Identification of Games of Incomplete Information with Multiple Equilibria and Common Knowledge Unobserved Heterogeneity

by

VICTOR AGUIRREGABIRIA and PEDRO MIRA

City University London, April 8, 2015

CONTEXT AND MOTIVATION

• Multiplicity of equilibria is a prevalent feature in many economic models [in particular, in static or dynamic discrete games often used in empirical IO].

• From an empirical perspective, there are two (not mutually exclusive) views on models with multiple equilibria:

(A) An imporant aspect needed to explain some features of data/reality.

(B) A nuisance: "Incompleteness" is a problem for inference/prediction.

• An **econometric model with multiple equilibria** is an incomplete model because it does not have a unique reduced form. This raises different issues in:

(a) Identification (b) Estimation (comput. issues) (c) Counterfactuals

• A substantial recent literature has popularized **two-step methods** for a class of games of incomplete information. The first step identifies nonparametrically the equilibrium, or equilibria, played in the data. In the second step, estimates of structural parameters are obtained by maximizing a criterion (e.g. a pseudo likelihood) based on best response or value functions evaluated at the equilibrium estimated in the first step. Though **computationally simple** and useful, two-step best response methods **have some limitations.**

1. Not statistically efficient, specially in small samples.

2a. Rely on consistent estimators of equilibrium strategies (CCP's) in first step. These estimators are simple to obtain only when behavior does not depend on common knowledge variables (payoff-relevant **or not**) unobserved by the econometrician.

2b. In particular, a key maintained assumption in almost every application: There is only one equilibrium type in the data. This rules out the possibility that multiplicity of equilibria in the data may explain features of real economies.

• Another feature of these games is that private information variables are independent across players. This implies that, conditional on common knowledge observable covariates x, the actions of individual players are statistically independent. This key implication of the structure of the game is testable and very likely to be rejected. Rejection has two interpretations: a) Multiple equilibria; b) Payoff-relevant unobservables which are common knowledge.

• THIS PAPER deals with identification in empirical games of incomplete/asymmetric information when there are **three sources of unobservables for the researcher**:

1. Payoff-Relevant variables, players' private information **(PI)**;
2. Payoff-Relevant variables, common knowledge to players **(PRU)**;
3. Multiple equilibria (**MEQ**) with equilibrium selection which is common knowlege to players but stochastic ;

• Previous studies have considered: only [**PI**]; or [**PI**] and [**PRU**]; or [**PI**] and [**MEQ**]; but not all three together.

RELATED LITERATURE: Games of incomplete information

• First papers (Aguirregabiria, 2004; Seim, 2006) include **only PI unobservables**. Multiple equilibria ("in the model") is not an issue for the identification of the model.

• Important limitation: ignoring **PR unobservables** eliminates (by assumption) any concern on endogeneity when estimating strategic interactions.

• Some papers (Grieco, 2012; Aguirregabiria and Mira, 2007; Arcidiacono and Miller, 2011) allow for **PR unobservables.** They still assume that, conditional on PR variables (observable and unobservable), the same equilibrium is selected;

• Other studies (Sweeting, 2008, De Paula and Tang, 2011; Bajari, Hahn, Hong and Ridder IER) allow for **MEQ**, but they assume that there are no **PR unobservables.**

• Otsu, Pesendorfer and Takahashi (2014) propose statistical tests for this assumption.

EXAMPLE (Based on Todd & Wolpin's "Estimating a Coordination Game within the Classroom")

• In a class, students and teacher choose their respective levels of effort. Each student has preferences on her own end-of-the-year knowledge. The teacher cares about the aggregate end-of-the-year knowledge of all the students.

• A production function determines end-of-the-year knowledge of a student: it depends on student's own effort, effort of her peers, teacher's effort, and exogenous characteristics.

• **PR-U unobs:** Class, school, teacher, and student characteristics that are known by the players but not to the researcher.

• **PI unobs:** Some student's and teacher's skills may be private info.

• **MEQ:** Coordination game with multiple equilibria. Classes with the same PR (human capital) characteristics may select different equilibria.

WHY IS IT IMPORTANT TO ALLOW FOR PR-U and MEQ ?

[1] Ignoring one type of heterogeneity typically implies that we over-estimate the contribution of the other.

- **Example**: In Todd and Wolpin, similar schools (in terms of observable inputs) have different outcomes mainly because they have different PR unobservables (e.g., cost of effort); or mainly because they have selected a different equilibrium.

[2] Counterfactuals: The two types of unobservables (PR and SS) enter differently in the model. They can generate very different counterfactual policy experiments.

RELATED LITERATURE: Games of complete information

• Bresnahan and Reiss (1990); Tamer (2003); Ciliberto and Tamer (2009); Bajari, Hong and Ryan (2009)

• Allow for **PRU** and, implicitly, for **MEQ**, but they do not try to identify separately their relative contribution.

• Tamer (2003): exclusion restrictions can (point) identify payoff function. [Bajari, Hong, Krainer and Nekipelov (2010) extend Tamer's result in games of incomplete information.].

CONTRIBUTION OF THE PAPER

• We study identification using cross-sectional choice data when the **three sources of unobservables may be present** in a semiparametric framework. As usual in this literature the distribution of private information is assumed known. However, the model is **nonparametric** for payoffs, equilibrium selection mechanism, and distribution of PR unobservables.

• Drawing on a statistical literature on finite mixtures, we first consider sequential identification which is needed for 2-step estimation approaches and their extensions. Under standard exclusion conditions for the estimation of games, we study how the payoff function, the distributions of PR heterogeneity and equilibrium selection can be identified.

• We study the relationship between sequential and joint identification. Some empirical models fail the conditions for sequential identification - when are they still identified?

OUTLINE OF PAPER

1. Discrete Games of Incomplete Information

2. Identification Results

3. Monte Carlo Experiments (illustrative, in progress)

4. Some comments on Estimation and Counterfactuals.

1. DISCRETE GAMES OF INCOMPLETE INFORMATION

• N players indexed by i. Each player has to choose an action, a_i, from a discrete set $\mathcal{A} = \{0, 1, ..., J\}$. to maximize his expected payoff.

• The payoff function of player i is:

$$\Pi_i = \pi_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x}, \omega) + \varepsilon_i(a_i)$$

- $\boldsymbol{a}_{-i} \in \mathcal{A}^{N-1}$ is a vector with choices of players other than i;

- $\mathbf{x} \in \mathcal{X}$ and $\omega \in \Omega$ are exogenous characteristics, common knowledge for all players. $\mathbf{x}$ is observable to the researcher, and ω is the **Payoff-Relevant (PR) unobservable**; $\mathcal{X}$ and Ω will be discrete.

- $\boldsymbol{\varepsilon}_i = \{\varepsilon_i(a_i) : a_i \in \mathcal{A}\}$ are private information variables for player i, and are unobservable to the researcher.

BAYESIAN NASH EQUILIBRIUM

• A **Bayesian Nash equilibrium (BNE)** is a set of strategy functions $\{\sigma_i(\mathbf{x}, \omega, \varepsilon_i) : i = 1, 2, ..., N\}$ such that any player maximizes his expected payoff given the strategies of the others:

$$\sigma_i(\mathbf{x}, \omega, \varepsilon_i) = \arg\max_{a_i \in \mathcal{A}} \ \mathbb{E}_{\boldsymbol{\varepsilon}_{-i}} \left(\ \pi_i(a_i, \ \sigma_{-i}(\mathbf{x}, \omega, \boldsymbol{\varepsilon}_{-i}), \ \mathbf{x}, \ \omega) \ \right) + \varepsilon_i(a_i)$$

• It will be convenient to represent players' strategies and BNE using **Conditional Choice Probability** (CCPs) **functions:**

$$P_i \left(a_i \mid \mathbf{x}, \omega \right) \equiv \int 1 \left\{ \sigma_i(\mathbf{x}, \omega, \varepsilon_i) = a_i \right\} \ dG_i(\varepsilon_i)$$

• Note: (a) The expectation on the RHS of the best response condition depends on the strategies of the other players only through their CCPs; (b) Equilibria depend on (x, ω) only through the payoffs $\pi_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x}, \omega)$.

• Let $\pi_{x,\omega}$ be a vector stacking the payoffs for all players and outcomes. Likewise, let P be a vector of CCPs. It can be shown that equilibrium CCPs at (x, ω) satisfy a fixed-point, best-response condition:

$$\mathbf{P}^*(\tilde{\boldsymbol{\pi}}_{(\mathbf{x},\omega)}) = \boldsymbol{\Psi}\left(\tilde{\boldsymbol{\pi}}_{(\mathbf{x},\omega)}, \mathbf{P}^*(\tilde{\boldsymbol{\pi}}_{(\mathbf{x},\omega)})\right)$$

• In this class of models, existence of at least a BNE is guaranteed. There may be **multiple equilibria**.

MULTIPLE EQUILIBRIA

• For some values of $(\mathbf{x}, \omega)$ the model has multiple equilibria. Let $\Gamma(\mathbf{x}, \omega)$ be the set of equilibria associated with $(\mathbf{x}, \omega)$.

• We assume that $\Gamma(\mathbf{x}, \omega)$ is a discrete and finite set (see Doraszelski and Escobar, 2010) for regularity conditions that imply this property.

• Each equilibria belongs to a particular "type" such that a marginal perturbation in the payoff function implies also a small variation in the equilibrium probabilities within the same type.

• We index equilibrium types by $\tau \in \{1, 2, ...\}$

EXAMPLE - COORDINATION GAME WITHIN THE CLASSROOM

- A much simplified version: Each student chooses high effort ($a_i = 1$) or low effort ($a_i = 0$).

- The teacher's combination of effort and skills is exogenous and common knowledge, represented by the scalar variable x.

- The payoff for student i is:

$$\Pi_i = \begin{cases} \alpha_0 + \beta_0 \, x + \gamma_0 \, x \, \left(\dfrac{1}{N-1} \sum_{j \neq i} a_j \right) + \varepsilon_i(0) & \text{if} \quad a_i = 0 \\ \\ \alpha_1 + \beta_1 \, x + \gamma_1 \, x \, \left(\dfrac{1}{N-1} \sum_{j \neq i} a_j \right) + \varepsilon_i(1) & \text{if} \quad a_i = 1 \end{cases}$$

where α_0, β_0, γ_0, α_1, β_1 and γ_1 are parameters. This specification establishes that a student's payoff depends on his own effort, the teacher's effort-skills, the average effort of the other students, and his own private information cost of effort (or skills).

• Suppose that $\varepsilon_i(0)$ and $\varepsilon_i(1)$ are normal random variables, independently distributed across students with zero mean and with $Var(\varepsilon_i(1) - \varepsilon_i(0)) = \delta^2$.

• The normalized payoff $\widetilde{\pi}_i(1, \boldsymbol{a}_{-i}, x, \omega) + \widetilde{\varepsilon}_i(1)$ is such that $\widetilde{\pi}_i(1, \boldsymbol{a}_{-i}, x, \omega) = \alpha + \beta\, x + \varphi\, \omega + \gamma\, x \left(\frac{1}{N-1} \sum_{j \neq i} a_j \right)$, with $\alpha \equiv (\alpha_1 - \alpha_0)/\delta$, $\beta \equiv (\beta_1 - \beta_0)/\delta$, $\varphi \equiv (\varphi_1 - \varphi_0)/\delta$, $\gamma \equiv (\gamma_1 - \gamma_0)/\delta$, and $\widetilde{\varepsilon}_i(1) \equiv (\varepsilon_i(1) - \varepsilon_i(0))/\delta$.

• Suppose that students are identical except for their private information variables and each student perceives the other student as identical and believes that all students have the same probability of high effort $P(x, \omega)$, i.e., we assume that the equilibrium is symmetric. Then, the best response probability function of each student in this model is:

$$\mathbf{\Psi}(1|\widetilde{\boldsymbol{\pi}}_{(x,\omega)},\ P) = \mathbf{\Phi}\left(\alpha + \beta\ x + \gamma\ x\ P(x, \omega)\right)$$

• Suppose that $x > 0$ and $\gamma > 0$ such that there positive synergies between the teacher's effort/skills and students' effort. Then, the model is a *Coordination Game*.

• Figures 1 and 2 come from this example when the parameter values are $\alpha = 2.0$, $\beta = -7.31$, and $\gamma = 6.75$, and the variable x that represents teacher's effort-skills is an index in the interval $[0, 1]$. Figure 1 presents the equilibrium mapping when teacher's effort is $x = 0.52$. For this level of teacher's effort the model has three equilibria with low, middle, and high probability of high students' effort.

Figure 1: Coordination Game. Three Types of Equilibria

Best response function: $\Psi(P) = \Phi(2.0 - 7.32\ x + 6.75\ x\ P)$

Teacher's effort: $x = 0.52$

Set of Equilibria: $\Upsilon(x = 0.52) = \{0.054,\ 0.521,\ 0.937\}$

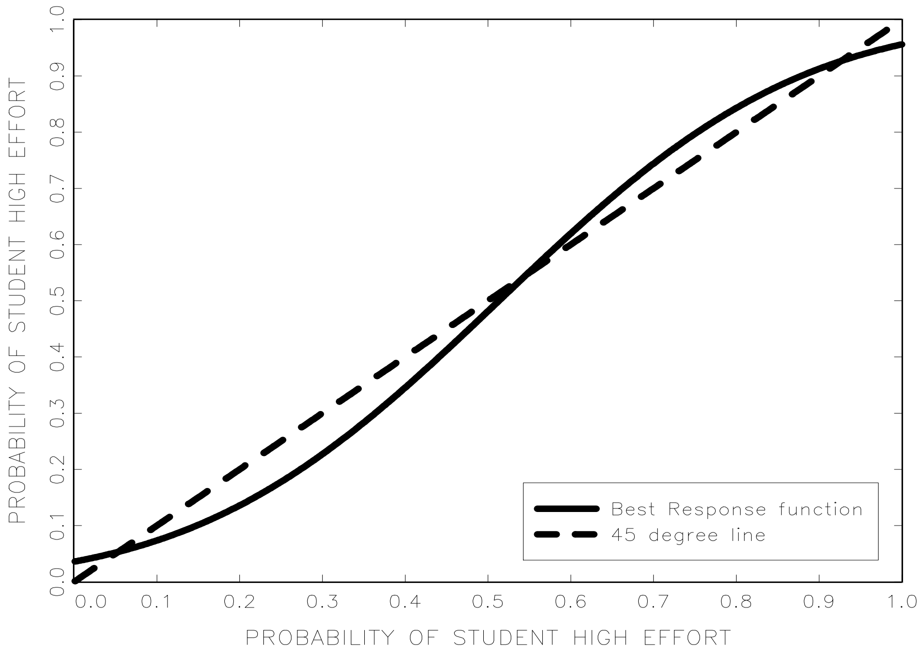
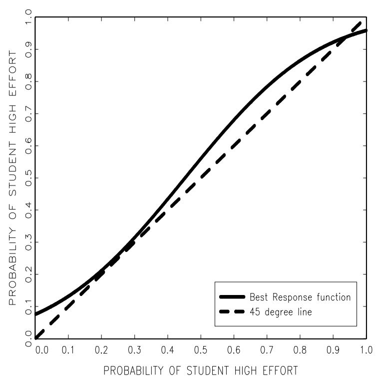


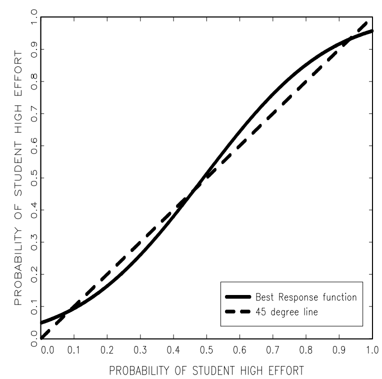
Figure 2: Coordination Game. Equilibrium Types

Best response function: $\Psi(P) = \Phi(2.0 - 7.32\ x + 6.75\ x\ P)$

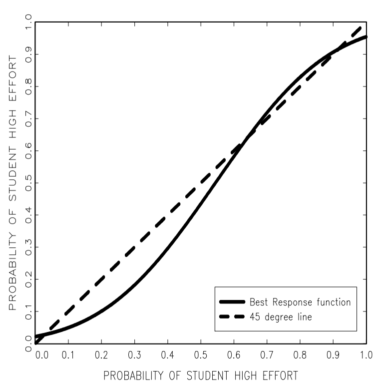
(A) $x = 0.47$. Equilibrium: 0.938



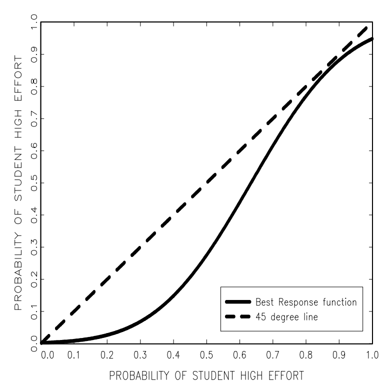
(B) $x = 0.50$. Equilibria: 0.086; 0.462; 0.932

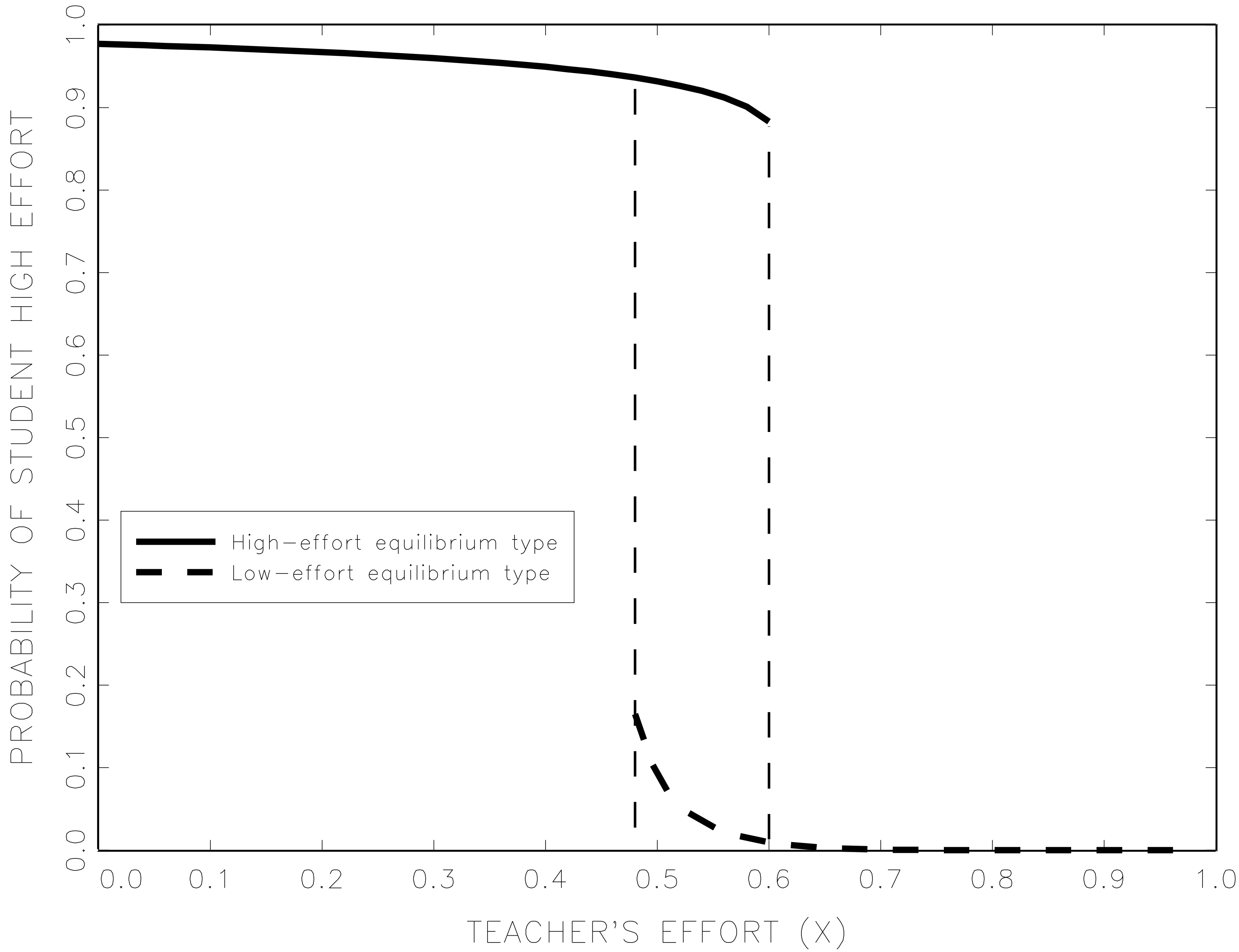


(C) $x = 0.55$. Equilibria: 0.028; 0.643; 0.917



(D) $x = 0.66$. Equilibrium: 0.001




PROBABILITY OF STUDENT HIGH EFFORT
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
High−effort equilibrium type
Low−effort equilibrium type
TEACHER'S EFFORT (X)

2. DATA, DGP, AND IDENTIFICATION

• The researcher observes M realizations of the game; e.g., M classes or markets.

$$Data = \{ \ a_{1m}, \ a_{2m}, ..., a_{Nm}, \mathbf{x}_m : m = 1, 2, ..., M \ \}$$

• DGP.

(A) $(\mathbf{x}_m, \omega_m) \sim i.i.d.$ draws from CDF $F_{x,\omega}$. Support of ω_m is discrete (finite mixture);

(B) The equilibrium type selected in observation m, τ_m, is a random draw from a probability distribution $\lambda(\tau|\mathbf{x}_m, \omega_m)$;

(C) $\boldsymbol{a}_m \equiv (a_{1m}, a_{2m}, ..., a_{Nm})$ is a random draw from a multinomial distribution such that:

$$\Pr \boldsymbol{a}_m \mid \mathbf{x}_m, \omega_m, \tau_m) = \prod_{i=1}^{N} P_i(a_{im} \mid \mathbf{x}_m, \omega_m, \tau_m)$$

2.1. IDENTIFICATION PROBLEM

• The researcher knows $G_i(\varepsilon_i)$, the probability distribution of private information.

• Let $Q(\boldsymbol{a}|\mathbf{x})$ be the probability distribution of observed players' actions conditional on observed exogenous variables: $Q(\boldsymbol{a}|\mathbf{x}) \equiv \Pr(\boldsymbol{a}_m = \boldsymbol{a} \mid \mathbf{x}_m = \mathbf{x})$.

• Under mild regularity conditions, $Q(.|.)$ is identified from our data.

• According to the model and DGP:

$$Q(\boldsymbol{a}|\mathbf{x}) = \sum_{\omega \in \Omega} \sum_{\tau \in \Upsilon(\mathbf{x},\omega)} F_\omega(\omega|\mathbf{x}) \ \lambda(\tau|\mathbf{x},\omega) \ \left[\prod_{i=1}^{N} P_i(a_i \mid \mathbf{x}, \omega, \tau; \boldsymbol{\pi}) \right] \quad (1)$$

• *DEFINITION: The model is (point) identified if given Q there is a unique value $\{\boldsymbol{\pi},\ F_\omega,\ \lambda\}$ that solves the system of equations (1).*

IDENTIFICATION QUESTIONS

• We focus on three main identification questions:

1. Sufficient conditions for the identification $\{\boldsymbol{\pi}, F_\omega, \lambda\}$;
2. Test of the null hypothesis of No PR unobservables;
3. Test of the null hypothesis of No MEQ;

• De Paula and Tang (2011) show how multiple equilibria in the data can identify strategic interactions. We would lime to know ex-ante if that inference is justified.

• Even with a rich nonparametric specification of PR unobservables, is it possible to reject the hypothesis of "No MEQ" and conclude that we need "multiple equilibria" to explain the data?

2.2. SEQUENTIAL (THREE-STEP) IDENTIFICATION APPROACH

• Most of our identification results are based on a three-step approach.

• Let $\kappa \equiv g(\omega, \tau)$ be a scalar discrete random variable that represents all the unobserved heterogeneity, both PR and SS. κ does not distinguish the source of this heterogeneity.

• Let $H(\kappa|\mathbf{x})$ be the PDF of κ (given $F_\omega(\omega|\mathbf{x})\ \lambda(\tau|\mathbf{x}, \omega)$),

$$H(\kappa|\mathbf{x}) = 1\{\kappa = g(\omega, \tau)\}\ F_\omega(\omega|\mathbf{x})\ \lambda(\tau|\mathbf{x}, \omega)$$

STEP 1. NP identification of $H(\kappa|\mathbf{x})$ and CCPs $P_i(a_i|\mathbf{x},\kappa)$ that satisfy restrictions:

$$Q(a_1,a_2,...,a_N \mid \mathbf{x}) = \sum_{\kappa} H(\kappa|\mathbf{x}) \left[\prod_{i=1}^{N} P_i(a_i \mid \mathbf{x},\kappa)\right]$$

STEP 2. Given the CCPs $\{P_i(a_i|\mathbf{x},\kappa)\}$ and the distribution of ε_i, it is possible to obtain the *differential-expected-payoff function* $\widetilde{\pi}_i^P(a_i,\mathbf{x},\kappa)$. This follows from well known invertibility properties of a broad class of discrete-choice econometric models. [see Lemma 2 based on McFadden, Hotz-Miller].

• $\widetilde{\pi}_i^P(a_i,\mathbf{x},\kappa)$ is the expected value for player i of choosing alternative a_i minus the expected value of choosing alternative 0 when players' strategies are P. By definition:

$$\widetilde{\pi}_i^P(a_i,\mathbf{x},\kappa) \equiv \sum_{a_{-i}} \left(\prod_{j \neq i} P_j(a_j|\mathbf{x},\kappa) \right) \; [\pi_i(a_i,\boldsymbol{a}_{-i},\mathbf{x},\omega) - \pi_i(0,\boldsymbol{a}_{-i},\mathbf{x},\omega)]$$

• Given this equation and the identified $\widetilde{\pi}_i^P$ and $\{P_j\}$, we study the identification of the payoff π_i.

• As is usual, we make the normalization $\pi_i(0,\boldsymbol{a}_{-i},\mathbf{x},\omega) = 0$ and we use $\widetilde{\pi}_i(a_i,\boldsymbol{a}_{-i},\mathbf{x},\omega)$ for $a_i > 0$ to denote normalized payoffs.

STEP 3. Given the identified payoffs π_i and the distribution $H(\kappa|\mathbf{x})$, we study the identification of the distributions $F_\omega(\omega|\mathbf{x})$ and $\lambda(\tau|\mathbf{x},\omega)$.

• Hypothesis like "No PR heterogeneity" or "No MEQ" can be expressed in terms of the rank of a matrix of payoffs. But this requires step 2 and exclusion restrictions!

• This three-step approach does not come without some loss of generality. Sufficient conditions of identification in step 1 can be 'too demanding'. We have examples of NP identified models that do not satisfy identification in step 1.

IDENTIFICATION IN STEP 1

• **Point-wise identification (for every value x)** of the NP finite mixture model:

$$Q(a_1,a_2,...,a_N \mid \mathbf{x}) = \sum_{\kappa} H(\kappa|\mathbf{x}) \left[\prod_{i=1}^{N} P_i(a_i \mid \mathbf{x}, \kappa) \right]$$

• Identification is based on the independence between players' actions once we condition on $(\mathbf{x}, \kappa)$.

• We exploit results by Hall and Zhou (2003), Hall, Neeman, Pakyari, and Elmore (2005), Allman et al (2009) and Kasahara and Shimotsu (2014).

IDENTIFICATION IN STEP 1 (II)

• Let L_{κ} be the number of "branches" (different values of κ) that we can identify in this NP finite mixture.

• A lower bound on the number L_{κ} is identified as long as $N \geq 2$. It might be a tight bound.

• The number of players N and choice alternatives $J + 1$ establish an upper bound on L_{κ}.

1. For $L_{\kappa} \geq 2$, we need at least 3 players;
2. With $N \geq 3$, we have that $L_{\kappa} \leq (J + 1)^{int[(N-1)/2]}$

IDENTIFYING THE NUMBER OF COMPONENTS - Detail:

Consider a partition of the set of players into two groups with N_1 and N_2 players, $N_1 + N_2 = N$. Let S_1 and S_2 be two random variables which summarize the outcome of one realization of our game for each subgroup of players given actions $\{a_i : i = 1, \ldots, N\}$. By construction, S_1 and S_2 are discrete and independent conditional on (x, κ). Let $C_{(S_1,S_2)}$ be a matrix describing the (population) joint distribution of (S_1, S_2). Then the rank of $C_{(S_1,S_2)}$ is a lower bound of the true number of mixture components L_κ. Furthermore, let $\widetilde{J}_1$ and $\widetilde{J}_2$ be the number of rows and columns of matrix $C_{(S_1,S_2)}$. Clearly, rank $(C_{(S_1,S_2)}) \leq \min[\widetilde{J}_1, \widetilde{J}_2]$. Then, the bound is tight and the number of components is exactly identified if the strict version of this inequality is satisfied.

• EXAMPLES:

(1) In an entry game with 3 players, the model is step 1 identified if the DGP has two mixture components, but no more.
We would set $S_1 = (a_1, a_2)$ and $S_2 = a_3$ so matrix $\mathbf{C}_{(S_1,S_2)}$ would be 4×2, its rank would be 2 and we would be able to tell that the number of components is at least 2.

(2) An entry game with 5 players is identified in step 1 with up to 4 mixture components, e.g., there might be a binary payoff-relevant unobservable with two different equilibria being played at each of the two values of the payoff-relevant unobservable.
In this case we might set $S_1 = (a_1, a_2, a_3)$, $S_2 = (a_4, a_5)$ and $\mathbf{C}_{(S_1,S_2)}$ would be 8×4. With 4 mixture components in the DGP its rank would be 4 and the researcher would obtain this as a lower bound on the unknown true number of components.

(3) Consider a game with $N = 5$, $J + 1 = 3$. The maximum number of components that can be identified is 9. If we set

$$S_1 = (1(a_1 \geq 1), 1(a_2 \geq 1), 1(a_3 \geq 1)), S_2 = (a_3, a_4)$$

then $\mathbf{C}_{(S_1,S_2)}$ is 8×9. If the DGP had 6 components the rank of $\mathbf{C}_{(S_1,S_2)}$ would be 6 which is smaller than $\min[9, 9]$ so the bound is tight and the researcher would know this to be the exact number of components.

IDENTIFICATION IN STEP 1 (III)

• So there are limits to the number of 'branches' L_κ we can identify using this sequential approach.

• A parametric specification for $H(\kappa|\mathbf{x})$, or exclusion restrictions on it, or independence between κ and $\mathbf{x}$ can allow for identification of unobserved heterogeneity with more points of support. But it may not be easy to justify such assumptions ...

IDENTIFICATION IN STEP 2

• We need to recover the normalized payoff function $\widetilde{\boldsymbol{\pi}}$ from the system of equations:

$$\widetilde{\pi}_i^P(a_i, \mathbf{x}) = \sum_{\boldsymbol{a}_{-i}} Q_{-i}(\boldsymbol{a}_{-i}|\mathbf{x}) \; \widetilde{\pi}_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x})$$

where $Q_{-i}(\boldsymbol{a}_{-i}|\mathbf{x}) \equiv \prod_{j \neq i} P_j(a_j|\mathbf{x})$ and $\widetilde{\pi}_i^P(a_i, \mathbf{x})$ is known by inversion.

• To lighten the notation a bit, I am using $\mathbf{x}$ as shorthand for (x, κ).

• If different values of κ correspond to multiple equilibria, then the payoff on the RHS would not vary with κ but the researcher does not know this ex-ante so we need to consider the "worst-case scenario" where different values of κ correspond to different values of ω.

• Because of strategic interactions there are multiple payoff values $\widetilde{\pi}_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x})$ for every $\widetilde{\pi}_i^P(a_i, \mathbf{x})$ that is identified from the data and Lemma 2, so a discrete game is potentially severely underidentified relative to a standard discrete choice - random utility model.

• Exclusion restrictions on payoffs are needed to restore identification. This is true even in the standard model with only PI $\varepsilon' s$, with no PRU and no MEQ.

• Suppose (for simplicity) that $\mathbf{x}$ has a discrete and finite support. We assume that $\mathbf{x}$ has the following structure:

$$\mathbf{x} = (\mathbf{x}^c, z_1, z_2, ..., z_N)$$

such that:

(i) $z_i \in \mathcal{Z}$ and the support $\mathcal{Z}$ has at least $J + 1$ points;

(ii) Exclusion restriction: $\pi_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x}) = \pi_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x}^c, z_i)$;

IDENTIFICATION IN STEP 2 **with exclusion restrictions**

- Then equations above can be written in matrix form as

$$\widetilde{\pi}_i^P(a_i, x^c, z_i) = \mathbf{Q}_{-i}(\mathbf{x}^c, z_i)\widetilde{\pi}_i(a_i, x^c, z_i)$$

where the matrix $Q_{-i}(x^c, z_i)$, with dimension $|Z|^{N-1} \times (J+1)^{N-1}$ is defined as

$$\begin{bmatrix} Q_{-i}(\boldsymbol{a}_{-i}^{(1)} \mid \mathbf{x}^c, z_i, \mathbf{z}_{-i}^{(1)}) & Q_{-i}(\boldsymbol{a}_{-i}^{(2)} \mid \mathbf{x}^c, z_i, \mathbf{z}_{-i}^{(1)}) & \cdots & Q_{-i}(\boldsymbol{a}_{-i}^{((J+1)^{N} \\ Q_{-i}(\boldsymbol{a}_{-i}^{(1)} \mid \mathbf{x}^c, z_i, \mathbf{z}_{-i}^{(2)}) & Q_{-i}(\boldsymbol{a}_{-i}^{(2)} \mid \mathbf{x}^c, z_i, \mathbf{z}_{-i}^{(2)}) & \cdots & Q_{-i}(\boldsymbol{a}_{-i}^{((J+1)^{N} \\ \vdots & \vdots & & \\ Q_{-i}(\boldsymbol{a}_{-i}^{(1)} \mid \mathbf{x}^c, z_i, \mathbf{z}_{-i}^{(|\mathcal{Z}|^{N-1})}) & Q_{-i}(\boldsymbol{a}_{-i}^{(2)} \mid \mathbf{x}^c, z_i, \mathbf{z}_{-i}^{(|\mathcal{Z}|^{N-1})}) & \cdots & Q_{-i}(\boldsymbol{a}_{-i}^{((J+1)^{N-1} \end{bmatrix}$$

We can recover the vector of payoffs $\widetilde{\pi}_i(a_i, x^c, z_i)$ as long as matrix $Q_{-i}(x^c, z_i)$ has full column rank.

IDENTIFICATION IN STEP 2: A matching problem

• In Step 1, identification of the distribution H and of CCPs P_i's is *up to label swapping*, and "pointwise" or separately for each subpopulation defined by observables $\mathbf{x}$.

• In order to implement Step 2, the researcher needs to be able to "match" mixture components which correspond to the same value of ω across different subpopulations of observables.

• If we use "unmatchable" assignments which (incorrectly) match mixture components corresponding to different values of ω then the system of equations which exploits exclusion restrictions is not satisfied at the true payoffs.

• Step 2 identification now requires distinguishing between unmatchable and matchable assignments and using one of the latter.

IDENTIFICATION IN STEP 2: Dealing with the matching problem

• Given an arbitrary label swap, let $\widetilde{\mathbf{Q}}_{-i}(x^c, z_i, \kappa)$ be the $Q_{-i}(x^c, z_i, \kappa)$ matrix augmented with column vector $\Pi_i^P(a_i, x^c, z_i)$ of expected payoffs.

LEMMA 3. Under exclusion restrictions, a matchable label assignment is identified in step 2 iff:

(a) Matrix $\mathbf{Q}_{-i}(\mathbf{x}^c, z_i, \kappa)$ has full column rank for all i in *any matchable label swap; and*

(b) For every unmatchable label swap, there is at least one player i for which the rank of augmented matrix $\widetilde{\mathbf{Q}}_{-i}(x^c, z_i, \kappa)$ is larger than the rank of matrix $\mathbf{Q}_{-i}(\mathbf{x}^c, z_i, \kappa)$.

- The number of label swaps is finite.

- If the researcher considers an unmatchable label assignment, then the system of equations $\mathbf{\Pi}_i^{\mathbf{P}}(a_i, \mathbf{x}^c, z_i, \kappa) = \mathbf{Q}_{-i}(\mathbf{x}^c, z_i, \kappa)\ \mathbf{\Pi}(a_i, \mathbf{x}^c, z_i, \kappa)$ will not have a solution.

- If a label swap leads to a system which has a unique solution, then that label assignment is matchable and the solution is the true vector of payoffs.

PROPOSITION 3. Assume the exclusion restriction, $\mathbf{x} = \{\mathbf{x}^c, z_i : i \in \mathcal{I}\}$ *where* $z_i \in \mathcal{Z}$ *and the set* $\mathcal{Z}$ *is discrete with at least* $J + 1$ *points, and* $\pi_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x}, \kappa)$ *depends on* $(\mathbf{x}^c, z_i, \kappa)$ *but not on* $\{z_j : j \neq i\}$. *Then:*

(A) The payoff functions are identified iff conditions (a) and (b) of Lemma 3 are satisfied.

(B) Payoff functions π_i *are identified if: (c) (rank condition)* **and**
(d) $H(\kappa|\mathbf{x})$ *depends on* $\mathbf{x}^c$ *but not on* $\{z_i : i \in \mathcal{I}\}$, *and* $H(\kappa|\mathbf{x}^c) \neq H(\kappa'|\mathbf{x}^c)$ *for any two values* κ *and* κ' *in the support of* $H(.|\mathbf{x}^c)$.

IDENTIFICATION IN STEP 3

• Suppose that the researcher has identified the distribution $H(\kappa|\mathbf{x})$ and the payoff functions $\widetilde{\pi}_i$. We want to identify the probability distributions $F_\omega(\omega|\mathbf{x})$ and $\lambda(\tau|(\boldsymbol{\pi}_{(\mathbf{x},\omega)})$. There are two sets of restrictions that we can exploit: (1) the payoff π_i depends on ω but not on τ; and (2) by definition, $H(\kappa|\mathbf{x}) = \sum\limits_{\omega,\tau} \mathbf{1}\{\kappa = g(\omega,\tau)\}\ F_\omega(\omega|\mathbf{x})\ \lambda(\tau|\boldsymbol{\pi}_{(\mathbf{x},\omega)})$.

• Let $\mathbf{\Pi}_i(\mathbf{x})$ be the matrix with dimension $J(J+1)^{N-1} \times L_\kappa(\mathbf{x})$ that contains all payoffs $\{\widetilde{\pi}_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x}, \kappa)\}$ for a given value of $\mathbf{x}$. More specifically, each column corresponds to a value of κ and it contains the payoffs $\pi_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x}, \kappa)$ for every value of $(a_i, \boldsymbol{a}_{-i})$ with $a_i > 0$. If two values of κ represent the same of value of ω, then the corresponding columns in the matrix $\mathbf{\Pi}_i(\mathbf{x})$ should be equal. Therefore, we can identify the number of mixture components $L_\omega(\mathbf{x})$ as:

$$L_\omega(\mathbf{x}) = rank\ (\ \mathbf{\Pi}_i(\mathbf{x})\ )$$

• The matrix $\mathbf{\Pi}_i(\mathbf{x})$ not only identifies the number of points in the support of the PR unobservables, but also the points of support themselves and, together with the set of restrictions (2), the distributions of these unobservables. Once we identify the columns of $\mathbf{\Pi}_i(\mathbf{x})$ that are different and the ones that are the same, we can label each column (i.e., each value of κ) with two values, one for ω and other for τ.

TESTING FOR "NO PR-U" OR "NO MEQ"

• At the end of step 1, before we recover payoffs, the hypothesis that there is no PR unobserved heterogeneity can tested as follows. Consider matrices $\mathbf{Q}_{-i}(\mathbf{x}^c, z_i)$ and $\widetilde{\mathbf{Q}}_{-i}(x^c, z_i)$ obtained when we stack vertically, for all κ, matrices $\mathbf{Q}_{-i}(\mathbf{x}^c, z_i, \kappa)$ and $\widetilde{\mathbf{Q}}_{-i}(x^c, z_i, \kappa)$. Under the null these matrices have the same rank which is the number of columns of $\mathbf{Q}_{-i}(\mathbf{x}^c, z_i, \kappa)$.

If we use the payoffs identified in step 2:

• Taking into account that $L_{\omega}(\mathbf{x}) = rank(\mathbf{\Pi}_i(\mathbf{x}))$ and that $L_{\kappa}(\mathbf{x}) = \text{cols}(\mathbf{\Pi}_i(\mathbf{x}))$, we have that testing for the null hypothesis of *"no MEQ"* is equivalent to testing for:

$$H_0: \text{ For every value of } \mathbf{x} \text{ the matrix } \mathbf{\Pi}_i(\mathbf{x}) \text{ has full column rank.}$$

• If there is not PR unobserved heterogeneity, then for any value of $\mathbf{x}$ in the sample the number of points in the support of ω should be equal to the $\mathbf{1}$. Therefore testing for the null hypothesis of *"no PR unobserved heterogeneity"* is equivalent to testing for:

$$H_0: \text{ For every value of } \mathbf{x} \text{ the matrix } \mathbf{\Pi}_i(\mathbf{x}) \text{ has rank equal to 1.}$$

• Therefore tests for these null hypotheses can be described in terms of tests of the rank of a matrix of statistics. This type of tests have been proposed and developed by Cragg and Donald (1993, 1996, 1997) and Robin and Smith (2000)

SEQUENTIAL IDENTIFICATION VS. JOINT IDENTIFICATION

• What if $N = 2$ or $N \geq 3$ but $L_\kappa > (J+1)^{(N-1)/2}$?

LEMMA 3: (A) Sequential identification implies joint identification. (B) The converse is not true in general. In particular, joint identification implies step 2 identification but it does not imply step 1 identification. (C) Sequential and joint identification are equivalent iff the Step 2 systems () are just identified.

• Equilibrium restrictions per se are insufficient to deliver identification. Models which are not amenable to the sequential approach can still be identified and estimated when exclusion restrictions are sufficiently overidentifying.

• The exclusion restriction makes the number of payoff parameters linear in the number of players N whereas the number of restrictions on the data remains exponential in N. An analysis of the order condition for joint identification shows that excluded variables with sufficiently large support should provide identification as long as $(J+1)^N > L_\kappa$.

• The 2x2 game is the worst-case scenario: A 2-player binary game with both PR-unobserved heterogeneity and multiple equilibria is not identified unless we impose restrictions on the equilibrium selection.

• With PR-U only, the 2x2 game is identified for $L_\omega = 2$ and $|\mathcal{Z}| \geq 4$, and for $L_\omega = 3$ and $|\mathcal{Z}| \geq 12$.

• *DEFINITIONS: Let* $\{\widetilde{\boldsymbol{\pi}}, H\}$ *be the structural parameters of the model and* $P \equiv \{P_i(a_i|x,\kappa) : a_i \in A - \{0\}; \text{ all } (x,\kappa)\}$ *the vector of true equilibrium CCP's.*

(1) $\{\widetilde{\boldsymbol{\pi}}, H\}$ *are jointly identified iff, given the population* $Q(\boldsymbol{a}|\mathbf{x})$, *there is a unique pair* $\{\widetilde{\boldsymbol{\pi}}, H\}$ *which satisfies the conditions:*

$$
\begin{aligned}
Q(\boldsymbol{a}|\mathbf{x}) &= \sum_{\kappa} H(\kappa|\mathbf{x}) \left[\prod_{i=1}^{N} P_i(a_i|\widetilde{\boldsymbol{\pi}}_{(\mathbf{x},\omega)}, \tau) \right] \\
\text{subject to } \mathbf{P}(\widetilde{\boldsymbol{\pi}}_{(\mathbf{x},\omega)}, \tau) &= \boldsymbol{\Psi}(\widetilde{\boldsymbol{\pi}}_{(\mathbf{x},\omega)}, \mathbf{P}(\widetilde{\boldsymbol{\pi}}_{(\mathbf{x},\omega)}, \tau))
\end{aligned}
$$

(2) $\{\boldsymbol{\pi}, H\}$ *are sequentially identified iff: (Step 1 identification) given the population* $Q(\boldsymbol{a}|\mathbf{x})$ *there is a unique pair* $\{H, \mathbf{P}\}$ *(up to label swapping) that satisfies conditions ()(a); and (Step 2 identification) given the true* CCP's $\mathbf{P}$ *and the expected payoffs* $\boldsymbol{\pi}^{\mathbf{P}}$ *obtained from them by Lemma 2, consider every possible label swap and the associated system of equations ()(b); then the system either has no solution, or else has the vector of payoffs* $\boldsymbol{\pi}$ *as its unique solution.*

• (More detail) The potential for exclusion restrictions to help with joint identification can also be seen in general order condition:

$$\left[(J+1)^N - 1\right] \cdot |\mathcal{Z}|^N \geq L_\omega \cdot N \cdot |\mathcal{Z}| \cdot (J+1)^{N-1} \cdot J + (L_\kappa - 1) \cdot |\mathcal{Z}|^N$$

where the left-hand-side counts the number of equations or probabilities in the $Q(a|z_1, \ldots, z_n)$ and the RHS is the sum of the number of payoffs and the number of mixing weights. We can ignore the existence of non-excluded regressors x_c without loss of generality because introducing them would just multiply each of the 3 terms of the order condition by the same factor, i.e., the cardinality of x_c. The exclusion restriction makes the number of payoff parameters linear in the number of players N whereas the number of restrictions on the data remains exponential in N. Excluded variables with sufficiently large support should provide identification as long as $(J+1)^N > L_\kappa$.

Table 1
Summary of DGPs in Monte Carlo Experiments

Common features in the two experiments

Payoff function:	$\pi_i = \alpha_i + \beta_i \; z_{im} + \omega_m + \delta_i \; \sum_{j \neq i} a_{jm}$
Distribution z_{im}:	i.i.d. Uniform $\left[\frac{0}{\lvert\mathcal{Z}\rvert-1}, \frac{1}{\lvert\mathcal{Z}\rvert-1}, \cdots \frac{\lvert\mathcal{Z}\rvert-1}{\lvert\mathcal{Z}\rvert-1}\right]$
Distribution ω_m:	Support $\{-0.75, +0.75\}$
	$F_{\omega}^{A}(\mathbf{z}) = f_0 + f_1 \; N^{-1} \left(\sum_{i=1}^{N} z_i\right)$
# equilibria in data:	1
# markets (M):	50 and 200 for each possible value of $\mathbf{z}$
# MC replications:	1,000

Experiment SEQ	**Experiment NONSEQ**
$N = 4$	$N = 2$
$\lvert\mathcal{Z}\rvert = 3$	$\lvert\mathcal{Z}\rvert = 5$
$\alpha_1 = -1.00$; $\alpha_2 = -0.80$; $\alpha_3 = -0.60$; $\alpha_4 = -0.40$	$\alpha_1 = -1.00$; $\alpha_2 = -0.40$
$\beta_1 = \beta_2 = \beta_3 = \beta_4 = 3.0$	$\beta_1 = \beta_2 = 3.0$
$\delta_1 = \delta_2 = \delta_3 = \delta_4 = -0.5$	$\delta_1 = \delta_2 = -0.5$
$f_0 = 0.20$ and $f_1 = 0.25$	$f_0 = 0.20$ and $f_1 = 0.25$

Table 2. Experiment SEQ

Sample Sizes: 50 and 200 markets per value of $\mathbf{z}$. Monte Carlo Simulations = 1,000

Panel A. Test of rank of $\mathbf{C}_{(S_1,S_2)}$ and empirical distribution of $\widehat{L_\kappa}(\mathbf{z})$

	% of cases		Estimate of $\widehat{L_\kappa}(\mathbf{z})$	
Tests	50 obs per z	200 obs per z	50 obs per z	200 obs per z
Reject $\det 4 = 0$	1.62%	0.35%	4	4
Accept $\det 4 = 0$ & reject $\det 3 = 0$	5.62%	2.72%	3	3
Accept $\det 4 = \det 3 = 0$ & reject $\det 2 = 0$	90.20%	95.70%	2	2
Accept $\det 4 = \det 3 = \det 2 = 0$	2.58%	1.23%	1	1

Panel B. Step 1: Estimation of $H(\kappa|\mathbf{z})$ and CCPs

	Bias (% true)		RMSE (% true)	
Parameter	50 obs per z	200 obs per z	50 obs per z	200 obs per z
$H(\kappa_A|\mathbf{z} = [0,0,0,0])$	0.0042 (2.1%)	0.0000 (0.0%)	0.1282 (64.0%)	0.0607 (30.3%)
$P_1^A(\mathbf{z} = [0,0,0,0])$	0.0003 (2.9%)	-0.0001 (-1.0%)	0.0070 (65.4%)	0.0034 (31.5%)
$P_1^B(\mathbf{z} = [0,0,0,0])$	0.0006 (0.4%)	-0.0011 (-0.8%)	0.0523 (40.6%)	0.0258 (20.1%)
$H(\kappa_A|\mathbf{z} = [1,1,1,1])$	-0.0041 (-0.9%)	0.0015 (0.3%)	0.1126 (25.0%)	0.0591 (13.1%)
$P_1^A(\mathbf{z} = [1,1,1,1])$	0.0069 (1.6%)	0.0004 (0.1%)	0.0837 (19.5%)	0.0405 (9.4%)
$P_1^B(\mathbf{z} = [1,1,1,1])$	0.0005 (0.1%)	-0.0027 (0.0%)	0.0522 (6.4%)	0.0276 (3.3%)

Panel C. Step 2: Estimation of players' payoffs

	Bias (% true)		RMSE (% true)	
Parameter	50 obs per z	200 obs per z	50 obs per z	200 obs per z
$\pi_1^A(\boldsymbol{a}_{-1} = [0,0,0], z_1 = 0.5)$	-0.0235 (-3.1%)	-0.0022 (-0.3%)	0.1438 (19.1%)	0.0717 (9.5%)
$\pi_1^A(\boldsymbol{a}_{-1} = [1,1,1], z_1 = 0.5)$	-0.1824 (-8.1%)	-0.0376 (-1.6%)	1.1030 (49.0%)	0.5074 (22.5%)
$\pi_1^B(\boldsymbol{a}_{-1} = [0,0,0], z_1 = 0.5)$	0.0153 (2.0%)	0.0047 (0.6%)	0.3518 (46.9%)	0.1684 (22.4%)
$\pi_1^B(\boldsymbol{a}_{-1} = [1,1,1], z_1 = 0.5)$	-0.0157 (-2.0%)	-0.0049 (-0.6%)	0.2006 (26.7%)	0.1005 (13.4%)

Table 3. Experiment NONSEQ

Sample Sizes: 50 and 200 markets per value of $\mathbf{z}$. Monte Carlo Simulations = $1,000$

	Bias (% true)		RMSE (% true)	
Parameter	**50 obs per z**	**200 obs per z**	**50 obs per z**	**200 obs per z**
$F_\omega^A(\mathbf{z}=[0,0])$	-0.0014 (0.7%)	0.0008 (0.4%)	0.0720 (36.0%)	0.0267 (13.3%)
$F_\omega^A(\mathbf{z}=[0,1])$	0.0015 (0.4%)	0.0006 (0.1%)	0.1010 (31.0%)	0.0554 (17.0%)
$F_\omega^A(\mathbf{z}=[1,0])$	-0.0087 (-2.6%)	0.0015 (0.4%)	0.1546 (49.1%)	0.0662 (20.3%)
$F_\omega^A(\mathbf{z}=[1,1])$	0.0020 (0.6%)	-0.0005 (-0.1%)	0.0639 (14.2%)	0.0270 (6.0%)
$\pi_1^A(a_2=0, z_1=0.4)$	-0.0115 (-2.1%)	0.0038 (0.7%)	0.1705 (31.0%)	0.0808 (14.7%)
$\pi_1^A(a_2=1, z_1=0.4)$	0.0514 (4.9%)	-0.0199 (-1.9%)	0.2426 (23.1%)	0.1071 (10.2%)
$\pi_1^B(a_2=0, z_1=0.4)$	0.0294 (3.1%)	0.0085 (0.9%)	0.3743 (39.4%)	0.1197 (12.6%)
$\pi_1^B(a_2=1, z_1=0.4)$	-0.0099 (-2.2%)	-0.0031 (-0.7%)	0.0936 (20.8%)	0.0414 (9.2%)

SOME THOUGHTS ON COUNTERFACTUALS

• A counterfactual scenario is described by x^*, $F_\omega(\omega|x^*), \widetilde{\pi}^*_{x^*}$.

• Of course, the interesting case is when there are multiple equilibria.

• There is tension between 'agnosticism' about the ESM and the desire to perform counterfactuals.

• Some additional assumptions are useful to make predictions. A "bottom-up" approach is possible if we can identify and estimate the ESM [i.e., test some additional assumptions?].

• Suppose that we obtain: a) All the equilibria associated with $(\widetilde{\pi}^*_{x^*})$; b) The subsample of observations in the data $\{\mathbf{x}_m\}$ which have the same "type" of equilibria as $\widetilde{\pi}^*_{x^*}$ given $\boldsymbol{\pi}$ for some ω. Let $X(\widetilde{\pi}^*_{x^*})$ be this subsample.

• Suppose we assume (and test?) that $\lambda(.|\mathbf{x}, \boldsymbol{\omega})$ is a smooth function of payoffs within the set of x, ω that have equilibria of the same type as $\widetilde{\pi}^*_{x^*}$ Then we may be able to estimate $\lambda^0(.|\mathbf{x}^*_t)$. Under standard regularity conditions on the bandwidth b_M and the kernel function $K(.)$, the following kernel estimator is a consistent estimator of the counterfactual distribution $\lambda(.|\boldsymbol{\pi}^*_{(\mathbf{x}^*,\omega)})$:

$$\widehat{\lambda}(\tau|\boldsymbol{\pi}^*_{(\mathbf{x}^*,\omega)}) = \frac{\sum_{m=1}^{M} \mathbf{1}\left\{\mathbf{x}_m \in \mathcal{X}(\boldsymbol{\pi}^*_{(\mathbf{x}^*,\omega)})\right\} \ \lambda(\tau|\boldsymbol{\pi}^0_{(\mathbf{x}_m,\omega)}) \ K\left(\frac{\boldsymbol{\pi}^0_{(\mathbf{x}_m,\omega)} - \boldsymbol{\pi}^*_{(\mathbf{x}^*,\omega)}}{b_M}\right)}{\sum_{m=1}^{M} \mathbf{1}\left\{\mathbf{x}_m \in \mathcal{X}(\boldsymbol{\pi}^*_{(\mathbf{x}^*,\omega)})\right\} \ K\left(\frac{\boldsymbol{\pi}^0_{(\mathbf{x}_m,\omega)} - \boldsymbol{\pi}^*_{(\mathbf{x}^*,\omega)}}{b_M}\right)} \quad (2)$$

MORE ON MULTIPLE EQUILIBRIA VS. PAYOFF-RELEVANT HETEROGENEITY

• The finite mixture model is the natural model in one case, in the other case it is an approximation ...

• Under reasonable assumptions (smoothness, support) payoff-relevant common knowledge unobservables would NOT lead to discontinuities in the distribution of outcomes $\widetilde{\Pi}^0(x)$ or in the number of components of the mixture, whereas multiple equilibria might.

• Additional data on payoffs would also should help ...

CONCLUSIONS

• We present new identification results for semiparametric games of incomplete information with two sources of common knowledge unobservables: PR-U and MEQ.

• If the number of action profiles in the game is sufficiently large relative to the number of mixture components then we can identify payoffs and the distributions of the two sources of unobservables under "standard" exclusion restrictions used to identify models without unobserved heterogeneity.

• However, implementation of a sequential identification/estimation approach requires that the researcher be able to match mixture components across games with different values of the excluded variables.

• Without using exclusion restrictions to identify payoffs in all mixture components it does not seem possible to distinguish ex-ante between PR-U and MEQ as explanations of the data.