Estimation of Games

Paolo Zacchia

Microeconometrics

Lecture 16

Measuring strategic interactions

- Many economic settings feature **strategic interactions**. It is interesting to *quantify* the extent of these interactions.

- In economics, strategic interactions are modeled as **games**.

- A prime example are models of **competition** between firms: both on the *intensive* margin (e.g. strategic price or quantity setting) and the *extensive* margin (e.g. entry in markets).

- How to estimate the **parameters** that govern these games? In Industrial Organization, this question has been posed and addressed primarily in the setting of **entry games**.

- This lecture reviews some key contributions on entry games, and highlights some **methodological takeaways** that can be generalized to other settings.

An overview of static games

Most of this lecture covers **static** entry games; *dynamic* ones are reviewed less extensively. The ensuing treatment of static games is structured as follows.

1. It sets foot with a classic: the model by Bresnahan and Reiss (1991) to study **entry** and **competition** in local markets.

2. It proceeds with another classic: Berry's (1992) entry model, which uses **simulations** to allow for **firm heterogeneity**.

3. It then addresses the issue of **equilibrium multiplicity** by overviewing the approach by Ciliberto and Tamer (2009), as well as the concept of **partial identification** they leverage.

4. Lastly, it summarizes the **incomplete information** static model of spatial competition by Seim (2006).

A seminal entry model (1/6)

The framework by Bresnahan and Reiss (1991) examines entry decisions in local markets in order to understand the structure of **competition** as a function of the number of equilibrium firms.

This seminal entry model displays features that partially overlap with those typical of demand estimation models (Lecture 14):

- estimation based on "aggregate" **market-level** data where econometricians observe the number N_i of *firms* selling some homogeneous good or service, and other local characteristics (but **no** prices, market shares, costs or margins);

- a latent variable specification that is derived from economic theory, and incorporates both **demand** and **supply**;

- the consequent specification of a multinomial model: in this particular case, an **ordered probit**.

A seminal entry model (2/6)

The **demand** function in market i is given by:

$$Q_i = D\left(P_i; \boldsymbol{y}_i, \boldsymbol{z}_i\right) = d\left(P_i; \boldsymbol{z}_i\right) S\left(\boldsymbol{y}_i\right)$$

where Q_i and P_i are clearly quantity and price, and:

- $d\left(P_i; \boldsymbol{z}_i\right)$ is the demand of a **representative consumer**;
- $S\left(\boldsymbol{y}_i\right)$ is the **number of consumers**;
- $\boldsymbol{y}_i$ and $\boldsymbol{z}_i$ are **demographic variables** that affect demand.

On the **supply** side, firms have:

- **fixed costs** $F\left(\boldsymbol{w}_i\right) + B$;
- **marginal costs** $MC\left(q, \boldsymbol{w}_i\right)$;
- **average variable costs** $AVC\left(q, \boldsymbol{w}_i\right)$;

given a **per-firm quantity** q and local **cost shifters** $\boldsymbol{w}_i$.

A seminal entry model (3/6)

The econometric model makes conclusions about competition by studying the relationship between the **number of firms** N_i and market **"size"** S_i.

To understand the underlying economic intuition, write a firm's **average profits** as:

$$\Pi_i = \left[P_{N_i} - AVC\left(q_{N_i}, \boldsymbol{w}_i\right)\right] d\left(P_{N_i}; \boldsymbol{z}_i\right) \frac{S\left(\boldsymbol{y}_i\right)}{N_i} - F\left(\boldsymbol{w}_i\right) - B_{N_i}$$

where some variables are indexed by N_i to highlight that these are affected by the structure of **competition** or successive entry. Then, define firms' **entry threshold** as:

$$s_{N_i} \equiv \frac{S_{N_i}\left(\boldsymbol{y}_i\right)}{N_i} = \frac{F\left(\boldsymbol{w}_i\right) + B_{N_i}}{\left[P_{N_i} - AVC\left(q_{N_i}, \boldsymbol{w}_i\right)\right] d\left(P_{N_i}; \boldsymbol{z}_i\right)}$$

that is, the market share that firms need to *at least break even* for given N_i ($\Pi_i = 0$). This function is **increasing** in N_i.

A seminal entry model (4/6)

To enable econometric estimation, a **parameterized** expression for profits is necessary. For $\boldsymbol{x}_i = (\boldsymbol{w}_i, \boldsymbol{z}_i)$:

$$\Pi_i = S(\boldsymbol{y}_i; \boldsymbol{\lambda}) V_{N_i}(\boldsymbol{w}_i, \boldsymbol{z}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}) - F_{N_i}(\boldsymbol{w}_i; \boldsymbol{\gamma}, \boldsymbol{\delta}) + \varepsilon_i$$

where $\varepsilon_i \sim \mathcal{N}(0, 1)$ and is independent across observations, and:

- the **market size** is specified as follows:

$$S(\boldsymbol{y}_i; \boldsymbol{\lambda}) = \boldsymbol{y}_i^{\mathrm{T}} \boldsymbol{\lambda}$$

- firms' **variable profits** (for $S = 1$) are specified as follows:

$$V_{N_i}(\boldsymbol{x}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}) = \alpha_1 - \sum_{n=2}^{N_i} \alpha_n + \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\beta}$$

- and firms' **fixed costs** are specified as follows.

$$F_{N_i}(\boldsymbol{w}_i; \boldsymbol{\gamma}, \boldsymbol{\delta}) = \gamma_1 + \sum_{n=2}^{N_i} \gamma_n + \boldsymbol{w}_i^{\mathrm{T}} \boldsymbol{\delta}$$

A seminal entry model (5/6)

- This yields in an estimable **ordered probit** where N_i is the outcome variable and Π_i is the latent variable.

- The ordered probit **thresholds** are the $\boldsymbol{\gamma}$ parameters; there are also "extra" thresholds $\boldsymbol{\alpha}$ that **interact** with $\boldsymbol{y}_i$. These parameters are meant to capture the effect of **competition**.

- To ensure identification, Bresnahan and Reiss normalize one element of $\boldsymbol{\lambda}$ (the parameter for total population) to one.

- The **entry thresholds** can be obtained from the estimates.

$$\widehat{S}_N = \frac{\widehat{\gamma}_1 - \sum_{n=2}^{N_i} \widehat{\gamma}_n + \bar{\mathbf{w}}^{\mathrm{T}} \widehat{\boldsymbol{\delta}}}{\widehat{\alpha}_1 - \sum_{n=2}^{N_i} \widehat{\alpha}_n + \bar{\mathbf{x}}^{\mathrm{T}} \widehat{\boldsymbol{\beta}}}$$

- These are obtained from sample averages $\bar{\mathbf{w}}$ and $\bar{\mathbf{x}}$, and allow to calculate $\widehat{s}_N = \widehat{S}_N / N$ for any integer N.

A seminal entry model (6/6)

- Bresnahan and Reiss estimate their model in several markets for local services: doctors, dentists, druggists (pharmacists), plumbers, tire dealers.

- They focus on several **isolated towns**. Note: heterogeneity between different doctors, dentists etc. is incorporated in ε_i.

- The results are meaningful: the econometric "thresholds" α_n and γ_n are hardly significant for $n = 3, 4, 5, \ldots$

- Suggestively, these markets approach **perfect competition** quite quickly as the number of oligopolists becomes modest.

- This is confirmed by the analysis and statistical tests about the ratios $\widehat{s}_N/\widehat{s}_M$ for $M < N$ (especially $M = 1$).

- This is a striking confirmation of **oligopoly theory**!

Simulating entry (1/6)

An early noteworthy extension of the Bresnahan-Reiss model was developed by Berry (1992). His central contribution was a richer treatment of **firm heterogeneity** via **simulations**.

His starting point was the following expression for the **profits** of firm f in market i:

$$\Pi_{if} = \boldsymbol{w}_{if}^{\mathrm{T}} \boldsymbol{\alpha} + \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\beta} + h\left(N_i; \boldsymbol{\delta}\right) + \rho \varepsilon_{i0} + \sigma \varepsilon_{if}$$

where:

- $\boldsymbol{w}_{if}$ are **firm-level** characteristics, possibly market-specific;
- $\boldsymbol{x}_i$ are **market-level** characteristics;
- $h\left(N_i; \boldsymbol{\delta}\right)$ is a function decreasing in N_i (**number of firms**);
- ε_{i0} and ε_{if} are two error terms (one market-specific and one firm-specific); both follow the **standard normal**.

Simulating entry (2/6)

Suppose that there are F_i firms that could *potentially* operate in market i. Under this representation of profits, a Nash equilibrium treatment of the entry game has N_i endogenously determined as:

$$N_i = \max_{0 \le n \le F_i} \left\{ n : \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\beta} + h\left(n; \boldsymbol{\delta}\right) + \rho \varepsilon_{i0} + \zeta_{in} \ge 0 \right\}$$

where ζ_{in} is the n-th element of the order sequence:

$$\zeta_{i1} > \zeta_{i2} > \ldots > \zeta_{iF_i}$$

with $\zeta_{if} \equiv \boldsymbol{w}_{if}^{\mathrm{T}} \boldsymbol{\alpha} + \sigma \varepsilon_{if}$ being the firm-specific profits component. Note that:

- the equilibrium value of N_i is **unique** if the errors are known;
- but computing the **conditional probability** $\mathbb{P}\left(N_i \middle| \boldsymbol{z}_{if}, \boldsymbol{x}_i \right)$ can be exceedingly difficult, hampering any MLE approaches to estimation of the parameters.

Simulating entry (3/6)

As an example, consider the case where $F_i = 2$, $\boldsymbol{\alpha} = \mathbf{0}$ and $\sigma = 1$. Also, denote the "total attractiveness" of market i by:

$$V_i\left(N_i; \boldsymbol{x}_i\right) \equiv \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\beta} + h\left(N_i; \boldsymbol{\delta}\right) + \rho \varepsilon_{i0}$$

which is also decreasing in N_i. Then:

$$\mathbb{P}\left(\left. N_i = 0 \right| \boldsymbol{x}_i\right) = \int_{-\infty}^{-V_i(1;\boldsymbol{x}_i)} \int_{-\infty}^{-V_i(1;\boldsymbol{x}_i)} \phi\left(\varepsilon_{i1}\right) \phi\left(\varepsilon_{i2}\right) d\varepsilon_{i1} d\varepsilon_{i2}$$

where $\phi\left(\cdot\right)$ is the standard normal's density. A similar expression can be derived for $\mathbb{P}\left(\left. N_i = 2 \right| \boldsymbol{x}_i\right)$. But, for $\boldsymbol{\varepsilon}_i = \left(\varepsilon_{i1}, \varepsilon_{i2}\right)$:

$$\mathbb{P}\left(\left. N_i = 1 \right| \boldsymbol{x}_i\right) = \mathbb{P}\left(\boldsymbol{\varepsilon}_i \in \mathbb{S}_{12}\right) + \mathbb{P}\left(\boldsymbol{\varepsilon}_i \in \mathbb{S}_{21}\right) - \mathbb{P}\left(\boldsymbol{\varepsilon}_i \in \left(\mathbb{S}_{12} \cap \mathbb{S}_{21}\right)\right)$$

where $\mathbb{S}_{k\ell}$ is defined as follows, for $k, \ell = 1, 2$.

$$\mathbb{S}_{k\ell} \equiv \left\{\boldsymbol{\varepsilon}_i : \varepsilon_{ik} \geq -V_i\left(1; \boldsymbol{x}_i\right) \quad \text{and} \quad \varepsilon_{i\ell} < -V_i\left(2; \boldsymbol{x}_i\right)\right\}$$

Simulating entry (4/6)

Some observations are in order.

- This issue stems from the underlying **multiple equilibria** (which firms enter and which don't for every N_i).

- For higher values of F_i, the computational burden obviously grows **exponentially**.

- If $\sigma = 0$ (quite a **restriction**) the only error term is ε_{i0}, and the ordered probit *à la* Bresnahan and Reiss can be adopted for a given ranking of the (observed) firm characteristics ζ_{if}.

- If $h(N_i; \delta) = 0$ and $\rho = 0$, the model would yield a **simple** firm-level **probit** about entering each market i, or not.

Being interested in a more general solution, Bresnahan proposed an approach based on the **Method of Simulated Moments**.

Simulating entry (5/6)

The MSM approach proceeds as follows. For a set of S simulated draws $\{\mathbf{u}_{is}\}_{s=1}^{S}$ from a multivariate standard normal distribution of dimension $F_i + 1$, calculate the unbiased **simulator**:

$$\widehat{N}_i\left(\boldsymbol{w}_{if}, \boldsymbol{x}_i; \boldsymbol{\theta}\right) = \frac{1}{S} \sum_{s=1}^{S} \max_{0 \leq n \leq F_i} \left\{ n : \sum_{k=1}^{F_i} \mathbb{1}\left[\widehat{\Pi}_{iks}\left(n, \mathbf{u}_{is}\right) \geq 0\right] \geq n \right\}$$

where $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \sigma)$, and, given $\mathbf{u}_{is} = (u_{is0}, u_{is1}, \dots, u_{isF_i})$:

$$\widehat{\Pi}_{iks}\left(n, \mathbf{u}_{is}\right) = \boldsymbol{w}_{if}^{\mathrm{T}} \boldsymbol{\alpha} + \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\beta} + h\left(n; \boldsymbol{\delta}\right) + \rho u_{is0} + \sigma u_{isf}$$

for $f = 1, \dots, F_i$. In words, $\widehat{N}_i\left(\boldsymbol{z}_{if}, \boldsymbol{x}_i; \boldsymbol{\theta}\right)$ expresses the average number of successful entrants in the simulation.

Thus, the simulated moments to feed into the MSM approach are obtained as follows, given a market-level **instrument vector** $\boldsymbol{z}_i$.

$$\mathbb{E}\left[N_i - \widehat{N}_i\left(\boldsymbol{w}_{if}, \boldsymbol{x}_i; \boldsymbol{\theta}\right) \middle| \, \boldsymbol{z}_i \right] = 0.$$

Simulating entry (6/6)

- In his application, Berry studies entry in **airline routes**: his definition of a market is a pair of airports (1219 in total).

- He uses an instrument vector $\boldsymbol{z}_i$ of dimension 24, which leads to overidentification, that is allegedly based on "functions of the exogenous data [variables]." However, the discussion on this point is rather nontransparent by modern standards.

- He specifies $h(N_i; \boldsymbol{\delta}) = -\delta \log(N_i)$ and sets $\sigma = \sqrt{1 - \rho^2}$: a **normalization** of the variance of the grand error term.

- The results from a less parametric treatment of $h(N_i; \boldsymbol{\delta})$ are allegedly similar (they are not reported in the paper).

- Berry makes a case for his MSM approach not so much based on the statistical properties of the model, but on the realism of its counterfactual predictions.

Game outcomes versus game equilibria

- In both papers reviewed thus far, the dependent variable is N_i: the number of entrants in a market. This is an **outcome** of the game that can arise from distinct equilibria.

- Can we learn anything from specifying an econometric model for **equilibrium selection**? Potentially, this can reveal the strength of the strategic dependence between players.

- The key problem is **equilibrium multiplicity**, which poses fundamental identification challenges.

- The "compromising" solution suggested and adopted in the literature is the **partial/set identification** approach.

- The next discussion: 1. first, outlines the **problem**; 2. then, summarizes the partial identification **framework**; 3. finally, reviews its **application** by Ciliberto and Tamer (2009).

Multiple equilibria (1/4)

To illustrate the problem, consider a simplified entry model with two potential entrants per market ($F_i = 2$) and per-firm profits:

$$\mathit{\Pi}_{fi} = \boldsymbol{x}_{fi}^{\mathrm{T}} \boldsymbol{\beta}_f + \delta_f Y_{(3-f)i} + \varepsilon_{fi}$$

for $f = 1, 2$. Here, $Y_{if} \in \{0, 1\}$ equals 1 if firm f enters market i, and 0 otherwise, whereas $\boldsymbol{x}_{fi}$ is a a vector of firm characteristics. Note that the $\boldsymbol{\beta}_f$ and δ_f parameters vary by firm.

Firm entry is determined by two simultaneous decisions:

$$Y_{1i} = \mathbb{1}\left[\boldsymbol{x}_{fi}^{\mathrm{T}} \boldsymbol{\beta}_1 + \delta_1 Y_{2i} + \varepsilon_{1i} \geq 0\right]$$
$$Y_{2i} = \mathbb{1}\left[\boldsymbol{x}_{fi}^{\mathrm{T}} \boldsymbol{\beta}_2 + \delta_2 Y_{1i} + \varepsilon_{2i} \geq 0\right]$$

which are obviously **strategically dependent**. Assume $\delta_f \leq 0$ for $f = 1, 2$ because of competition.

Multiple equilibria (2/4)

The **reduced form** representation of the model has the outcome vector (Y_{1i}, Y_{2i}) expressed as a function of **exogenous** observable and unobservable variables.

Adopt pure strategy **Nash equilibrium** as the solution concept. This yields the following reduced form.

$$\left\{\boldsymbol{x}_{1i}^{\mathrm{T}}\boldsymbol{\beta}_1 + \varepsilon_{1i} < 0\right\} \wedge \left\{\boldsymbol{x}_{2i}^{\mathrm{T}}\boldsymbol{\beta}_2 + \varepsilon_{2i} < 0\right\} \Rightarrow \begin{pmatrix} Y_{1i} \\ Y_{2i} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\left\{\boldsymbol{x}_{1i}^{\mathrm{T}}\boldsymbol{\beta}_1 + \delta_1 + \varepsilon_{1i} < 0\right\} \wedge \left\{\boldsymbol{x}_{2i}^{\mathrm{T}}\boldsymbol{\beta}_2 + \varepsilon_{2i} \geq 0\right\} \Rightarrow \begin{pmatrix} Y_{1i} \\ Y_{2i} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\left\{\boldsymbol{x}_{1i}^{\mathrm{T}}\boldsymbol{\beta}_1 + \varepsilon_{1i} \geq 0\right\} \wedge \left\{\boldsymbol{x}_{2i}^{\mathrm{T}}\boldsymbol{\beta}_2 + \delta_2 + \varepsilon_{2i} < 0\right\} \Rightarrow \begin{pmatrix} Y_{1i} \\ Y_{2i} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

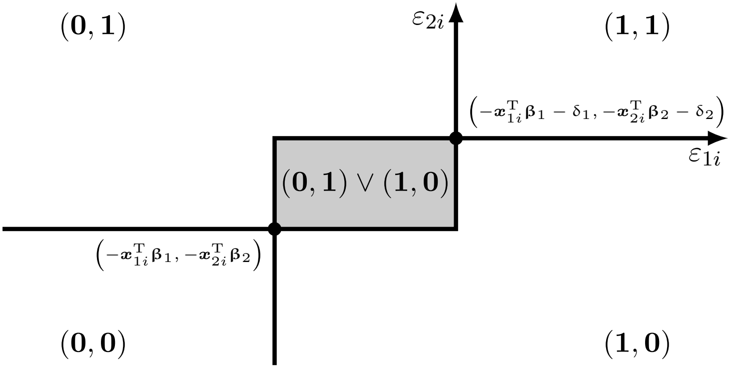
$$\left\{\boldsymbol{x}_{1i}^{\mathrm{T}}\boldsymbol{\beta}_1 + \delta_1 + \varepsilon_{1i} \geq 0\right\} \wedge \left\{\boldsymbol{x}_{2i}^{\mathrm{T}}\boldsymbol{\beta}_2 + \delta_2 + \varepsilon_{2i} \geq 0\right\} \Rightarrow \begin{pmatrix} Y_{1i} \\ Y_{2i} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Multiple equilibria (3/4)

Note that for the combinations of exogenous variables satisfying:

$$\left\{0 \leq \boldsymbol{x}_{1i}^{\mathrm{T}}\boldsymbol{\beta}_1 + \varepsilon_{1i} < -\delta_2\right\} \wedge \left\{0 \leq \boldsymbol{x}_{2i}^{\mathrm{T}}\boldsymbol{\beta}_2 + \varepsilon_{2i} < -\delta_1\right\}$$

there are **two equilibria**, i.e. $(0,1)$ and $(1,0)$. This is represented as the shaded region of the $(\varepsilon_{1i}, \varepsilon_{2i})$ plane below.



Multiple equilibria (4/4)

- This posits an obvious **identification problem**: the model *cannot predict the endogenous outcomes* for certain values of the error terms, conditional on the exogenous variables.

- This prevents the construction of unambiguous conditional probabilities **for each equilibrium** that would be functions of the parameters, given the observed data.

- Hence, MLE (but also GMM, SML, MSM) approaches are **inapplicable**.

- The problem becomes obviously **worse** with more players F_i. Similar issues occur if $\delta_f \geq 0$ (due to complementarities).

- If one is willing to work with the notion of **identified set**: **all** parameters that are **equally capable** to best explain the outcomes given some criterion, a partial solution is available.

Set identification in a nutshell (1/6)

It is helpful to develop a summary of set identification and of its associated estimation framework. This discussion draws liberally from Chernozhukov, Hong and Tamer (2007; CHT).

Recall some ideas on M-Estimation (Lecture 11): given some i.i.d. random vectors $\boldsymbol{x}_i$ and a criterion function $\mathcal{Q}_0(\theta) = \mathbb{E}[q(\boldsymbol{x}_i; \theta)]$, the true parameter vector θ_0 is point identified if $-\mathcal{Q}_0(\theta)$ has a unique minimizer. This criterion is typically informed by a set of **moment equalities** of the kind $\mathbb{E}[\boldsymbol{g}(\boldsymbol{x}_i; \theta_0)] = \mathbf{0}$ with a unique solution θ_0 (even in say the MLE case).

But what if assumptions lead to **moment inequalities**, like:

$$\mathbb{E}[\boldsymbol{g}(\boldsymbol{x}_i; \theta_0)] \leq \mathbf{0}$$

or even with reverse sign? The set $\boldsymbol{\Theta}_I = \{\theta_0 : \mathbb{E}[\boldsymbol{g}(\boldsymbol{x}_i; \theta_0)] \leq \mathbf{0}\}$ is very likely **multivalued**, i.e. $|\boldsymbol{\Theta}_0| \geq 1$.

Set identification in a nutshell (2/6)

Here are two examples of moment inequalities (from CHT).

1. **Missing data**: suppose one aims to estimate $\theta = \mathbb{E}[Y_i]$ but the variable Y_i is *missing*: one only observes some "brackets" $\underline{Y}_i$ and $\overline{Y}_i$ such that $Y_i \in [\underline{Y}_i, \overline{Y}_i]$. Then, $\mathbb{E}[\underline{Y}_i] \leq \theta \leq \mathbb{E}[\overline{Y}_i]$.

$$\boldsymbol{g}\left(\underline{Y}_i, \overline{Y}_i; \theta\right) = \begin{pmatrix} \underline{Y}_i - \theta \\ \theta - \overline{Y}_i \end{pmatrix}$$

2. **Regression for missing outcomes**: now let Y_i be missing as above, and $\mathbb{E}[Y_i | \boldsymbol{x}_i] = \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\beta}$. Hence, for a given $\boldsymbol{z}_i \in \mathbb{R}^K$:

$$\mathbb{E}\left[\boldsymbol{z}_i \underline{Y}_i\right] \leq \mathbb{E}\left[\boldsymbol{z}_i \boldsymbol{x}_i^{\mathrm{T}}\right] \boldsymbol{\beta} \leq \mathbb{E}\left[\boldsymbol{z}_i \overline{Y}_i\right]$$

with moment functions adapted to "instruments" $\boldsymbol{z}_i$.

$$\boldsymbol{g}\left(\underline{Y}_i, \overline{Y}_i, \boldsymbol{x}_i, \boldsymbol{z}_i; \boldsymbol{\beta}\right) = \begin{pmatrix} +\boldsymbol{z}_i \left(\underline{Y}_i - \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\beta}\right) \\ -\boldsymbol{z}_i \left(\overline{Y}_i - \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\beta}\right) \end{pmatrix}$$

Set identification in a nutshell (3/6)

Let $\|x\|_+ = \|\max(0, x)\|$ and $\|x\|_- = \|\min(0, x)\|$. Also let:

$$\mathcal{Q}_0(\boldsymbol{\theta}) \equiv \left\| \left(\mathbb{E}\left[\boldsymbol{g}(\boldsymbol{x}_i; \boldsymbol{\theta}) \right] \right)^{\mathrm{T}} \mathbf{A}_0^{\frac{1}{2}} \right\|_+^2$$

where $\mathbf{A}_0$ is a diagonal matrix with a strictly positive diagonal. Then, the **identified set** $\boldsymbol{\Theta}_I$ in this framework is appropriately defined as the (likely multivalued) set of solutions of $\mathcal{Q}_0(\boldsymbol{\theta})$.

$$\boldsymbol{\Theta}_I = \{ \boldsymbol{\theta}_0 : \mathcal{Q}_0(\boldsymbol{\theta}_0) = 0 \}$$

Let $\widehat{\mathcal{Q}}_N(\boldsymbol{\theta})$ be the **sample analog** of $\mathcal{Q}_0(\boldsymbol{\theta})$, for some $\mathbf{A}_N \xrightarrow{p} \mathbf{A}_0$.

$$\widehat{\mathcal{Q}}_N(\boldsymbol{\theta}) \equiv \left\| \left(\frac{1}{N} \sum_{i=1}^{N} \boldsymbol{g}(\boldsymbol{x}_i; \boldsymbol{\theta}) \right)^{\mathrm{T}} \mathbf{A}_N^{\frac{1}{2}} \right\|_+^2$$

The objective is to leverage function $\widehat{\mathcal{Q}}_N(\boldsymbol{\theta})$ to construct suitable **set estimates** $\widehat{\boldsymbol{\Theta}}_{SE}$ of the identified set $\boldsymbol{\Theta}_I$.

Set identification in a nutshell (4/6)

CHT define the **set estimator** as the set $\widehat{\boldsymbol{\Theta}}_{SE}(c_N)$ which, note, is a function of some sequence c_N, as follows:

$$\widehat{\boldsymbol{\Theta}}_{SE}(c_N) = \left\{\boldsymbol{\theta} : \widehat{\mathcal{Q}}_N(\boldsymbol{\theta}) \leq c_N\right\}$$

that is, the set of $\boldsymbol{\theta}$ values that solve the **inequality** $\widehat{\mathcal{Q}}_N(\boldsymbol{\theta}) \leq c_N$ for a given value of N.

- Given the definition of the identified set $\boldsymbol{\Theta}_I$, this is a fairly intuitive application of the analogy principle.
- Buy why *not*, even more intuitively, $c_N = 0$ uniformly? CHT allow for this, e.g. in the two previous examples.
- However, for reasons elaborated soon, a less restrictive c_N is more appropriate under the asymptotics developed in CHT.

Set identification in a nutshell (5/6)

Consider the mathematical notion of *Hausdorff distance* between two sets $\mathbb{X}$ and $\mathbb{Y}$, denoted as $d_H(\mathbb{X}, \mathbb{Y})$:

$$d_H(\mathbb{X}, \mathbb{Y}) = \max \left\{ \sup_{x \in \mathbb{X}} d(x, \mathbb{Y}), \sup_{y \in \mathbb{Y}} d(\mathbb{X}, y) \right\}$$

where $d(\cdot, \cdot)$ is a more "conventional" (e.g. Euclidean) distance. Intuitively, $d_H(\mathbb{X}, \mathbb{Y})$ is the largest distance between $\mathbb{X}$ and $\mathbb{Y}$.

CHT provide conditions under which $\widehat{\boldsymbol{\Theta}}_{SE}(c_N)$ is **consistent** in the following sense.

$$d_H\left(\widehat{\boldsymbol{\Theta}}_{SE}(c_N), \boldsymbol{\Theta}_I\right) \xrightarrow{p} 0$$

- The most important condition is that $\boldsymbol{\Theta}_I$ is **compact**.
- CHT show that if $c_N = (\log N)/N$, the speed of convergence is close to the conventional $\sqrt{N}$ rate.

Set identification in a nutshell (6/6)

How to conduct **inference** on $\widehat{\boldsymbol{\Theta}}_{SE}(c_N)$? For point estimators, inference usually leads to confidence intervals/sets, but $\widehat{\boldsymbol{\Theta}}_{SE}(c_N)$ is already a set.

One possibility is to construct $\widehat{\boldsymbol{\Theta}}_{SE}(c_N)$ *directly* as a **confidence region** that covers $\boldsymbol{\Theta}_I$ with probability equal to some confidence level $\alpha \in (0,1)$, by choosing c_N appropriately.

$$\mathbb{P}\left(\boldsymbol{\Theta}_I \subseteq \widehat{\boldsymbol{\Theta}}_{SE}(c_N)\right) = \alpha$$

How to do this in practice?

- CHT propose a **subsampling** algorithm akin to a bootstrap: after initializing $c_0 = \inf_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \widehat{\mathcal{Q}}_N(\boldsymbol{\theta})$, the proper c_N obtains as the α-quantile of $\sup_{\widehat{\boldsymbol{\Theta}}_{SE}(c_0)} \widehat{\mathcal{Q}}_N(\boldsymbol{\theta})$ over the subsamples.
- The algorithm is predicated on conditions specified by CHT. In some particular cases it simplifies to a modified bootstrap.

Set identification for entry games (1/7)

Ciliberto and Tamer (2009) provide a famous application of set identification to the problem of multiple equilibria in entry games.

The key idea is to acknowledge that each parameter set does not identify unique probabilities for each game outcome to occur, but these probabilities can be **bounded** and treated as inequalities.

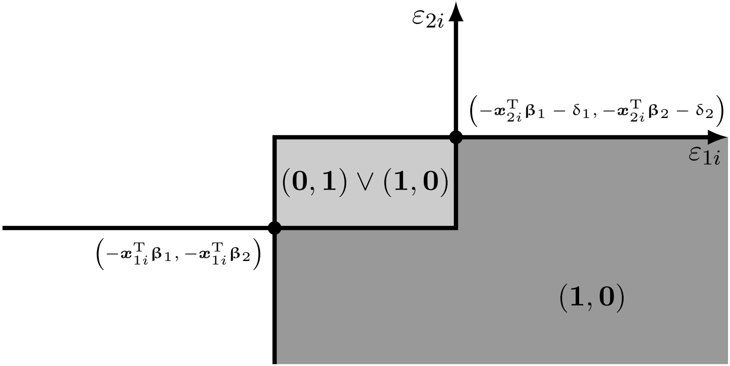
Consider for example the outcome $(Y_{1i}, Y_{2i}) = (1, 0)$ in the entry game with two potential entrants. Given values of $\boldsymbol{\theta}_f = (\boldsymbol{\beta}_f, \delta_f)$ for $f = 1, 2$ and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$, the **conditional** probability that this outcome occurs is bounded as follows:

$$\begin{aligned}\mathbb{P}\left((\varepsilon_{1i}, \varepsilon_{2i}) \in \mathbb{S}_1 \middle| \boldsymbol{x}_{1i}, \boldsymbol{x}_{2i}; \boldsymbol{\theta}\right) &\leq \mathbb{P}\left((1,0) \middle| \boldsymbol{x}_{1i}, \boldsymbol{x}_{2i}; \boldsymbol{\theta}\right) \\ &\leq \mathbb{P}\left((\varepsilon_{1i}, \varepsilon_{2i}) \in \mathbb{S}_1 \middle| \boldsymbol{x}_{1i}, \boldsymbol{x}_{2i}; \boldsymbol{\theta}\right) \\ &\quad + \mathbb{P}\left((\varepsilon_{1i}, \varepsilon_{2i}) \in \mathbb{S}_2 \middle| \boldsymbol{x}_{1i}, \boldsymbol{x}_{2i}; \boldsymbol{\theta}\right)\end{aligned}$$

where $\mathbb{S}_1$ and $\mathbb{S}_2$ are two subsets of $\mathbb{R}^2$.

Set identification for entry games (2/7)

In this example, $\mathbb{S}_1$ is the **dark gray** area in the figure, while $\mathbb{S}_2$ is the **light gray** area.



A symmetric analysis applies to the outcome $(Y_{1i}, Y_{2i}) = (0, 1)$. The logic can be extended to fairly low values of $F_i > 2$. However, for moderate-to-high values of F_i, enumerating the equilibria and finding the multiplicity regions becomes too cumbersome.

Set identification for entry games (3/7)

Write $\boldsymbol{x}_i = (\boldsymbol{x}_{1i}, \dots, \boldsymbol{x}_{Fi})$ as the collection of all the firm variables in market i and $\boldsymbol{\varepsilon}_i = (\varepsilon_{1i}, \dots, \varepsilon_{Fi})$ as the collection of all the error terms. Similarly, $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_F)$ collects all the parameters.

For a generic F (assumed equal across all markets i) the **moment inequalities** that are associated with the 2^F possible equilibria of the game of the kind $\boldsymbol{y}_i \in \{0,1\}^F$ can be expressed as follows.

$$\underbrace{\begin{pmatrix} \underline{b}_1(\boldsymbol{x}_i; \boldsymbol{\theta}) \\ \vdots \\ \underline{b}_{2^F}(\boldsymbol{x}_i; \boldsymbol{\theta}) \end{pmatrix}}_{\equiv \underline{\mathbf{b}}(\boldsymbol{x}_i; \boldsymbol{\theta})} \leq \underbrace{\begin{pmatrix} \mathbb{P}(\boldsymbol{y}_{1i} | \boldsymbol{x}_i; \boldsymbol{\theta}) \\ \vdots \\ \mathbb{P}(\boldsymbol{y}_{2^F i} | \boldsymbol{x}_i; \boldsymbol{\theta}) \end{pmatrix}}_{\equiv \mathbf{p}(\boldsymbol{x}_i; \boldsymbol{\theta})} \leq \underbrace{\begin{pmatrix} \overline{b}_1(\boldsymbol{x}_i; \boldsymbol{\theta}) \\ \vdots \\ \overline{b}_{2^F}(\boldsymbol{x}_i; \boldsymbol{\theta}) \end{pmatrix}}_{\equiv \overline{\mathbf{b}}(\boldsymbol{x}_i; \boldsymbol{\theta})}$$

The functions $\underline{\mathbf{b}}(\boldsymbol{x}_i; \boldsymbol{\theta})$ and $\overline{\mathbf{b}}(\boldsymbol{x}_i; \boldsymbol{\theta})$ characterize the **bounds** on the "true" conditional probabilities $\mathbf{p}(\boldsymbol{x}_i; \boldsymbol{\theta})$ of the 2^F equilibria. In general, these bounds are difficult to derive analytically.

Set identification for entry games (4/7)

Let $\boldsymbol{\varepsilon}_i$ be independent of $\boldsymbol{x}_i$ and drawn from a distribution $H(\boldsymbol{\varepsilon}_i)$ whose parameters, if there are any, are appended to $\boldsymbol{\theta}$. Let $G(\mathbf{x}_i)$ be the distribution that generates $\boldsymbol{x}_i$, which has support $\mathbb{X}$. Then:

$$\mathcal{Q}_0(\boldsymbol{\theta}) = \int_{\mathbb{X}} \Big[\left\| \mathbf{p}(\mathbf{x}_i; \boldsymbol{\theta}) - \underline{\mathbf{b}}(\mathbf{x}_i; \boldsymbol{\theta}) \right\|_- + \\ + \left\| \mathbf{p}(\mathbf{x}_i; \boldsymbol{\theta}) - \overline{\mathbf{b}}(\mathbf{x}_i; \boldsymbol{\theta}) \right\|_+ \Big] dG(\mathbf{x}_i)$$

gives the objective function associated with this set identification problem. Here, $\|\cdot\|_+$ and $\|\cdot\|_-$ are taken pointwise over vectors.

- The objective is to estimate $\boldsymbol{\Theta}_I = \{\boldsymbol{\theta}_0 : \mathcal{Q}_0(\boldsymbol{\theta}_0) = 0\}$. Thus, Ciliberto and Tamer resort to the methodology by CHT.

- Ciliberto and Tamer show that identification is improved if $\boldsymbol{x}_i$ includes variables that are unique to each firm ("exclusion restrictions") as in classical SEMs (Lecture 9).

Set identification for entry games (5/7)

To be implemented in practice, this approach requires knowledge of $\underline{\mathbf{b}}(\boldsymbol{x}_i;\theta)$, $\overline{\mathbf{b}}(\boldsymbol{x}_i;\theta)$ as well as $\mathbf{p}(\boldsymbol{x}_i;\theta)$. To overcome this issue, Ciliberto and Tamer propose the following **simulation** approach.

- Draw S vectors $\mathbf{u}_s = \{u_{1is},\dots,u_{Fis}\}_{i=1}^{M}$ from $H(\boldsymbol{\varepsilon}_i)$, where M is the number of markets, for $s = 1,\dots,S$.

- For **all** potential equilibria $\mathbf{y}_{\ell i}$ with $\ell = 1,\dots,2^F$, calculate **profits** $\pi_{\ell fis}(\mathbf{y}_{\ell i},\mathbf{x}_i,\mathbf{u}_{is};\theta)$ for every firm $f = 1,\dots,F$ and market $i = 1,\dots,M$, given a draw $\mathbf{u}_{is} = (u_{1is},\dots,u_{Fis})$.

- For every s, calculate the vector $\mathbf{i}_{is}(\mathbf{x}_i,\mathbf{u}_{is};\theta)$ that expresses whether any outcome $\mathbf{y}_{\ell i}$ in market i is an equilibrium.

$$\mathbf{i}_{is}(\mathbf{x}_i,\mathbf{u}_{is};\theta) = \begin{pmatrix} \prod_{f=1}^{F} \mathbb{1}\left[\pi_{1fis}(\mathbf{y}_{1i},\mathbf{x}_i,\mathbf{u}_{is};\theta) \geq 0\right] \\ \vdots \\ \prod_{f=1}^{F} \mathbb{1}\left[\pi_{2^F fis}(\mathbf{y}_{2^F i},\mathbf{x}_i,\mathbf{u}_{is};\theta) \geq 0\right] \end{pmatrix}$$

Set identification for entry games (6/7)

- Additionally, calculate the vector $\mathbf{i}_{is}^{m}(\mathbf{x}_i, \mathbf{u}_{is}; \boldsymbol{\theta})$ that denotes whether any outcome $\mathbf{y}_{\ell i}$ is a **unique** equilibrium.

$$\mathbf{i}_{is}^{*}(\mathbf{x}_i, \mathbf{u}_{is}; \boldsymbol{\theta}) = \mathbf{i}_{is}(\mathbf{x}_i, \mathbf{u}_{is}; \boldsymbol{\theta}) \cdot \mathbb{1}\left[\boldsymbol{\iota}^{\mathrm{T}} \mathbf{i}_{is}(\mathbf{x}_i, \mathbf{u}_{is}; \boldsymbol{\theta}) = 1\right]$$

- Both vectors are **binary**: their entries are either 0 or 1.

- The bounds are then *simulated* as follows, given $\mathbf{x}_i$ and $\boldsymbol{\theta}$.

$$\widehat{\underline{\mathbf{b}}}(\mathbf{x}_i; \boldsymbol{\theta}) = \frac{1}{S} \sum_{s=1}^{S} \mathbf{i}_{is}^{*}(\mathbf{x}_i, \mathbf{u}_{is}; \boldsymbol{\theta})$$

$$\widehat{\overline{\mathbf{b}}}(\mathbf{x}_i; \boldsymbol{\theta}) = \frac{1}{S} \sum_{s=1}^{S} \mathbf{i}_{is}(\mathbf{x}_i, \mathbf{u}_{is}; \boldsymbol{\theta})$$

- These simulators are **consistent** if S grows alongside M.

- More simply, $\widehat{\mathbf{p}}(\boldsymbol{x}_i; \boldsymbol{\theta})$ can be estimated nonparametrically.

Set identification for entry games (7/7)

- Using their framework, Ciliberto and Tamer revisit Berry's airline entry problem. They study $M = 2472$ airport pairs.

- They focus on $F = 4$ firms: American Airlines, Delta, United Airlines, and Southwest. Others are treated as *nonstrategic*.

- The $\boldsymbol{x}_i$ controls vary at the market, airline, or market-airline level. If continuous, they are *discretized* to obtain $\widehat{\mathbf{p}}(\boldsymbol{x}_i; \theta)$.

- Their baseline estimates of δ_f are similar across firms. More nuanced specifications where $\delta_{ff'}$ varies at the *firm pair* level suggests that Southwest is more susceptible to competition.

- They also study specifications of profits where $\boldsymbol{y}_i$ and key $\boldsymbol{x}_i$ variables *interact*. The resulting estimates suggest that the higher a firm's presence at a given airport, the stronger the negative competitive effect on other firms for that airport.

A game of incomplete information (1/7)

- The last paper covered in this Lecture's review (Seim, 2006), is notable in a number of ways.

- First and foremost, it introduces **incomplete information** in the entry game. This makes the game more realistic, but also somewhat counterintuitively *easier to estimate*.

- This paper introduces a **spatial** dimension of competition, which is applied to a very local (though nowadays obsolete) "urban" market: that of videotape rental stores.

- In the model, not only is entry endogenous, but the **location** of a store within a city also is. This introduces **endogenous product differentiation** into an entry model.

- These elements are held together via an elegant, fairly simple estimation framework.

A game of incomplete information (2/7)

Seim assumes a profit function for firms that also varies along L_i possible **locations** within market i, indexed as $\ell = 1, \dots, L_i$.

$$\Pi_{i\ell f} = \boldsymbol{x}_{i\ell}^{\mathrm{T}} \boldsymbol{\beta} + \sum_{b=0}^{B} \gamma_b N_{i\ell b} + \xi_i + \varepsilon_{\ell f}$$

In this expression:

- the regressors $\boldsymbol{x}_{i\ell}$ can vary at the market-location level, with $\boldsymbol{X}_i = (\boldsymbol{x}_{i1}, \dots, \boldsymbol{x}_{iL_i})$ collecting all covariates of market i;

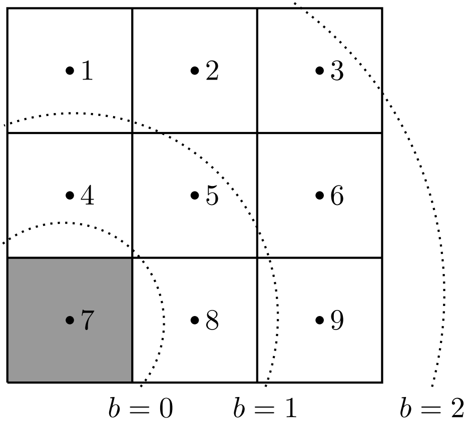
- the error term is split between a **market-level** component ξ_i and a shock $\varepsilon_{\ell f}$ that is **location-firm** specific;

- spatial competition is introduced via a linear expression that depends on $N_{i\ell b}$, the total **number** of firms that are located at **distance band** b relative to location ℓ in market i;

- distance bands are **discrete** and indexed as $b = 0, 1, \dots, B$.

A game of incomplete information (3/7)

An example from Seim herself helps clarify the setup. Here, the market i features $L_i = 9$ locations arranged in a square; if a firm locates in cell 7, its immediate competitors lie in cells 4, 5 and 8, while the farther competitors are in cells 1, 2, 3, 6 and 9.



A game of incomplete information (4/7)

Crucially, Seim assumes **incomplete information** with regard to the shocks $\varepsilon_{\ell f}$ of potential competitors. Yet firms form **beliefs** about the probability of competitors' choices. Write:

$$p_{i\ell g} \equiv \Pr\left(\ell \in \arg \max_{l=1,\dots,L_i} \mathbb{E}\left[\mathit{\Pi}_{ilg}\right] \middle| \boldsymbol{X}_i, \xi_i, E_i \right)$$

as the probability that firm g (a potential competitor of f) settles in location $\ell = 1, \dots, L_i$. This expression is **conditional** on the **number of entrants** E_i in market i, which for the moment shall be taken as given (to be later endogenized).

Obviously, $\mathbb{E}\left[\mathit{\Pi}_{ilg}\right]$, that is, the expected profits of firm g in every location l, depend in turn on the choices of all other firms.

This problem is recursive, but thanks to incomplete information, each firm f treats the entry probabilities of other firms g in every location as **symmetric**.

A game of incomplete information (5/7)

Let $\mathbf{p}_{ig} = (p_{i1g}, \dots, p_{iL_ig})$. By a **symmetry conjecture** (which is verified in equilibrium) it is $\mathbf{p}_{ig} = \mathbf{p}_{if} = \mathbf{p}^*$ for any two firms f and g. Hence, the expected profits of firm f in location ℓ are:

$$\underbrace{\mathbb{E}\left[\Pi_{i\ell f}\right] = \boldsymbol{x}_{i\ell}^{\mathrm{T}}\boldsymbol{\beta} + \gamma_0 + (E_i - 1)\sum_{b=1}^{B}\gamma_b \sum_{l \neq \ell} p_{il}^* \mathbb{1}\left[\mathcal{B}(\ell, l) = b\right] + \xi_i}_{\equiv \Pi_{i\ell}^*\left(\boldsymbol{x}_{i\ell}, \mathbf{p}^*; \boldsymbol{\theta}\right)}$$

where $\mathcal{B}(\ell, l)$ identifies the "distance band" between ℓ and l, and $\boldsymbol{\theta} = (\boldsymbol{\beta}; \gamma_0, \gamma_1, \dots, \gamma_B)$ collects all the model's parameters. Here, γ_0 captures the fact that firm f locates in ℓ (e.g. "cell 7").

If the shocks $\varepsilon_{f\ell}$ are i.i.d. standard Gumbel, all the L_i elements of the $\mathbf{p}^*$ vector have the familiar multinomial logit form.

$$p_{i\ell}^* = \frac{\exp\left(\Pi_{i\ell}^*\left(\boldsymbol{x}_{i\ell}, \mathbf{p}^*; \boldsymbol{\theta}\right)\right)}{\sum_{m=1}^{L_i} \exp\left(\Pi_{im}^*\left(\boldsymbol{x}_{im}, \mathbf{p}^*; \boldsymbol{\theta}\right)\right)}$$

A game of incomplete information (6/7)

It remains to endogenize entry. One can derive $\mathcal{P}_i = \mathcal{P}\left(\boldsymbol{X}_i, \xi_i; \boldsymbol{\theta}\right)$, that is, the conditional probability of entering market i, as a logit probability based on the normalized expected *maximum* payoff.

$$\mathcal{P}_i = \frac{\exp\left(\xi_i\right) \sum_{\ell=1}^{L_i} \exp\left(\mathit{\Pi}_{i\ell}^{*}\left(\boldsymbol{x}_{i\ell}, \mathbf{p}^{*}; \boldsymbol{\theta}\right)\right)}{1 + \exp\left(\xi_i\right) \sum_{\ell=1}^{L_i} \exp\left(\mathit{\Pi}_{i\ell}^{*}\left(\boldsymbol{x}_{i\ell}, \mathbf{p}^{*}; \boldsymbol{\theta}\right)\right)}$$

Hence, $E_i = F_i \mathcal{P}_i$; where F_i is, as in previous use of notation in this lecture, the total number of *potential* entrants.

By an inversion argument similar to the one by Berry (1994) for demand estimation (Lecture 14), one can solve for ξ_i.

$$\xi_i = \log\left(E_i\right) - \log\left(F_i - E_i\right) - \log\left(\sum_{\ell=1}^{L_i} \exp\left(\mathit{\Pi}_{i\ell}^{*}\left(\boldsymbol{x}_{i\ell}, \mathbf{p}^{*}; \boldsymbol{\theta}\right)\right)\right)$$

Seim leverages this expression in estimation, replacing E_i by the actually observed number of entrants in each market i: e_i.

A game of incomplete information (7/7)

To close the model, let $\xi_i \sim \mathcal{N}(\mu, \sigma^2)$. This allows to specify the following likelihood function computed over all M markets:

$$\mathcal{L}\left(\theta, \mu, \sigma^2\right) = \prod_{i=1}^{M} f_{\boldsymbol{d}_i}\left(\mathbf{d}_i \middle| \mathbf{X}_i, e_i, \xi_i; \theta\right) \frac{1}{\sigma} \phi\left(\frac{\xi_i - \mu}{\sigma} \middle| \mathbf{X}_i, e_i; \mu, \sigma^2\right)$$

where $\phi(\cdot)$ is the standard normal p.d.f., $\boldsymbol{d}_i$ is the random vector of length F_i (with realization $\mathbf{d}_i$) that denotes the choices of all firms about market i (entry – and in what location – or not) and $f_{\boldsymbol{d}_i}(\cdot|\cdot)$ is the conditional density of $\boldsymbol{d}_i$ induced by the model.

- Maximization of this likelihood function nests a fixed-point algorithm to solve for $\mathbf{p}^*$ at every iteration.
- Seim estimates the model on 151 cities (markets), splitting all of them into proper locations or "cells." The results show decaying γ_b coefficients and are robust to the choice of F_i.

A brief introduction to dynamic games

- Econometric models for *static* games reviewed thus far come with a major limitation: the observed outcomes are taken at their face value, ignoring the *history* that led to them.

- Relevant economic choices are strategic *and* forward-looking, **dynamic**: think of firms that invest, or enter markets.

- This led to developing the econometrics of **dynamic games**, bridging the Bresnahan-Reiss-Berry tradition with dynamic frameworks *à la* Rust (1987) and Hotz and Miller (1993).

- There is no way to make justice of this large literature within this self-contained, very introductory discussion.

- Therefore, this introduction must inevitably be selective; see Aguirregabiria, Collard-Wexler and Ryan (2021) for a recent, comprehensive survey.

Markov Perfect Nash Equilibrium (1/7)

Econometric models for dynamic games rely on the game solution concept called **Markov Perfect Nash Equilibrium** (in short, MPNE) by Ericson and Pakes (1994), and a general setup.

- A game is played over an **infinite horizon**.
- **Time** is discrete as indexed as $t = 1, \ldots, \infty;$.
- There are N **players** indexed as $i = 1, \ldots, N$.
- In each period t player i chooses an **action** $A_i \in \mathbb{A}_i$. Each set $\mathbb{A}_i$ is **discrete** with dimension $\mathcal{A}_i = |\mathbb{A}_i|$. Let $\mathbb{A} \equiv \times_{i=1}^{N} \mathbb{A}_i$.
- At time t, a player i is also subject to a **state** (variable) $\boldsymbol{x}_{it}$ with **discrete** support $\mathbb{X}_i$. This state is *publicly observable.*
- A player receives *privately observed* **choice-specific shocks** $\boldsymbol{\varepsilon}_{it} = (\varepsilon_{1it}, \ldots, \varepsilon_{\mathcal{A}_i it}) \in \mathbb{R}^{\mathcal{A}_i}$, independent across players i.

Markov Perfect Nash Equilibrium (2/7)

- The **period-specific** payoff for player i (at time t) is:

  $$\mathit{\Pi}_{it}\left(\boldsymbol{a}_t, \boldsymbol{x}_t, \boldsymbol{\varepsilon}_{it}; \theta_1\right) = \pi\left(\boldsymbol{a}_t, \boldsymbol{x}_t; \theta_1\right) + \sum_{a \in \mathbb{A}_i} \varepsilon_{ait} \cdot \mathbb{1}\left[A_{it} = a\right]$$

  for $\boldsymbol{a}_t = (A_{1t}, \ldots, A_{Nt})$ as played at t, $\boldsymbol{x}_t = (\boldsymbol{x}_{1t}, \ldots, \boldsymbol{x}_{Nt})$.

- Given a discount factor $\beta \in [0, 1]$, players aim at maximizing the following expected present value of future payoffs.

  $$\sum_{\tau=t}^{\infty} \beta^{\tau-t} \mathbb{E}\left[\left. \mathit{\Pi}_{i\tau}\left(\boldsymbol{a}_\tau, \boldsymbol{x}_\tau, \boldsymbol{\varepsilon}_{i\tau}; \theta_1\right) \right| \boldsymbol{x}_t, \boldsymbol{\varepsilon}_{it}\right]$$

- Implicit in this model is a specification of the probabilities to transition across states (like Rust, 1987): $q\left(\left. \boldsymbol{x}_{t+1} \right| \boldsymbol{x}_t, \boldsymbol{a}_t; \theta_2\right)$.

- The parameters of this model are collected as $\theta = (\theta_1, \theta_2)$.

- A *Bayesian perfect equilibrium* is too complex as a solution concept for a high-dimensional game like this one.

Markov Perfect Nash Equilibrium (3/7)

- The MPNE is a **simpler** solution concept that in each period makes strategies dependent only upon the *current* state.
- Let players hold **beliefs** $\sigma_i(\boldsymbol{a}_t|\boldsymbol{x}_t)$: a function representing the *subjective probability* that $\boldsymbol{a}_t$ is realized given state $\boldsymbol{x}_t$.
- For $\mathbb{X} \equiv \times_{i=1}^{N} \mathbb{X}_i$, let $\boldsymbol{\sigma}_i$ be the collection of player i's beliefs.

$$\boldsymbol{\sigma}_i \equiv \left\{ \left\{ \sigma_i(\boldsymbol{a}_t|\boldsymbol{x}_t) \right\}_{\boldsymbol{a}_t \in \mathbb{A}} \right\}_{\boldsymbol{x}_t \in \mathbb{X}}$$

- Define the **ex-ante value function** as:

$$V(\boldsymbol{x}_t, \boldsymbol{\sigma}_i; \theta) \equiv \sum_{\boldsymbol{a}_t \in \mathbb{A}} \sigma_i(\boldsymbol{a}_t|\boldsymbol{x}_t) \Bigg[\pi(\boldsymbol{a}_t, \boldsymbol{x}_t; \theta_1) + \\ + \mathbb{E}[\varepsilon_{it}|\boldsymbol{x}_t, \boldsymbol{a}_t] + \beta \sum_{\boldsymbol{\xi} \in \mathbb{X}} q(\boldsymbol{\xi}|\boldsymbol{x}_t, \boldsymbol{a}_t; \theta_2) V(\boldsymbol{\xi}, \boldsymbol{\sigma}_i; \theta) \Bigg]$$

where $\mathbb{E}[\varepsilon_{it}|\boldsymbol{x}_t, \boldsymbol{a}_t] \equiv \sum_{a \in \mathbb{A}_i} \mathbb{E}[\varepsilon_{ait}|\boldsymbol{x}_t] \cdot \mathbb{1}[A_{it} = a]$.

Markov Perfect Nash Equilibrium (4/7)

- A collection $\{(A_{it}^*; \boldsymbol{\sigma}_i^*)\}_{i=1}^N$ is a MPNE if for all the players, their own action is a *best response* given the actions of the others, and beliefs are *consistent in equilibrium.*

- Define here the **choice-specific mean value function** as:

$$\mathcal{U}\left(A_i, \boldsymbol{x}, \boldsymbol{\sigma}_i; \theta\right) \equiv \sum_{\boldsymbol{a}_{-i} \in \mathbb{A}_{-i}} \sigma_i\left(\boldsymbol{a}_{-i} \middle| \boldsymbol{x}\right) \Bigg[\pi\left(A_i, \boldsymbol{a}_{-i}, \boldsymbol{x}; \theta_1\right) + \\ + \beta \sum_{\boldsymbol{\xi} \in \mathbb{X}} q\left(\boldsymbol{\xi} \middle| \boldsymbol{x}, \boldsymbol{a}; \theta_2\right) V\left(\boldsymbol{\xi}, \boldsymbol{\sigma}_i; \theta\right)\Bigg]$$

  where $\mathbb{A}_{-i} \equiv \times_{j \neq i} \mathbb{A}_j$ is the set of all opponent strategies.

- In a MPNE, an action $A_{it}^* \in \mathbb{A}_i$ is a best response for a player $i = 1, \ldots, N$ at time t if the following holds.

$$A_{it}^* = \underset{a \in \mathbb{A}_i}{\arg\max} \left(\mathcal{U}\left(a, \boldsymbol{x}_t, \boldsymbol{\sigma}_i^*; \theta\right) + \varepsilon_{ait}\right)$$

Markov Perfect Nash Equilibrium (5/7)

- To characterize consistency of beliefs, some more notation is necessary. Write the *ex ante* probability that player i chooses $A_{it} \in \mathbb{A}_i$, given *any* beliefs $\boldsymbol{\sigma}$ and a state $\boldsymbol{x}_t$, as follows.

$$\psi\left(A_{it}, \boldsymbol{x}_t, \boldsymbol{\sigma}; \theta\right) \equiv \mathbb{P}\left(A_{it} = \underset{a \in \mathbb{A}_i}{\arg\max}\left(\mathcal{U}\left(a, \boldsymbol{x}_t, \boldsymbol{\sigma}; \theta\right) + \varepsilon_{ait}\right)\right)$$

- For $j = 1, \dots, N$, aggregate these probabilities as follows.

$$\Psi\left(\boldsymbol{a}_t, \boldsymbol{x}_t, \boldsymbol{\sigma}_j; \theta\right) = \prod_{i=1}^{N} \psi\left(A_{it}, \boldsymbol{x}_t, \boldsymbol{\sigma}_j; \theta\right)$$

- Beliefs $\{\boldsymbol{\sigma}_i^*\}_{i=1}^{N}$ are consistent in equilibrium if $\boldsymbol{\sigma}_i^* = \boldsymbol{\sigma}^*$ for $i = 1, \dots, N$ and equate all the model-predicted probabilities for all possible action profiles $\boldsymbol{a}_t \in \mathbb{A}$ and states $\boldsymbol{x}_t \in \mathbb{X}$.

$$\sigma_i^*\left(\boldsymbol{a}_t \middle| \boldsymbol{x}_t\right) = \Psi\left(\boldsymbol{a}_t, \boldsymbol{x}_t, \boldsymbol{\sigma}^*; \theta\right) = \mathbb{P}\left(\boldsymbol{a}_t \middle| \boldsymbol{x}_t\right)$$

- Call $\mathbb{P}\left(\boldsymbol{a}_t \middle| \boldsymbol{x}_t\right)$ the *choice probability* of $\boldsymbol{a}_t$ *conditional* on $\boldsymbol{x}_t$.

Markov Perfect Nash Equilibrium (6/7)

Let $\mathbb{A} = \{\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_L\}$ and $\mathbb{X} = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_Q\}$. Write:

$$\mathbf{p} = \begin{pmatrix} \mathbb{P}\left(\boldsymbol{\alpha}_1 | \boldsymbol{\xi}_1\right) \\ \vdots \\ \mathbb{P}\left(\boldsymbol{\alpha}_L | \boldsymbol{\xi}_1\right) \\ \\ \vdots \\ \\ \mathbb{P}\left(\boldsymbol{\alpha}_1 | \boldsymbol{\xi}_Q\right) \\ \vdots \\ \mathbb{P}\left(\boldsymbol{\alpha}_L | \boldsymbol{\xi}_Q\right) \end{pmatrix} \qquad \boldsymbol{\psi}\left(\boldsymbol{\sigma}_i; \theta\right) = \begin{pmatrix} \Psi\left(\boldsymbol{\alpha}_1, \boldsymbol{\xi}_1, \boldsymbol{\sigma}_i; \theta\right) \\ \vdots \\ \Psi\left(\boldsymbol{\alpha}_L, \boldsymbol{\xi}_1, \boldsymbol{\sigma}_i; \theta\right) \\ \\ \vdots \\ \\ \Psi\left(\boldsymbol{\alpha}_1, \boldsymbol{\xi}_Q, \boldsymbol{\sigma}_i; \theta\right) \\ \vdots \\ \Psi\left(\boldsymbol{\alpha}_L, \boldsymbol{\xi}_Q, \boldsymbol{\sigma}_i; \theta\right) \end{pmatrix}$$

where $\boldsymbol{\sigma}_i$ is a vector of length LQ which, for $i = 1, \ldots, N$, arrays all the values of $\sigma_i\left(\boldsymbol{\alpha}_\ell | \boldsymbol{\xi}_q\right)$ for $\ell = 1, \ldots, L$ and $q = 1, \ldots, Q$ just as $\mathbf{p}$ does. A MPNE thus admits a **fixed point** representation.

$$\mathbf{p} = \boldsymbol{\psi}\left(\mathbf{p}; \theta\right)$$

Markov Perfect Nash Equilibrium (7/7)

This characterization of a MPNE was studied by Pesendorfer and Schmidt-Dengler (2008). Here are a few observations.

- Conditional choice probabilities and beliefs involving $\boldsymbol{\alpha}_0$ are omitted in the above as they are residually determined.
- By Brower's fixed point theorem, a MPNE always **exists**.
- However, **multiple equilibria** are possible in this setting.
- The MPNE illustrated here is **symmetric**, as functions $\pi(\cdot)$ and $q(\cdot|\cdot)$ are identical across all players. Symmetry reduces the dimensionality of the problem, but may be relaxed.
- A MPNE easily generalizes to **continuous state variables**.
- Generalizing to a **continuous action space** $\mathbb{A}$ entails more complications, especially regarding estimation.

Estimation of dynamic games (1/6)

To estimate $\boldsymbol{\theta}$ using data $\{(\mathbf{a}_{1t}, \mathbf{x}_{1t}), \dots, (\mathbf{a}_{Mt}, \mathbf{x}_{Mt})\}_{t=1}^{T}$ from M *markets*, one typically proceeds in two steps as in Rust (1987).

As in Rust (1987), the model is semi-parametrically identified if net of the discount factor β; the latter is typically calibrated.

In the first step, $\boldsymbol{\theta}_2$ and, when necessary, $\mathbf{p}$ are estimated directly. The second step about $\boldsymbol{\theta}_1$ allows for two main approaches.

1. In **full solution methods**, $\boldsymbol{\theta}_1$ is estimated by maximizing a *partial* log-likelihood function:

$$\widehat{\boldsymbol{\theta}}_1 = \underset{\boldsymbol{\theta}_1 \in \boldsymbol{\Theta}_1}{\arg\max} \sum_{t=1}^{T} \sum_{m=1}^{M} \log \Psi \left(\mathbf{a}_{mt}, \mathbf{x}_{mt}, \mathbf{p}; \boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2 \right)$$

   where $\Psi(\cdot)$ is evaluated via (usually expensive) nested fixed point algorithms. With multiple equilibria, this is a QMLE.

Estimation of dynamic games (2/6)

2. In a **conditional choice probability** approach, estimates of $\mathbf{p}$ from the first step inform the following GMM problem:

$$\widehat{\theta}_1 = \underset{\theta_1 \in \Theta_1}{\arg\min} \left[\mathbf{p} - \boldsymbol{\psi}\left(\mathbf{p}; \theta_1, \widehat{\theta}_2\right)\right]^{\mathrm{T}} \mathbf{W} \left[\mathbf{p} - \boldsymbol{\psi}\left(\mathbf{p}; \theta_1, \widehat{\theta}_2\right)\right]$$

   where $\mathbf{W}$ is some suitable GMM weighting matrix. Given $\mathbf{p}$, $\boldsymbol{\psi}(\cdot)$ may be evaluated for example via forward simulation.

Both approaches benefit from parametric assumptions about the unobserved shocks $\boldsymbol{\varepsilon}_{it}$. If these are all i.i.d. standard Gumbel:

$$\Psi\left(\mathbf{a}_{mt}, \mathbf{x}_{mt}, \mathbf{p}; \theta_1, \widehat{\theta}_2\right) = \prod_{i=1}^{N} \frac{\exp\left(\mathcal{U}\left(a_{imt}, \mathbf{x}_{mt}, \mathbf{p}; \theta_1, \widehat{\theta}_2\right)\right)}{\sum_{a \in \mathbb{A}_i} \exp\left(\mathcal{U}\left(a, \mathbf{x}_{mt}, \mathbf{p}; \theta_1, \widehat{\theta}_2\right)\right)}$$

which can be evaluated if function $\mathcal{U}(A_{it}, \boldsymbol{x}_t, \boldsymbol{\sigma}_i; \theta)$ is known; this in turn depends on the ex-ante value function $V(\boldsymbol{x}_t, \boldsymbol{\sigma}_i; \theta)$.

Estimation of dynamic games (3/6)

Like in the single-agent dynamic logit, $V(\mathbf{x}_t, \mathbf{p}; \boldsymbol{\theta})$ can be derived algebraically if $\mathbf{p}$ is known (or estimated) and $\mathbb{X}$ is discrete. Let:

$$\boldsymbol{\pi}(\boldsymbol{x}_t; \boldsymbol{\theta}_1) = \begin{pmatrix} \pi(\boldsymbol{\alpha}_0, \boldsymbol{x}_t; \boldsymbol{\theta}_1) + \mathbb{E}[\varepsilon_{it} | \boldsymbol{\alpha}_0, \boldsymbol{x}_t] \\ \vdots \\ \pi(\boldsymbol{\alpha}_L, \boldsymbol{x}_t; \boldsymbol{\theta}_1) + \mathbb{E}[\varepsilon_{it} | \boldsymbol{\alpha}_L, \boldsymbol{x}_t] \end{pmatrix}$$

where expectations have a closed form in the Gumbel case; and:

$$\mathbf{v}(\mathbf{p}, \boldsymbol{\theta}) = \begin{pmatrix} V(\boldsymbol{\xi}_1, \mathbf{p}; \boldsymbol{\theta}) \\ \vdots \\ V(\boldsymbol{\xi}_Q, \mathbf{p}; \boldsymbol{\theta}) \end{pmatrix} \qquad \boldsymbol{\pi}(\boldsymbol{\theta}_1) = \begin{pmatrix} \boldsymbol{\pi}(\boldsymbol{\xi}_1; \boldsymbol{\theta}_1) \\ \vdots \\ \boldsymbol{\pi}(\boldsymbol{\xi}_Q; \boldsymbol{\theta}_1) \end{pmatrix}$$

and:

$$\widehat{\mathbf{Q}}(\boldsymbol{x}_t) \equiv \begin{pmatrix} q\left(\boldsymbol{\xi}_1 \middle| \boldsymbol{x}_t, \boldsymbol{\alpha}_0; \widehat{\boldsymbol{\theta}}_2\right) & \dots & q\left(\boldsymbol{\xi}_Q \middle| \boldsymbol{x}_t, \boldsymbol{\alpha}_0; \widehat{\boldsymbol{\theta}}_2\right) \\ \vdots & \ddots & \vdots \\ q\left(\boldsymbol{\xi}_1 \middle| \boldsymbol{x}_t, \boldsymbol{\alpha}_L; \widehat{\boldsymbol{\theta}}_2\right) & \dots & q\left(\boldsymbol{\xi}_Q \middle| \boldsymbol{x}_t, \boldsymbol{\alpha}_L; \widehat{\boldsymbol{\theta}}_2\right) \end{pmatrix}$$

...

Estimation of dynamic games (4/6)

…and:

$$\widehat{\mathbf{p}}\left(\boldsymbol{x}_t\right) = \begin{pmatrix} \widehat{\mathbb{P}}\left(\left.\boldsymbol{\alpha}_0\right|\boldsymbol{x}_t\right) \\ \vdots \\ \widehat{\mathbb{P}}\left(\left.\boldsymbol{\alpha}_L\right|\boldsymbol{x}_t\right) \end{pmatrix} \qquad \widehat{\mathbf{p}} = \begin{pmatrix} \widehat{\mathbf{p}}\left(\boldsymbol{\xi}_1\right) \\ \vdots \\ \widehat{\mathbf{p}}\left(\boldsymbol{\xi}_Q\right) \end{pmatrix}$$

and lastly:

$$\widehat{\boldsymbol{\Psi}} = \begin{pmatrix} \widehat{\mathbf{p}}^{\mathrm{T}}\left(\boldsymbol{\xi}_1\right) & \dots & \mathbf{0}^{\mathrm{T}} \\ \vdots & \ddots & \vdots \\ \mathbf{0}^{\mathrm{T}} & \dots & \widehat{\mathbf{p}}^{\mathrm{T}}\left(\boldsymbol{\xi}_Q\right) \end{pmatrix} \qquad \widehat{\mathbf{Q}} = \begin{pmatrix} \widehat{\mathbf{Q}}\left(\boldsymbol{\xi}_1\right) \\ \vdots \\ \widehat{\mathbf{Q}}\left(\boldsymbol{\xi}_Q\right) \end{pmatrix}$$

then, similarly as in the single-agent model, one can express the *equilibrium* ex-ante value function as a system of Q equations:

$$\mathbf{v}\left(\widehat{\mathbf{p}}, \boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2\right) = \widehat{\boldsymbol{\Psi}}\boldsymbol{\pi}\left(\boldsymbol{\theta}_1\right) + \beta\widehat{\boldsymbol{\Psi}}\widehat{\mathbf{Q}}\mathbf{v}\left(\widehat{\mathbf{p}}, \boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2\right)$$

which, when solved, returns the values that are sought after.

$$\mathbf{v}\left(\widehat{\mathbf{p}}, \boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2\right) = \left[\mathbf{I} - \beta\widehat{\boldsymbol{\Psi}}\widehat{\mathbf{Q}}\right]^{-1}\widehat{\boldsymbol{\Psi}}\boldsymbol{\pi}\left(\boldsymbol{\theta}_1\right)$$

Estimation of dynamic games (5/6)

In addition to methods derived from the dynamic logit tradition, Bajari, Benkard and Levin (2007) developed one which is closer to the literature on static games. In a nutshell, let:

$$\widetilde{\mathcal{V}}\left(A_{it}, \boldsymbol{x}_t, \boldsymbol{\sigma}_i; \theta\right) \equiv \mathcal{U}\left(A_{it}, \boldsymbol{x}_t, \boldsymbol{\sigma}_i; \theta\right) + \sum_{a \in \mathbb{A}_i} \varepsilon_{ait} \cdot \mathbb{1}\left[a = A_{it}\right]$$

and observe that a MPNE implies the following set of inequalites:

$$\widetilde{\mathcal{V}}\left(A_{it}^*, \boldsymbol{x}_t, \boldsymbol{\sigma}_i^*; \theta\right) \geq \widetilde{\mathcal{V}}\left(A_{it}, \boldsymbol{x}_t, \boldsymbol{\sigma}_i^*; \theta\right)$$

for all players i and their suboptimal actions $A_{it} \in \mathbb{A}_i$, $A_{it} \neq A_{it}^*$, and all $\boldsymbol{x}_t \in \mathbb{X}$. This leads to the following population criterion:

$$\mathcal{Q}_0\left(\theta\right) = \int_{\mathbb{X}} \sum_{i=1}^{N} \sum_{a \neq A_{it}^*} \left\| \widetilde{\mathcal{V}}\left(A_{it}^*, \mathbf{x}_t, \mathbf{p}; \theta\right) - \widetilde{\mathcal{V}}\left(a, \mathbf{x}_t, \mathbf{p}; \theta\right) \right\|_{-} dG\left(\mathbf{x}_t\right)$$

where $G\left(\mathbf{x}_t\right)$ is a suitable distribution for $\boldsymbol{x}_t$. Depending on issues of identification (induced e.g. by multiple equilibria), this allows for both point and set estimation.

Estimation of dynamic games (6/6)

The estimation of dynamic games can entail **complications** on the computational side, a traditional challenge to this literature.

- The framework based on the MPNE is tractable so long as N is small. For the many-players case the literature has focused on a solution concept called **oblivious equilibrium**, where payoffs only depend upon *statistics* of the opponents' actions.

- With N small enough and a **continuous action space** $\mathbb{A}$, which is popular in applications, a MPNE-based framework is still tractable. However, the conditional choice probability approach then requires to evaluate the *empirical distribution function* of $\boldsymbol{a}_t$ given $\boldsymbol{x}_t$. Issues also apply to other methods.

- In general, extensions such as **asymmetric information**, **unobserved heterogeneity**, and **behavioral biases** may lead to increased computational complexity.

Dynamic games: empirical applications (1/2)

This introduction to dynamic games concludes with an overview of selected empirical applications. Most study models of *dynamic oligopoly*, and extend the basic framework to various degrees.

- Benkard (2004) studies competition between the three main aircraft manufacturers: Boeing, Airbus and Lockheed. With a dynamic model and *learning-by-doing* one can rationalize why airplanes are initially sold below marginal cost.

- Ryan (2012) assesses the effect of environmental regulations in the U.S. cement industry, where producers are subject to *capacity constraints* that are costly to adjust. The estimates suggest consumer welfare losses due to the regulations.

- Lee (2013) studies *network effects* in the video game console industry, where the value of a console depends on the number of exclusive games that become available for it over time.

Dynamic games: empirical applications (2/2)

- Goettler and Gordon (2011) analyze the PC microprocessor duopoly, dominated by Intel and AMD. Microprocessors are subject to *endogenous technological obsolescence*: consumers like to replace them with more innovative releases. Through counterfactual analysis, the paper suggests that a monopoly run by Intel would lead to faster innovation (due to stronger incentives for R&D) but also lower overall consumer welfare.

- Hashmi and Biesebroek (2016) show that in the automobile industry, higher market power (estimated via BLP) provides more incentives for firms to climb the *quality ladder* faster.

- Igami (2017) looks instead at competition in the market for hard drives, where adopting newer technological standards is cheaper for entrants than for incumbents because of *product cannibalization*. The model allows to evaluate the incentives that incumbents have for avoiding cannibalization and entry.