ECO 2901
EMPIRICAL INDUSTRIAL ORGANIZATION

Lecture 1: Introduction to the course / Market Entry: Complete Information

Victor Aguirregabiria (University of Toronto)

January 9, 2020

Organization of the Course

- **Class Meetings:** ???
- **Office hours:** Tuesdays and Thursdays 2:00-3:00pm
- **Evaluation:** Problem Set (50%); Final Exam (50%)
- **I expect that you:** (1) attend every class meeting; (2) read the papers/material before each lecture; (3) participate in class; (4) go through class notes and understand them; (5) do the problem set on time; (6) prepare for the final exam.

A General Description of this Course

- This courses deals with **models, methods, and empirical applications** in Industrial Organization (IO).
- IO deals with the behavior and competition of firms in markets. In Empirical IO, we use data and models to understand the factors that determine firm behavior in markets.
- This course emphasizes the need to combine **data, models, and econometric techniques** to understand how markets operate**.**

Topics

- The topics covered in the course are divided in three parts that correspond to three general forms of competition between firms.

- **PART I: STATIC MODELS OF MARKET ENTRY AND SPATIAL COMPETITION**
  Competition at the extensive margin: entry and adoption decisions.

- **PART II: DYNAMIC GAMES OF OLIGOPOLY COMPETITION**
  Competition on investment decisions that affect future profits

Topics: Part I

STATIC MODELS OF MARKET ENTRY AND SPATIAL COMPETITION

- 1. Market entry models: complete information
- 2. Market entry models: incomplete information
- 3. Market entry and spatial competition
- 4. Relaxing assumptions on information structure in entry games
- 5. Static entry games with non-equilibrium beliefs

Topics: Part II

DYNAMIC GAMES OF OLIGOPOLY COMPETITION

- 6. Dynamic games of oligopoly competition: Models
- 7. Dynamic games of oligopoly competition: Estimation
- 8. Entry, exit, preemption, and cannibalization in retail industries
- 9. Uncertainty and firms' investment decisions
- 10. Networks and product positioning
- 11. Dynamic games of innovation
- 12. Dynamic games with non-equilibrium beliefs and learning

Today's Lecture

1. Basic concepts on empirical games of market entry
    1.1. What is a model of market entry?
    1.2. Why do we estimate models of market entry?
2. Entry models of complete information

1. Basic concepts in empirical games of market entry

Main features of a model of market entry

- **(1)** The dependent variable is **a firm decision to operate or not in a market**.
  - Entry in a market can be understood in a broad sense
  e.g., entry in an industry; opening a new store; introducing a new product; adopting a new technology; release of a new movie; participate in an auction, etc.

- **(2)** There is a **fixed sunk cost** of entry in the market.

- **(3)** The payoff of being active in the market depends on the number (and the characteristics) of other firms active in the market.
  i.e., the model is a **game**.

Main features of a model of market entry [2]

- Consider a market where there are N firms that potentially may to enter in the market.
- $a_i \in \{0, 1\}$ is a binary variable that represents the decision of firm i of being active in the market $(a_i = 1)$ or not $(a_i = 0)$.
- Profit of not being in the market is zero.
- Profit of being active is: $V_i(n) - F_i$ where $V_i(.)$ is the variable profit, n is the number of firms active, and F_i is the entry cost.
- The number of active firms, n, is endogenous: $n = \sum_{i=1}^{N} a_i$

Main features of a model of market entry [3]

- Under Nash assumption, every firm takes as given the decision of the other firms and makes a decision that maximizes its own profit.
- The best response of firm i under Nash equilibrium is:

$$a_i = \begin{cases} 1 & \text{if} \quad V_i \left(1 + \sum_{j \neq i} a_j\right) - F_i \geq 0 \\ & \\ 0 & \text{if} \quad V_i \left(1 + \sum_{j \neq i} a_j\right) - F_i < 0 \end{cases}$$

where $1 + \sum_{j \neq i} a_j$ represents firm i's Nash-conjecture about the number of active firms.

Two-stage game

- Where does the variable profit $V_i(n)$ comes from?
- A model of market entry is the first stage of a **two stage game**.
- **First stage:** N potential entrants simultaneously choose whether to enter or not in a market.
- **Second stage:** entrants compete (e.g., in prices or quantities) and the profits $V_i(n)$ of each firm are determined.
- **Example**: Cournot competition with linear demand $P = A - B\ Q$ and constant MCs, c, implies:

$$V_i(n) = \frac{1}{B} \left(\frac{A - c}{n + 1} \right)^2$$

Why do we estimate models of market entry?

- **[1] Identification of entry costs parameters.**
  - These parameters are important in the determination of firms profits, market structure, and market power.
  - Fixed costs do not appear in demand or in Cournot or Bertrand equilibrium conditions, so they cannot be estimated in those models.
- **[2] Data on prices and quantities may not be available.**
  - Sometimes all the data we have are firms' entry decisions. These data can reveal information about profits and competition.
- **[3] Dealing with endogenous entry/exit in PF estimation and demand estimation.**

2. Entry models of complete information

Data

- Consider an industry with N potential entrants.
  - For instance, the airline industry: AA, AC, UA, SW, etc.
- We observe M different local markets.
  - Different routes or city-pairs: Chicago-New York, Boston-Atlanta, etc.
- We index firms with i and markets with m.

$$\text{Data} = \{a_{im}, x_{im} : i = 1, 2, ..., N; \ m = 1, 2, ..., M\}$$

a_{im}: entry decision of firm i in market m;
x_{im} : exogenous market characteristics (population, income, input prices); some may vary across firms (the airline presence in the origin and destination airports of the route).

Model

- The (indirect) profit function of a firm is:

$$\Pi_{im} = \begin{cases} \pi_i\left(a_{-im}, x_{im}\right) + \varepsilon_{im} & \text{if} \quad a_{im} = 1 \\ \\ 0 & \text{if} \quad a_{im} = 0 \end{cases}$$

$a_{-im} \equiv \{a_{jm} : j \neq i\}$: Entry decisions of the other firms.

ε_{im} : unobservable to the researcher.

Model [2]

- The specification of the profit function may be:

$$\pi_i \left(a_{-im}, x_{im}\right) = V_i \left(a_{-im}, x_{im}^v\right) - EC_i \left(x_{im}^f\right)$$

- $V_i \left(a_{-im}, x_{im}^v\right)$: Variable profit function. May come from a model of price/quantity competition.

- $EC_i \left(x_{im}^f\right)$: Fixed cost + Entry cost.

- Many times the specification is "less structural" but flexible such as:

$$\pi_i \left(a_{-im}, x_{im}\right) = x_{im}\ \beta_i + \sum_{j \neq i} a_{jm}\ \delta_{ij}$$

- Parameters δ_{ij} are key parameters of interest because they capture competition effects.

Model [3]

- For concreteness, suppose that:

$$\pi_i \left(a_{-im}, x_{im}\right) = x_{im}\ \beta_i + \textstyle\sum_{j \neq i} a_{jm}\ \delta_{ij}$$

Assumptions:
(1) ε_{im} and x_{im} are independent.
(2) ε_{im} is independently distributedover markets with known distribution, e.g., $N(0, \sigma_i)$.

- As usual in discrete choice models, we can identify parameters up to scale: $\left\{ \dfrac{\beta_i}{\sigma_i}, \dfrac{\delta_{ij}}{\sigma_i} : \text{for any } i, j \right\}$.

Nash Equilibrium with complete information

- For the moment, we assume that all the profit functions – and therefore, all the variables $(x_{1m}, x_{2m}, ..., x_{Nm})$ and $(\varepsilon_{1m}, \varepsilon_{2m}, ..., \varepsilon_{Nm})$ – are **common knowledge for the firms**.

- We also assume that the observe decisions in any market m are the result of a Nash equilibrium in that market.

- A **Nash equilibrium** is an N-tuple $(a_{1m},\ a_{2m},\ ...,\ a_{Nm})$ such that for any player i takes the actions of the other players as given and chooses its best response:

$$a_{im} = 1\left\{\ \pi_i\left(a_{-im},\, x_{im}\right) + \varepsilon_{im} \geq 0\ \right\}$$

  where $1\{.\}$ is the indicator function.

Econometric Challenges

- The econometric model is a **system of simultaneous equations** where the dependent variables are binary:

$$
\begin{array}{lcl}
a_{1m} & = & 1\left\{ \ x_{1m}\ \beta_1 + \sum_{j\neq 1} a_{jm}\ \delta_{1j} + \varepsilon_{1m} \geq 0 \ \right\} \\
a_{2m} & = & 1\left\{ \ x_{2m}\ \beta_2 + \sum_{j\neq 2} a_{jm}\ \delta_{2j} + \varepsilon_{2m} \geq 0 \ \right\} \\
\vdots & & \vdots \\
a_{Nm} & = & 1\left\{ \ x_{Nm}\ \beta_N + \sum_{j\neq N} a_{jm}\ \delta_{Nj} + \varepsilon_{Mm} \geq 0 \ \right\}
\end{array}
$$

- There are two main econometric issues:
  (1) endogenous explanatory variables: a_{jm}
  (2) multiple equilibria.

Endogeneity of other players' actions

- The econometric model is a simultaneous equation model where the endogenous variables are binary.

$$\begin{array}{ccl} a_{1m} & = & 1\left\{ \ x_{1m}\ \beta_1 + \sum_{j\neq 1} a_{jm}\ \delta_{1j} + \varepsilon_{1m} \geq 0 \ \right\} \\ \vdots & & \vdots \\ a_{Nm} & = & 1\left\{ \ x_{Nm}\ \beta_N + \sum_{j\neq N} a_{jm}\ \delta_{Nj} + \varepsilon_{Nm} \geq 0 \ \right\} \end{array}$$

- There are two sources of endogeneity or correlation between a_{jm} and ε_{im}:

  (a) Simultaneity;
  (b) Correlation between ε_{im} and ε_{jm}.

Simultaneity

- Consider the model with $N = 2$:

$$
\begin{aligned}
a_1 &= 1\left\{x_1\beta_1 + \delta_1 \ a_2 + \varepsilon_1 \geq 0\right\} \\
a_2 &= 1\left\{x_2\beta_2 + \delta_2 \ a_1 + \varepsilon_2 \geq 0\right\}
\end{aligned}
$$

- In equilibrium, a_1 and a_2 depend on all $(x_1, x_2, \varepsilon_1, \varepsilon_2)$.
- Therefore, a_2 and ε_1 are not independent, and a_2 is an endogenous exp. var. in the best response equation of firm 1.
- In a market entry game (with δ's < 0), simultaneity generates negative correlation between ε_1 and a_2:

$$\uparrow \varepsilon_1 \ \longrightarrow \ \downarrow a_2$$

- **Downward bias in δ_1. Over-estimation of competition effects.**

Correlation between firms' unobservables

$$\begin{aligned} a_1 &= 1\left\{x_1\beta_1 + \delta_1 \; a_2 + \varepsilon_1 \geq 0\right\} \\ a_2 &= 1\left\{x_2\beta_2 + \delta_2 \; a_1 + \varepsilon_2 \geq 0\right\} \end{aligned}$$

- ε_{1m} and ε_{2m} can share a common market effect:

$$\varepsilon_{im} = \omega_m + u_{im}$$

- Positive correlation between ε_{1m} and ε_{2m} generates positive correlation between a_{2m} and ε_{1m}:

$$\uparrow \omega \quad \rightarrow \quad \text{both} \uparrow \varepsilon_1 \quad \text{and} \quad \uparrow a_2)$$

- **Upward bias in δ_1.Under-estimation of competition effects.**

How do we deal with this endogeneity problem?

- The intuition is that we could use IV:
  if there are variables in x_{2m} that do not enter in x_{1m}, we can use the variables as IVs for the endogenous regressor a_{2m} in the estimation of the best response equation for firm 1.
- Unfortunately, IV (or the Rivers-Vuong method) is not consistent in binary choice models with endogenous binary exp. vars.
- But there are consistent estimation methods that are in the same spirit as IV / Control function approaches.

How do we deal with this endogeneity problem? [2]

- Alternatively, we could consider Maximum Likelihood estimation (MLE) to deal with endogeneity.
- Maximize the log-likelihood function

$$l(\theta) = \sum_{m=1}^{M} \ln \Pr(a_{1m}, a_{2m}, ..., a_{Nm} \mid x_m, \ \theta)$$

- Unfortunately, the model have multiple equilibria and the **likelihood is NOT a function** but a **correspondence.**
- Standard Max. Likelihood is infeasible.

Multiple equilibria

- Consider a two-player entry game. The payoff matrix is:

		Firm 2	
		$a_2 = 0$	$a_2 = 1$
Firm 1	$a_1 = 0$	$[0\ ,\ 0]$	$[0\ ,\ \pi_2^{MON}]$
	$a_1 = 1$	$[\pi_1^{MON}, 0]$	$[\pi_1^{DUO}\ ,\ \pi_1^{DUO}]$

where π_i^{MON} and π_i^{DUO} represents firm i's profit under monopoly and duopoly, respectively.

- Choices $(a_1, a_2) = (0, 1)$ and $(1, 0)$ are both NE if:

$$\pi_1^{MON} \geq 0,\ \pi_2^{MON} \geq 0,\ \pi_1^{DUO} < 0, \text{ and } \pi_2^{DUO} < 0$$

Multiple equilibria

- Consider the model:

$$a_1 = 1\{\ x_1\beta_1 + \delta_1\ a_2 + \varepsilon_1 \geq 0\ \}$$

$$a_2 = 1\{\ x_2\beta_2 + \delta_2\ a_1 + \varepsilon_2 \geq 0\ \}$$

- The reduced form of the model is:

$$\{x_1\beta_1 + \varepsilon_1 < 0\}\ \&\ \{x_2\beta_2 + \varepsilon_2 < 0\} \Rightarrow (a_1, a_2) = (0, 0)$$

$$\{x_1\beta_1 + \delta_1 + \varepsilon_1 \geq 0\}\ \&\ \{x_2\beta_2 + \delta_2 + \varepsilon_2 \geq 0\} \Rightarrow (a_1, a_2) = (1, 1)$$

$$\{x_1\beta_1 + \delta_1 + \varepsilon_1 < 0\}\ \&\ \{x_2\beta_2 + \varepsilon_2 \geq 0\} \Rightarrow (a_1, a_2) = (0, 1)$$

$$\{x_1\beta_1 + \varepsilon_1 \geq 0\}\ \&\ \{x_2\beta_2 + \delta_2 + \varepsilon_2 < 0\} \Rightarrow (a_1, a_2) = (1, 0)$$

Multiple equilibria (2)

- We can see that if

$$\{ -x_1\beta_1 < \varepsilon_1 \leq -x_1\beta_1 - \delta_1 \} \text{ AND } \{ -x_2\beta_2 < \varepsilon_2 \leq -x_2\beta_2 - \delta_2 \}$$

- Then, the model has two equilibria:

$$(a_1, a_2) = (0, 1) \text{ AND } (a_1, a_2) = (1, 0)$$

- The model does not provide a unique prediction for the probabilities $\Pr((a_1, a_2) = (0, 1) \mid \theta)$ and $\Pr((a_1, a_2) = (1, 0) \mid \theta)$. There is a likelihood function.

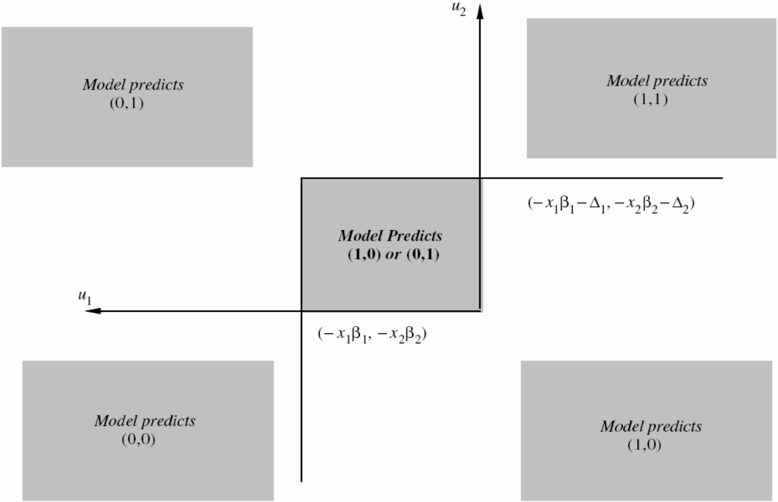


FIGURE 1

Incomplete model with multiple equilibria

Multiple equilibria (3)

- With $N > 2$ there are more possibilities for multiple equilibria.
- In general, for any value of the observables X and the parameters (β, δ), there are hyper-rectangles in the space of $(\varepsilon_1, \varepsilon_2, ..., \varepsilon_J)$ that imply multiple equilibrium outcomes for $(a_1, a_2, ..., a_J)$.
- The model does not have a likelihood function $\ln \Pr(a_1, a_2, ..., a_J | X, \beta, \delta)$ but a likelihood correspondence.
- Similarly, other sample criterion functions (e.g., GMM) based on $\Pr(a_1, a_2, ..., a_J | X, \beta, \delta)$ will be correspondences.
- How to define/construct a consistent estimator in this context? Is the model identified?

Tradional approach: restrictions to avoid ME

- **(a) Homogeneous firms [Bresnahan & Reiss, JPE-91].** $\beta_i = \beta$ and $\delta_{ij} = \delta$ for any i, j. If $\delta \leq 0$, this model has a unique equilibrium.
- **(b) Triangular system [Heckman, ECMA-78]** $\delta_{1j} = 0$ for any j; $\delta_{2j} = 0$ for any $j > 1$; $\delta_{3j} = 0$ for any $j > 2$; ... This model has a unique equilibrium.
- **(c) Restrictions on order of entry [Berry, ECMA-93].** $\delta_{ij} = \delta_i$ such that profit is $x_{im}\ \beta_i + \delta_i\ n_m + \varepsilon_{im}$. The multiple equilibria have the same value for the number of entrants n_m. Then, the identity of the entrants is solved with an assumption on the order of entry: firm 1 enters first; then firm 2; etc

Multiple equilibria & Identification

- The common wisdom was that Multiple Equilibria was an identification problem and that we need to impose restrictions in the model to eliminate Multiple Equilibria and obtain identification.
- **This common wisdom was wrong**. Multiple equilibria and Identification are two different problems and, in most models with multiple equilibria, we do not need to impose equilibrium uniqueness to identify structural parameters.
- We start with a general but stylized framework to study multiple equilibria and identification.

Multiple equilibria & Identification: A general framework

- Let θ be the vector of parameters of the model. Let P represent the vector that represents the probability distribution $\Pr(a_1, a_2, ..., a_N | x)$, i.e., the prediction of the model.
- Let Θ be the parameter space. And let $\mathcal{P}$ be the space of the probabilities $\Pr(a_1, a_2, ..., a_N | x)$.
- A model can be described as a mapping , $M(\theta) : \Theta \rightarrow \mathcal{P}$.
- **Multiple equilibria** means that the mapping $M(.)$ is a correspondence.
- **No identification** means that, at the true P^0 in the population, the inverse mapping M^{-1} is a correspondence.

Multiple equilibria & Identification (3)

Figure 2
Just Identification

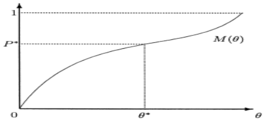


Figure 3
Non Identification

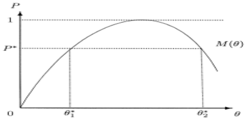
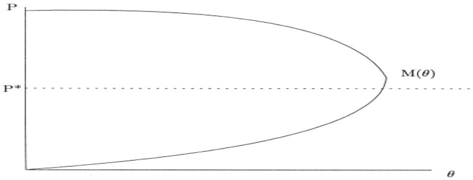


Figure 4
Multiple Equilibria



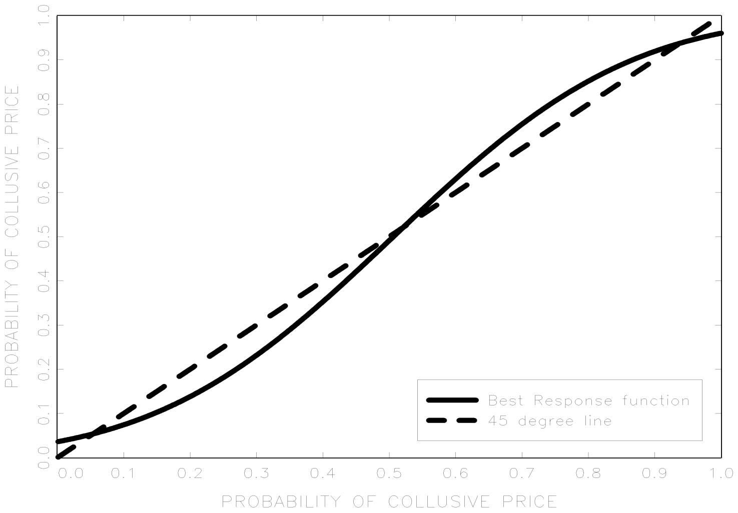
Multiple equilibria & Identification (4)

- **Example:** Consider a simple equilibrium model where both θ and P are scalars, the model $M(\theta)$ is defined as the set of probabilities P that solves the fixed point problem:

$$P = \Phi\left(-1.8 + \theta \; P\right)$$

where $\Phi\left(.\right)$ is the CDF of the standard normal.

- For instance, for $\theta_0 = 3.5$, the set of equilibria is : $M(\theta_0) = \{\ P^{(A)}(\theta_0) = 0.054,\ P^{(B)}(\theta_0) = 0.551,\ P^{(C)}(\theta_0) = 0.924\}$.


PROBABILITY OF COLLUSIVE PRICE
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
PROBABILITY OF COLLUSIVE PRICE
Best Response function
45 degree line

Identification and multiple equilibria Example

- Let P_0 the probability that we observe in the population. P_0 can be either P_A or P_B or P_C, and the researcher does not know it.
- We now show that θ_0 is uniquely identified given P_0
- P_0 is an equilibrium associated with θ_0 and therefore:

$$P_0 = \Phi\left(-1.8 + \theta_0 \; P_0\right)$$

- Since $\Phi(.)$ is an invertible function, we have that:

$$\theta_0 = \frac{\Phi^{-1}\left(P_0\right) + 1.8}{P_0}$$

- Given $P_0 \neq 0$, θ_0 is uniquely determined (identified).

Some general identification results: Lemma 1

- Consider an econometric model that implies the following restriction:

$$P(x) = g\left(x'\theta\right)$$

  - $P(x)$ is conditional moment or probability (e.g., $P(x) = \Pr(Y = 1|X = x)$) that is identified from the data;
  - $g\,(.)$ is a strictly monotonic and continuous function, and it is known to the researcher;
  - x is a vector of exogenous variables. $\mathbb{E}\left[xx'\right]$ is full column rank.

- Then, the vector of parameters θ is identified, i.e.,

$$\theta = \left(\mathbb{E}\left[xx'\right]\right)^{-1}\ \mathbb{E}\left[x\ g^{-1}\left(P(x)\right)\right]$$

Semiparametric extension of Lemma 1

- Consider an econometric model that implies: $P(x) = g\left(x'\theta\right)$
- with the same interpretation as in Lemma 1 but now:
  - $g(.)$ is unknown to the researcher except for a location restriction, e.g., $g(0) = 0$.
  - θ_1 is restricted to be 1. x_1 is a continuos variable with support $\mathbb{R}$
- Then, θ and $g(.)$ are identified. Matzkin (Econometrica, 1992).

Some general identification results: Lemma 2

- Consider an econometric model that implies:

$$P(x_1, x_2) = g_1\left(x_1'\theta_1\right) + g_2\left(x_2'\theta_2\right)$$

- $P(x_1, x_2)$ is a moment or probability identified from the data;
- $g_1(.)$ and $g_2(.)$ are strictly monotonic and continuous functions that are known to the researcher;
- x_1, x_2 are exogenous variables, linearly independent.
- Normalization (fixing intercept):$x_1^0 : g(x_1^{0\prime}\theta_1) = 0$.

- Then, (γ_1,γ_2) is identified:

$$\theta_1 = \left(\mathbb{E}\left[\left(x_1 - x_1^0\right)\ \left(x_1 - x_1^0\right)'\right]\right)^{-1}$$

$$\mathbb{E}\left[\left(x_1 - x_1^0\right)\ g_1^{-1}\left(P(x_1, x_2^0) - P(x_1^0, x_2^0)\right)\right]$$

Semiparametric extension of Lemma 2

- Consider an econometric model that implies:

$$P(x_1, x_2, z) = g_1\left(x_1'\theta_1\right) + g_2\left(x_2'\theta_2\right)$$

- Same restriction as in Lemma 2 but now:
  - $g_1(.)$ and $g_2(.)$ are unknown to the researcher expect for $g_1(0) = g_2(0) = 0$
  - $\theta_1 = \theta_2 = 1$. x_{11} and x_{21} are continuous variables with support $\mathbb{R}$

- Then, functions $g_1(.)$ and $g_2(.)$ and parameters θ_1, θ_2 are identified. Matzkin (Econometrica, 1992).

Point - Identification in discrete choice game

- Consider the 2-player binary choice game:

$$
\begin{aligned}
a_1 &= 1\{\ x_1'\beta_1 + a_2 x_1'\delta_1 + \varepsilon_1 \geq 0\ \} \\
a_2 &= 1\{\ x_2'\beta_2 + a_1 x_2'\delta_2 + \varepsilon_2 \geq 0\ \}
\end{aligned}
$$

- Suppose that $\delta_1 \leq 0$, $\delta_2 \leq 0$, and $(\varepsilon_1, \varepsilon_{2m})$ are independent (for the moment) standard normals.

$$
\begin{aligned}
P(0,0|x_1,x_2) &= \Phi(-x_1\beta_1)\ \Phi(-x_2\beta_2) \\
P(1,1|x_1,x_2) &= [1-\Phi(-x_1'[\beta_1+\delta_1])]\ \ [1-\Phi(-x_2'[\beta_2+\delta_2])]
\end{aligned}
$$

Point - Identification [2]

- The first equation implies:

$$\ln P\left(0,0|x_1,x_2\right) = \ln \Phi\left(-x_1\beta_1\right) + \ln \Phi\left(-x_2\beta_2\right)$$

- By Lemma 2, we have identification of β_1, β_2.
- The second equation implies:

$$\ln P\left(1,1|x_1,x_2\right) = \begin{array}{l} \ln\left[1-\Phi\left(-x_1'\left[\beta_1+\delta_1\right]\right)\right] \\ +\ln\left[1-\Phi\left(-x_2'\left[\beta_2+\delta_2\right]\right)\right] \end{array}$$

- By Lemma 2, we have identification of $[\beta_1 + \delta_1]$, $[\beta_2 + \delta_2]$.
- Combining the two conditions we have identification of β_1, β_2, δ_1, δ_2.

Point - Identification [3]

- The previous model includes several restrictions that can be relaxed and still keeping point identification.

$$cov\left(\varepsilon_1, \varepsilon_2\right) = 0$$

- Note that we are not exploiting an important restriction of the model.

- The model implies an upper bound and a lower bound on the probability $P\left(0, 1 | x_1, x_2, z\right)$.

$$L(x_1, x_2, z; \theta) \leq P\left(0, 1 | x_1, x_2, z\right) \leq L(x_1, x_2, z; \theta)$$

  where the bounds $L(x_1, x_2, z; \theta)$ and $U(x_1, x_2, z; \theta)$ are known function (up to θ) provided by the model.

- We now study how to incorporate these restrictions in an efficient estimation of the model.

Complete info games: Estimation

- Tamer (REStud, 2003) and Ciliberto & Tamer (ECMA, 2009).
- Consider the discrete choice game:

$$\begin{array}{ccl} a_{1m} & = & 1\{\ x_{1m}\ \beta_1 + \sum_{j\neq 1} a_{jm}\ \delta_{1j} + \varepsilon_{1m} \geq 0\ \} \\ \vdots & & \vdots \\ a_{Nm} & = & 1\{\ x_{Nm}\ \beta_N + \sum_{j\neq N} a_{jm}\ \delta_{Nj} + \varepsilon_{Nm} \geq 0\ \} \end{array}$$

- Let $P_0(\mathbf{a}_m|\mathbf{x}_m) \equiv P_0(a_{1m}, ..., a_{Nm}|x_{1m}, ..., x_{Nm})$ be the true probabillity in the population.
- $P_0(\mathbf{a}_m|\mathbf{x}_m)$ is nonparametrically identified from the data.

Complete info: Estimation [2]

- For every data point $(\mathbf{a}_m, \mathbf{x}_m)$ and vector of parameters θ, the model implies a lower bound (strictly greater than 0) and an upper bound (strictly lower than 1) for the probability $P_0(\mathbf{a}_m|\mathbf{x}_m)$:

$$L(\mathbf{a}_m|\mathbf{x}_m;\theta) \leq P_0(\mathbf{a}_m|\mathbf{x}_m) \leq U(\mathbf{a}_m|\mathbf{x}_m;\theta)$$

- The bound probabilities $L(\mathbf{a}_m|\mathbf{x}_m;\theta)$ and $U(\mathbf{a}_m|\mathbf{x}_m;\theta)$ are functions that can be obtained by integrating over the distribution of ε in the model.

Complete info: Estimation [3]

- For instance, for the two-player game:

$$L(0,0|\mathbf{x};\theta) = U(0,0|\mathbf{x};\theta)$$
$$= \Pr(\varepsilon_1 < -x_1\beta_1 \;\&\; \varepsilon_2 < -x_2\beta_2)$$

$$L(1,1|\mathbf{x};\theta) = U(1,1|\mathbf{x};\theta)$$
$$= \Pr(\varepsilon_1 \geq -x_1\beta_1 - \delta_1 \;\&\; \varepsilon_2 \geq -x_2\beta_2 - \delta_2)$$

$$U(0,1|\mathbf{x};\theta) = \Pr(\varepsilon_1 < -x_1\beta_1 - \delta_1 \;\&\; \varepsilon_2 \geq -x_2\beta_2);$$
$$L(0,1|\mathbf{x};\theta) = U(0,1|\mathbf{x};\theta) - \text{"Ambiguous rectangle"}$$

$$U(1,0|\mathbf{x};\theta) = \Pr(\varepsilon_1 \geq -x_1\beta_1 \;\&\; \varepsilon_2 < -x_2\beta_2 - \delta_2);$$
$$L(1,0|\mathbf{x};\theta) = U(1,0|\mathbf{x};\theta) - \text{"Ambiguous rectangle"}$$

Estimation: Tamer (2003)

- Tamer (2003) proposes the following Likelihood criterion function and estimator:

$$\widehat{\theta}_{MLE} = \arg\max_{\theta} \sum_{m=1}^{M} \ln P^{0}\left(\mathbf{a}_m \mid \mathbf{x}_m\right)$$

$$\text{subject to:} \quad L(\mathbf{a}_m|\mathbf{x}_m;\theta) \leq P_0(\mathbf{a}_m|\mathbf{x}_m) \leq U(\mathbf{a}_m|\mathbf{x}_m;\theta) \quad \text{for any } m$$

- That can be represented as $(\widehat{\theta}_{MLE}, \widehat{\lambda}_{MLE}) = \arg\max_{\theta,\lambda} Q(\theta,\lambda)$, with

$$\begin{aligned} Q(\theta,\lambda) = & \sum_{m=1}^{M} \ln P^{0}\left(\mathbf{a}_m \mid \mathbf{x}_m\right) \\ & +\lambda_m^U \max\left\{0 \; ; \; \ln P^{0}\left(\mathbf{a}_m \mid \mathbf{x}_m\right) - \ln U(\mathbf{a}_m|\mathbf{x}_m;\theta)\right\} \\ & +\lambda_m^L \max\left\{0 \; ; \; \ln L\left(\mathbf{a}_m \mid \mathbf{x}_m\right) - \ln P^{0}(\mathbf{a}_m|\mathbf{x}_m;\theta)\right\} \end{aligned}$$

Estimation: Ciliberto & Tamer (2009)

- Tamer (2003)'s criterion function is highly dimensional because the Kuhk-Tucker multipliers.
- Chernozukov, Hong, and Tamer (2007), and Ciliberto and Tamer (2009) propose the following criterion (penalty) function and estimator:

$$\widehat{\theta} = \arg\min_{\theta} \quad \begin{array}{l} \sum_{m=1}^{M} \max\left\{0\ ;\ P^0\left(\mathbf{a}_m \mid \mathbf{x}_m\right) - U(\mathbf{a}_m|\mathbf{x}_m;\theta)\right\}^2 \\ + \sum_{m=1}^{M} \max\left\{0\ ;\ L(\mathbf{a}_m|\mathbf{x}_m;\theta) - P^0\left(\mathbf{a}_m \mid \mathbf{x}_m\right)\right\}^2 \end{array}$$

Estimation: Ciliberto & Tamer (2009)

- The method proceeds in two-steps.
- **Step 1:** Nonparametric estimator of $P^0\left(\mathbf{a}_m \mid \mathbf{x}_m\right)$ at every data point $\left(\mathbf{a}_m, \mathbf{x}_m\right)$.
- **Step 2:** Given estimates $\widehat{P}_0(\mathbf{a}_m \mid x_m)$, we estimate of θ by minimizing the penalty function:

$$\widehat{\theta} = \arg\min_{\theta \in \Theta} \; Q\left(\theta, \widehat{\mathbf{P}}_0\right)$$

with

$$\begin{aligned} Q\left(\theta, \widehat{\mathbf{P}}_0\right) &= \sum_{m=1}^{M} \max\left\{ L(\mathbf{a}_m \mid x_m; \theta) - \widehat{P}_0(\mathbf{a}_m \mid x_m) \; , \; 0 \right\}^2 \\ &+ \sum_{m=1}^{M} \max\left\{ \widehat{P}_0(\mathbf{a}_m \mid x_m) - U(\mathbf{a}_m \mid x_m; \theta) - \; , \; 0 \right\}^2 \end{aligned}$$