Moment Inequalities for Entry Games with Heterogeneous Types

C. Bontemps[1], R. Kumar[2] and M. Lesellier[3]

1: Toulouse School of Economics & Ecole Nationale Aviation Civile
2: Indian Institute of Technology Delhi
3: Université de Montréal

ESEM

Barcelona, August 2023

Outline

Motivation I

- Entry games are very popular in the empirical IO literature, mainly because they can estimate features of some industry while observing the decision of firms to enter or not in independent markets and their characteristics (see Berry and Reiss, 2007, for a survey).
- Many applications : **competition between airlines** (Ciliberto and Tamer, 2009, Chen, X., Christensen, T. M., and Tamer, ,2018, Berry, 1992), **retail industry** (Cleeren et al., 2010, Aradillas-Lopez and Rosen, 2022, Andrews, Berry and Jia, 2004, Grieco, 2014), **motels** (Mazzeo, 2002), and many others.
- Static entry games are games with multiple equilibria. There is no longer uniqueness of the model predictions. There are regions for profit shocks in which two or more outcomes can be predicted without the model telling us which outcome should be the right one. As a result, we can not estimate them with standard econometric procedures without imposing further assumptions.

Motivation II

- There are *incomplete* models because the selection mechanism is indeed unknown in the regions of multiple equilibria. Standard solutions often used are the following ones:
  - Postulating some selection mechanism (Bjorn and Vuong, 1984, Berry, 1992, Cleeren et al., 2010, among others ...),
  - Working from an outcome which is invariant, like the number of active firms at the equilibrium, e.g. Berry (1992), or in Cleeren et al. (2010)
  - Using the recent literature on moment inequalities like in Ciliberto and Tamer, 2009, Beresteanu et al., 2011, Galichon and Henry, 2011, Kline and Tamer, 2016, Chesher and Rosen, 2019, Bontemps and Kumar, 2020, Aradillas-Lopez and Rosen, 2022, Magnolfi and Roncoroni, 2022, Cox and Shi, 2022 and Kaido and Molinari, 2022.
- Remark that multiple equilibria does not necessary imply set identification.

Main challenges

There exists various methods to estimate a static entry game with multiple equilibria (and it does not depend on the equilibrium concept). Many challenges:

- Which inequalities to test ? The problem becomes quickly untractable numerically because the number of moment inequalities necessary to sharply characterize a valid parameter is exponentially increasing.
- Usually, people address this issue by "selecting" some of these inequalities like the min max approach used by Ciliberto and Tamer (2009) (see also Chesher and Rosen, 2019). As a result, the identified set is (often) not sharply characterized.
- Which test statistic and which critical value to use ? Necessity to have competitive critical values which are not too complicated to calculate/simulate because it should be performed for each parameter θ tested (we invert a test).
- How to incorporate continuous explanatory variables ? They are often discretized to alleviate the numerical burden.

Our contribution I

1. In this paper, we contribute to these three issues under a slightly restrictive model because we assume that firms are pooled by types and are homogeneous within each type.
2. The collection of moment inequalities being linked to the structure of the multiple equilibria regions, we propose an algorithm to catalog the set of moment inequalities which characterize sharply the "identified" set. As we invert a test, our procedure is (uniformly ?) valid even if the model is point identified (see Tamer, 2003 or Aradillas-Lopez and Rosen, 2022). The algorithm can be stopped at any order to provide an outer set which is closer and closer to the true identified set.
    - It eliminates redundant inequalities (see also Galichon and Henry, 2011, Chesher and Rosen, 2019, Bontemps and Kumar, 2020, Ponomarev, 2022)
    - It allows to calculate the different quantities and, therefore, avoids additional numerical errors, often ignored in the applications

Our contribution II

3. Once a set of moment inequalities has been provided, we propose a testing procedure which is asymptotically pivotal by smoothing the set defined by these moment inequalities. Smoothing leads to a loss of sharpness but this loss can be controlled by an accurate choice of the smoothing parameter, and is asymptotically vanishing by letting the parameter tending to $+\infty$ at the right speed.
4. We do the same for models with continuous covariates and provide a two-step procedures in which the first step consists in estimating the conditional probabilities of the different outcomes that we plug into the test statistic. This test statistic is asymptotically normal with a correction term due to the first step estimation (like in Newey, 1994).
5. (later) We apply it to supermarket data in France and would like to evaluate the impact of the Loi Raffarin on the market structure of the French retail industry.

Outline

Outline

The general model I

- We consider a generic entry game with T types: firms are homogeneous within types, known ex-ante and exogeneous (Full Service Carriers/Low-Cost Carriers or Hypermarkets/Supermarkets/Local Retailers/ Discounters). Mazzeo (2002) lets firms to choose their types but the inequalities derived are similar.
- Free entry: firms enter if their long-run profit is non negative
- Firms have complete information and play pure strategies
- The profit of players of type t in market i writes:

$$\Pi_{t,m} = \pi_t(X_{t,m}, N_{t,m}, \mathbf{N}_{-t,m}; \omega) + \varepsilon_{t,m},$$

With:

- $X_{t,m}$ vector of exogenous variables which characterize market m,
- $N_{t,m}$ number of firms of the type t on market m,
- $\mathbf{N}_{-t,m}$ vector containing the number of firms of type $t' \neq t$ on market m,
- $\pi_t(\cdot;\omega)$ function parametrized by ω', the parameter of interest,

The general model II

- $\varepsilon_{t,m}$ unobserved stochastic shock.

- Profit is decreasing with respect to the number of competitors
- $(\varepsilon_{1,m}, \ldots, \varepsilon_{T,m}) \sim F_\eta(\cdot)$ known up to η and independent of $X_{t,m}$
- In the following, we denote by θ all the parameters to be estimated.

Example

- Static entry game with 2 types (maximum two per type), linear profit functions and no covariates:

$$\Pi_1 = \beta_1 - \delta_{1,1} N_1 - \delta_{2,1} N_2 + \varepsilon_1$$
$$\Pi_2 = \beta_2 - \delta_{1,2} N_1 - \delta_{2,2} N_2 + \varepsilon_2$$

-N_t is the number of firms of type t=1,2

-ε_t for $t = 1, 2$ idiosyncratic shocks

- $\delta's$ are positive and capture the competitive effect

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

- Parameter of interest: $\theta = (\beta_1, \beta_2, \delta_{1,1}, \delta_{2,1}, \delta_{1,2}, \delta_{2,2})$

Equilibrium analysis

- Given N_2, firms of type 1 enter until $\Pi_1(N_1+1, N_2) < 0$ (resp. for firm 2)
- An outcome $y = (N_1, N_2)$ is a NE of this game if and only if:

$$\Pi_1(N_1, N_2) \geq 0 \text{ and } \Pi_1(N_1+1, N_2) < 0$$
$$\Pi_2(N_1, N_2) \geq 0 \text{ and } \Pi_2(N_1, N_2+1) < 0$$

  Equilibrium region in the space of shocks:

$$-\beta_1 + \delta_{1,1}N_1 + \delta_{2,1}N_2 \leq \varepsilon_1 < -\beta_1 + \delta_{1,1}(N_1+1) + \delta_{2,1}N_2$$
$$-\beta_2 + \delta_{1,2}N_1 + \delta_{2,2}N_2 \leq \varepsilon_2 < -\beta_2 + \delta_{1,2}N_1 + \delta_{2,2}(N_2+1)$$

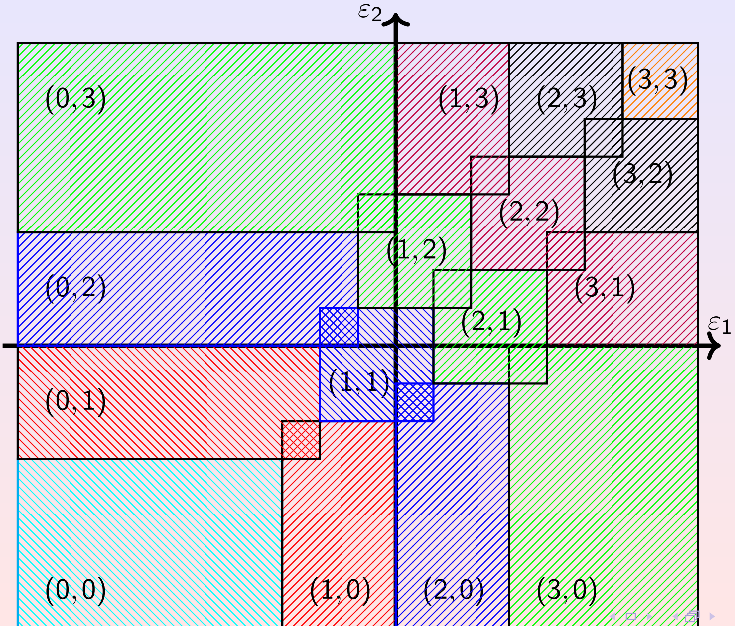
  We call this region $\mathcal{R}_\omega(N_1, N_2)$.

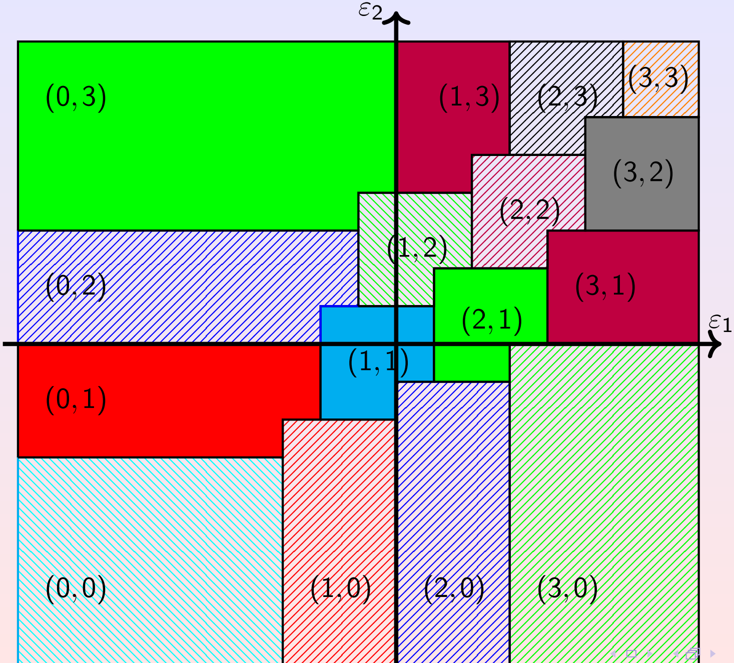
- For "a lot of" values of θ, there are regions of ε with multiple equilibria

  $\rightarrow$ no one-to-one mapping between the space of outcomes and the space of shocks ε

  $\rightarrow$ prevents the use of standard identification arguments (MLE, GMM)

Illustration of the equilibrium structure

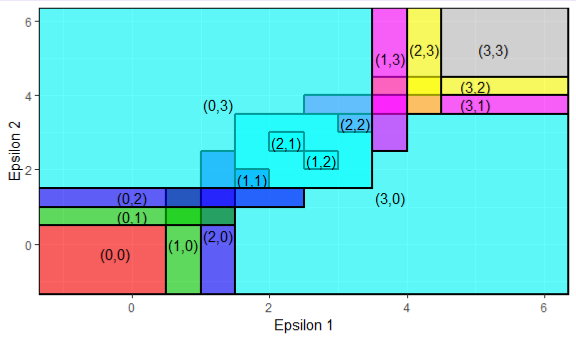




ε_2
ε_1
(0, 3)
(1, 3)
(2, 3)
(3, 3)
(3, 2)
(2, 2)
(1, 2)
(0, 2)
(3, 1)
(2, 1)
(1, 1)
(0, 1)
(0, 0)
(1, 0)
(2, 0)
(3, 0)

More "complicated" market structure

Figure: Equilibrium structure for $\delta_{11} = 0.7$, $\delta_{12} = 1.3$, $\delta_{21} = 0.9$ and $\delta_{22} = 1.2$



Possibility to simplify with "economic" restrictions based on the positioning of firms but it is complicated with more than two types.

Outline

Characterizing the Identified set I

Unrefined characterization of the Identified set

- We omit the dependence in X and introduce the problem with two types. An outcome is any pair (N_1, N_2) and we would like to calculate the probability to observe each outcome.
- When we observe such an outcome, we know that necessarily $\varepsilon \in \mathcal{R}_\omega(N_1, N_2)$. It gives an upper bound on the probability of each outcome.
  Therefore, for each outcome, y

$$P_0(Y = y) \leq \int_{\mathcal{R}_\omega(y)} dF_\eta(\varepsilon). \tag{1}$$

We can generalize it to union of outcomes.
For each pair of outcomes, $A = \{y_1; y_2\}$:

$$P_0(Y = y_1 \text{ or } y_2) \leq \int_{\mathcal{R}_\omega(A)} dF_\eta(\varepsilon). \tag{2}$$

Characterizing the Identified set II

Unrefined characterization of the Identified set

We need to know whether $\mathcal{R}_\omega(y_1)$ and $\mathcal{R}_\omega(y_2)$ overlap to calculate the upper bound.

- if they overlap,

$$P_\eta(\varepsilon \in \mathcal{R}_\omega(A)) = P_\eta(\varepsilon \in \mathcal{R}_\omega(y_1)) + P_\eta(\varepsilon \in \mathcal{R}_\omega(y_2)) - P_\eta(\varepsilon \in \mathcal{R}_\omega(y_1) \cap \mathcal{R}_\omega(y_2)),$$

- if they don't overlap, we don't need to consider this pair.

and, so on.

- Therefore, in order to bound each outcome or union of outcomes we need to "know" (or simulate like in CT, 2009) the multiple equilibria structure.

Core determining set

- The collection of moment inequalities generated from bounding all outcomes and union of outcomes characterizes the identified set sharply (see Beresteanu et al., 2011) but leads generally to too many inequalities.
- For example, if $N_{\max} = 3$, we have 16 outcomes possible from $(0,0)$ to $(3,3)$ in our two-type model. As a result, it generates $2^{16} - 1 = 65535$ inequalities.
  $\rightarrow$ In general, brute force can not be a solution.
- Often many of these inequalities are "redundant" in the sense that imposing a subset of them is sufficient to characterize sharply the set. In some examples (like in Galichon and Henry, 2011), the non redundant inequalities are much less numerous.
- This set is defined as the core determining set and we know it is linked to the characterization of subgraphs which are connected or not (see also Galichon and Henry, 2011, Chesher and Rosen, 2019, Bontemps and Kumar, 2020, Ponomarev, 2022). We propose a (recursive) algorithm which computes in the same steps the multiple equilibria structure and eliminates the redundant inequalities.

Algorithm I

Our algorithm is executed as follows.

1. Compute the regions $\mathcal{R}_\omega(y)$ for all single outcomes of $\mathcal{Y}$. Collect the moment inequalities generated from the upper bound of the probability of each outcome.
2. Check all the pairs $(K = 2)$ to see if the two outcomes y_1 and y_2 of each pair have their equilibrium regions $\mathcal{R}_\omega(y_1)$ and $\mathcal{R}_\omega(y_2)$ which overlap. If they overlap, for each type t,

$$\max\left(-\pi_t(N_t, \mathbf{N}_{-t}; \omega), -\pi_t(\bar{N}_t, \bar{\mathbf{N}}_{-t}; \omega)\right)$$
$$< \min\left(-\pi_t(N_t + 1, \mathbf{N}_{-t}; \omega), -\pi_t(\bar{N}_t + 1, \bar{\mathbf{N}}_{-t}; \omega)\right)$$

3 $\rightarrow$ Eliminate all moment inequalities generated from pairs for which this is not the case.
For the remaining pairs $A = (y_1, y_2)$, compute the sharp upper bound of $P(Y \in A)$:

$$P(Y \in A) \leq P_\eta(\varepsilon \in \mathcal{R}_\omega(y_1)) + P_\eta(\varepsilon \in \mathcal{R}_\omega(y_2)) - P_\eta(\varepsilon \in \mathcal{R}_\omega(y_1) \cap \mathcal{R}_\omega(y_2)).$$

Algorithm II

Add these moment inequalities to the set of inequalities generated by the single outcomes.

3. $K = 3$. We now check all triplets (y_1, y_2, y_3) given that their equilibrium regions might overlap if and only if they overlap two by two. In other words, we focus on the remaining pairs to select our "triplet candidates". Again, if the three regions overlap we have a connected subset of three elements, otherwise we do not keep the triplet and do not consider the moment inequality generated by an eliminated triplet.

For the remaining connected subsets of three elements $A = (y_1, y_2, y_3)$, compute the sharp upper bound of $P(Y \in A)$:

$$
\begin{aligned}
P(Y \in A) \leq \; & P_\eta(\varepsilon \in \mathcal{R}_\omega(y_1)) + P_\eta(\varepsilon \in \mathcal{R}_\omega(y_2)) + P_\eta(\varepsilon \in \mathcal{R}_\omega(y_3)) \\
& - P_\eta(\varepsilon \in \mathcal{R}_\omega(y_1) \cap \mathcal{R}_\omega(y_2)) - P_\eta(\varepsilon \in \mathcal{R}_\omega(y_1) \cap \mathcal{R}_\omega(y_3)) - P_\eta(\varepsilon \in \mathcal{R}_\omega(y_2) \cap \mathcal{R} \\
& + P_\eta(\varepsilon \in \mathcal{R}_\omega(y_1) \cap \mathcal{R}_\omega(y_2) \cap \mathcal{R}_\omega(y_3)).
\end{aligned}
$$

Algorithm III

4. $K = 4$. Check now for all connected subsets of four elements given that each subset of three elements must be in the remaining connected subsets of three elements and so forth

Graph and inequalities

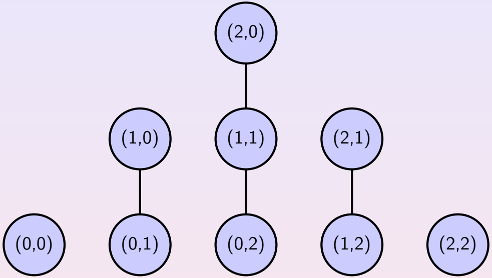


Figure: $\Gamma(\theta)$ for $\beta_1 = 3$, $\beta_2 = 2$, $\delta_{11} = \delta_{22} = 1.5$ and $\delta_{12} = \delta_{21} = 0.75$

14 inequalities (instead of 2^9) are sufficient. If they are satisfied, the 2^9 inequalities are satisfied (remark).

Outline

Moment inequalities I

The identified set is the collection of parameters θ which are observationally equivalent (to the true value), i.e., such that

$$P_\theta(Y) = P_0(Y),\ X\ a.s.$$

$$\Theta_I = \{\theta \in \Theta \mid P_0(Y \in A) \leq P_\eta(\mathcal{R}_\omega(A)) \ \ \forall A \in \mathcal{C}(\theta)\}$$

$\rightarrow$ The identified set can be rewritten in terms of moment inequalities:

moments: $m(Y_i, A, \theta) = P_\eta(\mathcal{R}_\omega(A)) - 1\{Y_i \in A\}$

Moment inequalities II

$$\begin{aligned}\theta \in \Theta_I &\iff \mathbb{E}[m(Y_i, A, \theta)] \geq 0 \quad \forall A \in \mathcal{C}(\theta), \\ &\iff q_j^\top P_0 \leq C_{\theta,j},\ j = 1, \ldots, p. \\ &\iff P_0 \in A(\theta),\ \text{a convex set.}\end{aligned}$$

Remark that the q_js are vectors with ones and zeros.

Moment inequalities III

Two strategies are possible:

- Select a test statistics (QLR, MMM, minimum of individual statistics) and use existing procedures to simulate a critical value (GMS, subsampling, etc.). They are usually quite conservative. Often critical values must be recomputed for each value θ tested.
- Use the fact that the moment inequalities define a convex set and that testing whether they are valid and test that the vector of probabilities belong to this convex set (support function, distance of the point to the set). Again, the critical value depends on whether the true point is a vertex, lies on an exposed faces, i.e.; depends on the number of binding moments

Moment inequalities IV

In our first set of simulations, it is not obvious to see which strategy seems to provide the most optimal procedure.

In the following, we focus on the minimum test statistic:

$$\xi_n(\theta) = \min_j \sqrt{n} \frac{C_{\theta,j} - q_j^\top P_n}{\sqrt{q_j^\top \Sigma_n q_j}},$$

in which P_n is the frequency estimator and $\Sigma_n = diag(P_n) - P_n P_n^\top$, an estimation of the variance of P_n.

Critical value for the min statistic

$$\xi_n(\theta) \xrightarrow[n\to\infty]{d} \min_{j\in\mathcal{J}(\theta)} \frac{q_j^\top Z}{\sqrt{q_j^\top \Sigma_0 q_j}},$$

in which Z follows a normal distribution with variance Σ_0 and $\mathcal{J}(\theta)$ is the collection of indices j corresponding to the binding moments.

This asymptotic distribution depends on the number and the identity of the binding moments, as expected. In the following, p^* denotes the number of binding moments, i.e., the cardinal of $\mathcal{J}(\theta)$.

A critical value can be computed after a first step estimation of the set of binding moments $\mathcal{J}(\theta)$ like in the GMS procedure of [Andrews and Soares, 2010]. Simulation methods (bootstrap and/or subsampling techniques) can be also considered to improve the accuracy of the critical value. [Chernozhukov et al., 2018] propose the following one:

$$c^*(\alpha) = \frac{\Phi^{-1}(\alpha/p)}{\sqrt{1-\Phi^{-1}(\alpha/p)^2/n}} \tag{3}$$

Smoothing the identified set I

In order to smooth the identified set, we replace the minimum operator by a smooth approximation. For $z = (z_1, z_2, ..., z_p) \in \mathbb{R}^p$, we have that a smooth approximation of the minimum between the elements of z and 0 writes:

$$g_\rho(z) = \frac{\sum_{j=1}^{p} z_j \exp(-\rho z_j)}{1 + \sum_{j=1}^{p} \exp(-\rho z_j)},$$

in which ρ, the smoothing parameter, controls the level of approximation. Following Chernozhukov et al. (2015), we have:

$$|min(0, z_1, z_2, \ldots, z_p) - g_\rho(z)| \leq \frac{1}{\rho} \log\left(\frac{p-1}{e}\right),$$

for $p > 10$.

Smoothing the identified set II

We define an outer set $\Theta_I^o(\rho)$ as follows:

$$\Theta_I^o(\rho) = \left\{ \theta \in \mathbb{R}^{\dim(\theta)} \middle| \; g_\rho(m_\theta) = \frac{\sum_{j=1}^{p} m_{\theta,j} e^{-\rho m_{\theta,j}}}{1 + \sum_{j=1}^{p} e^{-\rho m_{\theta,j}}} \geq 0 \right\}, \tag{4}$$

where $m_{\theta,j} = \mathbb{E} m_j(Y,\theta) = C_{\theta,j} - q_j^\top P_0$.

Proposition

The following statements hold

1. *For any* $\rho > 0$, $\Theta_I \subset \Theta_I^o(\rho)$
2. $\lim_{\rho \to +\infty} d_H(\Theta_I, \Theta_I^o(\rho)) = 0$, *where* d_H *is the Hausdorff distance used in set theory.*

Smoothing the identified set III

Proposition

Let $\Gamma_0(\theta) = -\frac{\sum_{j\in\mathcal{J}(\theta)} q_j}{1+p^*}$ *and* $\Sigma_0 = diag(P_0) - P_0 P_0^\top$ *and . Let* ρ_n *a divergent sequence of positive number such that* $\rho_n = O(n^a)$, $0 < a < 1/2$. *Then,*

$$\sqrt{n}(g_{\rho_n}(m_{n,\theta}) - g_{\rho_n}(m_\theta)) \xrightarrow[n\to\infty]{d} N(0, \Gamma_0(\theta)\Sigma_0\Gamma_0(\theta)^\top). \quad (5)$$

A consistent estimator of $\Gamma_0(\theta)$ is $(m_{n,\theta,j} = \mathbb{E}_n m_j(Y,\theta) = C_{\theta,j} - q_j^\top P_n)$:

$$\nabla g_{\rho_n}(m_{n,\theta}) = -\sum_{j=1}^{p} \frac{e^{-\rho_n m_{n,\theta,j}}(1 - \rho_n m_{n,\theta,j} + \rho_n g_{\rho_n}(m_{n,\theta}))}{1 + \sum_{j=1}^{p} e^{-\rho_n m_{n,\theta,j}}} q_j.$$

Intuition: asymptotic normality of the estimator of P_0 jointly with continuous differentiability of the function g_ρ (it is infinitely differentiable).

Smoothing the identified set IV

As a result, let

$$\xi_n(\theta) = \sqrt{n}\frac{g_{\rho_n}(m_{n,\theta})}{\sqrt{\nabla g_{\rho_n}(m_{n,\theta})^T \Sigma_n \nabla g_{\rho_n}(m_{n,\theta})}}$$

with Σ_n a consistent estimator of Σ_0. Our confidence region of confidence level $1-\alpha$ is defined as follows:

$$\mathrm{CR}_n(1-\alpha) = \{\xi_n(\theta) \geq z_\alpha\}$$

in which z_α is the α-quantile of the standard normal distribution.

Smoothing the identified set V

Proposition (Validity and consistency of the confidence region)

The confidence region $\mathrm{CR}_n(1-\alpha)$ *defined above is asymptotically valid and consistent, i.e.,*

- *Asymptotic validity:* $\liminf_{n\to\infty} \inf_{\theta\in\Theta_I} P_0(\theta \in \mathrm{CR}_n(1-\alpha)) \geq 1-\alpha.$
- *Consistency:* $\forall\theta \notin \Theta_I$, $P_0(\theta \in \mathrm{CR}_n(1-\alpha)) \underset{n\to\infty}{\longrightarrow} 0.$

Remarks We achieve this subsection with a few remarks.

- The quantity $e^{-\rho_n m_{n,\theta,j}}$ "selects" the binding moments.
- We speak later about the choice of ρ_n. It is derived to balance the "bias", i.e., the distance between the outer set and the true identified set and the accuracy of the first order expansion (by controlling the next term).
- There exist other smoother of the minimum function. The Log Sum Exp function, used in Machine Learning, requires to preestimate the number of binding moments before implementing an equivalent procedure.

Outline

Identification with covariates

$$\theta \in \Theta_I \iff P_0(X) \in A(\theta, X) \ X \ a.s.$$
$$\iff D_X P_0(X) \leq C_\theta(X), \ X \ a.s.$$
$$\iff \mathbb{E}(m_j(Y, X, \theta)|X) = \left(C_{\theta,j}(X) - q_j^\top P_0(X)\right) \geq 0, \forall\, j = 1, \ldots, p_X \ X \ a.s.$$

- Additional difficulty in the characterization: the core determining class $\mathcal{C}(\theta, x)$ must be computed for each $x \in \mathcal{X}$
- Assumption additive profit shifters: $\pi_t(X, N_t, \mathbf{N}_{-t}; \omega) = \kappa_t(X; \omega_1) + \phi_t(N_t, \mathbf{N}_{-t}; \omega_2), \quad \forall t = 1, \ldots, T.$
- Proposition: under the previous assumption,

$$\forall x \in \mathcal{X}, \quad \mathcal{C}(\theta, x) = \mathcal{C}(\theta)$$

$\rightarrow$ We can compute the core determining class only once for each θ.

The traditional estimation procedure I

$$\Theta_I = \{\theta \in \Theta \; | \mathbb{E}(m_j(Y,X,\theta)|X) = \left(C_{\theta,j}(X) - q_j^\top P_0(X)\right) \geq 0, \; \forall j \in \{1,\dots,p\}, \; X a.s.\}.$$

- X is continuous $\Longrightarrow$ the sharp identified set is characterized by an infinite number of inequalities.
- conditional moments are non-parametric objects that are harder to estimate and display non-standard asymptotic properties (eg: no CLT, curse of dimensionality...).

The traditional estimation procedure II

Various methods have been proposed in the literature: [Andrews and Shi, 2013], [Armstrong and Chan, 2016],[Armstrong, 2014],...

→ leading method in [Andrews and Shi, 2013]) consists in transforming the conditional moment inequalities into unconditional ones.
The identified set can be rewritten in terms of unconditional moment inequalities.

$$\bar{\Theta}_I = \{\theta \in \Theta \,|\mathbb{E}(m_j(Y,X,\theta)|X)g(X) = \left(C_{\theta,j}(X)g(X) - q_j^\top P_0(X)g(X)\right) \geq 0,\ \forall j \in \{1,\dots,p\},\ \forall$$

Alternative approach

Let us define the following moment, which serves as the basis of our estimation procedure:

$$m(X,\theta) = \min\left(0, \min_{j=1,\dots,p}\left\{C_{\theta,j}(X) - q_j^\top P_0(X)\right\}\right).$$

The sharp identified set can easily be expressed in terms of the previous moment, from a choice of $g(\cdot)$ which is positive, smooth, and does not vanish on the support of X.

Proposition

$$\begin{aligned}\theta \in \Theta_I &\iff m(X,\theta) = 0 \ X \ a.s \\ &\iff \mathbb{E}[m(X,\theta)g(X)] = 0.\end{aligned}$$

Implementation I

Now $P_0(X) = \mathbb{E}(Y|X)$ is estimated non parametrically. We define

$$\hat{m}_\theta(X_i) = (C_{\theta,j} - q_j^\top \hat{P}_0(X_i))_j.$$

Let $W_i = (X_i, Y_i)$.
Under regular (usual) assumptions about the non parametric estimation of $P_0(X)$, let γ be the rate of convergence of $\hat{P}_0(\cdot)$ toward $P_0(\cdot)$ and take $\rho_n = cn^\alpha$ with $\alpha < 2\gamma - \frac{1}{2}$.
We have:

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} g_{\rho_n}(\hat{m}_\theta(X_i)) - \mathbb{E}[g_{\rho_n}(m_\theta(X_i))]\right)$$
$$= \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} g_{\rho_n}(m_\theta(X_i)) - \mathbb{E}[g_{\rho_n}(m_\theta(X_i))] + \alpha(W_i)\right) + o_p(1),$$

Implementation II

with $\alpha(W_i) = \frac{\partial g_{\rho_n}(m_\theta(X_i))}{\partial m}^\top Q^\top(\mathbf{1}\{Y_i = y\} - h_0(X_i))$.

• $\alpha(\cdot)$ is the adjustment term that arises because of the first stage estimation of $P_0(\cdot)$.

Now, we can define the confidence region.

$$\xi_n(\theta) = \sqrt{n}\frac{\frac{1}{n}\sum_{i=1}^{n} g_{\rho_n}(\hat{m}_\theta(X_i))}{\sqrt{V_n}}$$

with V_n an estimator of the asymptotic variance of $g_\rho(m_\theta(X_i)) + \alpha(W_i)$. The confidence region of level $1-\alpha$ is simply define as follows.:

$$\mathrm{CR}_n(1-\alpha) = \{\xi_n(\theta) \geq z_\alpha\}$$

Implementation III

Proposition

Let $\rho_n = cn^{\alpha}$ *with* $\alpha < 2\gamma - \frac{1}{2}$ *and* $c > 0$ *a constant, let* $\hat{m}_\theta$ *a non-parametric estimator satisfying the usual regularity assumptions. Then,* $\mathrm{CR}_n(1-\alpha)$ *is asymptotically valid and consistent, i.e.,*

- *Asymptotic validity:* $\liminf_{n\to\infty} \inf_{\theta\in\Theta_I} \Pr(\theta \in \mathrm{CR}_n(1-\alpha)) \geq 1-\alpha$.
- *Consistency:* $\forall\theta \notin \Theta_I$, $\Pr(\theta \in \mathrm{CR}_n(1-\alpha)) \to 0$.

Outline

Monte Carlo simulations - setup.

$$\Pi_1 = \beta_1 - \delta_{1,1} N_1 - \delta_{2,1} N_2 + \varepsilon_1$$
$$\Pi_2 = \beta_2 - \delta_{1,2} N_1 - \delta_{2,2} N_2 + \varepsilon_2,$$

$\beta_1 = 3$, $\beta_2 = 2$, $\delta_{11} = \delta_{22} = 1.5$ and $\delta_{12} = \delta_{21} = 0.5$. To make the exposition easier, we assume that the econometrician knows β_1 and β_2 and that $\delta_{11} = \delta_{22}$ as well as $\delta_{12} = \delta_{21}$.

- The sample size is $n = 1,000$.
- The number of Monte Carlo replications is 1,000.
- For each sample, we compute the decision to reject or not $\theta \in \Theta_I$ for a 5% level of significance.
- The grid tested is composed by values from 1 to 2 with a tick of 0.02 for δ_{ii} and values from 0.4 to 1.4 with a tick of 0.02 for δ_{ij}, i.e., 2601 points tested in total.
- Given the DGP, $P_0 = [0.021, 0.074, 0.256, 0.047, 0.131, 0.421, 0.012, 0.034, 0.0004]^\top$.

Without covariates - normal shocks I

Sample size		500		1000	
Inference method	Inequality set	Coverage	Volume	Coverage	Volume
AS	min-max	0.962	95	0.941	48
AS	core 2	0.977	98	0.975	51
AS	core 3	0.977	99	0.975	51
AS	core 5	0.977	99	0.975	51
AS	all inequalities	0.945	80	0.941	42
CCK	min-max	0.988	129	0.971	66
CCK	core 2	0.982	115	0.98	59
CCK	core 3	0.982	115	0.98	59
CCK	core 5	0.982	115	0.98	59
CCK	all inequalities	0.993	136	0.992	69

Without covariates - normal shocks II

Sample size		500		1000	
smooth min rho=10	min-max	1	508	1	399
smooth min rho=10	core 2	1	288	1	230
smooth min rho=10	core 3	1	320	1	252
smooth min rho=10	core 5	1	320	1	252
smooth min rho=10	all inequalities	1	454	1	379
smooth min rho=50	min-max	1	258	1	191
smooth min rho=50	core 2	1	172	1	124
smooth min rho=50	core 3	1	177	1	126
smooth min rho=50	core 5	1	177	1	126
smooth min rho=50	all inequalities	1	227	1	183
smooth min rho=100	min-max	0.976	103	0.986	68
smooth min rho=100	core 2	0.932	73	0.965	46
smooth min rho=100	core 3	0.932	73	0.965	46
smooth min rho=100	core 5	0.932	73	0.965	46
smooth min rho=100	all inequalities	0.868	53	0.936	38

Results with one covariate I

	Coverage	$\min \delta_{11}$	$\max \delta_{11}$	$\min \delta_{12}$	$\max \delta_{12}$	Nb. points
Smooth-*core* P_0, $\rho = 20$	0.9840	1.3179	1.6515	0.6362	1.0011	255
Smooth-min $-$ max P_0, $\rho = 20$	0.9980	1.2957	1.6694	0.6089	1.0248	333
Smooth-min $-$ max $+$ equalities, $\rho = 20$	0.9910	1.3346	1.6457	0.6330	0.9772	235
Smooth-*core* P_0, $\rho = 50$	0.6990	1.3865	1.5963	0.6703	0.8948	100
Smooth-min $-$ max P_0, $\rho = 50$	0.7360	1.3991	1.5951	0.6542	0.8759	99
Smooth-min $-$ max $+$ equalities, $\rho = 50$	0.6440	1.4137	1.5858	0.6729	0.8647	78

Choice of ρ_n I

Two opposite forces for the choice of ρ_n. A small ρ_n makes the asymptotic normality of the test statistic more accurate. A high ρ_n limits the bias induced by the estimation of an outer set.

$$\sqrt{n} g_{\rho_n}(m_{\theta,n}) = \underbrace{\sqrt{n} g_{\rho_n}(m_\theta)}_{\text{Outer set bias}} + \underbrace{\nabla g_{\rho_n}(m_\theta)\sqrt{n}(m_{\theta,n} - m_0)}_{\text{First order approximation}}$$
$$+ \underbrace{\frac{\rho_n}{\sqrt{n}} \sqrt{n}(m_{\theta,n} - m_\theta)^\top \frac{H_{\rho_n}(\tilde{m}_{\theta,n})}{\rho_n} \sqrt{n}(m_{\theta,n} - m_\theta)}_{\text{Rest in Taylor's expansion}}$$

$$\left|\mathbb{E}\left[\text{Bias} + \text{Accuracy terms}\right]\right| \leq (p - J_0) \frac{\frac{1}{\rho_n} e^{-1}}{1 + J_0} + \frac{\rho_n}{\sqrt{n}} K_0$$

Choice of ρ_n II

with $K_0 > 0$ a constant that evaluates the expectation of the quadratic term in the expansion. Thus, the choice ρ_n^* that minimizes this upper bound is equal to:

$$\rho_n^* = n^{1/4}\sqrt{\frac{(p - J_0)e^{-1}}{(1 + J_0)K_0}}$$

To be validated by simulations. We see that the "optimal" choice of ρ_n increases with the number of non-binding moments and decreases with the number of binding moments and the variance of these moments. Let us observe that the optimal speed of divergence $\alpha^* = \frac{1}{4}$ is also contained in $(0, 1/2)$

Outline

Conclusion

- In this paper, we develop new tools to simplify the estimation of entry games when the Equilibrium selection mechanism is unrestricted
- We develop an algorithm which allows us to recursively select a subset of inequalities and to compute the theoretical upper bounds
- We propose a new estimation strategy, which is based on smoothing the identified set to recover a pivotal asymptotic distribution.
- We show that this new procedure can seamlessly accommodate covariates
- What we plan to do next:
  - Conduct full scale Monte Carlo simulations to assess the performance of our new estimation procedure
  - Apply our new tools to study competition in the French retail industry
  - Extend this estimation approach to other models characterized by moment inequalities

References I

Andrews, D. W. and Shi, X. (2013).
Inference based on conditional moment inequalities.
Econometrica, 81(2):609–666.

Andrews, D. W. and Soares, G. (2010).
Inference for parameters defined by moment inequalities using generalized moment selection.
Econometrica, 78(1):119–157.

Armstrong, T. B. (2014).
Weighted ks statistics for inference on conditional moment inequalities.
Journal of Econometrics, 181(2):92–116.

Armstrong, T. B. and Chan, H. P. (2016).
Multiscale adaptive inference on conditional moment inequalities.
Journal of Econometrics, 194(1):24–43.

References II

Chernozhukov, V., Chetverikov, D., and Kato, K. (2018).
Inference on causal and structural parameters using many moment inequalities.
The Review of Economic Studies, 86(5):1867–1900.