Supplement to "A Geometric Approach to Inference in Set-Identified Entry Games"

by Christian Bontemps and Rohit Kumar

Supplementary Appendix

This Supplementary Appendix contains the proofs and the algorithms of the paper "A Geometric Approach to Inference in Set-Identified Entry Games".[20] The three player example is detailed and some additional Monte Carlo simulations are provided.

A.1 Proof of Proposition 1

First, observe that, by a revealed preference argument, the region of ε that corresponds to an outcome $y_j^{(K)} = (a_{j,1}, \ldots, a_{j,N})^\top$ with K active firms in equilibrium is included in the region:

$$\mathcal{R}(y_j^{(K)}) = \left\{ \varepsilon = (\varepsilon_1, \ldots, \varepsilon_N) : \begin{array}{ll} \varepsilon_i \leq -\beta_i - K\alpha_i & \text{if } a_{j,i} = 0 \\ \varepsilon_i > -\beta_i - (K-1)\alpha_i & \text{if } a_{j,i} = 1 \end{array} \right\}.$$

Following the last result, we show that there is no region of ε which predicts two outcomes with different numbers of active firms. Let $y_j^{(K)}$ and $y_{j'}^{(K')}$ two outcomes with $K < K'$. There is at least one firm i which is not active in the first case and active in the second case , i.e., $a_{j,i} = 0$ for $y_j^{(K)}$ and 1 for $y_j^{(K')}$. The necessary conditions above imply that the profit shock for this firm is less or equal to $-\beta_i - K\alpha_i$ in the first case and strictly above $-\beta_i - (K'-1)\alpha_i \geq -\beta_i - K\alpha_i$, in the second case. Consequently, there is no intersection between $\mathcal{R}(y_j^{(K)})$ and $\mathcal{R}(y_{j'}^{(K')})$.

Now, assume, without loss of generality, that $S = \{y_1^{(K)}, \ldots, y_m^{(K)}\}$ is a collection of outcomes in multiplicity. We first characterize the region of ε that generates this set of outcomes. First, we define three subsets of $\{1, \ldots, N\}$. N_0 is the set of indices for which the action of player i is 0 for all outcomes in S, and N_1 is the set of indices i for which the action of player i is 1 for all outcomes in S. The remaining set N_s corresponds to the players who play actions 0 or 1 across the outcomes of S. Without loss of generality, we assume that $N_0 = \{1, 2, \ldots, n_0\}$, $N_1 = \{n_0+1, n_0+2, \ldots, n_0+n_1\}$, and $N_s = \{n_0+n_1+1, \ldots, N\}$. We now prove that $\mathcal{R}_S^{(K)}(\theta)$, the region of ε that predicts all

[20]https://doi.org/10.106/j.jeconom.2020.04.021

outcomes in S, is defined as follows:

$$\mathcal{R}_S^{(K)}(\theta) = \left\{ \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{pmatrix} : \begin{pmatrix} \varepsilon_i \leq -\beta_i - (K-1)\cdot\alpha_i & i \leq n_0 \\ \varepsilon_i > -\beta_i - K\cdot\alpha_i & n_0 < i \leq n_0 + n_1 \\ -\beta_i - (K-1)\cdot\alpha_i < \varepsilon_i \leq -\beta_i - K\cdot\alpha_i & i > n_0 + n_1 \end{pmatrix} \right\}.$$

First, take ε in the region defined by the right hand side. Each firm 1 to n_0 is only profitable when it has $K-2$ competitors, each firm n_0+1 to n_0+n_1 is profitable with K competitors, and each of the remaining firms is profitable with $K-1$ competitors. In a situation with complete information, firms n_0+1 to n_0+n_1 enter the market, firms 1 to n_0 do not enter, and $K-n_1$ firms out of the last $n_s = N-n_0-n_1$ enter. There are therefore $\binom{N-n_0-n_1}{K-n_1}$ possibilities.

Conversely, consider ε in $\mathcal{R}_S^{(K)}(\theta)$, i.e. assume that there is a region which predicts all the outcomes in S as possible equilibria. Obviously, this region is contained in $\bigcap_{1\leq j\leq m} \mathcal{R}(y_j^{(K)})$. If, for $i \leq n_0$, one of the profit shocks ε_i were between $-\beta_i-(K-1)\cdot\alpha_i$ and $-\beta_i-K\cdot\alpha_i$, the corresponding firms could enter to replace one of the last n_s firms, which is in contradiction with the fact that the model predicts all the outcomes in S only. Thus, in fact, $\varepsilon_i \leq -\beta_i-(K-1)\cdot\alpha_i$, $i \leq n_0$. Similarly, $\varepsilon_i > -\beta_i - K\cdot\alpha_i$ for $n_0 < i \leq n_0+n_1$. This proves the reverse inclusion. The cardinality of S is therefore $\binom{N-n_0-n_1}{K-n_1}$.

A.2 Proof of Proposition 2

Following Proposition 1, any set $S = \{y_1^{(K)}, \ldots, y_m^{(K)}\}$ of outcomes in multiplicity is such that there are n_0 firms that never enter, n_1 firms that always enter and $n_s = N-n_0-n_1$ that enter in some outcomes and do not in others, with there being in total $K-n_1$ entering among these n_s firms for each outcome (thus, $n_s > K-n_1$). Obviously, $n_1 \leq K-1$ because S contains at least two different outcomes. There are $\binom{N}{n_1}$ choices for these n_1 firms. Among the remaining $N-n_1$, n_0 never enter and $n_s = N-n_0-n_1$ "switch" across outcomes. For each value of n_0, there are $\binom{N-n_1}{n_0}$ choices for each choice of the n_1 firms. As $n_s \geq K-n_1+1$, n_0 is therefore bounded above by $N-K-1$. The number of multiple equilibria regions is equal to:

$$s_K = \sum_{n_1=0}^{K-1} \sum_{n_0=0}^{N-K-1} \binom{N}{n_1}\binom{N-n_1}{n_0}.$$

A.3 Proof of Proposition 3

The convexity of the set $A(\theta)$ can be easily proved from the expression of the $P_j^{(K)}(\theta,\eta)$'s in Equation (3) for any $K=0,\dots,N$ and any $j=1,\dots,d_K$. Let $\lambda\in[0,1]$, $\eta_1(\cdot)$ and $\eta_2(\cdot)$ two selection mechanisms and $P_1=P(\theta,\eta_1)$ and $P_2=P(\theta,\eta_2)$, two vectors of choice probabilities. First, $\eta(\cdot)=\lambda\eta_1(\cdot)+(1-\lambda)\eta_2(\cdot)$ is also a selection mechanism. Second, for any $K=0,\dots,N$ and any $j=1,\dots,d_K$,

$$P_j^{(K)}(\theta,\eta)=P_j^{(K)}(\theta,\lambda\eta_1+(1-\lambda)\eta_2)=\lambda P_j^{(K)}(\theta,\eta_1)+(1-\lambda)P_j^{(K)}(\theta,\eta_2).$$

Consequently, $P(\theta,\eta)=\lambda P_1+(1-\lambda)P_2$.

For the cartesian product, consider two different ε_1 and ε_2 in $\mathcal{R}_S^{(K_1)}(\theta)$ and $\mathcal{R}_{S'}^{(K_2)}(\theta)$, for $K_1\neq K_2$; the equilibrium selection mechanism is equal to zero when $y\in\mathcal{Y}_{K_1}$ and $\varepsilon=\varepsilon_2$ or when $y\in\mathcal{Y}_{K_2}$ and $\varepsilon=\varepsilon_1$.

A.4 Proof of Proposition 4

We introduce some useful additional notation. For any $S\in S^{(K)}$ and any $j\le d_K$, we define

$$u_j(S)=\frac{\int_{\mathcal{R}_S^{(K)}(\theta)}\eta(y_j^{(K)}|\varepsilon,\theta)dF(\varepsilon;\theta)}{\int_{\mathcal{R}_S^{(K)}(\theta)}dF(\varepsilon;\theta)}$$

and set $u_j(S)=0$ when $S\notin S^{(K)}$. Note that for all j such that $y_j^{(K)}\notin S$, $u_j(S)=0$ because a ε in $\mathcal{R}_S^{(K)}(\theta)$ does not predict $y_j^{(K)}$ as a potential outcome. By construction, $0\le u_j(S)\le 1$ and

$$\sum_{j|\,y_j^{(K)}\in S}u_j(S)=1.$$

for any $S\in S^{(K)}$. We also define the possibility set for $u_j(S)$, $j=1,\dots,d_K$ as

$$U^{(K)}(S)=\left\{u_j(S)\in[0;1],j=1,\dots,d_K,\ \text{such that}\ \sum_{j|\,y_j^{(K)}\in S}u_j(S)=1\ \text{and}\ u_j(S)=0,\ \text{if}\ y_j^{(K)}\notin S\right\},$$

and

$$\mathcal{U}^{(K)}=\left\{U^{(K)}(S),S\in S^{(K)}\right\}.$$

Based on this additional notation, we can define the set $B_K(\theta)$ as

$$B_K(\theta) = \left\{ P^{(K)} : P_j^{(K)} = \Delta_j^{(K)}(\theta) + \sum_{S \in S_j^{(K)}} u_j(S)\Delta_S^{(K)}(\theta), \ \ u_j(S) \in U^{(K)}(S), j = 1, \ldots, d_K, S \in S^{(K)} \right\}. \tag{A.1}$$

Following the definition of the support function:

$$\begin{aligned}
\delta\left(q_K; B_K(\theta)\right) &= \sup_{P^{(K)} \in B_K(\theta)} q_K^\top P^{(K)} \\
&= \sum_{j=1}^{d_K} q_{j,K}\Delta_j^{(K)}(\theta) + \sup_{u_j(S) \in \mathcal{U}^{(K)}} \sum_{j=1}^{d_K} \left(q_{j,K} \sum_{S \in S_j^{(K)}} u_j(S)\Delta_S^{(K)}(\theta) \right) \\
&= \sum_{j=1}^{d_K} q_{j,K}\Delta_j^{(K)}(\theta) + \sup_{u_j(S) \in \mathcal{U}^{(K)}} \sum_{j=1}^{d_K} \left(q_{j,K} \sum_{S \in S^{(K)}} u_j(S)\Delta_S^{(K)}(\theta) \right).
\end{aligned}$$

The last equality (the sum is indexed by $S \in S^{(K)}$ instead of $S \in S_j^{(K)}$) is true because $u_j(S)$ is equal to zero when $S \notin S_j^{(K)}$.

Consequently,

$$\begin{aligned}
\sup_{u_j(S) \in \mathcal{U}^{(K)}} \sum_{j=1}^{d_K} \left(q_{j,K} \sum_{S \in S^{(K)}} u_j(S)\Delta_S^{(K)}(\theta) \right) &= \sup_{u_j(S) \in \mathcal{U}^{(K)}} \sum_{S \in S^{(K)}} \Delta_S^{(K)}(\theta) \left(\sum_{j=1}^{d_K} q_{j,K} u_j(S) \right) \\
&= \sum_{S \in S^{(K)}} \Delta_S^{(K)}(\theta) \sup_{u_j(\cdot) \in U^{(K)}(S)} \left(\sum_{j=1}^{d_K} q_{j,K} u_j(S) \right) \\
&= \sum_{S \in S^{(K)}} \Delta_S^{(K)}(\theta) \left(\max_{j | y_j^{(K)} \in S} q_{j,K} \right)
\end{aligned}$$

Thus,

$$\delta\left(q; B_K(\theta)\right) = \sum_{j=1}^{d_K} q_{j,K}\Delta_j^{(K)}(\theta) + \sum_{S \in S^{(K)}} \Delta_S^{(K)}(\theta) \left(\max_{j | y_j^{(K)} \in S} q_{j,K} \right)$$

We can therefore reorder according to the new partition $\mathcal{O}_{i_1}^{(K)}, \mathcal{O}_{i_2}^{(K)}, \ldots, \mathcal{O}_{i_{d_K}}^{(K)}$.

$$\delta\left(q; B_K(\theta)\right) = \sum_{j=1}^{d_K} q_{j,K}\Delta_j^{(K)}(\theta) + \sum_{j=1}^{d_K} q_{i_j,K} \left\{ \sum_{S \in \mathcal{O}_{i_j}^{(K)}} \Delta_S^{(K)}(\theta) \right\}.$$

A.5 Proof of Proposition 6

We only need to prove the following result:

$$\forall K \in \{0,1,2,\dots,N\},\ \ \forall q_K \in \mathcal{Q}_K,\ q_K^\top P_0^{(K)} \leq \delta^*(q_K; B_K(\theta)) \Longrightarrow \forall q \in \mathbb{R}^{2^N},\ q^\top P_0 \leq \delta^*(q; A(\theta)).$$

Let q_K be a direction of $\mathbb{R}^{d_K}$ and assume that its components are ranked in the following order

$$q_{i_1,K} \geq q_{i_2,K} \geq \ldots \geq q_{i_{d_K},K}.$$

Let the d_K directions of $\mathbb{R}^{d_K}$, $e_{i_1}^{(K)}$, $e_{i_1,i_2}^{(K)}$,..., $e_{i_1,i_2,\dots,i_{d_K-1}}^{(K)}$, $e_{i_1,i_2,\dots,i_{d_K}}^{(K)}$, where the components are equal to 1 when the indices are subscripts of $e^{(K)}$ and 0 otherwise. Obviously, these directions belong to $\mathcal{Q}_K$.[21] We can write q_K as a function of these directions with non-negative weights:

$$q_K = (q_{i_1,K} - q_{i_2,K})e_{i_1}^{(K)} + (q_{i_2,K} - q_{i_3,K})e_{i_1,i_2}^{(K)} + \ldots + (q_{i_{d_K-1},K} - q_{i_{d_K},K})e_{i_1,i_2,\dots,i_{d_K-1}}^{(K)} + q_{i_{d_K},K}e_{1,2,\dots,d_K}^{(K)}.$$

Assume that the inequalities $\tilde{q}_K^\top P_0^{(K)} \leq \delta^*(\tilde{q}_K; B_K(\theta))$ are satisfied for any direction $\tilde{q}_K \in \mathcal{Q}_K$. We have:

$$\begin{aligned}
q_K^\top P_0^{(K)} &= (q_{i_1,K} - q_{i_2,K})(e_{i_1}^{(K)})^\top P_0^{(K)} + (q_{i_2,K} - q_{i_3,K})(e_{i_1,i_2}^{(K)})^\top P_0^{(K)} + \ldots + q_{i_{d_K},K}(e_{1,2,\dots,d_K}^{(K)})^\top P_0^{(K)} && \text{(A.2)}\\
&\leq (q_{i_1,K} - q_{i_2,K})\delta^*(e_{i_1}^{(K)}; B_K(\theta)) + (q_{i_2,K} - q_{i_3,K})\delta^*(e_{i_1,i_2}^{(K)}; B_K(\theta)) + \ldots + q_{i_{d_K},K}\delta^*(e_{1,2,\dots,d_K}^{(K)}; B_K(\theta)) && \text{(A.3)}\\
&\leq \delta^*((q_{i_1,K} - q_{i_2,K})e_{i_1}^{(K)}; B_K(\theta)) + \delta^*((q_{i_2,K} - q_{i_3,K})e_{i_1,i_2}^{(K)}; B_K(\theta)) + \ldots + \delta^*(q_{i_{d_K},K}e_{1,2,\dots,d_K}^{(K)}; B_K(\theta)) && \text{(A.4)}\\
&\leq \delta^*((q_{i_1,K} - q_{i_2,K})e_{i_1}^{(K)} + (q_{i_2,K} - q_{i_3,K})e_{i_1,i_2}^{(K)} + \ldots + q_{i_{d_K},K}e_{1,2,\dots,d_K}^{(K)}; B_K(\theta)) = \delta^*(q_K; B_K(\theta)) && \text{(A.5)}
\end{aligned}$$

Inequality (A.3) comes from the fact that the directions $e_{i_1,\dots}^{(K)}$ belong to $\mathcal{Q}_K$; inequality (A.4) holds because the support function is positive homogeneous; inequality (A.5) is due to the subadditivity of the support function. Consequently:

$$q^\top P_0 = \sum_{K=0}^{N} q_K^\top P_0^{(K)} \leq \sum_{K=0}^{N} \delta^*(q_K; B_K(\theta)) = \delta^*(q; A(\theta)).$$

[21] Observe that $e_{i_1,i_2,\dots,i_{d_K}}^{(K)} = e_{1,2,\dots,d_K}^{(K)}$.

A.6 Proof of Proposition 7

Note that $\mathcal{Y}_K$ is well connected (the empty set is connected). Thus, $\mathcal{Y}_K \in \Omega_K$ for $K = 0, 1, \ldots, N$. If the inequality holds for all well-connected subsets, then

$$\mathbb{P}(\mathcal{Y}_K) \leq \mathcal{L}(\mathcal{Y}_K) \tag{A.6}$$

However, we also know that

$$\mathcal{L}(\mathcal{Y}_K \cup \mathcal{Y}_{K'}) = \mathcal{L}(\mathcal{Y}_K) + \mathcal{L}(\mathcal{Y}_{K'})$$

because $\Gamma_{\mathcal{Y}_K}$ is a component of $\Gamma_{\mathcal{Y}}$, i.e., there is no multiplicity between $\mathcal{Y}_K$ and $\mathcal{Y}_{K'}$. As $\mathcal{L}(\mathcal{Y}) = 1$, we have

$$\sum_{K=0}^{N} \mathcal{L}(\mathcal{Y}_K) = 1 \tag{A.7}$$

From Equations (A.6) and (A.7), we have

$$\begin{aligned} \mathbb{P}(\mathcal{Y}_K) &= 1 - \sum_{i \neq K} \mathbb{P}(\mathcal{Y}_i) \\ &\geq 1 - \sum_{i \neq K} \mathcal{L}(\mathcal{Y}_i) = \mathcal{L}(\mathcal{Y}_K), \end{aligned}$$

where the second line uses Inequality (A.6) and the last equality comes from (A.7). Finally, we have

$$\mathbb{P}(\mathcal{Y}_K) = \mathcal{L}(\mathcal{Y}_K), \tag{A.8}$$

for $K = 0, 1, \ldots, N$.

• We first show that if the inequality holds for all well connected subset, then for any subset $C \subset \mathcal{Y}$, $\mathbb{P}(C) \leq \mathcal{L}(C)$.

Assume that $C = \cup_{K=1}^{N} C_K$, where $C_K \subset \mathcal{Y}_K$. Following Corollary 9, $\mathcal{L}(C) = \sum_{K=1}^{N} \mathcal{L}(C_K)$. Thus, if the inequality holds for each C_K, then it clearly holds for C. Without loss of generality, we now assume that there is one K such that $C \subset \mathcal{Y}_K$.

If C is not well connected, then $\mathcal{Y}_K \setminus C$ is not connected in $\Gamma_{\mathcal{Y}_K \setminus C}$. Therefore, $\mathcal{Y}_K \setminus C$ is a disjoint union of p components $\{W_i\}_{i=1}^{p}$ of the graph $\Gamma_{\mathcal{Y}_K \setminus C}$. Define, for each i in $1, \ldots, p$

$$B_i = C \cup W_1 \cup \ldots \cup W_{i-1} \cup W_{i+1} \cup \ldots \cup W_p$$

$\mathcal{Y}_{\mathcal{K}} \backslash B_i = W_i$ is connected in $\Gamma_{\mathcal{Y}_{\mathcal{K}} \backslash B_i} = \Gamma_{W_i}$ because W_i is a component of the graph $\Gamma_{\mathcal{Y}_K \backslash C}$. B_i is therefore well connected, and thus, $B_i \in \Omega_K$. Therefore, by definition,

$$\mathbb{P}(B_i) \leq \mathcal{L}(B_i).$$

We can now impose a lower bound on $\mathbb{P}(W_i)$ using $\mathbb{P}(W_i) + \mathbb{P}(B_i) = \mathbb{P}(\mathcal{Y}_K)$.

$$\mathbb{P}(W_i) \geq \mathcal{L}(\mathcal{Y}_K) - \mathcal{L}(B_i),$$

because $\mathbb{P}(\mathcal{Y}_K) = \mathcal{L}(\mathcal{Y}_K)$ from Equation (A.8) and the inequality above.

We can now impose an upper bound on $\mathbb{P}(C)$:

$$\begin{aligned}
\mathbb{P}(C) &= \mathbb{P}(\mathcal{Y}_K) - \mathbb{P}(\mathcal{Y}_K \backslash C) \\
&= \mathbb{P}(\mathcal{Y}_K) - \sum_{i=1}^{k} \mathbb{P}(W_i) \\
&\leq \mathcal{L}(\mathcal{Y}_K) - \sum_{i=1}^{p} \left[\mathcal{L}(\mathcal{Y}_K) - \mathcal{L}(B_i)\right]
\end{aligned}$$

We finally prove that the last term is $\mathcal{L}(C)$. For each i, following the definition of the Choquet capacity, $\mathcal{L}(\mathcal{Y}_K) - \mathcal{L}(B_i)$ is the sum of probabilities of the unique regions of outcomes of W_i and multiplicity regions only involving outcomes of W_i. Since W_i is not connected to W_j in $\Gamma_{\mathcal{Y}_K \backslash C}$, $\sum_{i=1}^{p} \left[\mathcal{L}(\mathcal{Y}_K) - \mathcal{L}(B_i)\right]$ is the probability of the unique region of outcomes of $\mathcal{Y}_K \backslash C$ and multiplicities only involving outcomes of $\mathcal{Y}_K \backslash C$. Hence, $\mathcal{L}(\mathcal{Y}_K) - \sum_{i=1}^{p} \left[\mathcal{L}(\mathcal{Y}_K) - \mathcal{L}(B_i)\right]$ is the sum of probabilities of unique regions of outcomes in C and multiplicity regions involving only outcomes in C. This is $\mathcal{L}(C)$. We therefore have

$$\mathbb{P}(C) \leq \mathcal{L}(C).$$

• We now prove that if a well-connected subset B of $\mathcal{Y}_K$ is not part of Ω_K, we can define a DGP where all inequalities $\mathbb{P}(C) \leq \mathcal{L}(C)$ hold except in B, thus violating the assumption that Ω is core determining for $\mathcal{L}$.

Assume that there are p elements in B, $y_1, \ldots, y_p$. We omit in the proof the superscript (K) for ease of exposition. For a given $\varepsilon > 0$, we consider the following probability outcome. It is an outcome in which we reallocate some of the predictions in the multiple equilibria regions from the first outcome y_1 and y_{p+1} to the $p-1$ outcomes $y_2, \ldots, y_p$.

$$
\begin{aligned}
\mathbb{P}(y_1) &= \mathcal{L}(y_1) - (p-2)\varepsilon \\
\mathbb{P}(y_2) &= \mathcal{L}(\{y_1, y_2\}) - \mathcal{L}(y_1) + \varepsilon \\
&\vdots \\
\mathbb{P}(y_{p-1}) &= \mathcal{L}(\{y_1, \ldots, y_{p-1}\}) - \mathcal{L}(\{y_1, \ldots, y_{p-2}\}) + \varepsilon \\
\mathbb{P}(y_p) &= \mathcal{L}(\{y_1, \ldots, y_p\}) - \mathcal{L}(\{y_1, \ldots, y_{p-1}\}) + \varepsilon \\
&- \\
\mathbb{P}(y_{p+1}) &= \mathcal{L}(\{y_1, \ldots, y_{p+1}\}) - \mathcal{L}(\{y_1, \ldots, y_p\}) - \varepsilon \\
\mathbb{P}(y_{p+2}) &= \mathcal{L}(\{y_1, \ldots, y_{p+2}\}) - \mathcal{L}(\{y_1, \ldots, y_{p+1}\}) \\
&\vdots \\
\mathbb{P}(y_{|\mathcal{Y}_K|}) &= \mathcal{L}(\{y_1, \ldots, y_{|\mathcal{Y}_K|}\}) - \mathcal{L}(\{y_1, \ldots, y_{|\mathcal{Y}_K|-1}\})
\end{aligned}
$$

Our goal is to show that the inequalities $\mathbb{P}(C) \leq \mathcal{L}(C)$ are satisfied for all elements of $\mathcal{Y}_\mathcal{K}$ but B for some adequate choice of ε.

First, note that the violation of the inequality for B is obvious because:

$$
\begin{aligned}
\mathbb{P}(B) &= \sum_{i=1}^{p} \mathbb{P}(y_i) \\
&= \sum_{i=2}^{p} \Big[\mathcal{L}(\{y_1, \ldots, y_i\}) - \mathcal{L}(\{y_1, \ldots, y_{i-1}\}) + \varepsilon\Big] + \mathcal{L}(y_1) - (p-2)\varepsilon \\
&= \mathcal{L}(\{y_1, \ldots, y_p\}) + \varepsilon \\
&= \mathcal{L}(B) + \varepsilon
\end{aligned}
$$

Now, we show that no other inequality is violated for this constructed probability under some condition on ε.

(i) Find r such that $y_2, \ldots, y_r \in B$ are directly connected to y_1 in the graph $\Gamma_{\mathcal{Y}_K}$ (this is possible because $\Gamma_{\mathcal{Y}_K}$ is connected). $y_2, \ldots, y_r$ can be divided into two subgroups: the subgroup $y_2, \ldots, y_{r_1}$ of elements such that each element of $y_{r+1}, \ldots, y_p$ is directly connected to some element of this subgroup and subgroup $y_{1+r_1}, \ldots, y_r$, which is not connected to any element from $y_{1+r}, \dots, y_p$ as shown in Figure 2. Note that $y_2, \dots, y_{r_1}$ and $y_{1+r_1}, \dots, y_r$ may have some connections. Henceforth, we assume that $r > r_1$. It is easy to adapt the proof to the case in which $r = r_1$. Note further that it may be the case that some elements of $y_{r+1}, \dots, y_p$ are not directly connected to any element of $y_2, \dots, y_{r_1}$. As B is well connected, they are connected to some elements of $y_{r+1}, \dots, y_p$, and we can also adapt the proof to this case by adding a layer on our tree, as shown in the right part of Figure 2. We assume henceforth that this is not the case, but again, the proof is similar.

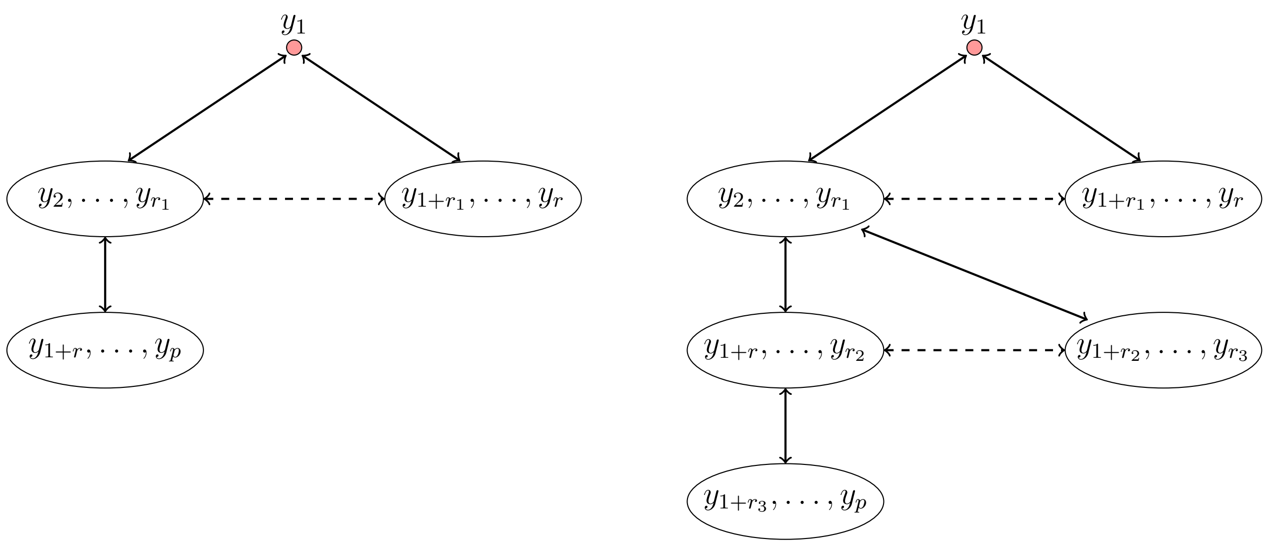


Figure 2: Construction of a tree from elements of B starting from y_1 and an additional layer in the tree if all $y_{1+r}, \dots, y_p$ may not be connected directly to some element $y_2, \dots, y_{1+r_1}$.

(ii) If S contains y_1, it is easy to prove that $\mathbb{P}(S) \leq \mathcal{L}(S)$ because we simply subtract some ε.

(iii) We have to prove it now for the subset S that does not contain y_1.

Let z_2 be an element of $\{y_2, \dots, y_{r_1}\}$, z_3 be an element of $\{y_{1+r_1}, \dots, y_r\}$ and z_4 be an element of $\{y_{1+r}, \dots, y_p\}$. First, y_1 and z_2 are connected. This means that there is at least one region of multiple equilibria that predicts y_1 and z_2 among other outcomes. We call Δ_2 the area of this region. Obviously, we have $\Delta_2 > 0$. Similarly, we have $\mathcal{L}(\{y_1, z_2\}) \leq \mathcal{L}(\{y_1\}) + \mathcal{L}(\{z_2\}) - \Delta_2$ because Δ_2 is counted in both $\mathcal{L}(\{y_1\})$ and $\mathcal{L}(\{z_2\})$ (and there may be other regions of multiple equilibria than the one considered here that predict these outcomes). We do the same for y_1 and z_3 with Δ_3 and the same for z_4 and one element of $\{y_2, \dots, y_{r_1}\}$ that we call z_2' with Δ_4.

z_2' may be z_2 or not. The construction is described in Figure 3.

(a) We have

$$\mathbb{P}(z_2) \leq \mathcal{L}(\{z_2\}) - \Delta_2 + \varepsilon.$$

If $z_2 = y_2$, the inequality expressed above yields the following:

$$\begin{aligned}\mathbb{P}(\{y_2\}) &= \mathcal{L}(\{y_1, y_2\}) - \mathcal{L}(\{y_1\}) + \varepsilon \\ &\leq \mathcal{L}(\{y_1\}) + \mathcal{L}(\{y_2\}) - \Delta_2 - \mathcal{L}(y_1) + \varepsilon.\end{aligned}$$

If $z_2 = y_3$, we can prove it similarly:

$$\begin{aligned}\mathbb{P}(y_3) &= \mathcal{L}(\{y_1, y_2, y_3\}) - \mathcal{L}(\{y_1, y_2\}) + \varepsilon \\ &\leq \mathcal{L}(\{y_1, y_2\}) + \mathcal{L}(\{y_3\}) - \Delta_2 - \mathcal{L}(\{y_1, y_2\}) + \varepsilon.\end{aligned}$$

and so forth (the last inequality holds because there is at least the region of area Δ_2 in multiplicity between $z_2 = y_3$ and $y_1 \in \{y_1, y_2\}$).

(b) Similarly, $\mathbb{P}(z_3) \leq \mathcal{L}(\{z_3\}) - \Delta_3 + \varepsilon$ and $\mathbb{P}(z_4) \leq \mathcal{L}(\{z_4\}) - \Delta_4 + \varepsilon$

(c) Again $\mathbb{P}(z_2, z_3) \leq \mathcal{L}(\{z_2, z_3\}) - \min(\Delta_2, \Delta_3) + 2\varepsilon$, $\mathbb{P}(z_3, z_4) \leq \mathcal{L}(\{z_3, z_4\}) - \min(\Delta_3, \Delta_4) + 2\varepsilon$, $\mathbb{P}(z_2, z_4) \leq \mathcal{L}(\{z_2, z_4\}) - \min(\Delta_2, \Delta_4) + 2\varepsilon$ and $\mathbb{P}(z_2, z_3, z_4) \leq \mathcal{L}(\{z_2, z_3, z_4\}) - \min(\Delta_2, \Delta_3, \Delta_4) + 3\varepsilon$.

(d) Therefore, if $3\varepsilon < \min_{i \in \{2,3,4\}} \Delta_i$, then $\mathbb{P}(S) \leq \mathcal{L}(S)$ for every $S \subset \left\{z_2, z_3, z_4\right\}$.

It is straightforward to extend the argument for any subset that contains elements of the type (z_2, z_3, z_4). We need to choose ε such that $(p-1)\varepsilon < \min_{S \in S^{(K)}} \Delta_S^{(K)}(\theta)$.

We therefore have $\mathbb{P}(S) \leq \mathcal{L}(S)$ for every $S \subset B$.

(iv) As $\mathbb{P}(S) \leq \mathcal{L}(S)$ for every $S \subset B$, it is easy to see that this is also satisfies for any union $S \cup C$, where $C \subseteq \mathcal{Y}_K \backslash B$. We still have to prove that the inequalities $\mathbb{P}(S) \leq \mathcal{L}(S)$ are satisfied for $S = B \cup C$, where $C \subseteq \mathcal{Y}_K \backslash B$. We will build a similar tree for $\mathcal{Y}_K \backslash B$. Select $y_{p+1} \in \mathcal{Y}_K \backslash B$. If C contains y_{p+1}, checking the inequality is straightforward. Now, we have to prove this when C does not contain y_{p+1}. The proof is similar to that above.

(v) Find s such that y_{p+1} is directly connected to each $y_{p+2}, \ldots, y_s \in \mathcal{Y}_K \backslash B$ in graph $\Gamma_{\mathcal{Y}_K \backslash B}$ (this is possible because $\mathcal{Y}_K \backslash B$ is connected in $\Gamma_{\mathcal{Y}_K \backslash B}$). $y_{p+2}, \ldots, y_s$ can be divided into two subgroups: the subgroup $y_{p+2}, \ldots, y_{s_1}$ such that each outcome $y_{1+s}, \ldots, y_{d_K}$ is directly connected to some element of this subgroup and the subgroup $y_{1+s_1}, \ldots, y_s$, which is not connected to any element from $y_{1+s}, \ldots, y_{d_K}$. Note further that not all $y_{1+s}, \ldots, y_{|\mathcal{Y}_K|}$ may be connected directly to some element $y_{p+2}, \ldots, y_{s_1}$, but if not, then we will only have an additional layer in the tree, and the proof can easily be modified to any additional layer. Two alternative, simplified trees are also built similar to the construction above (see Figure 4). If $\varepsilon < \min_{S \in S^{(K)}} \Delta_S^{(K)}(\theta)$, a similar argument to that above proves that the inequalities are satisfied for any C.

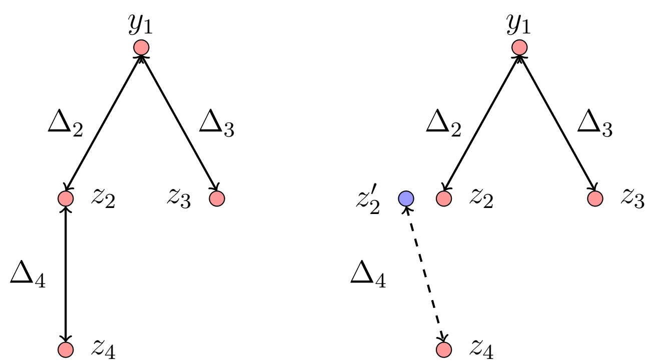


Figure 3: Two simplified trees.

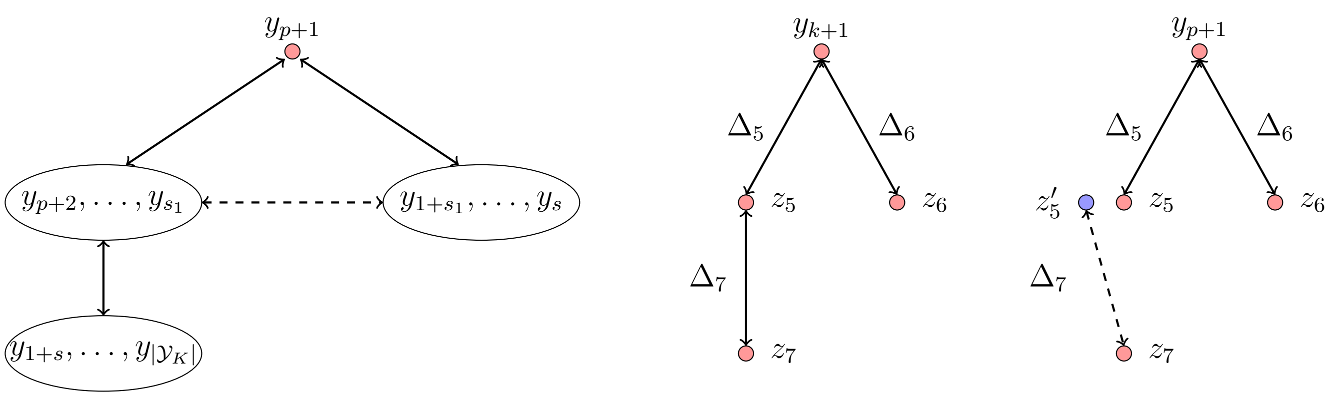


Figure 4: Construction of a tree from elements of $\mathcal{Y}_K \backslash B$ starting from y_{p+1} and two simplified trees.

A.7 Constructing the Core Determining Class

Proposition 7 permits to build an algorithm to construct the core determining class. The algorithm is decomposed in four steps:

1. First, we collect the subsets $B \subseteq \mathcal{Y}_K$ such that B is not connected in Γ_B. We call this collection $\mathcal{D}_K$.

2. Second, we define $\mathcal{D}_K^* = \left\{\mathcal{Y}_K \backslash C : C \in \mathcal{D}_K\right\}$, which is the collection of non well connected subsets of $\mathcal{Y}_K$. As a matter of fact, for any $B \in \mathcal{D}_K^*$, there exists $C \in \mathcal{D}_K$, such that $B = \mathcal{Y}_K \backslash C$ and C, due to the first step, is not connected in Γ_C.

3. Third, we define $\Omega_K = \mathcal{P}^*(\mathcal{Y}_K) \backslash \mathcal{D}_K^*$, which gathers all well connected subsets of $\mathcal{Y}_K$.

4. Finally, we define $\Omega = \left\{\Omega_K : K = 1, \ldots, N\right\}$, the well connected subsets of $\mathcal{Y}$.

A.7.1 The algorithm details of the first step.

The first step is the one which needs more details.

Find $\mathcal{D}_K = \left\{B \subseteq \mathcal{Y}_K : \ B \text{ is not connected in } \Gamma_B\right\}$. For simplification, we denote by $\mathcal{P}^*(C)$, for any set C, the collection of all non empty subsets of C and by an abuse of notation $\mathcal{P}^*(\mathcal{C})$, for any collection of sets $\mathcal{C}$, the collection of $\mathcal{P}^*(C)$ for all the elements of $\mathcal{C}$. We also define the concatenation $\oplus$ of two collections C and B is defined as

$$C \oplus B = \bigcup_{c \in C} \left\{c \cup b : b \in B\right\}$$

For example,

$$\left\{\{y_1^{(K)}\}; \{y_2^{(K)}\}\right\} \oplus \left\{\{y_3^{(K)}\}; \{y_4^{(K)}\}\right\} = \left\{\{y_1^{(K)}, y_3^{(K)}\}; \{y_1^{(K)}, y_4^{(K)}\}; \{y_2^{(K)}, y_3^{(K)}\}; \{y_2^{(K)}, y_4^{(K)}\}\right\}.$$

We denote by $S^{(K)}(h)$, the elements of $S^{(K)}$ with h outcomes. The intuition behind the algorithm is to start from the pairs which are not in multiple equilibria and to extend the sequence by sequentially increasing the set with k tuples k increasing from 3 to d_K.

- For any j, consider the set $C_j^{(K)}$ of outcomes which are not in multiple equilibria with $y_j^{(K)}$ and call $C_j^{+(K)}$ the union of $\{y_j^{(K)}\}$ and $C_j^{(K)}$. We now prove that $C_j^{+(k)}$ is the largest subset of $\mathcal{Y}_K$ such that $y_j^{(K)}$ is an isolated node in the graph generated by itself.

  Proof. If $y_j^{(K)}$ would be connected to one another node in the graph generated by $C_j^{+(k)}$, there would exist outcomes $y_{i_1}^{(K)}, \ldots, y_{i_m}^{(K)} \in C_j^{(K)}$ such that $y_j^{(K)}, y_{i_1}^{(K)}, \ldots, y_{i_m}^{(K)}$ are in multiplicity. Following the characterization of the multiple equilibria in Proposition 1, we define the values n_1 of indices where firms always play 1 across all the outcomes, n_0 the indices of the firms which always play 0, and n_s the indices of firms which switch. The series of outcomes gathering all the possible switching values, there exist at least one outcome $y_{i_p}^{(K)}$ among $y_{i_1}^{(K)}, \ldots, y_{i_m}^{(K)} \in C_j^{(K)}$ which differentiate from $y_j^{(K)}$ only from two switcher firms, one switching from 0 to 1 and one switching from 1 to 0 when going from $y_j^{(K)}$ to $y_{i_p}^{(K)}$. This is in contradiction with the definition of $C_k^{(K)}$ which collects all outcomes which can't be in multiplicity with $y_j^{(K)}$.

  There $y_j^{(K)}$ is isolated in $C_j^{+(k)}$ and any other outcome outside this set being in multiple equilibria with $y_j^{(K)}$ can't be added to this set. ■

  Therefore, we initialize our construction of the set $\mathcal{D}_K$ by collecting across j all subsets of $\mathcal{Y}_K$ which contain $y_j^{(K)}$ and any part of $C_j^{(K)}$:

$$\mathcal{S}_{K,1} = \bigcup_{j=1}^{d_K} \left\{ y_j^{(K)} \oplus \mathcal{P}^* \left(C_j^{(K)} \right) \right\}.$$

- Now, we extend the construction. We first consider any pair $\left\{y_i^{(K)}, y_j^{(K)}\right\}$ in multiplicity. We can show that $C = \left\{y_i^{(K)}, y_j^{(K)}\right\} \cup \left\{C_i^{(K)} \cap C_j^{(K)}\right\}$, is not connected in Γ_C. The proof is similar than above. We can therefore augment $\mathcal{S}_{K,1}$ by all the possible combinations of the previous type:

$$\mathcal{S}_{K,2} = \mathcal{S}_{K,1} \cup \left\{ \bigcup_{i,j \text{ s.t. } \{y_i^{(K)}, y_j^{(K)}\} \in S_K(2) \text{ and } C_i^{(K)} \cap C_j^{(K)} \neq \emptyset} \left\{ \{y_i^{(K)}, y_j^{(K)}\} \oplus \mathcal{P}^* \left(C_i^{(K)} \cap C_j^{(K)} \right) \right\} \right\}.$$

- and so on, with triples, ...until $h = d_K$.

$$\mathcal{S}_{K,h} = \mathcal{S}_{K,h-1} \cup \left\{ \bigcup_{\left\{a \in \mathcal{P}^*(S_K(h)) \setminus \mathcal{P}^*(S_K(h-1)) : \bigcap_{j, y_j^{(K)} \in a} C_j^{(K)} \neq \emptyset \right\}} \left\{ a \oplus \mathcal{P}^* \left(\bigcap_{j, y_j^{(K)} \in a} C_j^{(K)} \right) \right\} \right\}.$$

Now $\mathcal{S}_{K,d_K} = \mathcal{D}_K$. Take any set B not connected in Γ_B. There exists a component C of B in Γ_B. Define n_1, n_0 and n_s like above (see Proposition 1), this set is picked in step $h = \binom{n_s}{K-n_1}$.

A.7.2 The Core Determining Class for $N = 4$

We now apply the previous construction for the entry game with four players. First, for $K \neq 2$, any subset of $\mathcal{Y}_K$ is in the core determining class because all series of outcomes are in multiplicity. Therefore, we only detail the case $K = 2$. There are six outcomes in $B_2(\theta)$.

$$\begin{aligned} y_1^{(2)} &= (1,1,0,0)^\top, \\ y_2^{(2)} &= (1,0,1,0)^\top, \\ y_3^{(2)} &= (1,0,0,1)^\top, \\ y_4^{(2)} &= (0,1,1,0)^\top, \\ y_5^{(2)} &= (0,1,0,1)^\top, \\ y_6^{(2)} &= (0,0,1,1)^\top. \end{aligned}$$

First, we apply proposition 1 to find all the elements of $S^{(2)}$, the outcomes in multiple equilibria

$$\begin{aligned} S^{(2)} = \Big\{ & \{y_1^{(2)}, y_2^{(2)}\}, \{y_1^{(2)}, y_3^{(2)}\}, \{y_1^{(2)}, y_4^{(2)}\}, \{y_1^{(2)}, y_5^{(2)}\}, \{y_2^{(2)}, y_3^{(2)}\}, \{y_2^{(2)}, y_4^{(2)}\}, \{y_2^{(2)}, y_6^{(2)}\}, \{y_3^{(2)}, y_5^{(2)}\}, \\ & \{y_3^{(2)}, y_6^{(2)}\}, \{y_4^{(2)}, y_5^{(2)}\}, \{y_4^{(2)}, y_6^{(2)}\}, \{y_5^{(2)}, y_6^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}, y_3^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}, y_4^{(2)}\}, \{y_1^{(2)}, y_3^{(2)}, y_5^{(2)}\}, \\ & \{y_1^{(2)}, y_4^{(2)}, y_5^{(2)}\}, \{y_2^{(2)}, y_3^{(2)}, y_6^{(2)}\}, \{y_2^{(2)}, y_4^{(2)}, y_6^{(2)}\}, \{y_3^{(2)}, y_5^{(2)}, y_6^{(2)}\}, \{y_4^{(2)}, y_5^{(2)}, y_6^{(2)}\}, \\ & \{y_1^{(2)}, y_2^{(2)}, y_3^{(2)}, y_4^{(2)}, y_5^{(2)}, y_6^{(2)}\} \Big\} \end{aligned}$$

- It happens that, for any j, $y_j^{(2)}$ and $y_{7-j}^{(2)}$ are never in multiplicity. So, following our algorithm, we have

$$\begin{aligned} \mathcal{S}_{2,1} &= \bigcup_{j=1}^{6} \left\{ y_j^{(2)} \oplus \mathcal{P}^* \left(y_{7-j}^{(2)} \right) \right\} \\ &= \left\{ \{y_1^{(2)}, y_6^{(2)}\}, \{y_2^{(2)}, y_5^{(2)}\}, \{y_3^{(2)}, y_4^{(2)}\} \right\}. \end{aligned}$$

- For $h = 2$, there is no pair $\{y_i^{(2)}, y_j^{(2)}\}$ in $S^{(2)}(2)$ such that $\{C_i^{(2)} \cap C_j^{(2)}\} \neq \emptyset$.
- For $h = 3$, this is the same ; there is no 3-tuple $\{y_i^{(2)}, y_j^{(2)}, y_l^{(2)}\}$ in $S^{(2)}(3)$ such that $\{C_i^{(2)} \cap C_j^{(2)} \cap C_l^{(2)}\} \neq \emptyset$.
- For $h = 4$ or 5, there is no element in $S^{(2)}(h)$.
- Finally for $h = 6$, there is only one element, $\{y_1^{(2)}, y_2^{(2)}, y_3^{(2)}, y_4^{(2)}, y_5^{(2)}, y_6^{(2)}\}$. But the intersection of the $C_j^{(2)}$ for all these elements is empty.

Therefore

$$\mathcal{D}_2 = \Big\{\{y_1^{(2)}, y_6^{(2)}\}, \{y_2^{(2)}, y_5^{(2)}\}, \{y_3^{(2)}, y_4^{(2)}\}\Big\}$$

and

$$\begin{aligned}\mathcal{D}_2^* &= \Big\{\{y_1^{(2)}, y_6^{(2)}\}, \{y_2^{(2)}, y_5^{(2)}\}, \{y_3^{(2)}, y_4^{(2)}\} \backslash D, D \in \mathcal{D}_2\Big\}, \\ &= \Big\{\{y_2^{(2)}, y_3^{(2)}, y_4^{(2)}, y_5^{(2)}\}, \{y_1^{(2)}, y_3^{(2)}, y_4^{(2)}, y_6^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}, y_5^{(2)}, y_6^{(2)}\}\Big\}.\end{aligned}$$

Among all the non empty subparts of $\mathcal{Y}_2$, i.e. 63 sets, only 3 are not in the core determining class. For example, Figure 5 draws the graph $(V_{\{y_3^{(2)},y_4^{(2)},y_5^{(2)},y_6^{(2)}\}}, E_{\{y_3^{(2)},y_4^{(2)},y_5^{(2)},y_6^{(2)}\}})$ from the knowledge of S_2 (there is no link between $y_3^{(2)}$ and $y_4^{(2)}$ because they don't occur in multiplicity involving only outcomes $\{y_3^{(2)}, y_4^{(2)}, y_5^{(2)}, y_6^{(2)}\}$). This graph is clearly connected, so $\{y_1^{(2)}, y_2^{(2)}\}$ is well connected and is part of the core determining class. A contrario, $\{y_3^{(2)}, y_4^{(2)}\}$ is not connected in $\Gamma_{\{y_3^{(2)},y_4^{(2)}\}}$ because these outcomes are not in multiplicity. Therefore $\{y_1^{(2)}, y_2^{(2)}, y_5^{(2)}, y_6^{(2)}\}$ is not well connected and is not part of the core determining class.

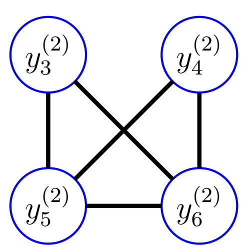


Figure 5: Graph $(V_{\{y_3^{(2)},y_4^{(2)},y_5^{(2)},y_6^{(2)}\}}, E_{\{y_3^{(2)},y_4^{(2)},y_5^{(2)},y_6^{(2)}\}})$ from the multiplicity.

The core determining class is

$$\Big\{\{y_1^{(2)}\},\{y_2^{(2)}\},\{y_3^{(2)}\},\{y_4^{(2)}\},\{y_5^{(2)}\},\{y_6^{(2)}\},\{y_1^{(2)},y_2^{(2)}\},\{y_1^{(2)},y_3^{(2)}\},\{y_1^{(2)},y_4^{(2)}\},\{y_1^{(2)},y_5^{(2)}\},\{y_1^{(2)},y_6^{(2)}\}$$
$$\{y_2^{(2)},y_3^{(2)}\},\{y_2^{(2)},y_4^{(2)}\},\{y_2^{(2)},y_5^{(2)}\},\{y_2^{(2)},y_6^{(2)}\},\{y_3^{(2)},y_4^{(2)}\},\{y_3^{(2)},y_5^{(2)}\},\{y_3^{(2)},y_6^{(2)}\},\{y_4^{(2)},y_5^{(2)}\},\{y_4^{(2)},y_6^{(2)}\}$$
$$\{y_5^{(2)},y_6^{(2)}\},\{y_1^{(2)},y_2^{(2)},y_3^{(2)}\},\{y_1^{(2)},y_2^{(2)},y_4^{(2)}\},\{y_1^{(2)},y_2^{(2)},y_5^{(2)}\},\{y_1^{(2)},y_2^{(2)},y_6^{(2)}\},\{y_1^{(2)},y_3^{(2)},y_4^{(2)}\},\{y_1^{(2)},y_3^{(2)},y_5^{(2)}\}$$
$$\{y_1^{(2)},y_3^{(2)},y_6^{(2)}\},\{y_1^{(2)},y_4^{(2)},y_5^{(2)}\},\{y_1^{(2)},y_4^{(2)},y_6^{(2)}\},\{y_1^{(2)},y_5^{(2)},y_6^{(2)}\},\{y_2^{(2)},y_3^{(2)},y_4^{(2)}\},\{y_2^{(2)},y_3^{(2)},y_5^{(2)}\},\{y_2^{(2)},y_3^{(2)},y_6^{(2)}\}$$
$$\{y_2^{(2)},y_4^{(2)},y_5^{(2)}\},\{y_2^{(2)},y_4^{(2)},y_6^{(2)}\},\{y_2^{(2)},y_5^{(2)},y_6^{(2)}\},\{y_3^{(2)},y_4^{(2)},y_5^{(2)}\},\{y_3^{(2)},y_4^{(2)},y_6^{(2)}\},\{y_3^{(2)},y_5^{(2)},y_6^{(2)}\},\{y_4^{(2)},y_5^{(2)},y_6^{(2)}\}$$
$$\{y_1^{(2)},y_2^{(2)},y_3^{(2)},y_4^{(2)}\},\{y_1^{(2)},y_2^{(2)},y_3^{(2)},y_5^{(2)}\},\{y_1^{(2)},y_2^{(2)},y_3^{(2)},y_6^{(2)}\},\{y_1^{(2)},y_2^{(2)},y_4^{(2)},y_5^{(2)}\},\{y_1^{(2)},y_2^{(2)},y_4^{(2)},y_6^{(2)}\}$$
$$\{y_1^{(2)},y_3^{(2)},y_4^{(2)},y_5^{(2)}\},\{y_1^{(2)},y_3^{(2)},y_5^{(2)},y_6^{(2)}\},\{y_1^{(2)},y_4^{(2)},y_5^{(2)},y_6^{(2)}\},\{y_2^{(2)},y_3^{(2)},y_4^{(2)},y_6^{(2)}\},\{y_2^{(2)},y_3^{(2)},y_5^{(2)},y_6^{(2)}\}$$
$$\{y_2^{(2)},y_4^{(2)},y_5^{(2)},y_6^{(2)}\},\{y_3^{(2)},y_4^{(2)},y_5^{(2)},y_6^{(2)}\},\{y_1^{(2)},y_2^{(2)},y_3^{(2)},y_4^{(2)},y_5^{(2)}\},\{y_1^{(2)},y_2^{(2)},y_3^{(2)},y_4^{(2)},y_6^{(2)}\}$$
$$\{y_1^{(2)},y_2^{(2)},y_3^{(2)},y_5^{(2)},y_6^{(2)}\},\{y_1^{(2)},y_2^{(2)},y_4^{(2)},y_5^{(2)},y_6^{(2)}\},\{y_1^{(2)},y_3^{(2)},y_4^{(2)},y_5^{(2)},y_6^{(2)}\},\{y_2^{(2)},y_3^{(2)},y_4^{(2)},y_5^{(2)},y_6^{(2)}\}$$
$$\{y_1^{(2)},y_2^{(2)},y_3^{(2)},y_4^{(2)},y_5^{(2)},y_6^{(2)}\}\Big\}$$

A.8 The geometric selection procedure

The geometric selection procedure consists in, first determining the local extreme point and, then, deriving the supporting hyperplanes at this extreme point. We adopt the convention of $e_{s_1 i_1,\ldots,s_k i_k}$ is the vector where the component i_j is 1 if $s_j = +1$ and -1 if $s_j = -1$.

Determining the local extreme point The procedure to determine the local extreme point is the following:

(1) Pick K and select the subvector $P_0^{(K)}$. This is a vector in a space of dimension d_K.

(2) For each component i, $i = 1, \ldots, d_K$, calculate the support function in direction e_i and e_{-i}. Calculate the width in direction e_i, i.e. $D_i = \delta^*(e_i; B_K(\theta)) + \delta^*(e_{-i}; B_K(\theta))$. Calculate the distance to the center point of the cube along the axis of component i: $x_i = \left\{ e_i^\top P_0^{(K)} - \delta^*(e_i; B_K(\theta)) + \frac{D_i}{2} \right\}$.

(3) Pick the coordinate i_1 of the highest values of $|x_i|$. If it is $x_{i_1} > 0$, i_1 is the highest index of the local extreme point, i.e. the local vertex is related to the order $i_1? \ldots ?$ where the remaining indices need to be found ; otherwise i_1 is the lowest index, i.e. the local vertex is related to the order $? \ldots ? i_1$.

(a) Assume $x_{i_1} > 0$. Pick e_{i_1}, and construct the orthogonal projection of $P_0^{(K)}$, $P_{i_1}^{(K)}$, onto the facet F_{i_1}, $x_{i_1} = \delta^*(e_{i_1}; B_K(\theta))$. Then restart Step 2 with now the second index, i'. For each $i' \neq i_1$, take $e_{i_1,i'}$ and $e_{i_1,-i'}$. Compute the width of the intersection of the facet and the set $B_K(\theta)$:

$$D_{i_1,i'} = \delta^*(e_{i_1,i'}; B_K(\theta)) - \delta^*(e_{i_1}; B_K(\theta)) + \delta^*(e_{i_1,-i'}; B_K(\theta)) - \delta^*(e_{i_1}; B_K(\theta)).$$

Calculate the distance to the center of the new cube which contains this intersection:

$$x_{i_1,i'} = \left(e_{i_1,i'}^\top P_{i_1}^{(K)} - \delta^*(e_{i_1,i'}; B_K(\theta)) + \frac{D_{i_1,i'}}{2} \right).$$

Pick the coordinate i_2 of the highest values of $|x_{i_1,i'}|$.

(b) If now $x_{i_1} \leq 0$. Pick e_{i_1}, and construct the orthogonal projection of $P_0^{(K)}$, $P_{i_1}^{(K)}$, onto the facet F_{-i_1}, $x_{i_1} = -\delta^*(e_{-i_1}; B_K(\theta))$. Then restart Step 2 with now the second index, i'. For each $i' \neq i_1$, take $e_{-i_1,i'}$ and $e_{-i_1,-i'}$. Compute the width of the intersection of the facet and the set $B_K(\theta)$:

$$D_{-i_1,i'} = \delta^*(e_{-i_1,i'}; B_K(\theta)) - \delta^*(e_{-i_1}; B_K(\theta)) + \delta^*(e_{-i_1,-i'}; B_K(\theta)) - \delta^*(e_{-i_1}; B_K(\theta)).$$

Calculate the distance to the center of the new cube which contains this intersection:

$$x_{-i_1,i'} = \left(e_{-i_1,i'}^\top P_{i_1}^{(K)} - \delta^*(e_{-i_1,i'}; B_K(\theta)) + \frac{D_{-i_1,i'}}{2} \right).$$

Pick the coordinate i_2 of the highest values of $|x_{-i_1,i'}|$.

(4) Repeat loop 2 and 3 until determining $i_1 i_2 \ldots i_{d_K}$. This is our local extreme point.

(5) Do this procedure for all the values K. Collect the local extreme points accordingly.

Finding the facets at one extreme point Assume the extreme point is $E_{i_1,i_2,\ldots,i_{d_K}}^{(K)}(\theta)$. We now want to determine the facets of $B_K(\theta)$ at this extreme point. The algorithm, detailed below, is based on the idea that, when multiple equilibria does not exist between a series of outcomes, the corresponding indices can be swapped if they are in consecutive ranks without changing the point $E_{i_1,i_2,\ldots,i_{d_K}}^{(K)}(\theta)$ in the space. It increases the number of inequalities that are binding at this extreme point.

1: Start with $\mathcal{L}_K = \big\{\{y_{i_1}^{(K)}\}\big\}$ and set $k = 2$.

k: Find the largest m such that $\big\{y_{i_k}^{(K)}, y_{i_{k+1}}^{(K)}, \dots, y_{i_m}^{(K)}\big\}$ are not in multiplicity with elements in $\big\{y_{i_j}^{(K)}\big\}$, $j \geq k$. Note $\big\{y_{i_k}^{(K)}, y_{i_{k+1}}^{(K)}, \dots, y_{i_m}^{(K)}\big\}$ can be in multiplicity with elements in $\big\{y_{i_j}^{(K)}\big\}$, $j \leq k-1$.

$$\mathcal{L}_K = \mathcal{L}_K \cup \Big\{\big\{y_{i_1}^{(K)}, y_{i_2}^{(K)}, \dots, y_{i_{k-1}}^{(K)}\big\} \oplus \mathcal{P}^*\big\{y_{i_k}^{(K)}, y_{i_{k+1}}^{(K)}, \dots, y_{i_m}^{(K)}\big\}\Big\}.$$

Then, update to $k+1$.[22]

R: Repeat the previous step for $k = 2, \dots, d_K$ steps and find

$$\mathcal{L}_K = \mathcal{L}_K \cap \Omega_K.$$

$\mathcal{L}_K$ can be converted into equivalent support directions. Any element C_K of $\mathcal{L}_K$ yields to the direction e_{C_K} following Equation (10).

We provide a simple illustration of this algorithm in section A.9. The local geometry of set $A(\theta)$ in the extreme point considered is

$$\mathcal{L} = \Big\{\mathcal{L}_K : K = 0, 1, \dots, N\Big\}. \tag{A.9}$$

Note that the composition of $\mathcal{L}$ is specific to each extreme point.

A.9 Determining the number of facets for $B_2(\theta)$ when K=4

Following the previous section, we now illustrate how to determine the number of facets in a given extreme point. Consider, for example, the extreme point $E_{1,2,3,4,5,6}^{(2)}(\theta)$ of $B_2(\theta)$. We now determine the number of facets. We know that it is at least 6 but, due to the fact that some outcomes are not in multiplicity, we know that this point is also the same point than $E_{1,2,4,3,5,6}^{(2)}(\theta)$. The procedure determines that. We show that, for this point:

[22] We define the concatenation $\oplus$ of two collections C and B as

$$C \oplus B = \bigcup_{c \in C} \Big\{c \cup b : b \in B\Big\}$$

and $\mathcal{P}^*$ is defined in Appendix A.7.1.

$$\mathcal{L}_K = \Big\{ \{y_1^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}, y_3^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}, y_4^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}, y_3^{(2)}, y_4^{(2)}\}, \\ \{y_1^{(2)}, y_2^{(2)}, y_3^{(2)}, y_4^{(2)}, y_5^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}, y_3^{(2)}, y_4^{(2)}, y_5^{(2)}, y_6^{(2)}\} \Big\}$$

It means that the inequalities which are binding in $E^{(2)}_{1,2,3,4,5,6}(\theta)$ are based on the following directions (that should be completed by zeros accordingly to give direction in $\mathbb{R}^{2^4}$): $e_1 = (1,0,0,0,0,0)^\top$, $e_{1,2} = (1,1,0,0,0,0)^\top$, $e_{1,2,3} = (1,1,1,0,0,0)^\top$, $e_{1,2,4} = (1,1,0,1,0,0)^\top$, $e_{1,2,3,4} = (1,1,1,1,0,0)^\top$, $e_{1,2,3,4,5} = (1,1,1,1,1,0)^\top$ and $e_{1,2,3,4,5,6} = (1,1,1,1,1,1)^\top$.

We now follow the steps of the algorithm introduced in section 4.3.1.

(1) Set $\mathcal{L}_K = \big\{\{y_1^{(2)}\}\big\}$.

(2) At step 2, find the largest m such that $\big\{y_2^{(2)}, y_3^{(2)}, \ldots, y_m^{(2)}\big\}$ are not in multiplicity even with outcomes in $\big\{y_2^{(2)}, y_3^{(2)}, \ldots, y_6^{(2)}\big\}$. Since $y_2^{(2)}$ and $y_3^{(2)}$ are in multiplicity, $m = 2$.

$$\begin{aligned} \mathcal{L}_K &= \Big\{\{y_1^{(2)}\}\Big\} \cup \Big\{\Big\{\{y_1^{(2)}\} \oplus \mathcal{P}^*(y_2^{(2)})\Big\} \Big\backslash \mathcal{D}_2^*\Big\} \\ &= \Big\{\{y_1^{(2)}\}\Big\} \cup \Big\{\{y_1^{(2)}, y_2^{(2)}\}\Big\} \\ &= \Big\{\{y_1^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}\}\Big\} \end{aligned}$$

(3) At step 3, we look for the largest m such that $\big\{y_3^{(2)}, \ldots, y_m^{(2)}\big\}$ are not in multiplicity even with outcomes in $\big\{y_3^{(2)}, \ldots, y_6^{(2)}\big\}$. Since $y_3^{(2)}$ and $y_4^{(2)}$ are not in multiplicity, but $y_3^{(2)}$ and $y_5^{(2)}$ are, $m = 4$.

$$\begin{aligned} \mathcal{L}_K &= \Big\{\{y_1^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}\}\Big\} \cup \Big\{\Big\{\{y_1^{(2)}, y_2^{(2)}\} \oplus \mathcal{P}^*\{y_3^{(2)}, y_4^{(2)}\}\Big\} \Big\backslash \mathcal{D}_2^*\Big\} \\ &= \Big\{\{y_1^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}, y_3^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}, y_4^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}, y_3^{(2)}, y_4^{(2)}\}\Big\} \end{aligned}$$

(4) At step 4, we look for the largest m such that $\big\{y_4^{(2)}, \ldots, y_m^{(2)}\big\}$ are not in multiplicity even with outcomes in $\big\{y_4^{(2)}, \ldots, y_6^{(2)}\big\}$. Since $y_4^{(2)}$ and $y_6^{(2)}$ are in multiplicity, $m = 5$.

$$\mathcal{L}_K = \Big\{ \{y_1^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}, y_3^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}, y_4^{(2)}\}, \{y_1^{(2)}, y_2^{(2)}, y_3^{(2)}, y_4^{(2)}\}, \\ \{y_1^{(2)}, y_2^{(2)}, y_3^{(2)}, y_4^{(2)}, y_5^{(2)}\} \Big\}$$

(5) Finally, add $\mathcal{Y}_2 = \left\{\{y_1^{(2)}, y_2^{(2)}, y_3^{(2)}, y_4^{(2)}, y_5^{(2)}, y_6^{(2)}\}\right\}$.

Observe that this does not correspond to the extreme point with the maximum number of facets. This happens in $E^{(2)}_{1,3,4,2,5,6}(\theta)$. In this case, we can start again the algorithm to determine that there are 8 facets defined by the following directions: $e_1 = (1,0,0,0,0,0)^\top$, $e_{1,3} = (1,0,1,0,0,0)^\top$, $e_{1,4} = (1,0,0,1,0,0)^\top$, $e_{1,3,4} = (1,0,1,1,0,0)^\top$, $e_{1,2,3,4} = (1,1,1,1,0,0)^\top$, $e_{1,3,4,5} = (1,0,1,1,5,0)^\top$, $e_{1,2,3,4,5} = (1,1,1,1,1,0)^\top$ and $e_{1,2,3,4,5,6} = (1,1,1,1,1,1)^\top$.

A.10 Proof of Proposition 8

We do the proof for $K = 1$ and it is similar for $K = 2$, which proves the global result.

Fix θ. The goal is to proof that, if a point $P_0^{(1)}$ does not belong to $B_1(\theta)$, a local selection procedure would detect it.

First, observe that any extreme point is linked to an order between the three possible equilibria. Each extreme point $E^{(1)}_{i_1,i_2,i_3}(\theta)$ has supporting hyperplanes with outer normal vectors, $e^{(1)}_{i_1}$, $e^{(1)}_{i_1,i_2}$ and $e^{(1)}_{i_1,i_2,i_3} = (1,1,1)^\top$.

There are three cases:

- If $P_0^{(1)}$ is outside the cube which contains $B_1(\theta)$. It means at least one of the values x_i is outside a bounded interval $[-D_i/2, D_i/2]$, where $D_i =$ is the width in direction e_i. The highest value of $|x_1|$, $|x_2|$, $|x_3|$ selects a face which separates $B_1(\theta)$ and $P_0^{(1)}$. Assume this is $|x_1|$ and that $x_1 > 0$. The first component of $P_0^{(1)}$ is above the largest value of the first component of any point of $B_1(\theta)$. The local extreme point is $E^{(1)}_{1??}(\theta)$. Whatever the next choice, the direction $e_1^{(1)}$ defines a supporting hyperplane of $B_1(\theta)$ at this extreme point which separates $B_1(\theta)$ and $P_0^{(1)}$. Consequently, $T_\infty(e_1^{(1)};\theta) < 0$.

- If $P_0^{(1)}$ is in the cube but not in $B_1(\theta)$. Whatever the choice of the extreme point of $B_1(\theta)$, the third direction $(1,1,1)^\top$ defines a supporting hyperplane which separates $B_1(\theta)$ and $P_0^{(1)}$. Consequently, $T_\infty(e^{(1)}_{1,2,3};\theta) < 0$. And, so forth for the other possibilities.

- If $P_0^{(1)} \in B_1(\theta)$, any choice of local extreme point is valid because, for any direction q, $T_\infty(q, K;\theta) \geq 0$.

When $P_0^{(1)} \in B_1(\theta)$, the procedure does not reject θ. When $P_0^{(1)} \notin B_1(\theta)$, the procedure does. It is therefore a valid and sharp characterization of $B_1(\theta)$.

A.11 Proof of Proposition 9

Under condition UI, following Lemma 3.1 of Romano and Shaikh (2008), we have

$$\sup_{P\in\mathcal{P}} \sup_{S\in\mathcal{L}} \left| \mathbb{P}\left(\sqrt{M}\left(\mu(P) - \hat{P}_M\right) \in S\right) - \Phi_{\Sigma(P)}(S) \right| \underset{M\to\infty}{\longrightarrow} 0,$$

where $\Phi_\Sigma(\cdot)$ is the cumulative distribution function of the centered multivariate normal distribution with variance Σ, $\mu(P) = \mathbb{E}_P(Y)$ and $\Sigma(P) = \text{diag}(\mu(P)) - \mu(P)\mu(P)^\top$ and $\mathcal{L}$ is a collection of convex sets with zero boundary.

Consider the directions q of $\mathcal{G}$ and relabel them $q_1, \ldots, q_m$. Then, define m convex sets in $\mathbb{R}^{2^N}$, D_1, ...,D_m such that,

$$\forall U \in D_i, \quad \frac{q_i^\top U}{\sqrt{q_i^\top \Sigma q_i}} \leq \min_{j\neq i} \frac{q_j^\top U}{\sqrt{q_j^\top \Sigma q_j}}.$$

Now, we can define, for a given $x \in \mathbb{R}$ the sets $S_1, ..., S_m$ ($S_i \subset D_i$) such that

$$\forall U \in S_i, \quad x \leq \frac{q_i^\top U}{\sqrt{q_i^\top \Sigma q_i}} \leq \min_{j\neq i} \frac{q_j^\top U}{\sqrt{q_j^\top \Sigma q_j}}.$$

Now, we have

$$\begin{aligned}\mathbb{P}\left(\inf_{q\in\mathcal{G}}\left(\sqrt{M}\frac{q^\top(\mu(P) - \hat{P}_M)}{\sqrt{q^\top \Sigma q}}\right) \geq x\right) &= \sum_{i=1}^m \mathbb{P}\left(\sqrt{M}(\mu(P) - \hat{P}_M) \in S_i\right) \\ &\underset{M\to\infty}{\longrightarrow} \sum_{i=1}^m \Phi_{\Sigma(P)}\left(S_i\right) = \sum_{i=1}^m \mathbb{P}(Z \in S_i),\end{aligned}$$

uniformly over $P \in \mathcal{P}$, for $Z \sim \mathcal{N}(0, \Sigma(P))$. Moreover,

$$\begin{aligned}\sum_{i=1}^m \mathbb{P}(Z \in S_i) =& \mathbb{P}\left(Z \in \bigcup_{i=1}^m S_i\right) \\ =& \mathbb{P}\left(\inf_{q\in\mathcal{G}} \frac{q^\top Z}{\sqrt{q^\top \Sigma q}} \geq x\right).\end{aligned}$$

So, uniformly over $P \in \mathcal{P}$,

$$\inf_{q \in \mathcal{G}} \left(\sqrt{M} \frac{q^\top (\mu(P) - \hat{P}_M)}{\sqrt{q^\top \Sigma q}} \right) \xrightarrow[M \to \infty]{d} \inf_{q \in \mathcal{G}} \frac{q^\top Z}{\sqrt{q^\top \Sigma q}}. \tag{A.10}$$

Following, Bontemps et al. (2012), proof of Proposition 10, we can now consider the two different cases:

- If P_0, the true choice probability vector belongs to $A(\theta)$, the set of minimizers of $T_M(q;\theta)$ tends to Q_θ, the set of minimizers of $T_\infty(q,\theta)$. This set may not be reduced to a singleton if P_0 is at the intersection of at least two facets. Therefore,

$$\begin{aligned}
\xi_M(\theta) &= \sqrt{M} \min_{q \in \mathcal{G}} \frac{T_M(q;\theta)}{\sqrt{q^\top \Sigma q}} \\
&= \sqrt{M} \inf_{q \in \mathcal{G}} \frac{T_M(q;\theta) - T_\infty(q;\theta) + T_\infty(q;\theta)}{\sqrt{q^\top \Sigma q}} \\
&= \inf_{q \in \mathcal{G}} \frac{\sqrt{M}(q^\top (P_0 - \hat{P}_M)) + \sqrt{M} T_\infty(q;\theta)}{\sqrt{q^\top \Sigma q}} \\
&= \inf_{q \in Q_\theta} \frac{\sqrt{M}(q^\top (P_0 - \hat{P}_M))}{\sqrt{q^\top \Sigma q}}
\end{aligned}$$

  The last equality holds because for any $q \in Q_\theta$, $T_\infty(q;\theta) = 0$ and $q \notin Q_\theta$, $T_\infty(q;\theta) > 0$. So asymptotically, the argmin belongs to Q_θ (remember that $\mathcal{G}$ is discrete). We conclude using the uniform convergence of Equation (A.10).

- If $P_0 \notin A(\theta)$, the value $T_\infty(q,\theta)$ is negative for any direction q. $T_M(q,\theta)$ converges uniformly in q, on the unit sphere, toward a strictly negative value and is therefore bounded away from zero uniformly. The rescaling by $\sqrt{M}$ makes the limit $-\infty$.

Now, we need to consider the fact that Σ is estimated. We need to use the following additional result to replace Σ by $\hat{\Sigma}$ in the proofs above:

$$\sup_{P \in \mathcal{P}} \left\| \hat{\Sigma}(P) - \Sigma(P) \right\| \xrightarrow{P} 0$$

where $\|.\|$ is the component-wise maximum of absolute value of each **element**. This follows from lemma S.7.1 in supplement to Romano and Shaikh (2012).

A.12 Proof of Corollary 11

In the proof of Proposition 9, we show that uniformly over $P \in \mathcal{P}$,

$$\xi_M(\theta) \xrightarrow[M\to\infty]{d} \inf_{(q)\in Q_\theta} \mathcal{N}(0, q^\top \Sigma(P)\, q)$$

if $P_0 \in A(\theta)$. Observe that the distribution depends on θ only through the minimizing set Q_θ, but θ doesn't affect the covariance of the distribution. Define

$$S = \cup_{\theta\in\Theta_I} Q_\theta$$

Since $S \subseteq \mathcal{G}$ the result follows.

A.13 Proof of Proposition 12

Following Proposition 2, we know that any subset of $\mathcal{Y}_1$ of cardinality greater than 2 corresponds to a multiple equilibria region. Consequently, $\Delta_S^{(1)}(\theta)$ for any subset $S \subseteq \mathcal{Y}_1$ is non-zero, and, following Proposition 4, any change in the order gives a different point. Let $i_1, i_2, \ldots, i_{d_1}$ be an order of the coordinates that defines an extreme point $E^{(1)}_{i_1,i_2,\ldots,i_{d_1}}(\theta)$ and $\mathcal{C}^{(1)}_{i_1,\ldots,i_{d_1}}$ be the cone of directions q such that $\delta^*(q; B_1(\theta)) = q^\top E^{(1)}_{i_1,\ldots,i_{d_1}}(\theta)$. Each direction defines a supporting hyperplane (or facet) of $B_1(\theta)$ at $E^{(1)}_{i_1,\ldots,i_{d_1}}(\theta)$.

Any direction q in the cone can be written as

$$q = (q_{i_1} - q_{i_2})e_{i_1} + (q_{i_2} - q_{i_3})e_{i_1,i_2} + \ldots + (q_{i_{d_1-1}} - q_{i_{d_1}})e_{i_1,i_2,\ldots,i_{d_1-1}} + q_{i_{d_1}} e_{i_1,i_2,\ldots,i_{d_1}}.$$

All the coefficients except the last one are positive. The cone is therefore generated by e_{i_1}, e_{i_1,i_2},..., $e_{i_1,i_2,\ldots,i_{d_1-1}}$, $e_{i_1,i_2,\ldots,i_{d_1}}$ or $-e_{i_1,i_2,\ldots,i_{d_1}}$.[23]

In other words, there are only d_1 supporting hyperplanes of $B_1(\theta)$ at this point, and it is sufficient to check the inequalities related to these hyperplanes/facets for a point locally around $E^{(1)}_{i_1,\ldots,i_{d_1}}(\theta)$.

A.14 Proof of Proposition 13

To find the upper bound on the number of facets in any extreme point for $1 < K < N - 1$, we first have to pack the indices which correspond to outcomes which are not in multiplicity.

[23] Remark that $e_{i_1,i_2,\ldots,i_{d_1}} = (1, 1, \ldots, 1)^\top$.

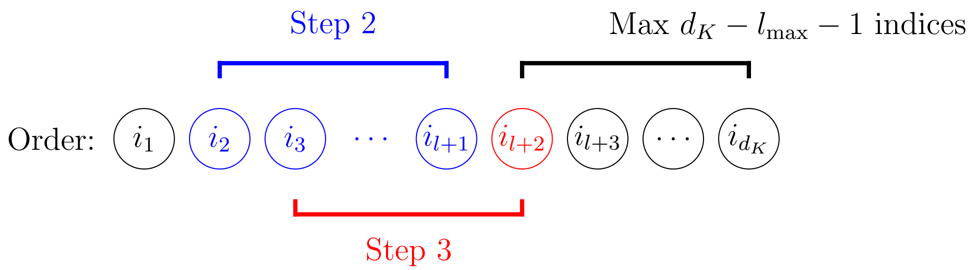
- Let us first define $l_{\max}$ the cardinality of the maximal subset $S \in \mathcal{Y}_K$ such that any pair of S is not in multiplicity. If we collect these indices from the second one, we can switch their order and still have the same point. For example, if $y_{i_2}^{(K)}$ and $y_{i_3}^{(K)}$ are not in multiplicity, the point $E_{i_1,i_2,i_3,???}^{(K)}$ is the same point than $E_{i_1,i_3,i_2,???}^{(K)}$ when the next orders don't change. Consequently, it defines at most one additional outer normal vector following the construction we used earlier in the case $K = 1$. The first outer normal vectors are $e_{i_1}^{(K)}$, $e_{i_1,i_2}^{(K)}$, $e_{i_1,i_3}^{(K)}$, $e_{i_1,i_2,i_3}^{(K)}$, etc. These series of indices are related to $2^{l_{\max}} - 1$ outer normal vector corresponding to all $e_{i_1,i_j}^{(K)}$, $j = 2..l_{\max} + 1$, $e_{i_1,i_j,i_k}^{(K)}$, j, k between 2 and $l_{\max} + 1$ up to $e_{i_1,i_2,\ldots,i_{l_{\max}+1}}^{(K)}$.

- Then we add indices according to the following rule. At each step k, starting from 3, the next index i_m is such that, if possible, $y_{i_k}^{(K)}, \ldots, y_{i_m}^{(K)}$ are not in multiplicity even with all remaining outcomes $\mathcal{Y}_K \backslash \{y_{i_1}^{(K)} \ldots, y_{i_{k-1}}^{(K)}\}$. Otherwise, we pick a random index and we go on. It adds at the maximum (if the added index is not in multiplicity with anybody before) $2^{l_{\max}-1}$ new supporting hyperplanes (as you can switch all orders between k and $k + l_{\max} - 1$, and count only those where the last index is not the last value, i.e. $2^{l_{\max}} - 2^{l_{\max}-1} = 2^{l_{\max}-1}$). See figure 6 below.

- After Step 1, it remains $d_K - l_{\max} - 1$ points after having chosen $i_1, i_2, \ldots, i_{l_{\max}+1}$.

- The maximum number of facets is therefore $\mathcal{L}^*_{\max} = 1 + 2^{l_{\max}} - 1 + (d_K - l_{\max} - 1)(2^{l_{\max}-1}) = 2^{l_{\max}} + (d_K - l_{\max} - 1)(2^{l_{\max}-1})$.

Observe that this bounds is a loose bound and can always be refined by brute force on a case to case basis. Nevertheless, it gives a sufficiently precise estimate of $c(\mathcal{L}^*)$ the true cut off value. When $N = 4$ and $K = 2$, $d_K = 6$ and $l_{max} = 2$ (see Section A.7.2). Applying the formula above we obtain, $\mathcal{L}^*_{\max} = 10$. The true $\mathcal{L}^*$ is equal to 8. However the two cut off values for a level of 5% (see Section 4.3) are $c(8) = -2.51$ and $c(10) = -2.58$. When $N = 5$ and $K = 2$, $d_K = 10$ and $l_{max} = 2$. $\mathcal{L}^*_{\max} = 18$ whereas $\mathcal{L}^* = 15$.



Step 2: we can switch any of these indices. $2^{l}_{\max} - 1$ facets.
Step 3: at max, we can switch any of these indices. $2^{l_{\max}-1}$ additional facets.
Note: l denotes $l_{\max}$ in the circles related to the order.

Figure 6: Counting the maximum number of facets at $E^{(K)}_{i_1,i_2,\ldots,i_{d_K}}(\theta)$.

K	$N=5$	$N=6$	$N=7$	$N=8$	$N=9$
2	2	3	3	4	4
3	2	4	7	8	10
4	1	3	7	14	14
5		1	3	8	14
6			1	4	10
7				1	4
8					1

Table VII: Value of $l_{\max}$ for $N \leq 9$

A.15 The simulation of the critical values

We propose an algorithm to simulate the critical values $c(\mathcal{G},\alpha)$ of Section 4.2. Observe that:

$$\min_{q\in\mathcal{G}} \frac{q^\top Z}{\sqrt{q^\top \hat{\Sigma} q}} = \min_{q\in\mathcal{G}} \frac{(\hat{\Sigma}^{1/2}q)^\top \hat{\Sigma}^{1/2}\tilde{Z}}{\sqrt{(\hat{\Sigma}^{1/2}q)^\top(\hat{\Sigma}^{1/2}q)}} + 0_p(1),$$

where $\tilde{Z} \sim N(0,\hat{\Sigma}^{-1})$. The matrix $\hat{\Sigma}^{1/2}$ rotates the quantities of interest but the direction q which minimizes the quantity are still the same.

1. Draw $\tilde{Z}$ from the normal distribution $\mathcal{N}(0,\hat{\Sigma}^{-1})$. Cut $\tilde{Z}$ in subvectors $\tilde{Z}^{(0)},\tilde{Z}^{(1)},\ldots,\tilde{Z}^{(N)}$.

2. For each $K\in\{0,1,\ldots,N\}$, order $\tilde{Z}_i^{(K)}$, $i=1,\ldots,d_K$ in the increasing order $\tilde{Z}_{i_1}^{(K)},\tilde{Z}_{i_2}^{(K)},etc.$ The direction q_K which minimizes $q_K^\top\hat{\Sigma}\tilde{Z}^{(K)}/\sqrt{q_K^\top\hat{\Sigma}q_K}$ is among the direction $e_{i_1}^{(K)}$, $e_{i_1,i_2}^{(K)}$, ..., $e_{i_1,i_2,\ldots,i_{d_K}}^{(K)}$. Calculate all the values $q_K^\top\hat{\Sigma}\tilde{Z}^{(k)}/\sqrt{q_K^\top\hat{\Sigma}q_K}$ for all the potential candidates and take the minimum one, called m_K.

3. Take $\underline{m} = \min_{K=0,\ldots,N} m_K$.

4. Repeat the previous steps, $S-1$ times to get S realizations of the distribution of the lower bound and take the α-quantile of this distribution. This is $c(\mathcal{G},\alpha)$.

A.16 The entry game with three players

In this section, we consider our entry game with three firms. The profit of firm i in market m, π_{im} is modeled as:

$$\pi_i = \beta_i + \alpha_i\Big(\sum_{j\neq i} a_j\Big) + \varepsilon_i, \tag{A.11}$$

where a_1 (resp. a_2, a_3) is equal to 1 when $\pi_1>0$ (resp. $\pi_2>0$, $\pi_3>0$), 0 otherwise. The joint distribution of $(\varepsilon_1,\varepsilon_2,\varepsilon_3)$, $F(\cdot;\gamma)$, is assumed to be known up to parameter and θ denote all parameters in the model. We also assume that α's are negative.

A.16.1 Multiple equilibria regions

There are, in this setup, eight regions of multiple equilibria, which correspond to the set of outcomes expressed in Table VIII. First note that $S^{(1)} = \{S_1,S_2,S_3,S_4\}$ and $S^{(2)} = \{S_5,S_6,S_7,S_8\}$.

$N=1$	$S_1=\begin{pmatrix}0,0,1\\0,1,0\end{pmatrix}$	$S_2=\begin{pmatrix}0,0,1\\1,0,0\end{pmatrix}$	$S_3=\begin{pmatrix}0,1,0\\1,0,0\end{pmatrix}$	$S_4=\begin{pmatrix}0,0,1\\0,1,0\\1,0,0\end{pmatrix}$
$N=2$	$S_5=\begin{pmatrix}0,1,1\\1,0,1\end{pmatrix}$	$S_6=\begin{pmatrix}0,1,1\\1,1,0\end{pmatrix}$	$S_7=\begin{pmatrix}1,0,1\\1,1,0\end{pmatrix}$	$S_8=\begin{pmatrix}0,1,1\\1,0,1\\1,1,0\end{pmatrix}$

Table VIII: All multiplicities in pure strategy Nash equilibrium of entry game with 3 players.

A.16.2 The set of predicted choice probabilities

Recall that the probability of each outcome can be written with the unknown selection mechanism $\eta(\cdot)$. For example,

$$P_{001} = P_1^{(1)}(\theta,\eta) = \Delta_1^{(1)}(\theta) + \sum_{S\in\{S_1,S_2,S_4\}} \int_{\mathcal{R}_S^{(K)}(\theta)} \eta((0,0,1)^\top|\varepsilon,\theta)dF(\varepsilon;\gamma),$$

Let $u_j(S_k)$ be defined like in Section A.4. The set $A(\theta)$ is the collection of points in $\mathbb{R}^8$ that can be written, for some specific choice of weights $u_j(S_k)$:

$$\begin{pmatrix} P_{000} \\ \hline P_{001} \\ P_{010} \\ P_{100} \\ \hline P_{011} \\ P_{101} \\ P_{110} \\ \hline P_{111} \end{pmatrix} = \begin{pmatrix} \Delta_1^{(0)}(\theta) \\ \hline \Delta_1^{(1)}(\theta) + u_1(S_1)\Delta_{S_1}^{(1)}(\theta) + u_1(S_2)\Delta_{S_2}^{(1)}(\theta) + u_1(S_4)\Delta_{S_4}^{(1)}(\theta) \\ \Delta_2^{(1)}(\theta) + u_2(S_1)\Delta_{S_1}^{(1)}(\theta) + u_2(S_3)\Delta_{S_3}^{(1)}(\theta) + u_2(S_4)\Delta_{S_4}^{(1)}(\theta) \\ \Delta_3^{(1)}(\theta) + u_3(S_2)\Delta_{S_2}^{(1)}(\theta) + u_3(S_3)\Delta_{S_3}^{(1)}(\theta) + u_3(S_4)\Delta_{S_4}^{(1)}(\theta) \\ \hline \Delta_1^{(2)}(\theta) + u_1(S_5)\Delta_{S_5}^{(2)}(\theta) + u_1(S_6)\Delta_{S_6}^{(2)}(\theta) + u_1(S_8)\Delta_{S_8}^{(2)}(\theta) \\ \Delta_2^{(2)}(\theta) + u_2(S_5)\Delta_{S_5}^{(2)}(\theta) + u_2(S_7)\Delta_{S_7}^{(2)}(\theta) + u_2(S_8)\Delta_{S_8}^{(2)}(\theta) \\ \Delta_3^{(2)}(\theta) + u_3(S_6)\Delta_{S_6}^{(2)}(\theta) + u_3(S_7)\Delta_{S_7}^{(2)}(\theta) + u_3(S_8)\Delta_{S_8}^{(2)}(\theta) \\ \hline \Delta_1^{(3)}(\theta) \end{pmatrix}, \tag{A.12}$$

with the constraint $\sum_{j|y_j^{(K)}\in S} u_j(S) = 1,\ \ 0 \le u_j(S) \le 1$ for $S \in S^{(K)}$. The partition here refers to different K (number of active firm in any outcome). This partition is very useful as the convex set decomposes into cartesian product of smaller dimension convex set. This convex set only need 18 directions to characterize it.

Figure 7 displays the set $B_1(\theta)$, its outer cube and the inequalities (in red) which are tested in our geometric selection procedure.

A.16.3 The directions used in the Monte Carlo experiment

Following Proposition 4, we have a closed-form expression for the support function. It is equal to:

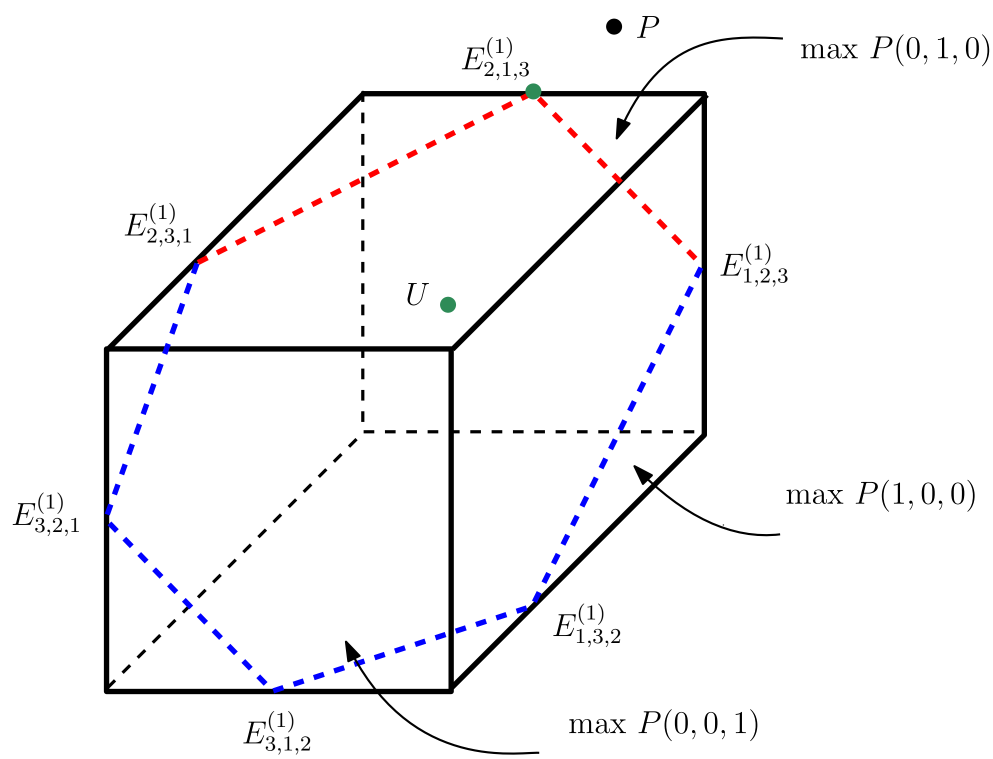


Figure 7: The geometric selection procedure for $B_1(\theta)$.

$$
\begin{aligned}
\delta^*\left(q; A(\theta)\right) = {} & q^\top \Delta(\theta) \\
& + \max(q_2, q_3)\Delta^{(1)}_{S_1}(\theta) + \max(q_2, q_4)\Delta^{(1)}_{S_2}(\theta) + \max(q_3, q_4)\Delta^{(1)}_{S_3}(\theta) + \max(q_2, q_3, q_4)\Delta^{(1)}_{S_4}(\theta) \\
& + \max(q_5, q_6)\Delta^{(2)}_{S_5}(\theta) + \max(q_5, q_7)\Delta^{(2)}_{S_6}(\theta) + \max(q_6, q_7)\Delta^{(2)}_{S_7}(\theta) + \max(q_5, q_6, q_7)\Delta^{(2)}_{S_8}(\theta),
\end{aligned}
$$

where $q = (q_1, \ldots, q_8)^\top$ and

$$
\Delta(\theta) = \left(\Delta^{(0)}_1(\theta), \Delta^{(1)}_1(\theta), \Delta^{(1)}_2(\theta), \Delta^{(1)}_3(\theta), \Delta^{(2)}_1(\theta), \Delta^{(2)}_2(\theta), \Delta^{(2)}_3(\theta), \Delta^{(3)}_1(\theta)\right)^\top.
$$

The identified set can be estimated by testing that the point P_0 belongs to $A(\theta)$ or, equivalently, by testing that

$$
\min_{q \in \mathcal{G}} \delta^*\left(q; A(\theta)\right) - q^\top P_0 \geq 0.
$$

The following sets of directions considered for $\mathcal{G}$ are given below. For "Ineq$_1$", we consider 16 inequalities derived from these 16 directions. Each direction is a column of the following set:

$$
\mathcal{G}_1 = \left\{
\begin{array}{cccccccccccccccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}
\right\}.
$$

Similarly, for "Ineq$_2$" we consider also 16 directions:

$$
\mathcal{G}_2 = \left\{
\begin{array}{cccccccccccccccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}
\right\}.
$$

"Ineq$_3$" takes the whole set of inequalities and equalities which define $B_0(\theta)$ $B_1(\theta)$ and $B_3(\theta)$

(the equality related to $B_2(\theta)$ being redundant is dropped):

$$\mathcal{G}_3 = \left\{ \begin{array}{cccccccccccccccccc} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{array} \right\}.$$

"Ineq$_4$" replaces the directions $e_{ij}^{(K)}$ in "Ineq$_3$ by the directions $-e_l^{(K)}$

$$\mathcal{G}_4 = \left\{ \begin{array}{cccccccccccccccccc} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{array} \right\}.$$

A.17 Additional Monte Carlo simulations

We compare our procedure with the results of the GMS procedure proposed by Andrews and Soares (2010) and its refinement by Romano et al. (2012). We use the DGP of the example with three player considered in table IV, i.e, $\beta = 0.35$ and $\alpha_1 = \alpha_2 = \alpha_3 = -0.4$. Table IX displays the mean rejection rate across simulations for a sequence of points on a curve displayed in Figure 8. The sequence of points contains points from the identified set and outside this identified set. First, for $M = 1000$, all rejection rates are fine with values close to 5% at the boundary. Second, the power of our local procedure is better than the GMS procedures because the critical value is smaller.

M=1000				M=2000			
G	L	AS	RSW	G	L	AS	RSW
0.889	0.923	0.914	0.899	0.998	0.999	0.999	0.999
0.826	0.881	0.865	0.841	0.995	0.997	0.997	0.996
0.756	0.820	0.802	0.773	0.984	0.992	0.990	0.988
0.679	0.750	0.731	0.698	0.964	0.978	0.975	0.971
0.595	0.675	0.656	0.616	0.927	0.952	0.950	0.941
0.512	0.594	0.571	0.532	0.875	0.913	0.907	0.897
0.434	0.516	0.492	0.453	0.797	0.854	0.844	0.826
0.365	0.442	0.419	0.383	0.718	0.775	0.764	0.742
0.304	0.375	0.353	0.315	0.625	0.694	0.681	0.657
0.246	0.316	0.294	0.257	0.522	0.603	0.589	0.559
0.196	0.259	0.240	0.206	0.413	0.494	0.479	0.446
0.155	0.209	0.193	0.163	0.319	0.392	0.378	0.351
0.121	0.170	0.157	0.128	0.241	0.300	0.289	0.265
0.095	0.135	0.125	0.101	0.180	0.229	0.218	0.200
0.077	0.112	0.102	0.082	0.127	0.172	0.161	0.145
0.066	0.095	0.086	0.069	0.095	0.124	0.118	0.107
0.056	0.083	0.076	0.057	0.070	0.094	0.088	0.078
0.048	0.073	0.065	0.049	0.052	0.071	0.067	0.060
0.045	0.065	0.059	0.045	0.042	0.058	0.054	0.048
0.043	0.062	0.054	0.043	0.035	0.048	0.045	0.040
0.040	0.059	0.053	0.040	0.033	0.043	0.040	0.036
0.038	0.057	0.049	0.038	0.031	0.041	0.038	0.035
0.037	0.054	0.048	0.037	0.030	0.039	0.036	0.033
0.036	0.054	0.046	0.036	0.030	0.038	0.035	0.031
0.036	0.053	0.045	0.035	0.030	0.038	0.035	0.031
0.036	0.052	0.044	0.035	0.029	0.038	0.035	0.030
0.035	0.052	0.044	0.035	0.029	0.038	0.035	0.030
0.035	0.052	0.043	0.035	0.029	0.038	0.035	0.029
0.035	0.052	0.043	0.035	0.029	0.038	0.035	0.029
0.035	0.052	0.043	0.035	0.029	0.038	0.035	0.028
0.035	0.052	0.044	0.035	0.029	0.038	0.035	0.029
0.035	0.052	0.045	0.035	0.029	0.038	0.035	0.030
0.035	0.052	0.046	0.035	0.029	0.038	0.035	0.030
0.036	0.054	0.048	0.036	0.030	0.038	0.036	0.032
0.043	0.064	0.058	0.044	0.039	0.052	0.048	0.043
0.050	0.075	0.069	0.051	0.056	0.079	0.071	0.064
0.069	0.103	0.094	0.072	0.103	0.140	0.132	0.118
0.152	0.208	0.189	0.158	0.316	0.390	0.375	0.346
0.104	0.145	0.133	0.108	0.196	0.244	0.236	0.215
0.223	0.289	0.269	0.235	0.468	0.541	0.527	0.504
0.300	0.382	0.355	0.314	0.600	0.680	0.669	0.641
0.394	0.471	0.449	0.410	0.729	0.793	0.781	0.758
0.491	0.569	0.548	0.512	0.836	0.882	0.874	0.856
0.600	0.676	0.656	0.618	0.921	0.947	0.945	0.934
0.708	0.774	0.754	0.722	0.963	0.976	0.974	0.970
0.800	0.855	0.841	0.812	0.988	0.993	0.992	0.990
0.877	0.916	0.905	0.888	0.996	0.997	0.997	0.996
0.933	0.958	0.952	0.940	0.998	1.000	0.999	0.999
0.968	0.979	0.976	0.972	1.000	1.000	1.000	1.000
0.985	0.993	0.991	0.988	1.000	1.000	1.000	1.000
0.996	0.998	0.997	0.997	1.000	1.000	1.000	1.000
0.999	1.000	0.999	0.999	1.000	1.000	1.000	1.000

The rejection rates for the points in the identified set are colored in grey.

Table IX: Rejection frequencies for points tested according to different inference methods. See Figure 8 for a plot of the points in the space (α_1, α_2).

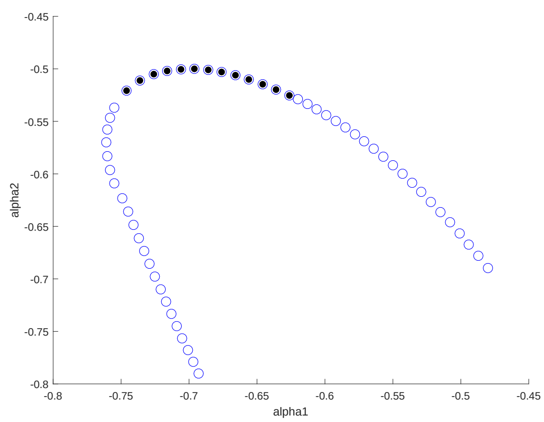


Figure 8: Sequence of points tested (the points of the identified set are colored).