m, p_b^m \leq 1, m = 1, ..., M \end{aligned}$$

- Denote the solution as $(\alpha^*, \beta^*, p_a^*, p_b^*)$
- Suppose we know (p_a^*, p_b^*), how do we recover (α^*, β^*)?

2-Step Methods: Recovering (α^*, β^*)

- Idea 1: Solve the BNE equations for (α^*, β^*)

$$\begin{aligned} p_a^* &= \Psi_a(p_b^*, x_a; \alpha, \beta) \\ p_b^* &= \Psi_b(p_a^*, x_b; \alpha, \beta) \end{aligned}$$

- Idea 2: Choose (α, β) to

$$\max_{\alpha,\beta} \log \mathcal{L}(\Psi_a(p_b^*, x_a; \alpha, \beta), \Psi_b(p_a^*, x_b; \alpha, \beta); X)$$

2-Step Methods: Recovering (α^*, β^*)

- Idea 1:
  - Step 1: Estimate $\hat{p} = (\hat{p}_a, \hat{p}_b)$ from the data
  - Step 2: Solve

$$\begin{aligned} \hat{p}_a &= \Psi_a(\hat{p}_a, x_a; \alpha, \beta) \\ \hat{p}_b &= \Psi_b(\hat{p}_b, x_b; \alpha, \beta) \end{aligned}$$

- Idea 2
  - Step 1: Estimate $\hat{p} = (\hat{p}_a, \hat{p}_b)$ from the data
  - Step 2: : Choose (α, β) to

$$\max_{\alpha,\beta} \log \mathcal{L}(\Psi_a(\hat{p}_b, x_a; \alpha, \beta), \Psi_b(\hat{p}_a, x_b; \alpha, \beta); X)$$

2-Step Methods: Potential Issues to be Addressed

- How do we estimate $\hat{p} = (\hat{p}_a, \hat{p}_b)$?
- Different methods give different $\hat{p}$
- One method is the frequency estimator:

$$
\begin{aligned}
\hat{p}_a &= \frac{1}{N}\sum_i^N I_{\{d_a^i=1\}} \\
\hat{p}_b &= \frac{1}{N}\sum_i^N I_{\{d_b^i=1\}}
\end{aligned}
$$

- if $(\hat{p}_a, \hat{p}_b) \neq (p_a^*, p_b^*)$ then $(\hat{\alpha}, \hat{\beta}) \neq (\alpha^*, \beta^*)$
- For a given $(\hat{p}_a, \hat{p}_b)$, there might not be a solution to the BNE equations

$$
\begin{aligned}
\hat{p}_a &= \Psi_a(\hat{p}_a, x_a; \alpha, \beta) \\
\hat{p}_b &= \Psi_b(\hat{p}_b, x_b; \alpha, \beta)
\end{aligned}
$$

2-Step Methods: Pseudo Maximum Likelihood

In 2-step methods

- Step 1: Estimate $\hat{p} = (\hat{p}_a, \hat{p}_b)$ from the data
- Step 2: Solve

$$\max_{\alpha,\beta,p_a,p_b} \quad \log \mathcal{L}(p_a, p_b; X)$$

subject to

$$\begin{aligned} p_a &= \Psi_a(\hat{p}_a, x_a; \alpha, \beta) \\ p_b &= \Psi_b(\hat{p}_b, x_b; \alpha, \beta) \\ 0 &\leq p_a^m, p_b^m \leq 1, m = 1, ..., M \end{aligned}$$

Or equivalently

- Step 1: Estimate $\hat{p} = (\hat{p}_a, \hat{p}_b)$ from the data
- Step 2: Solve

$$\max_{\alpha,\beta} \quad \log \mathcal{L}(\Psi_a(\hat{p}_b, x_a; \alpha, \beta), \Psi_b(\hat{p}_a, x_b; \alpha, \beta); X)$$

Least Square Estimators

Pesendofer and Schmidt-Dengler (2008)

- Step 1: Estimate $\hat{p} = (\hat{p}_a, \hat{p}_b)$ from the data
- Step 2: Solve

$$\min_{\alpha,\beta} \left\{ (\hat{p}_a - \Psi_a(\hat{p}_b, x_a; \alpha, \beta))^2 + (\hat{p}_b - \Psi_b(\hat{p}_a, x_b; \alpha, \beta); X))^2 \right\}$$

For dynamic games, Markov perfect equilibrium conditions are characterized by

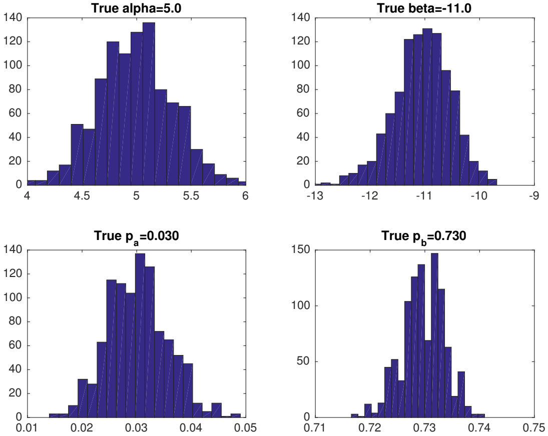
$$p = \Psi(p, \theta)$$

- Step 1: Estimate $\hat{p}$ from the data
- Step 2: Solve

$$\min_{\alpha,\beta} [\hat{p} - \Psi(\hat{p}; \theta)]' W [\hat{p} - \Psi(\hat{p}; \theta)]'$$

Static Game Example: 2-Step PML

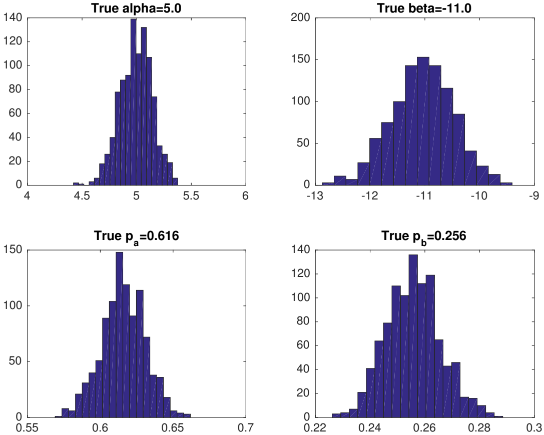
Figure: Data generated from equilibrium 1



- Pakes, Ostrovsky, and Berry (2007): PML 2-step estimator and can lead to large bias in finite samples.
- In this example, however, it seems to work pretty OK. Why?

Static Game Example: 2-Step PML

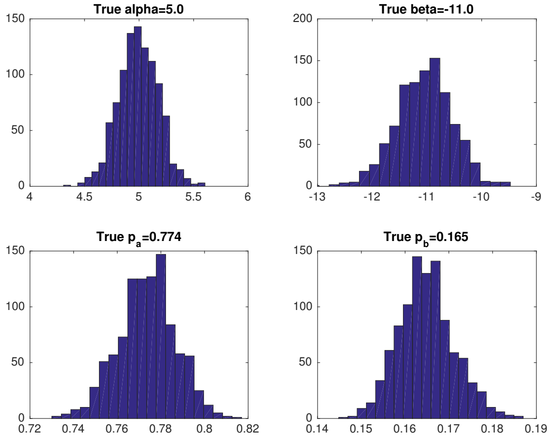
Figure: Data generated from equilibrium 2



- Pakes, Ostrovsky, and Berry (2007): PML 2-step estimator and can lead to large bias in finite samples.
- In this example, however, it seems to work pretty OK. Why?

Static Game Example: 2-Step PML

Figure: Data generated from equilibrium 3



- Pakes, Ostrovsky, and Berry (2007): PML 2-step estimator and can lead to large bias in finite samples.
- In this example, however, it seems to work pretty OK. Why?

Nested Pseudo Likelihood (NPL): Aguirregabiria and Mira (2007)

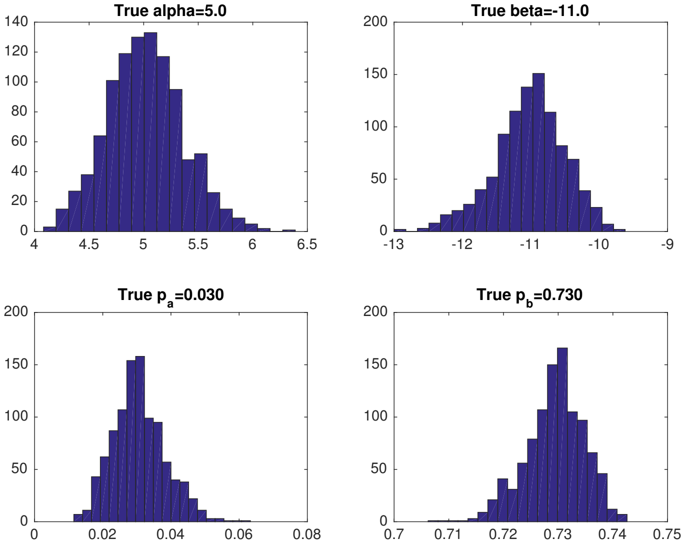
NPL iterates on the 2-step methods

1. Step 1: Estimate $\hat{p}^0 = (\hat{p}_a^0, \hat{p}_b^0)$ from the data, set $k = 0$
2. Step 2:
   REPEAT
   1. Solve
   $$\alpha^{k+1}, \beta^{k+1} = \arg\max_{\alpha,\beta} \quad \log \mathcal{L}(\Psi_a(\hat{p}_b^k, x_a; \alpha, \beta), \Psi_b(\hat{p}_a^k, x_b; \alpha, \beta); X)$$
   2. One best-reply iteration on $\hat{p}^k$
   $$\begin{aligned} \hat{p}_a^{k+1} &= \Psi_a(\hat{p}_a^k, x_a; \alpha^{k+1}, \beta^{k+1}) \\ \hat{p}_a^{k+1} &= \Psi_b(\hat{p}_b^k, x_b; \alpha^{k+1}, \beta^{k+1}) \end{aligned}$$
   3. Let k:=k+1;

   UNTIL convergence in (α^k, β^k) and $(\hat{p}_a^k, \hat{p}_b^k)$

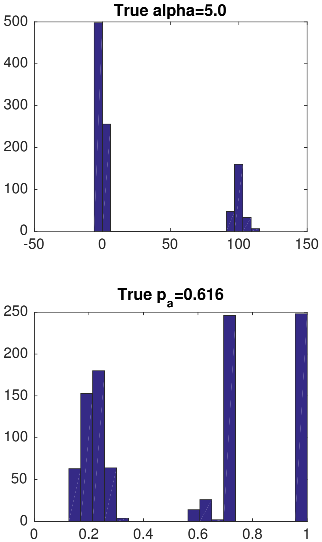
Monte Carlo Results: NPL with Eq 1

Figure: Equilibrium 1 - $\hat{p}_j = 1/N \sum_i I(d_j = 1)$



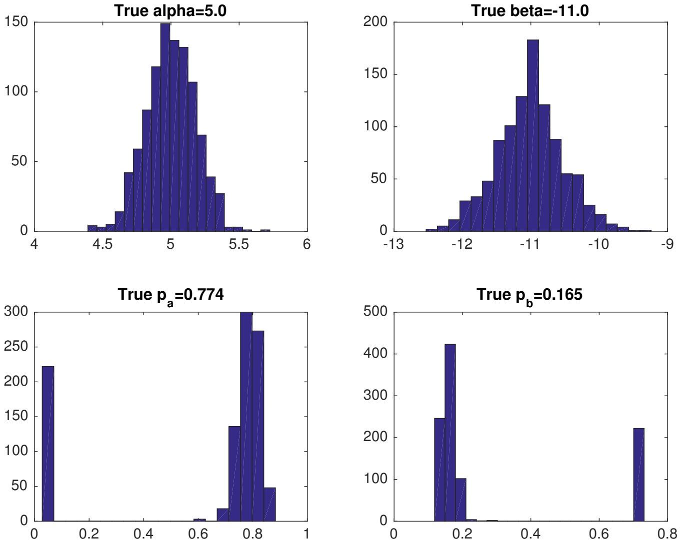
Monte Carlo Results: NPL with Eq 2

Figure: Equilibrium 2 - $\hat{p}_j = 1/N \sum_i I(d_j = 1)$



Monte Carlo Results: NPL with Eq 3

Figure: Equilibrium 3 - $\hat{p}_j = 1/N \sum_i I(d_j = 1)$



Conclusions

- NFXP/MPEC implementations of MLE is statistically efficient, but computational daunting.
- Two step estimators - computationally fast, but inefficient and biased in small samples.
- NPL (Aguirregabiria and Mira 2007) should bridge this gab, but does not seem to be an appropriate method for estimating games with multiple equilibria.
- Estimation of dynamic games is an interesting but challenging computational optimization problem
  - Multiple equilibria leads makes likehood function discontinuous $\rightarrow$ non-standard inference and computational complexity
  - Multiple equilibria leads to indeterminacy problem and identification issues.
- All these problems are amplified by orders of magnitude when we move to Dynamic models

NEXT

Estimation of dynamic games of incomplete information

- Estimation dynamic game with NPL: Agurregabiria and Mira (2012)
- Estimation of dynamic discrete choice games of incomplete information using MPEC - Egesdal, Lai and Su (2015)
- All solution algorithms necessary for NFXP: Development of all solution algorithms for solving games with Multiple Equilibria (Iskhakov et al. 2016)

PART I (b)

Structural Estimation of Dynamic Games of Incomplete Information

Bertel Schjerning

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

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. Empirical results
10. Simulation results

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

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

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

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

- $f_{\mathcal{X}}^{\mathbf{P}}(\mathbf{x}'|\mathbf{x})$: 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

$$\begin{array}{rl} \max\limits_{(\theta,\mathbf{P},\mathbf{V})} & \mathcal{L}(\mathbf{Z},\mathbf{P}) \\ \text{subject to} & \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V},\mathbf{P},\theta) \\ & \mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V},\mathbf{P},\theta) \end{array}$$

- 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
- Are you convinced? 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 $\beta = .095$; the scale parameter of the type I extreme value distribution, $\sigma = 1$
- The common-knowledge component of the per-period payoff

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

- $\theta = (\theta^{RS}, \theta^{RN}, \theta^{FC}, \theta^{EC})$: 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)	0.383	0.133	0.127	0.073	0.062
on log(market size)[a]	(0.043)	(0.019)	(0.024)	(0.020)	(0.018)
Regression log(firm size)	−0.019	0.153	−0.066	0.223	0.097
on log(market size)[b]	(0.034)	(0.082)	(0.050)	(0.081)	(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.

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.
- However, survival is more likely in gas stations than in the other industries.

What explains these facts?

- Economies of scale
  (smaller fixed cost for restaurants?)
- Sunken entry costs
  (smaller for restaurants?)
- Strategic Interactions
  (is product differentiation possible for gas stations?)

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
$\frac{\theta_{FC}}{\theta_{RS}\ln(S_{Med})}$	0.590	0.716	0.852	0.772	0.742
$\frac{\theta_{EC}}{\theta_{RS}\ln(S_{Med})}$	0.357	0.585	0.299	0.311	0.542
$100\frac{\theta_{RN}\ln(2)}{\theta_{RS}\ln(S_{Med})}$	7.1%	10.9%	5.9%	10.1%	11.4%
$\frac{\sigma_\omega^2}{\theta_{RS}^2\operatorname{var}(\ln(S))+\sigma_\omega^2+1}$	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
$\frac{\theta_{\mathrm{FC}}}{\theta_{RS}\ln(S_{\mathrm{Med}})}$	0.590	0.716	0.852	0.772	0.742
$\frac{\theta_{\mathrm{EC}}}{\theta_{RS}\ln(S_{\mathrm{Med}})}$	0.357	0.585	0.299	0.311	0.542
$100\frac{\theta_{RN}\ln(2)}{\theta_{RS}\ln(S_{\mathrm{Med}})}$	7.1%	10.9%	5.9%	10.1%	11.4%
$\frac{\sigma_\omega^2}{\theta_{RS}^2\,\mathrm{var}(\ln(S))+\sigma_\omega^2+1}$	0.278	0.436	0.235	0.423	0.642

- Fixed operating costs are a very important component of total profits.
- Sunken entry costs are statistically significant in the five industries.
- As expected, 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)

- Economies of scale are smaller in the restaurant industry and this is the main factor that explains the large number of restaurants.
- Second, strategic interactions are particularly small among restaurants and among bookstores, which might be due to more product differentiation in those industries.
- This also contributes to explain the large number of restaurants.
- Economies of scale seem very important in the bookstore industry. However, the number of bookstores is, in fact, larger than the number of gas stations or the number of shoe shops. The reason is that negative strategic interactions are weak in this industry.
- Fourth, industry-specific investments, i.e., sunken entry costs, are significant in the five industries. However, these costs are smaller than annual fixed operating costs.
- Gas stations is the industry with largest sunken costs, but the magnitude of these costs does not result in a particularly small number of firms in this industry. However, it does contribute to explain the lower turnover for gas stations.

Experiment 1: Aguirregabiria and Mira (2007) Example

- Market size transition matrix is

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

- Market size: $|\mathcal{S}| = 5$ with $\mathcal{S} = \{1, 2, ..., 5\}$
- True values of structural parameters $\theta_0^{FC} = (1.9, 1.8, 1.7, 1.6, 1.5)$ and $\theta_0^{EC} = 1$
- Consider two sets of true parameter values for θ^{RS} and θ^{RN}
- Two cases:
  $(\theta^{RS}, \theta^{RN}) = (2, 1)$
  $(\theta^{RS}, \theta^{RN}) = (4, 2)$
- The first one is the experiment 3 in Aguirregabiria and Mira (2007)

Experiment 2

- Consider two sets of market size values:

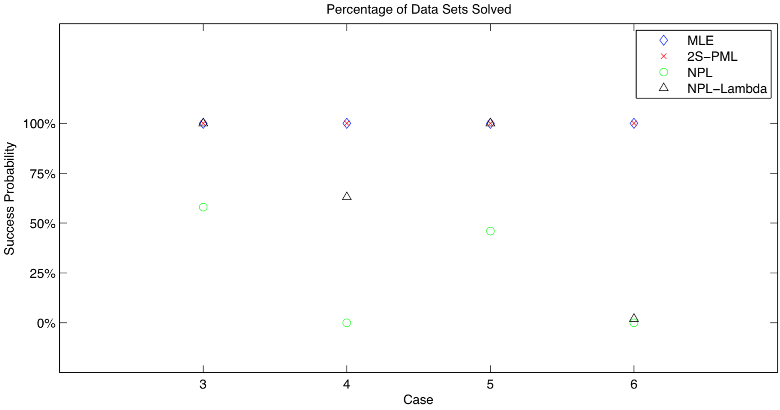
  $|\mathcal{S}| = 10$ with $\mathcal{S} = \{1, 2, ..., 10\}$
  .... i.e. the ML estimator solves the constrained optimization problem with 4,800 constraints and 4,808 variables.
  $|\mathcal{S}| = 15$ with $\mathcal{S} = \{1, 2, ..., 15\}$
  .... ie. the ML estimator solves the constrained optimization problem with 7,200 constraints and 7,208 variables.

- All other specifications remain the same as those in the first in Experiment 1

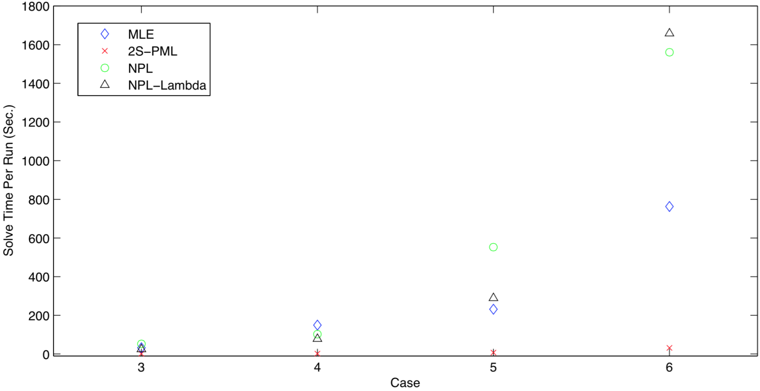
Data Simulation and Algorithm Implementation

- In each dataset: $M = 400$ and $T = 10$
- For Case 3 and 4 in Experiment 1
  - Construct 100 data sets for each case
  - MPEC: 10 starting points for each data set
- For Cases 5 and 6 in Experiments 2
  - Construct 50 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 is 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.