ECO 2901
EMPIRICAL INDUSTRIAL ORGANIZATION

Lecture 5:
Market entry models

Victor Aguirregabiria (University of Toronto)

February 7th, 2019

Today's Lecture (Topic 5)

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

3. Entry models of incomplete 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** associated with being active 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?
- It is useful to see **a model of market entry as part of a two stage game**.
- In a **First stage**, N potential entrants simultaneously choose whether to enter or not in a market.
- In a **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] Explaining market structure.**
  - Why different industries (and different markets within the same industry) have different number of active firms?
- **[2] 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.
- **[3] 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 about the nature of competition.
- **[4] Dealing with endogenous entry/exit in PF estimation and endogenous product presence in demand estimation.**

2. Entry models of complete information

Heterogeneous firms & complete information

- **[Data]** Consider an industry with N potential entrants. For instance, the airline industry.
- We observe M different local markets, e.g., different routes, Toronto-New York, Montreal-Washington, etc.
- We index firms with i and markets with m.

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

- x_{im} could include only market characteristics, but it may include also market-firm characteristics.

Heterogeneous firms & complete information (2)

- **[Model]** The (indirect) profit function of a firm is:

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

- $a_{-im} \equiv \{ a_{jm} : j \neq i \}$; and ε_{im} is an unobservable that is common knowledge to all firms (complete information).

- A **Nash equilibrium** is an N-tuple $a_m^* = \left(a_{1m}^*, a_{2m}^*, ..., a_{Nm}^* \right)$ such that for any player i:

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

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

Heterogeneous firms & complete information (3)

- The specification of the profit function may be:

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

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

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

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

Heterogeneous firms & complete information (4)

- For concreteness, suppose that:

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

- ε_{im} is independent of x_{im} and independently distributed over markets with a known distribution, e.g., $N(0, \sigma_i)$.

- We are interested in the estimation of the vector of parameters $\theta = \left\{ \dfrac{\beta_i}{\sigma_i}, \dfrac{\delta_{ij}}{\sigma_i} : \text{for any } i, j \right\}$.

- 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.
- The simultaneous equations are the players' best response functions:

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

- There are two sources of endogeneity or correlation between a_{jm} and ε_{im}:
  (a) Simultaneity;
  (b) Correlation between ε_{im} and ε_{jm}.

Endogeneity of other players' actions (2)

- **Simultaneity.** Consider the model with $N = 2$:

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

- In equilibrium (i.e., in the reduced form of the model) a_2 depends on $(x_1, x_2, \varepsilon_1, \varepsilon_2)$.
- 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 an entry game, simultaneity generates positive correlation between error ε_1 and a_2 ($\uparrow \varepsilon_1 \rightarrow \downarrow a_2$) —> **downward bias in δ_1, i.e., we over-estimate competition effects.**

Endogeneity of other players' actions (2)

- **Correlation between firms' unobservables.**

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

- For instance, ε_{im} can have a common market effect:

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

- Positive correlation between ε_{1m} and ε_{2m} generates negative correlation between a_{2m} and ε_{1m} ($\uparrow \omega \rightarrow \uparrow \varepsilon_1$ and $\uparrow a_2$): **upward bias in δ_1, i.e., we under-estimate competition effects.**

Endogeneity of other players' actions (3)

- **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 we will end up using a method "in the spirit of IV".

Endogeneity of other players' actions (4)

- **How do we deal with this endogeneity problem?**
- Alternatively, we could suggest using Maximum Likelihood as the approach 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 this **likelihood function** does not exist, i.e., it is a **likelihood correspondence.**
- Standard Max. Likelihood is unfeasible.

Multiple equilibria

- Consider the model:

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

- The reduced form of the model is:

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

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.

TAMER BIVARIATE MODEL WITH MULTIPLE EQUILIBRIA 149

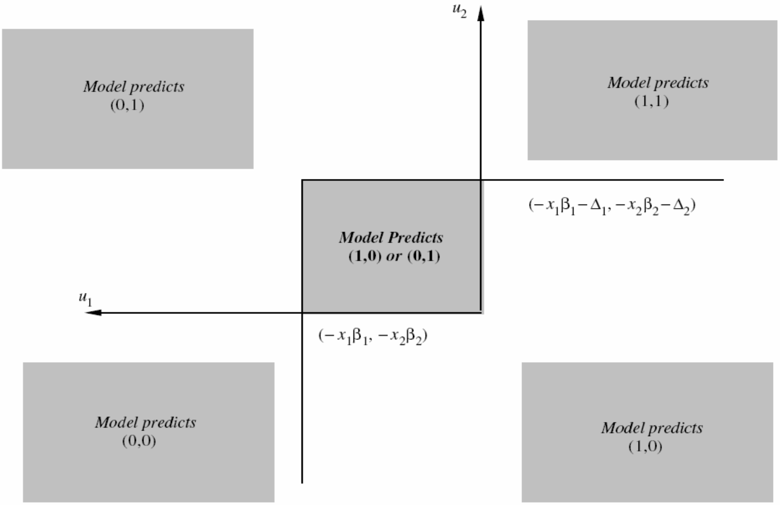


FIGURE 1
Incomplete model with multiple equilibria

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?

Avoiding multiple equilibria

- The first approaches to "deal with" Multiple Equilbria in the estimation of games consisted in avoiding the problem: impose restrictions on the model to avoid multiple equilibria.
- **(a) Homogeneous firms [Bresnahan & Reiss, JPE-91].** $x_{im}\ \beta_i + \sum_{j \neq i} a_{jm}\ \delta_{ij} + \varepsilon_{im} = x_m\ \beta + \delta\ n_m + \varepsilon_m$. If $\delta \leq 0$, this model has a unique equilibrium; a unique $\Pr(n_m | x_m)$.
- **(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].** Model with utilities $x_{im}\ \beta_i + \delta_i\ n_m + \varepsilon_{im}$ is such that in a region of ε's with multiple equilibria, all of them have the same 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

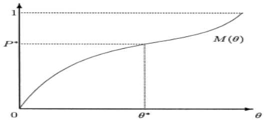


FIGURE 3
Non Identification

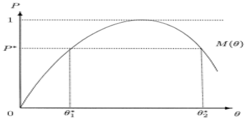
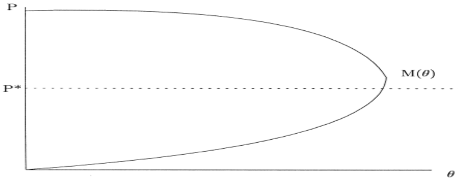


FIGURE 4
Multiple Equilibria



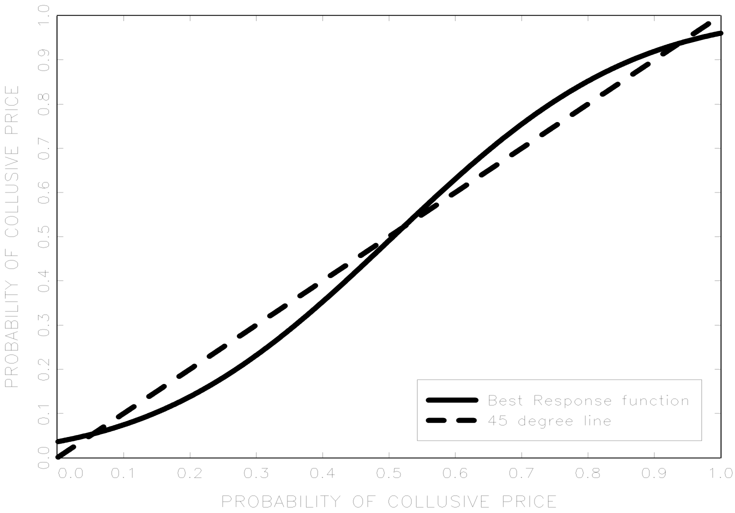
Multiple equilibria & Identification (4)

- **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;
- $g\left(.\right)$ 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:

$$\begin{aligned} \theta_1 &= \left(\mathbb{E}\left[\left(x_1 - x_1^0\right)\ \left(x_1 - x_1^0\right)'\right]\right)^{-1} \\ & \quad \mathbb{E}\left[\left(x_1 - x_1^0\right)\ g_1^{-1}\left(P(x_1, x_2^0) - P(x_1^0, x_2^0)\right)\right] \end{aligned}$$

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:

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

- 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:

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

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

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

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

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

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

3. Entry models of incomplete information

Entry models with incomplete information

- A market with N potential entrants. If firm i is active in the market ($a_{im} = 1$), its profit is:

$$\Pi_{im} = x_{im}\ \beta_i + \omega_m + \varepsilon_{im} + \sum_{j \neq i} \delta_{ij}\ a_{jm}$$

- $\mathbf{x}_m = (x_{1m}, x_{2m}, ..., x_{Nm})$ is **common knowledge** to firms and observable to the researcher.

- ω_m is is **common knowledge** to firms but unobservable to the researcher.

- ε_{im} is **private information** of firm i, independent across firms, independent of $(\mathbf{x}_m, \omega_m)$, and unobservable to the researcher. For concreteness, $\varepsilon_{im} \sim iid\ N(0, 1)$.

Bayesian Nash Equilibrium (BNE)

- The information of firm i is $(\mathbf{x}_m, \omega_m, \varepsilon_{im})$.
- A player's strategy depends on the variables in his information set.
- Let $\alpha_i(\mathbf{x}_m, \omega_m, \varepsilon_{im})$ be a strategy function for firm i such that $\alpha_i : X \times \Omega \times \mathbb{R} \rightarrow \{0, 1\}$.
- We can define a **Bayesian Nash Equilibrium (BNE)** in terms of the strategy functions $\alpha_i(\mathbf{x}_m, \omega_m, \varepsilon_{im})$.
- It will be convenient to represent a BNE in terms of **Choice Probabilities**.

Conditional choice probabilities (CCPs)

- **Players' choice probabilities.**
- Given a strategy function $\alpha_i(\mathbf{x}_m, \omega_m, \varepsilon_{im})$, the associated choice probability is the result of integration this strategy function over the distribution of the player's private information

$$P_i(\mathbf{x}_m, \omega_m) \equiv \int \alpha_i(\mathbf{x}_m, \omega_m, \varepsilon_{im}) \; d\Phi_i(\varepsilon_{im})$$

- It represents the expected behavior of player i from the point of view of the other players who do not know the private information ε_{im}.

BNE in terms of CCPs

- Firm i's expected profit is:

$$\Pi_{im}^{e} = x_{im}\ \beta_i + \omega_m + \varepsilon_{im} + \sum_{j \neq i} \delta_{ij}\ P_j(\mathbf{x}_m, \omega_m)$$

- Firm i's best response is:

$$\{a_{im} = 1\} \Leftrightarrow \left\{ \varepsilon_{im} < x_{im}\ \beta_i + \omega_m + \sum_{j \neq i} \delta_{ij}\ P_j(\mathbf{x}_m, \omega_m) \right\}$$

- And firm i's **best response probability function** is:

$$\Pr\left(a_{im} = 1 | \mathbf{x}_m, \omega_m, P_{j \neq i}\right) \quad = \quad \Phi\left(x_{im}\ \beta_i + \omega_m + \sum_{j \neq i} \delta_{ij}\ P_j(\mathbf{x}_m, \omega_m) \right)$$

BNE in terms of CCPs [2]

- Given $(\mathbf{x}_m, \omega_m)$, a **Bayesian Nash equilibrium (BNE)** is a vector of probabilities $\mathbf{P}(\mathbf{x}_m, \omega_m) \equiv \{P_i(\mathbf{x}_m, \omega_m) : i = 1, 2, ..., N\}$ that solves the fixed point problem:

$$P_i(\mathbf{x}_m, \omega_m) \quad = \quad \Phi\left(x_{im} \ \beta_i + \omega_m + \sum_{j \neq i} \delta_{ij} \ P_j(\mathbf{x}_m, \omega_m) \right)$$

- In a BNE, firms' beliefs about their opponents' entry probabilities are the opponents' best responses to their own beliefs.
- By Brower FP Theorem, the model has at least one BNE.
- The equilibrium may not be unique.

Entry models with incomplete information

- We study the following topics in the econometrics discrete choice games of I.I.

1. Identification
2. Problems with standard estimation methods (MLE, GMM)
3. Two-step and K-step Pseudo ML estimators

Entry models with incomplete information (6)

- The first entry models and empirical applications with incomplete information assumed that the only unobservables for the researcher where the private information variables ε_{im}. That is, they assume that $\omega_m = 0$.

- This restriction simplifies very substantially the identification and estimation of this type of models.

- However, it is quite unrealistic and it can be easily rejected by the data. This restriction implies that:

$$\Pr(a_{1m}, a_{2m}, ..., a_{Nm} \mid x_m) = \prod_{i=1}^{N} \Pr(a_{im} \mid x_m)$$

- Ignoring ω_m can induce substantial biases in the estimation of the parameters δ that measure players' strategic interactions.

Identification: Assumptions

- Suppose that we have a random sample of markets and we observe:

$$\{\mathbf{x}_m, a_{im} : m = 1, 2, ..., M;\ i = 1, 2, ..., N\}$$

- **Assumption 1:** ω_m is independent of $\mathbf{x}_m$ and it has a finite mixture distribution: $\omega_m \in \{c_1, c_2, ..., c_L\}$ with $\Pr(\omega_m = c_k) \equiv \lambda_k$.

- **Assumption 2:** $\{P_i^0(\mathbf{x}_m, \omega_m)\}$ is such that two markets, m and m', with the same common knowledge variables $(\mathbf{x}_m, \omega_m)$ select the same type of equilibrium.

- Under these assumptions, and standard rank conditions, we can identify the model parameters θ.

Identification: Step 1

- The proof of identification proceeds in two steps.
- First, we show that the probabilities $P_i^0(\mathbf{x}, \omega)$ are nonparametrically identified.
- This is obvious in the model with $\omega_m = 0$ because:

$$P_i^0(\mathbf{x}) = \mathbb{E}(a_{im} \mid \mathbf{x}_m = \mathbf{x})$$

- In the model with $\omega_m = 0$, the nonparametric identification of $P_i^0(\mathbf{x}, \omega)$ is based on the identification of **nonparametric finite mixture model**.

Identification: Step 1 (Nonparametric finite mixture)

- With common knowledge unobs, ω_m, the estimation of choice probabilties is more complicated. But there are many recent results (Hall & Zhou, 2003, Kasahara & Shimotsu, 2009, 2013).

- The model is:

$$\Pr(\mathbf{a}_m = a \mid \mathbf{x}_m = x) = \sum_{k=1}^{L} \lambda_k \left[\prod_{i=1}^{N} P_i^0(x, c_k) \right]$$

- Different results show the NP identification of $\lambda_k' s$ and $P_i^0(x, c_k)$'s.

- The key identification assumption is the independence of players' a_{im} conditional on $(\mathbf{x}_m, \omega_m)$.

Identification: Step 2

- Given $P_i^0(\mathbf{x}_m, \omega)$ for every market m and type ω, we can represent our model as a linear regression-like model:

$$\Phi^{-1}\left(P_i^0(\mathbf{x}_m)\right) = x_{im}\ \beta_i + \sum_{j \neq i} \delta_{ij}\ P_j^0(\mathbf{x}_m)$$

- Define $Y_{im} \equiv \Phi^{-1}\left(P_i^0(\mathbf{x}_m)\right)$; $Z_{im} \equiv \left(x_{im}, P_j^0(\mathbf{x}_m) : j \neq i\right)$; and $\theta_i \equiv \left(\beta_i, \delta_{ij} : j \neq i\right)$. Then,

$$Y_{im} = Z_{im}\ \theta_i + e_{im}$$

- θ_i is identified iff $\mathbb{E}\left(Z_{im}' Z_{im}\right)$ has full column rank. For this, we need exclusion restrictions, i.e., player specific variables in x_{im}. [or functional form identification].

Maximum likelihood estimation (1)

- Suppose that the **only unobservables for the researcher are the private information variables** ε_{im}.

- If the model had unique equilibrium, then we could estimate θ by MLE:

$$\hat{\theta}_{MLE} = \arg\max_{\theta} \sum_{m=1}^{M} \sum_{i=1}^{N} a_{im} \ \ln P_i(\mathbf{x}_m, \theta) + (1 - a_{im}) \ln(1 - P_i(\mathbf{x}_m, \theta))$$

  where $P_i(\mathbf{x}_m, \theta)$ is the unique equilibrium probability of player i given $(\mathbf{x}_m, \theta)$.

- However, when the model has multiple equilibria, the likelihood is not a function but a correspondence.

Maximum likelihood estimation (2)

- We still can define the MLE in a model with multiple equilibria.
- For any (θ, P), define the **extended likelihood function** is:

$$
\begin{aligned}
Q(\theta, P) &= \sum_{m=1}^{M} \sum_{i=1}^{N} a_{im} \ \ln \Phi \left(x_{im} \ \beta_i + P_{-i}(\mathbf{x}_m) \ \delta_i \right) \\
&+ \ (1 - a_{im}) \ \ln \Phi \left(-x_{im} \ \beta_i - P_{-i}(\mathbf{x}_m) \ \delta_i \right)
\end{aligned}
$$

where $P_{-i}(\mathbf{x}_m) = \{P_j(\mathbf{x}_m) : j \neq i\}$ and $\delta_i = \{\delta_{ij} : j \neq i\}$.

- This is a well-defined **function** for any values of (θ, P).

Maximum likelihood estimation (3)

- The MLE is defined as:

$$\hat{\theta}_{MLE} = \arg\max_{\theta} \left\{ \begin{array}{l} \max_{P} \;\; Q(\theta, P) \\ \text{subject to:} \\ P_i(x_m) = \Phi\left(x_{im}\ \beta_i + P_{-i}(\mathbf{x}_m)\ \delta_i\right) \ \text{ for every } i,m \end{array} \right\}$$

- This estimator has all the good properties of MLE under standard regularity conditions.
- However, it can be very difficult to implement in practice.
- It requires optimization with respect to P which is a high dimensional vector. Many local maxima.
- Judd and Su (2012). MPEC method.

Two-step Pseudo ML estimation (1)

- Let $\mathbf{P}^0$ be the vector of choice probabilities (for each i and x_m) in the population.
- It is possible to show that the true θ^0 uniquely maximizes $Q_\infty(\theta, \mathbf{P}^0)$.
- The two-step Pseudo ML estimator of θ^0 is defined as the sample counterpart of θ^0.
- That is:

$$\hat{\theta} = \arg\max Q_M(\theta, \widehat{\mathbf{P}^0})$$

where $\widehat{\mathbf{P}^0}$ is a consistent nonparametric estimator of P^0.

Two-step Pseudo ML estimation (2)

- The first-step can be just a Nadaraya-Watson Kernel estimator of the choice probabilities: $\widehat{P}_i(\mathbf{x})$.
- The second step is just a standard Probit model with likelihood:

$$\sum_{m=1}^{M}\sum_{i=1}^{N} a_{im} \ln \Phi\left(x_{im}\beta_i + \sum_{j\neq i} \delta_{ij} \ \widehat{P}_j(x_m) \right)$$
$$+(1-a_{im}) \ln \Phi\left(-x_{im}\beta_i - \sum_{j\neq i} \delta_{ij} \ \widehat{P}_j(x_m) \right)$$

- It can be generalized to deal with unobserved heterogeneity ω_m.

K-Step Estimator

- The first-step nonparametric estimator can have large variance and finite sample bias because the curse of dimensionality in NP estimation.
- This translates into the two-step estimator of θ that can have also large variance and finite sample bias.
- The K-step estimator is a solution to this problem.
- Let $\hat{\theta}_i^{(1)}$ be the two-step estimator.
- Given $\hat{\theta}_i^{(1)}$ and $\widehat{P^{(0)}}$, we can construct new choice probabilities, $\widehat{P^{(1)}}$, that now are parametric and exploit part of the structure of the model:

$$\widehat{P^{(1)}}(\mathbf{x}_m) \quad = \quad \Phi\left(x_i \; \hat{\beta}_i^{(1)} + \textstyle\sum_{j \neq i} \hat{\delta}_{ij}^{(1)} \; \widehat{P^{(0)}}(\mathbf{x}_m)\right)$$

- Under some regularity conditions (Kasahara & Shimotsu, 2009), $\widehat{P^{(1)}}$ has smaller variance and finite sample bias than $\widehat{P^{(0)}}$.

K-Step Estimator [2]

- Given the new estimator $\widehat{P^{(1)}}$, we can obtain a new estimator of θ:

$$\widehat{\theta}^{(2)} = \arg\max_{\theta} Q_M(\theta, \widehat{P^{(1)}})$$

with $Q_M(\theta, \widehat{P^{(1)}}) = \sum_{m=1}^{M} \sum_{i=1}^{N} a_{im} \ln \Phi\left(x_{im}\beta_i + \sum_{j\neq i} \delta_{ij} \; \widehat{P}_j^{(1)}(x_m) \right) +$
$(1 - a_{im}) \ln \Phi\left(-x_{im}\beta_i - \sum_{j\neq i} \delta_{ij} \; \widehat{P}_j^{(1)}(x_m) \right)$

- We can also apply this procedure recursively to define a $K - step$ estimator.

- Under some regularity conditions (Kasahara & Shimotsu, 2009), $\widehat{\theta^{(K)}}$ with $K > 1$ has smaller variance and finite sample bias than $\widehat{\theta^{(1)}}$.