ECO 3901
EMPIRICAL INDUSTRIAL ORGANIZATION

Lecture 2

Data and Identification of structural dynamic games

Victor Aguirregabiria (University of Toronto)

January 20, 2022

Lecture 2: Data & Identification of dynamic games Outline

1. **Datasets in applications**
2. **The identification problem**
3. **Basic assumptions and Non-identification result**
4. **Basic positive identification result**
5. **Relaxing restrictions in basic identification result**

1. Datasets in Applications

Type of Data in most Empirical Applications

- Panel data of M geographic markets, over T periods, and N firms.

$$Data = \{\boldsymbol{a}_{mt}, \ \mathbf{x}_{mt} : m = 1, 2, ..., M; \ t = 1, 2, ..., T\}$$

- **Example 1**: Major airlines in US ($N = 10$), in the markets/routes defined by all the pairs of top-50 US airports ($M = 1,275$), over $T = 20$ quarters (5 years).

- **Example 2**: Supermarket chains in Ontario ($N = 6$), in the geographic markets defined by census tracts ($M > 1k$), over $T = 24$ months.

Type of Data in most Empirical Applications [2]

- This data structure applies to industries characterized by **many geographic markets**, where a separate (dynamic) game is played in each market: e.g., retail industries, services, airline markets, procurement auctions, ...

- However, there are many manufacturing industries where **competition is more global**: a single national or even international market: e.g., microchips.

- For these "global" industries, applications rely on sample variability that comes from a **combination of modest N, M, and T**.

- Some other industries are characterized by a **large number of heterogeneous firms** (large N), e.g., NYC taxis.

2. The identification problem

Our identification problem

- The **primitives of the model** are:

$$\{\pi_i(.),\ \delta_i,\ F_x(.):\ i \in \mathcal{I}\}$$

- Empirical applications assume that these primitives are known to the researcher up to a **vector of structural parameters θ**.

- The **identification problem** consists in using the data and the restrictions of the model to:

  - uniquely determine the value of $\boldsymbol{\theta}$ **(point identification)**
  - or to obtain bounds on $\boldsymbol{\theta}$ **(partial / set identification)**.

- This is a **revealed preference identification approach**: under the assumption that firms' are maximizing profits, their actions reveal information about the structure of their profit functions.

3. Basic Assumptions and Non-Identification Result

Basic Assumptions

- Set of assumptions used in many applications in this literature.

ID.1 **No common knowledge unobservables**. The researcher observes $\mathbf{x}_t$. The only unobservables are ε_{it}.

ID.2 **Single equilibrium in the data**. Every observation (i, m, t) in the data comes from the same MPE.

ID.3 **Additive unobservables**. The unobservables ε_{it} enter additively in the payoff function: $\pi_i(\boldsymbol{a}_t, \mathbf{x}_t) + \varepsilon_{it}(a_{it})$.

ID.4 **Known distribution of unobservables**. The distribution of ε_{it} is completely known to the researcher.

ID.5 **Conditional independence**. Conditional on $(\boldsymbol{a}_t, \mathbf{x}_t)$ the distribution of $\mathbf{x}_{t+1}$ does not depend on $\boldsymbol{\varepsilon}_t$.

A Positive Identification Result but Not for Primitives

- Under Assumptions [ID.1] and [ID.2], the vector of equilibrium CCPs in the population, $\mathbf{P}^0$, is identified from the data. For every $(i, a_i, \mathbf{x}$:

$$P_i^0(a_i|\mathbf{x}) = \mathbb{E}\left(1\{a_{imt} = a_i\} \mid \mathbf{x}_{mt} = \mathbf{x}\right)$$

- Given CCPs and under assumptions [ID.3] to [ID.5], **Hotz-Miller Inversion Theorem** implies the identification of **conditional-choice value function** relative to a baseline alternative (say 0):

$$\widetilde{v}_i^{\mathbf{P}}(a_i, \mathbf{x}) \equiv v_i^{\mathbf{P}}(a_i, \mathbf{x}) - v_i^{\mathbf{P}}(0, \mathbf{x})$$

- For instance, when ε is Type I extreme value:

$$\widetilde{v}_i^{\mathbf{P}}(a_i, \mathbf{x}) = \ln P_i^0(a_i|\mathbf{x}) - \ln P_i^0(0|\mathbf{x})$$

A Negative Identification Result (on Primitives)

- Unfortunately, the identification of function $\widetilde{v}_i^{\mathbf{P}}(a_i, \mathbf{x})$ is not sufficient to identify the primitive preference function $(\pi_i(\boldsymbol{a}, \mathbf{x}),\ \delta_i)$. See **Rust (1994, Handbook)**, **Magnac & Thesmar (2002, ECMA)**

- There are three identification issues.

P1. **Non innocuous normalizations**. In contrast to static models, normalizing $\pi_i(0, \mathbf{x}) = 0$ has implications on important empirical questions.

P2. **No identification of discount factor**.

P3. **No identification competition effects**. $\widetilde{v}_i^{\mathbf{P}}(a_i, \mathbf{x})$ does not have a_{-it} as an argument, but we are interested in the effect of a_{-it} on π_i.

4. Basic Positive Identification Result

Additional Assumptions Restrictions

ID.6 **Normalization of payoff of one choice alternative**. $\pi_i(a_i = 0, \boldsymbol{a}_{-i}, \mathbf{x}) = 0$ for every $(i, \boldsymbol{a}_{-i}, \mathbf{x})$.

Example: If a firm decides not being in the market its profit is zero, regardless she is a potential entrant or an incumbent (i.e., no scrap value of exit costs).

ID.7 **Known discount factor**. δ_i is known to the researcher.

Example: With annual frequency, $\delta_i = 0.95$ for every firm.

ID.8 **Exclusion restriction in profit function**. $\mathbf{x}_t = (\mathbf{x}_t^c), z_{it} : i \in \mathcal{I}$ such that $\pi_i(\boldsymbol{a}_t, \mathbf{x}^c, z_{it})$ does not depend on z_{jt}) for $j \neq i$.

Example: In a game of market entry-exit, firm i's profit depends on the current entry decisions of competitors ($\boldsymbol{a}_{-it}$), and on the own incumbency status ($a_{i,t-1}$) but there is not a (direct) effect of the competitors' incumbency status ($\boldsymbol{a}_{-i,t-1}$).

Positive Identification Result

- Under Assumptions [ID.1] to [ID.8], the profit functions $\pi_i(\boldsymbol{a}, \mathbf{x})$ are **nonparametrically identified** from the conditional-choice values $\widetilde{v}_i^{\mathbf{P}}(a_i, \mathbf{x})$.

  Proposition 3 in Pesendorfer & Schmidt-Dengler (REStud, 2008).

- Assumptions [ID.1] to [ID.8] are very common in empirical applications of dynamic games in IO. In most cases, they are combined with parametric restrictions on function π_i.

5. Relaxing Restrictions Basic Identification Result

Relaxing Restrictions

- Serially correlated unobservables
- Multiple equilibria in the data
- Relaxing normalization restrictions
- Identification of discount factors
- Non-additive unobservables
- Nonparametric distribution of unobservables
- Non-equilibrium beliefs