Moment Inequality and Partial Identification in Industrial Organization

Handbook of IO, Volume 4, Chapter 5

Presented by: Hang XU, Xinrui LIU

Hong Kong University of Science and Technology

March 1, 2025

Outline

Point Identification

Point Identification

Estimating model parameters as a single value, e.g., $\hat{\theta} = 0.5$.

Point Identification

Point Identification

Estimating model parameters as a single value, e.g., $\hat{\theta} = 0.5$.

- In traditional econometrics, we often try to find point identification **with a cost** of adding more assumptions (to improve identifying power).
- Common assumptions include: Certain distributions of error terms (e.g., logit model); Independence of error terms.
- If assumptions are satisfied, one can always obtain the true value of the parameter by using an infinite amount of data.
- But in reality, these assumptions are often violated or not well justified by economic theory (e.g., correlation between error terms and regressor).
- **Problem**: Two researchers using the same data may reach different conclusions due to different assumptions.

Partial Identification

- Need to find out what can be learned from the data **without imposing strong assumptions** – May give up point identification and turn to partial identification.

Partial Identification

- Need to find out what can be learned from the data **without imposing strong assumptions** – May give up point identification and turn to partial identification.

Partial Identification

Estimating model parameters as a set, e.g., $\hat{\theta} \in [0.3, 0.7]$.

Partial Identification

- Need to find out what can be learned from the data **without imposing strong assumptions** – May give up point identification and turn to partial identification.

Partial Identification

Estimating model parameters as a set, e.g., $\hat{\theta} \in [0.3, 0.7]$.

- Cannot know the true value of the parameter even with an infinite dataset, but can still reveal some insights about the object of interest.
- Usually, the partial identification of the parameter is derived from a set of inequalities, which are called **moment inequalities**.
- The inequalities are derived from optimizing behavior based on economic theory (e.g., revealed preferences) or simply the statistical properties from the data.

Roadmap

- Partial identification is widely used in empirical IO, especially in
  - Models with measurement errors or unobserved heterogeneity,
  - Models based on revealed preference,
  - Strategic interaction models among multiple agents (e.g., market entry, auction),
  - Models with multiple equilibria or incomplete information.
- Partial identification also gives challenges for estimation and inference:
  - Estimation: How do we estimate a set? What is a "good" estimate of set?
  - Inference: How to test an hypothesis about true parameters with partial identification?

Acknowledgement

This presentation is based on

- Chapter 5: *Moment Inequality and Partial Identification in Industrial Organization* from the Handbook of Industrial Organization, Volume 4,
- *Introduction to Partial Identification* from C. Bontemps' lecture notes.

Outline

A Warm-up: Frisch (1934): Measurement Error Models

- **Example**: Consider a simple linear regression model $y^* = \beta x^* + u$.
- If we assume $E(x^* u) = 0$, we can estimate a consistent β as $\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i^* - \bar{x})(y_i^* - \bar{y})}{\sum_{i=1}^{n}(x_i^* - \bar{x})^2}$.
- Now, suppose both x^* and y^* are measured with errors:

$$x = x^* + v_x$$
$$y = y^* + v_y$$

  Where v_x, v_y are unobserved measurement errors with i.i.d. distribution.
- How to estimate β in the true model by only observing (x, y)?

A Warm-up: Frisch (1934): Measurement Error Models (Cont'd)

The true model implies

$$y = \beta x + \underbrace{u + v_y - \beta v_x}_{w}$$

and if we regress y on x:

$$\hat{\beta} = \beta + \frac{Cov(x, w)}{Var(x)}$$

- This leads to inconsistency of the estimation.
- Traditional methods (e.g., IV) can get point identification of β if conditions hold.
- Alternatively, we can use partial identification to get a set of $\hat{\beta}$ by finding some moment inequalities.

A Warm-up: Frisch (1934): Measurement Error Models (Cont'd)

The model imposes some equations in second order moments of observables:

$$Var(x) = Var(x^*) + Var(v_x) \tag{1}$$

$$Var(y) = Var(y^*) + Var(v_y) \geq \beta^2 Var(x^*) + Var(v_y) \tag{2}$$

$$Cov(x, y) = Cov(x^* + v_x, y^* + v_y) = \beta Var(x^*) \tag{3}$$

$$(2) + (3) \rightarrow Var(y) \geq \beta Cov(x, y) + Var(v_y) \tag{4}$$

Two inequalities can be derived based on the statistical properties of variance:

- $Var(v_x) \geq 0 \rightarrow \beta \geq \frac{Cov(x,y)}{Var(x)}$
- $Var(v_y) \geq 0 \rightarrow \beta \leq \frac{Var(y)}{Cov(x,y)}$
- Then, we can get a set of $\hat{\beta} \in [\frac{Cov(x,y)}{Var(x)}, \frac{Var(y)}{Cov(x,y)}]$

Outline

Revealed Preference

- Revealed preference theory is a method to infer consumer preferences from observed choices.
- The theory is based on the assumption that consumers' choices are consistent with their preferences.
- The theory can be used to derive moment inequalities for partial identification.
- Formally, let $\pi(d_i, d_{-i}, y_i, \theta)$ be the pay-off agent i receives from choosing action d_i given its competitors chose d_{-i} and other determinants y_i, we have:

$$E[\pi(d_i, d_{-i}, y_i, \theta)|I_i] \geq E[\pi(d, d_{-i}, y_i, \theta)|I_i] \quad \forall d \in \mathbb{D}_i$$

  Where I_i is the information set, and $\mathbb{D}_i$ is the choice set available to agent i.
- The expectation is taken since competitors actions or other determinants might be random at the time of decision.

Katz (2007): Shoppers Driving Time Costs

- Katz (2007) studies the costs shoppers assign to driving to a supermarket, which is important for zoning laws and public transportation policy.

- This problem involves a two-stage decision process: agent i's decision $d_i = (s_i, b_i)$,
  - First, choose a supermarket (store) s_i.
  - Then, choose products to purchase (a complex choice set of possible bundles), say the basket b_i.

Katz (2007): Shoppers Driving Time Costs (Cont'd)

- The utility functions of the agent with z_i characteristics are additively separable functions of:
  - The utility of the basket bought $u(b_i, z_i)$,
  - The expenditure on the basket $e(b_i, s_i)$,
  - Driving time to the store $\theta_i dt(s_i, z_i)$.

Utility function is:

$$\pi(d_i, z_i, \theta) = u(b_i, z_i) - e(b_i, s_i) - \theta_i dt(s_i, z_i)$$

Where $\theta_i = \theta_0 + v_i$ is the parameter of interest (the dollar value of driving time), and v_i is the unobserved preference for driving time.

Katz (2007): Shoppers Driving Time Costs (Cont'd)

- By using revealed preference, we compare the actual choice, d_i, to the alternative d_i^f of purchasing:
  - The same basket b_i of goods,
  - At a store s_i^f that is farther away from the agent's home than the actual store s_i.

  Thus, we have:

$$\begin{aligned}\Delta\pi(d_i, d_i^f, z_i, \theta) &= \pi(d_i, z_i, \theta) - \pi(d_i^f, z_i, \theta) \\ &= -\Delta e(b_i, s_i, s_i^f) - (\theta_0 + v_i)\Delta dt(s_i, s_i^f, z_i) \geq 0\end{aligned}$$

- Notice that choosing a different basket at d_i^f just reinforces the inequality.
- Then we have:

$$\theta_0 + v_i \quad \geq \quad -\frac{\Delta e(b_i, s_i, s_i^f)}{\Delta dt(s_i, s_i^f, z_i)}$$

Katz (2007): Shoppers Driving Time Costs (Cont'd)

- Assume that v_i follows a distribution with mean zero: $E(v_i) = 0$.
- Then we can get the moment inequality by aggregating over all agents:

$$\theta_0 \geq -\frac{E[\Delta e(b_i, s_i, s_i^f)]}{E[\Delta dt(s_i, s_i^f, z_i)]}$$

- Similar approach can be applied to a store that is closer to the agent's home: s_i^c.

$$\theta_0 \leq -\frac{E[\Delta e(b_i, s_i, s_i^c)]}{E[\Delta dt(s_i, s_i^c, z_i)]}$$

- This approach eliminates the needs to specify:
  - The specific distribution assumption of v_i (where MNL model is often used),
  - The choice set of the basket b_i and store s_i for each agent,
  - The two-stage decision problem.

Outline

Strategic Interactions

- The previous case is a single-agent model, where the agent's decision depends on the agent's own characteristics.
- In many IO models, agents' decisions depend on the decisions of others, which is called strategic interactions, or multi-agent models.
- The payoffs in multi-agent models are often functions of actions of all agents, and the agents' actions are interdependent.
- Models with strategic interactions may have multiple equilibria, incomplete information, or other complications where MNL with MLE are not applicable.
- In such models, the revealed preference approach tends to use only a subset of conditions implied by the model, thereby leading to learning less about the parameters (may have a wide set of estimates).
- Thus, other conditions (e.g., equilibrium conditions, equilibria selections, etc.) are needed to narrow down the set of estimates.

A Market Entry Game with Complete Information

- Two firms $i = 1, 2$ decide whether to enter a market: $d_i = 1, 0$ with 0 to be not. The payoff of firm i is:

$$\pi_1 = d_1 \times (x_1\beta_1 + \Delta_1 d_2 + \epsilon_1)$$
$$\pi_2 = d_2 \times (x_2\beta_2 + \Delta_2 d_1 + \epsilon_2)$$

Table 5 A parametric two-player, two-action game in normal form.

	$d_2 = 0$	$d_2 = 1$
$d_1 = 0$	$(0, 0)$	$(0, x_2\beta_2 + \epsilon_2)$
$d_1 = 1$	$(x_1\beta_1 + \epsilon_1, 0)$	$(x_1\beta_1 + \Delta_1 + \epsilon_1, x_2\beta_2 + \Delta_2 + \epsilon_2)$

- Firms have complete information about the other firm's profit function,
- Assume both Δ_1 and Δ_2 are negative, a competition effect,
- Only focus on pure strategy Nash equilibrium.

A Market Entry Game with Complete Information (Cont'd)

Shocks $\epsilon_i \perp x_i$ are known to the firm but not to the econometrician:

- Firms observe all components of the payoff, including ϵ_i, so their decisions satisfy:

$$d_i = \mathbf{1}\{\pi_i \geq 0\} \quad for \quad i = 1, 2$$

- The econometrician can only observe the entry decisions d_i: a pair of indicators (d_1, d_2) and the characteristics x_i.
  - If $(d_1, d_2) = (1, 0)$ is observed (firm 1 is the monopolist), $\Rightarrow x_1\beta_1 + \epsilon_1 \geq 0$ and $x_2\beta_2 + \Delta_2 + \epsilon_2 \leq 0$, $\Rightarrow$ A necessary condition for $(1, 0)$ to be the equilibrium;
  - From a reverse logic, when $x_1\beta_1 + \epsilon_1 \geq 0$ and $x_2\beta_2 + \epsilon_2 \leq 0$, $\Rightarrow$ $(1, 0)$ is a dominant strategy equilibrium, $\Rightarrow$ A sufficient condition for $(1, 0)$ to be the equilibrium.
- The aim is to learn the vector of parameters $\theta = (\beta_1, \beta_2, \Delta_1, \Delta_2)$.

A Market Entry Game with Complete Information (Cont'd)

- It seems we can obtain a point identification of θ by maximizing the likelihood function of the observed data if assuming the distributions of ϵ_i.
- i.e., choose θ such that we match the observed four choice probabilities $p_{ij} = P(d_1 = i, d_2 = j)$ as good as possible.
- However, since multiple equilibria exists in some regions, the choice probability for some outcomes cannot be written as a function of θ (the pay-off functions), even with the distributional assumption of ϵ_i:
  - Within certain range of ϵ, the outcome is still random since which equilibrium is selected is random.

A Market Entry Game with Complete Information (Cont'd)

- Specify the choice probabilities as (assuming Δ_1 and Δ_2 are negative):

$$
\begin{aligned}
p_{00} &= P(d_1 = 0, d_2 = 0|x) = P\left(\epsilon_1 \leq -x_1\beta_1, \epsilon_2 \leq -x_2\beta_2\right) \\
p_{11} &= P(d_1 = 1, d_2 = 1|x) = P\left(\epsilon_1 \geq -x_1\beta_1 - \Delta_1, \epsilon_2 \geq -x_2\beta_2 - \Delta_2\right) \\
p_{10} &= P(d_1 = 1, d_2 = 0|x) = P\left(\epsilon_1 \geq -x_1\beta_1, \epsilon_2 \leq -x_2\beta_2 - \Delta_2, \epsilon \notin S_\beta\right) \\
&\qquad\qquad\qquad\qquad\quad + P\left(d_1 = 1, d_2 = 0|x, \epsilon \in S_\beta\right) \times P\left(\epsilon \in S_\beta\right) \\
p_{01} &= P(d_1 = 0, d_2 = 1|x) = 1 - p_{00} - p_{11} - p_{10}
\end{aligned}
$$

- Where $S_\beta = \{\epsilon : -x_1\beta_1 \leq \epsilon_1 \leq -x_1\beta_1 - \Delta_1, -x_2\beta_2 \leq \epsilon_2 \leq -x_2\beta_2 - \Delta_2\}$ are the regions where the equilibrium is not unique: $(1, 0)$ and $(0, 1)$ are both equilibria.
- For p_{00} and p_{11}, we can write them as functions of θ and x (the pay-off functions can uniquely reflect the realized data), but not for p_{10} and p_{01} if no more assumptions on multiple equilibria selections are made.

A Market Entry Game with Complete Information (Cont'd)

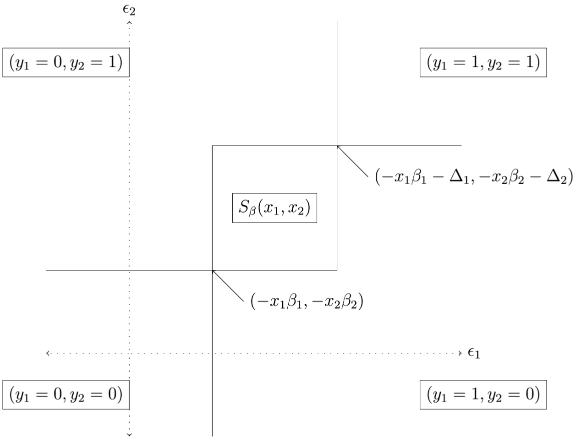

FIGURE 1

Mapping from ϵ to outcomes.

A Market Entry Game with Complete Information (Cont'd)

- More assumptions are needed to specify the equilibrium selection rules to get a point identification of θ;
- Instead, without any further assumptions, we can still get a set of θ by using partial identification approach:
- The inequality conditions are coming from the statistical property of the probability. We apply this inequality condition on the ones we do not have a function of θ: $P(d_1 = 1, d_2 = 0|x, \epsilon \in S_\beta) \in [0, 1]$.
- Thus, we find the set of θ that satisfies the inequality conditions:

$$H_L^{(1,0)} \leq P(d_1 = 1, d_2 = 0|x) \leq H_U^{(1,0)}$$

- Where $H_L^{(1,0)} = P(\epsilon_1 \geq -x_1\beta_1, \epsilon_2 \leq -x_2\beta_2 - \Delta_2, \epsilon \notin S_\beta)$,
- $H_U^{(1,0)} = P(\epsilon_1 \geq -x_1\beta_1, \epsilon_2 \leq -x_2\beta_2 - \Delta_2)$.

A Market Entry Game with Complete Information (Cont'd)

- Similar approach for $p_{01} = p(d_1 = 0, d_2 = 1|x)$ to find $H_L^{(0,1)}$ and $H_U^{(0,1)}$.
- For p_{00} and p_{11}, we have equality conditions, which can be written as functions of θ: $H^{(0,0)} = P(\epsilon_1 \leq -x_1\beta_1, \epsilon_2 \leq -x_2\beta_2)$ and $H^{(1,1)} = P(\epsilon_1 \geq -x_1\beta_1 - \Delta_1, \epsilon_2 \geq -x_2\beta_2 - \Delta_2)$.
- Finally, the identified set of θ is the intersection of the sets from all conditions:

$$
\begin{aligned}
\Theta = \{\theta : & H_L^{(1,0)} \leq P(d_1 = 1, d_2 = 0|x) \leq H_U^{(1,0)}, \\
& H_L^{(0,1)} \leq P(d_1 = 0, d_2 = 1|x) \leq H_U^{(0,1)}, \\
& P(d_1 = 0, d_2 = 0|x) = H^{(0,0)}, \\
& P(d_1 = 1, d_2 = 1|x) = H^{(1,1)}\}
\end{aligned}
$$

Other Strategic Interaction Models: Incomplete Information

A market entry game with incomplete information:

- With incomplete information, firms form beliefs about the other firm's action based on what they know:
- The payoffs are now functions of the actions and the beliefs of the other firm:

$$\pi_1 = d_1 \times (x_1\beta_1 + \Delta_1\omega_1(d_2 = 1|x) + \epsilon_1)$$
$$\pi_2 = d_2 \times (x_2\beta_2 + \Delta_2\omega_1(d_1 = 1|x) + \epsilon_2)$$

- Where $\omega_1(d_2 = 1|x)$ is the belief that firm 1 holds about the probability that firm 2 will enter the market.
- By using economic theory, one can find the upper and lower bounds of $\omega_i(d_{-i} = 1|x)$, and then find the set of θ that satisfies the moment inequalities.

Other Strategic Interaction Models: Auctions

- In a typical auction model, bidder i with valuation θ_i has the pay-off: $\theta_i x_i - t_i$ where:
  - t_i is the actual transfer he made,
  - $x_i \in \{0, 1\}$ gives the allocation state of the object,
  - bidder i bids with b_i.
- Haile and Tamer (2003) proposes two assumptions in auction settings:
  - Bids are weakly less than corrsponding valuation: $b_i(\theta_i) \leq \theta_i$;
  - A bidder that loses the auction has a valuation that makes it unprofitable to beat the winning bid: $\theta_i \leq b^* + \Delta$
  - Where b^* is winning bid, $\Delta \geq 0$ is the minimum bid increment.
- These two assumptions generate upper and lower bounds on the distributions of valuation θ_i, a partial identification approach.

Revisit the Roadmap

- Partial identification is widely used in empirical IO, especially in
  - Models with measurement errors or unobserved heterogeneity,
  - Models based on revealed preference,
  - Strategic interaction models among multiple agents (e.g., market entry, auction),
  - Models with multiple equilibria or incomplete information.
- Partial identification also gives challenges for estimation and inference:
  - Estimation: How do we estimate a set? What is a "good" estimate of set?
  - Inference: How to test an hypothesis about true parameters with partial identification?

Outline

Introduction

Example 1: Measurement Error Models

Example 2: Models Based on Revealed Preference

Examples with Partial Identification: Strategic Interactions

Estimation and Inference in Partial Identification