Dynamic Discrete Choice Structural Models in Empirical IO

Lecture 2: Dynamic Games of Oligopoly Competition

Victor Aguirregabiria (University of Toronto)

Carlos III, Madrid

June 27, 2017

Dynamic Games: Outline

1. Structure of empirical dynamic games

2. Data and Identification

3. Estimation

4. Dealing with unobserved heterogeneity

5. Network Competition in the Airline Industry: Aguirregabiria & Ho (2012)

Examples of Empirical Applications

• Land use regulation (entry cost) and entry-exit and competition in the hotel industry: Suzuki (IER; 2013);

• Subsudies to entry in small markets for the dentist industry: Dunne et al. (RAND, 2013);

• Environmental regulation and entry-exit and capacity choice in cement industry: Ryan (ECMA, 2012);

• Demand uncertainty and firm investment in the concrete industry: Collard-Wexler (ECMA, 2013);

• Time-to-build, uncertainty, and dynamic competition in the shipping industry: Kalouptsidi (AER, 2014);

Examples of Empirical Applications [2]

• Fees for musical performance rights and the choice of format (product design) of radio stations: Sweeting (ECMA, 2013);

• Hub-and-spoke networks, route entry-exit and competition in the airline industry: Aguirregabiria and Ho (JoE, 2012);

• Competition in R&D and product innovation between Intel and AMD: Goettler and Gordon (JPE, 2011);

• Product innovation of incumbents and new entrants in the hard drive industry: Igami (JPE, 2017);

• Cannibalization and preemption strategies in the Canadian fast-food industry: Igami and Yang (QE, 2016)

Examples of Empirical Applications [3]

- Rrelease date of a movie: Einav (EI, 2010).

- Dynamic price competition: Kano (IHIO, 2013); Ellickson, Misra, and Nair (JMR, 2012).

- Endogenous mergers: Jeziorski (RAND, 2014).

- Explotation of a common natural resource (foshing): Huang and Smith (AER, 2014).

Dynamic Games: Basic Structure

• Time is discrete and indexed by t. The game is played by N firms that we index by i.

• Following the standard structure in the Ericson-Pakes (1995) framework, firms compete in two different dimensions: a static dimension and a dynamic dimension.

• We denote the dynamic dimension as the "investment decision".

• Here I focus on the dynamic part.

Dynamic Games: Basic Structure (2)

• Let a_{it} be the variable that represents the investment decision of firm i at period t.

• This investment decision can be an entry/exit decision, a choice of capacity, investment in equipment, R&D, product quality, other product characteristics, etc.

• Every period, given their capital stocks that can affect demand and/or production costs, firms compete in prices or quantities in a static Cournot or Bertand model. Let p_{it} be the static decision variables (e.g., price) of firm i at period t.

Dynamic Games: Basic Structure (3)

• The action is taken to maximize the expected and discounted flow of profits in the market,

$$E_t \left(\sum_{r=0}^{\infty} \delta^r \ \Pi_{it+r} \right)$$

where $\delta \in (0,1)$ is the discount factor, and Π_{it} is firm i's profit at period t.

• The profits of firm i at time t:

$$\Pi_{it} = \pi_i \left(a_{it}, \mathbf{a}_{-it}, \mathbf{x}_t \right) + \varepsilon_{it}(a_{it})$$

$\mathbf{a}_{-it}$ = Vector with opponnets´ actions
$\mathbf{x}_t$ = Vector of common knowledge state variables
$\varepsilon_{it} = (\varepsilon_{it}(0), ..., \varepsilon_{it}(J))$ = Vector of variables that are private information of player i

State variables

- The vector $\mathbf{x}_t$ includes both endogenous and exogenous state variables.

- For instance, consider a model of competition in quality, i.e., a_{it} represents quality choice.

- In this case, $\mathbf{x}_t = (a_{1,t-1}, ..., a_{N,t-1}, z_{1t}, ...z_{Nt})$, where $(a_{1,t-1}, ..., a_{N,t-1})$ represents the endogenous state variables, and $\mathbf{z}_t = (z_{1t}, ...z_{Nt})$ the exogenous.

State variables (2)

• The specification of the model is completed with the transition rules of these state variables.

• (1) Exogenous state variables follow an exogenous Markov process with transition probability function $f_z(\mathbf{z}_{t+1}|\mathbf{z}_t)$.

• (2) Private information shock ε_{it} is i.i.d. over time and independent across firms.

Markov Perfect Equilibrium

• Most of the recent literature in IO studying industry dynamics focuses on studying a Markov Perfect Equilibrium (MPE), as defined by Maskin and Tirole (Econometrica, 1988).

• The key assumption in this solution concept is that players' strategies are functions of only payoff-relevant state variables.

• In this model, the payoff-relevant state variables for firm i are $(\mathbf{x}_t, \varepsilon_{it})$.

• Allowing strategies to depend on history $\mathbf{x}_{t-1}$, $\mathbf{x}_{t-2}$, ...

Markov Perfect Equilibrium (2)

• Let $\boldsymbol{\alpha} = \{\alpha_i(\mathbf{x}_t, \varepsilon_{it}) : i \in \{1, 2, ..., N\}\}$ be a set of strategy functions, one for each firm.

• A MPE is a set of strategy functions $\boldsymbol{\alpha}^*$ such that every firm is maximizing its value given the strategies of the other players.

• For given strategies of the other firms, the decision problem of a firm is a single-agent dynamic programming (DP) problem.

Markov Perfect Equilibrium (3)

• Let $V_i^{\boldsymbol{\alpha}}(\mathbf{x}_t, \varepsilon_{it})$ be the value function of the DP problem that describes the best response of firm i to the strategies $\boldsymbol{\alpha}_{-i}$ of the other firms.

• This value function is the unique solution to the Bellman equation:

$$V_i^{\boldsymbol{\alpha}}(\mathbf{x}_t, \varepsilon_{it}) = \max_{a_{it}} \left\{ \begin{array}{l} \Pi_i^{\boldsymbol{\alpha}}(a_{it}, \mathbf{x}_t) - \varepsilon_{it}(a_{it}) \\ \\ +\delta \int V_i^{\boldsymbol{\alpha}}(\mathbf{x}_{t+1}, \varepsilon_{it+1}) \ f_\varepsilon(\varepsilon_{it+1}) \ f_i^{\boldsymbol{\alpha}}(\mathbf{x}_{t+1} | a_{it}, \mathbf{x}_t) \ d\varepsilon_{it+1} \ d \end{array} \right.$$

where $\Pi_i^{\boldsymbol{\alpha}}(a_{it}, \mathbf{x}_t)$ and $F_i^{\boldsymbol{\alpha}}(\mathbf{x}_{t+1} | a_{it}, \mathbf{x}_t)$ are the expected one-period profit and the expected transition of the state variables, respectively, for firm i given the strategies of the other firms.

MPE and Conditional Choice Probabilities

• Given a strategy function $\alpha_i(\mathbf{x}_t, \varepsilon_{it})$, we can define the corresponding *Conditional Choice Probability (CCP)* function as :

$$\begin{aligned} P_i(a|\mathbf{x}) &\equiv \Pr\left(\alpha_i(\mathbf{x}_t, \varepsilon_{it}) = a \mid \mathbf{x}_t = \mathbf{x}\right) \\ &= \int 1\{\alpha_i(\mathbf{x}_t, \varepsilon_{it}) = a\} \; f_\varepsilon(\varepsilon_{it+1}) \; d\varepsilon_{it+1} \end{aligned}$$

• The expected profit function and the transition probability function, $\Pi_i^{\boldsymbol{\alpha}}$ and $f_i^{\boldsymbol{\alpha}}$, can be represented in terms of CCPs, $\Pi_i^{\mathbf{P}}$ and $f_i^{\mathbf{P}}$.

MPE in terms of CCPs

- Based on the concept of CCP, we can represent the equilibrium mapping and a MPE in way that is particularly useful for the econometric analysis.

- This representation has two main features:
  (1) a MPE is a vector of CCPs;
  (2) a player's best response is an optimal response not only to the other players' strategies but also to his own strategy in the future.

MPE in terms of CCPs (2)

• A MPE is a vector of CCPs, $\mathbf{P} \equiv \{P_i(a_i|\mathbf{x}) : \text{for any } (i, a_i, x)\}$, such that:

$$P_i(a_i \mid \mathbf{x}) = \Pr\left(a_i = \arg\max_j \left\{v_i^{\mathbf{P}}(j, \mathbf{x}) - \varepsilon_i(j)\right\} \mid \mathbf{x}\right)$$

• $v_i^{\mathbf{P}}(a_i, \mathbf{x})$ = Value of firm i if the firm chooses alternative a_i today:
- firm i behaves according to CCP P_i in the future;
- all the other firms behave according to their respective CCPs in $\mathbf{P}$ now and in the future.

• The Representation Lemma in Aguirregabiria and Mira (2007) shows that every MPE in this dynamic game can be represented using this mapping.

MPE in terms of CCPs (3)

- The form of this equilibrium mapping depends on the distribution of ε_i.

- For instance, in the entry/exit model, if ε_i is $N(0, \sigma_\varepsilon^2)$:

$$P_i(1|\mathbf{x}) = \Phi\left(\frac{v_i^{\mathbf{P}}(1,\mathbf{x}) - v_i^{\mathbf{P}}(0,\mathbf{x})}{\sigma_\varepsilon}\right)$$

- In the model with endogenous quality choice, if $\varepsilon_i(a)$'s are extreme value type 1 distributed:

$$P_i(a|\mathbf{x}) = \frac{\exp\left\{\dfrac{v_i^{\mathbf{P}}(a,\mathbf{x})}{\sigma_\varepsilon}\right\}}{\sum_{a'=0}^{A}\exp\left\{\dfrac{v_i^{\mathbf{P}}(a',\mathbf{x})}{\sigma_\varepsilon}\right\}}$$

Computing values and best response probs

• Let $V_i^{\mathbf{P}}(\mathbf{x})$ be the value function of firm i when all the firms behave according to their CCPs in $\mathbf{P}$. By definition:

$$V_i^{\mathbf{P}}(\mathbf{x}_t) = \sum_{a_{it}=0}^{A} P_i(a_{it}|\mathbf{x}_t) \left[\Pi_i^{\mathbf{P}}(a_{it}, \mathbf{x}_t) + \delta \sum_{\mathbf{x}_{t+1}} V_i^{\mathbf{P}}(\mathbf{x}_{t+1}) \; f_i^{\mathbf{P}}(\mathbf{x}_{t+1}|a_{it}, \mathbf{x}_t) \right]$$

with

$$\Pi_i^{\mathbf{P}}(a_{it}, \mathbf{x}_t) = \sum_{a_{-it}} P_{-i}(a_{-it}|\mathbf{x}_t) \; \pi_i(a_{it}, a_{-it}, \mathbf{x}_t)$$

$$f_i^{\mathbf{P}}(\mathbf{x}_{t+1}|a_{it}, \mathbf{x}_t) = \sum_{a_{-it}} P_{-i}(a_{-it}|\mathbf{x}_t) \; f_x(\mathbf{x}_{t+1}|a_{it}, a_{-it}, \mathbf{x}_t)$$

Computing values and best response probs [2]

• When the space $\mathcal{X}$ is discrete we can obtain $V_i^{\mathbf{P}}$ as the solution of a system of linear equations:

$$\mathbf{V}_i^{\mathbf{P}} = \left[\sum_{a_{it}=0}^{A} \mathbf{P}_i(a_{it}) * \Pi_i^{\mathbf{P}}(a_{it}) \right] + \delta \left[\sum_{a_i=0}^{A} \mathbf{P}_i(a_{it}) * \mathbf{F}_i^{\mathbf{P}}(a_i) \right] \mathbf{V}_i^{\mathbf{P}}$$

such that:

$$\mathbf{V}_i^{\mathbf{P}} = \left(\mathbf{I} - \delta \left[\sum_{a_i=0}^{A} \mathbf{P}_i(a_{it}) * \mathbf{F}_i^{\mathbf{P}}(a_i) \right] \right)^{-1} \left[\sum_{a_{it}=0}^{A} \mathbf{P}_i(a_{it}) * \Pi_i^{\mathbf{P}}(a_{it}) \right]$$

• The main computational cost in this valuation comes from the inverse matrix.

Computing values and best response probs [3]

- The conditional choice value function is:

$$v_i^{\mathbf{P}}(a_{it}, \mathbf{x}_t) = \Pi_i^{\mathbf{P}}(a_{it}, \mathbf{x}_t) + \delta \sum_{\mathbf{x}_{t+1}} V_i^{\mathbf{P}}(\mathbf{x}_{t+1})\ f_i^{\mathbf{P}}(\mathbf{x}_{t+1} | a_{it}, \mathbf{x}_t)$$

- And the best response mapping in the space of CCPs becomes:

$$P_i(a_{it} \mid \mathbf{x}_t) = \Pr\left(a_{it} = \arg\max_j \left\{ v_i^{\mathbf{P}}(j, \mathbf{x}_t) + \varepsilon_i(j) \right\} \mid \mathbf{x}\right)$$

- This is a continuous mapping in the space of CCPs.

- An equilbrium exists. In general, there are multiple equilibria.

Data

• The researcher observes a random sample of M markets, indexed by m, over T periods of time, where the observed variables consists of players' actions and state variables.

• For the moment, we consider that the industry and the data are such that:

(a) each firm is observed making decisions in every of the M markets;

(b) the researcher knows all the payoff relevant market characteristics that are common knowledge to the firms, $\mathbf{x}$.

Data

• With this type of data we can allow for rich firm heterogeneity that is fixed across markets and time by estimating firm-specific structural parameters, $\boldsymbol{\theta}_i$.

• This 'fixed-effect' approach to deal with firm heterogeneity is not feasible in data sets where most of the competitors can be characterized as *local players*, i.e., firms specialized in operating in a few markets.

• Condition (b) rules out the existence of unobserved market heterogeneity. Though it is a convenient assumption, it is also unrealistic for most applications in empirical IO. Later I present estimation methods that relax conditions (a) and (b) and deal with unobserved market and firm heterogeneity.

Data

• Suppose that we have a random sample of M local markets, indexed by m, over T periods of time, where we observe:

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

• We want to use these data to estimate the model parameters in the population that has generated this data: $\boldsymbol{\theta}^0 = \{\boldsymbol{\theta}_i^0 : i \in I\}$.

Identification

• A significant part of this literature has considered the following identification assumptions.

Assumption (ID 1): Single equilibrium in the data. Every observation in the sample comes from the same Markov Perfect Equilibrium, i.e., for any observation (m, t), $\mathbf{P}_{mt}^{0} = \mathbf{P}^{0}$.

Assumption (ID 2): No unobserved common-knowledge variables. The only unobservables for the econometrician are the private information shocks ε_{imt} and the structural parameters $\boldsymbol{\theta}$.

Identification (2)

- Under assumptions ID-1 & ID-2, the equilibrium that has generated the data, $\mathbf{P}^0$, can be estimated consistently and nonparametrically from the data. For any $(i, a_i, \mathbf{x})$:

$$P_i^0(a_i|\mathbf{x}) = \Pr(a_{imt} = a_i \mid \mathbf{x}_{mt} = \mathbf{x})$$

For instance, we can estimate consistently $P_i^0(a_i|\mathbf{x})$ using the following simple kernel estimator:

$$P_i^0(a_i|\mathbf{x}) = \frac{\sum_{m,t} 1\{a_{imt} = a_i\}\ K\left(\dfrac{\mathbf{x}_{mt} - \mathbf{x}}{b_n}\right)}{\sum_{m,t} K\left(\dfrac{\mathbf{x}_{mt} - \mathbf{x}}{b_n}\right)}$$

Identification (3)

- Note that under the single-equilibrium-in-the-data assumption, the multiplicity of equilibria in the model does not play any role in the identification of the structural parameters.

- The single-equilibrium-in-the-data assumption is a sufficient for identification but it is not necessary.

- Sweeting (2013) and Aguirregabiria and Mira (2015) present conditions for the point-identification of games of incomplete information when there are multiple equilibria in the data.

Identification (4)

• Given that $\mathbf{P}^0$ is identified, the payoff function $\pi_i(a_{it}, a_{-it}, \mathbf{x}_t)$ is nonparametrically identified under the following assumptions (sufficient conditions):

Assumption (ID 3): The distribution of the unobservables ε is known to the researcher.

Assumption (ID 4): The discount factor δ is known to the researcher.

Assumption (ID 5): Normalization: $\pi_i(a_{it}, a_{-it}, \mathbf{x}_t) = 0$

Assumption (ID 6): Exclusion restriction: $\mathbf{x}_t = (s_{1t}, s_{2t}, ..., s_{Nt}, \mathbf{w}_t)$ such that:

$$\pi_i(a_{it}, a_{-it}, \mathbf{x}_t) = \pi_i(a_{it}, a_{-it}, s_{it}, \mathbf{w}_t)$$

Estimation

- I will describe the following estimators:

  1. *Two-step Pseudo MLE*
  2. *Recursive K-step Pseudo MLE*

- For illustration, I focus on a binary probit model with:

$$\pi_i \left(a_{it}, a_{-it}, \mathbf{x}_t \right) = z \left(a_{it}, a_{-it}, \mathbf{x}_t \right) \; \boldsymbol{\theta}_i$$

  where $z \left(a_{it}, a_{-it}, \mathbf{x}_t \right)$ is a vector of known function and $\boldsymbol{\theta}_i$ is a vector of unknown parameters.

Estimation (2)

- The equilibrium mapping is:

$$P_i(a_{it} \mid \mathbf{x}_t, \boldsymbol{\theta}_i) = \Phi\left(\left[\tilde{z}_i^{\mathbf{P}}(1, \mathbf{x}_t) - \tilde{z}_i^{\mathbf{P}}(0, \mathbf{x}_t)\right] \ \boldsymbol{\theta}_i\right)$$

- $\tilde{z}_i^{\mathbf{P}}(1, \mathbf{x}_t)$ and $\tilde{z}_i^{\mathbf{P}}(0, \mathbf{x}_t)$ present values can be calculated by solving a system of inear equations (see valuation operator above) for given vector of CCPs and the function $z()$.

Pseudo Likelihood Function

- For the description of the different estimators, it is convenient to define the following **Pseudo Likelihood function**:

$$
\begin{aligned}
Q(\boldsymbol{\theta}, \mathbf{P}) \quad = \quad & \sum_{m=1}^{M} \sum_{i=1}^{N} \sum_{t=1}^{T} a_{imt} \;\; \ln \Phi \left(\left[\tilde{z}_i^{\mathbf{P}}(1, \mathbf{x}_{mt}) - \tilde{z}_i^{\mathbf{P}}(0, \mathbf{x}_{mt}) \right] \;\; \boldsymbol{\theta}_i \right) \\
& + (1 - a_{imt}) \ln \left[1 - \Phi \left(\left[\tilde{z}_i^{\mathbf{P}}(1, \mathbf{x}_{mt}) - \tilde{z}_i^{\mathbf{P}}(0, \mathbf{x}_{mt}) \right] \;\; \boldsymbol{\theta}_i \right) \right]
\end{aligned}
$$

- This pseudo likelihood function treats firms' beliefs $\mathbf{P}$ as parameters to estimate together with $\boldsymbol{\theta}$.

- Note that for given $\mathbf{P}$, the function $Q(\boldsymbol{\theta}, \mathbf{P})$ is the likelihood of a Probit model.

Two-step methods

- Suppose that we knew the equilibrium in the population, $\mathbf{P}^0$.
- Given $\mathbf{P}^0$ we can construct the variables $\tilde{z}_i^{\mathbf{P}^0}(1, \mathbf{x}_{mt}) - \tilde{z}_i^{\mathbf{P}^0}(0, \mathbf{x}_{mt})$ and then obtain a very simple estimator of $\boldsymbol{\theta}^0$.

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \; Q(\boldsymbol{\theta}, \mathbf{P}^0)$$

- This estimator is root-M consistent and asymptotically normal under the standard regularity conditions. It is not efficient because it does not impose the equilibrium constraints (only asymptotically).
- While equilibrium probabilities are not unique functions of structural parameters, the best response probabilities that appear in $Q(\boldsymbol{\theta}, \mathbf{P})$ are unique functions of structural parameters and players' beliefs.

Two-step methods (2)

- The previous method is infeasible because $\mathbf{P}^0$ is unknown.
- However, under the Assumptions **"No-unobserved-market-heterogeneity"** and **"One-MPE-in-the-data"** we can estimate $\mathbf{P}^0$ consistently and at with a convergence rate such that the two-step estimator $\hat{\boldsymbol{\theta}}$ is root-M consistent and asymptotically normal.
- For instance, a kernel estimator of $\mathbf{P}^0$ is:

$$\widehat{P}_i^0(\mathbf{x}) = \frac{\sum_{m=1}^{M} \sum_{t=1}^{T} a_{imt} \; K\left(\dfrac{\mathbf{x}_{mt} - \mathbf{x}}{b}\right)}{\sum_{m=1}^{M} \sum_{t=1}^{T} K\left(\dfrac{\mathbf{x}_{mt} - \mathbf{x}}{b}\right)}$$

Two-step methods: Finite sample properties (1)

- The most attractive feature of two-step methods is their relative simplicity.
- However, they suffer of a potentially important problem of finite sample bias.
- The finite sample bias of the two-step estimator of θ^0 depends very importantly on the properties of the first-step estimator of $\mathbf{P}^0$. In particular, it depends on the rate of convergence and on the variance and bias of $\widehat{\mathbf{P}^0}$.
- It is well-known that there is a **curse of dimensionality in the NP estimation** of a regression function such as $\mathbf{P}^0$.

Two-step methods: Finite sample properties (2)

- In particular, the rate of convergence of $\widehat{\mathbf{P}^0}$ declines, and the variance and bias increase, very quickly as the number of conditioning regressors increases.
- In our simple example, the vector x_{mt} contains only three variables: the binary indicators $a_{im,t-1}$ and the (continuous) market size S_{mt}. In this case, the NP estimator of $\mathbf{P}^0$ has a relatively high rate of convergence an its variance and bias can be small even with relatively small sample.
- However, there are applications with more than two (heterogeneous) players and where firm size, capital stock or other predetermined continuos firm-specific characteristics are state variables.
- Even with binary state variables ($a_{im,t-1}$), when the number of players is relatively large (e.g., more than 10) .....

Recursive K-step estimator

- K-step extension of the 2-step estimator. Given an initial consistent (NP) estimator $\widehat{\mathbf{P}^0}$, the sequence of estimators $\{\widehat{\boldsymbol{\theta}^K}, \widehat{\mathbf{P}^K} : K \geq 1\}$ is defined as:

$$\widehat{\boldsymbol{\theta}^{K+1}} = \arg\max_{\boldsymbol{\theta}} \; Q\left(\boldsymbol{\theta}, \widehat{\mathbf{P}^K}\right)$$

where:

$$\widehat{P_i^K}(\mathbf{x}) = \Phi\left(\left[\tilde{z}_i^{\widehat{\mathbf{P}^{K-1}}}(1,\mathbf{x}) - \tilde{z}_i^{\widehat{\mathbf{P}^{K-1}}}(0,\mathbf{x})\right] \; \widehat{\boldsymbol{\theta}_i^K}\right)$$

- Aguirregabiria and Mira (2002, 2007) present Monte Carlo experiments which illustrate how this recursive estimators can have significantly smaller bias than the two-step estimator.
- Kasahara and Shimotsu (2008) derive a second order approximation to the bias of these K-stage estimators. They show that, if the equilibrium in the population is stable, then this recursive procedure reduces the bias.

Dealing with the curse of dimensionality

- The main computation cost comes from the valuation operator that requires solving a system of linear equations with dimension $|\mathcal{X}|$.
- The cost of solving this system is $O(|\mathcal{X}|^3)$.
- When players are heterogeneous (state variables specific of each player), $|\mathcal{X}| = |\mathcal{S}|^N$.
- Examples: Four state variables per player; each state variable takes 10 values (i.e., $|\mathcal{S}| = 10^4$); $N = 5$ players: then, $|\mathcal{X}| = |\mathcal{S}|^5 = 10^{20}$.

Dealing with the curse of dimensionality

- Different approaches have been proposed / applied to deal with this curse of dimensionality:

    1. *No permanent heterogeneity across players / firms.*
    2. *Oblivious equilibrium.*
    3. *Monte Carlo simulation methods.*
    4. *Euler equations / Finite dependence*

Counterfactual Experiments with Estimated Model (1)

- One of the most attractive features of structural models is that they can be used to predict the effects of new policies or changes in parameters (counterfactuals).
- However, this a challenging exercise in a model with multiple equilibria.
- The data can identify the "factual" equilibrium. However, under the counterfactual scenario, which of the multiple equilibria we should choose?

Counterfactual Experiments with Estimated Model (2)

- Different approaches have been implemented in practice.
- Select the equilibrium to which we converge by iterating in the (counterfactual) equilibrium mapping starting with the factual equilibrium $\mathbf{P}^0$
- Select the equilibrium with maximum total profits (or alternatively, with maximum welfare).
- Homotopy method: Aguirregabiria and Ho (2007)

Counterfactual Experiments: Homotopy method

- Let $\boldsymbol{\theta}$ be the vector of structural parameters in the model. An let $\Psi(\boldsymbol{\theta}, \mathbf{P})$ be the equilibrium mapping such that an equilibrium associated with $\boldsymbol{\theta}$ can be represented as a fixed point:

$$\mathbf{P} = \Psi(\boldsymbol{\theta}, \mathbf{P})$$

- The model could be completed with an equilibrium selection mechanism: i.e., a criterion that selects one and only one equilibrium for each possible $\boldsymbol{\theta}$.
- Suppose that there is a "true" equilibrium selection mechanism in the population under study, but we do not know that mechanism.
- Our approach here (both for the estimation and for counterfactual experiments) is completely agnostic with respect to the equilibrium selection mechanism.

Counterfactual Experiments: Homotopy method

- We only assume that there is such a mechanism, and that it is a smooth function of $\boldsymbol{\theta}$.
- Let $\boldsymbol{\pi}(\boldsymbol{\theta})$ be the (unique) selected equilibrium, for given $\boldsymbol{\theta}$, if we apply the "true" selection mechanism.
- Since we do not know the mechanism, we do not know $\boldsymbol{\pi}(\boldsymbol{\theta})$ for every possible $\boldsymbol{\theta}$.
- However, we DO know $\boldsymbol{\pi}(\boldsymbol{\theta})$ at the true $\boldsymbol{\theta}_0$ because we know that:

$$\mathbf{P}_0 = \boldsymbol{\pi}(\boldsymbol{\theta}_0)$$

and both $\mathbf{P}_0$ and $\boldsymbol{\theta}_0$ are identified.

Counterfactual Experiments: Homotopy method

- Let θ_0 and $\mathbf{P}_0$ be the the population values. Let $(\hat{\theta}_0, \hat{\mathbf{P}}_0)$ be our consistent estimator.
- We do not know the function $\pi(\theta)$. All what we know is that the point $(\hat{\theta}_0, \hat{\mathbf{P}}_0)$ belongs to the graph of this function π.
- Let θ^* be the vector of parameters under a counterfactual scenario.
- We want to know the counterfactual equilibrium $\pi(\theta^*)$.

Counterfactual Experiments: Homotopy method

- A Taylor approximation to $\pi(\theta^*)$ around our estimator $\hat{\theta}_0$ implies that:

$$
\begin{aligned}
\pi(\theta^*) &= \pi\left(\hat{\theta}_0\right) + \frac{\partial \pi\left(\hat{\theta}_0\right)}{\partial \theta'}\left(\theta^* - \hat{\theta}_0\right) + O\left(\left\|\theta^* - \hat{\theta}_0\right\|^2\right) \\
&= \hat{\mathbf{P}}_0 + \frac{\partial \pi\left(\hat{\theta}_0\right)}{\partial \theta'}\left(\theta^* - \hat{\theta}_0\right) + O\left(\left\|\theta^* - \hat{\theta}_0\right\|^2\right)
\end{aligned}
$$

- To get a first-order approximation to $\pi(\theta^*)$ we need to know $\frac{\partial \pi\left(\hat{\theta}_0\right)}{\partial \theta'}$.

Counterfactual Experiments: Homotopy method

- We know that $\pi\left(\hat{\theta}_0\right) = \Psi(\hat{\theta}_0, \hat{\mathbf{P}}_0)$, and this implies that:

$$\frac{\partial \pi\left(\hat{\theta}_0\right)}{\partial \theta'} = \left(I - \frac{\partial \Psi(\hat{\theta}_0, \hat{\mathbf{P}}_0)}{\partial \mathbf{P}'}\right)^{-1} \frac{\partial \Psi(\hat{\theta}_0, \hat{\mathbf{P}}_0)}{\partial \theta'}$$

- Then, $\pi(\theta^*) =$

$$\hat{\mathbf{P}}_0 + \left(I - \frac{\partial \Psi(\hat{\theta}_0, \hat{\mathbf{P}}_0)}{\partial \mathbf{P}'}\right)^{-1} \frac{\partial \Psi(\hat{\theta}_0, \hat{\mathbf{P}}_0)}{\partial \theta'} \left(\theta^* - \hat{\theta}_0\right) + O\left(\left\|\theta^* - \hat{\theta}_0\right\|^2\right)$$

- Therefore, $\hat{\mathbf{P}}_0 + \left(I - \frac{\partial \Psi(\hat{\theta}_0, \hat{\mathbf{P}}_0)}{\partial \mathbf{P}'}\right)^{-1} \frac{\partial \Psi(\hat{\theta}_0, \hat{\mathbf{P}}_0)}{\partial \theta'} \left(\theta^* - \hat{\theta}_0\right)$ is a first-order approximation to the counterfactual equilibrium $\mathbf{P}^*$.

Motivation

- This paper:
- 1. Proposes a **dynamic game of network competition** in the airline industry.
- 2. Proposes methods to solve, estimate, and perform counterfactual experiments using the model.
- 3. Uses the model to study empirically the role of **strategic entry deterrence** as a factor to explain why many companies in the US airline industry operate using **hub-and-spoke networks**.

Motivation: The Model

- Structural games of competition in the airline industry take into account the existence of **network effects**, but they treat them as **exogenous factors**.
- These models do not specify explicitly the links between the costs and benefits of an airline at different city-pairs.
- To answer important policy questions, we need to take into account these links and **to endogenize airlines networks**.
- We build on and extend the work of Hendricks et al (1995, 1999) to present a **dynamic game of airlines' network competition** that can be estimated using publicly available data.

Motivation: Methods

- By combining simplifying assumptions (decentralizing the decision problem; inclusive-values) and Monte Carlo simulation, we develop a method to solve and to estimate this dynamic game.

- We propose a method to implement counterfactual experiments using the estimated model and taking into account the existence of multiple equilibria in the model.

Motivation Empirical Application

- We study empirically the contribution of demand, cost, and strategic factors to explain why many companies in the US airline industry operate using **hub-and-spoke networks**.
- We place particular attention to the role of **strategic entry deterrence.**
- We estimate the model and use counterfactual experiments to obtain the contribution of different factors (and in particular of entry deterrence) to explain hub-and-spoke networks.

Hub-and-Spoke Networks

- What is a hub-and-spoke network?
- Since the deregulation of the US airline industry in 1978, most airlines have adopted network structures that concentrate their operation in a few airports.
- Southwest Airlines is an important exception: Point-to-Point network.
- **Strategic Entry Deterrence**: Hendricks, Piccione and Tan (1997).

'Hubbing' in the US Airline Industry: Year 2004

Airline (Code)	1st largest hub (# connections)	CR1	2nd largest hub (# connections)	CR2
Southwest (WN)	Las Vegas (35)	9.3	Phoenix (33)	18.2
American (AA)	Dallas (52)	22.3	Chicago (46)	42.0
United (UA)	Chicago (50)	25.1	Denver (41)	45.7
Delta (DL)	Atlanta (53)	26.7	Cincinnati (42)	48.0
Continental (CO)	Houston (52)	36.6	New York (45)	68.3
Northwest (NW)	Minneapolis (47)	25.6	Detroit (43)	49.2
US Airways (US)	Charlotte (35)	23.3	Philadelphia (33)	45.3

Model: Airlines, Cities, and Routes

- N airlines and C cities, exogenously given.
- Given the C cities, there are $M \equiv C(C-1)/2$ **non-directional city-pairs** (or markets).
- For each city-pair, an airline decides whether to operate non-stop flights.
- A **route** (or path) is a **directional round-trip between 2 cities**. A route may or may not have stops.
- A route-airline is a product, and there is a demand for each route-airline product.
- Airlines choose prices for each route they provide.

Model: Networks

- We index city-pairs by m, airlines by i, and time (quarters) by t.
- $x_{imt} \in \{0,1\}$ is a binary indicator for the event "airline i operates non-stop flights in city-pair m"
- $\mathbf{x}_{it} \equiv \{x_{imt} : m = 1, 2, ..., M\}$ is the network of airline i at period t.
- The network $\mathbf{x}_{it}$ describes all the routes (products) that the airline provides, and whether they are non-stop or stop routes.
- Industry network: $\mathbf{x}_t \equiv \{\mathbf{x}_{it} : i = 1, 2, ..., N\}$

Model: Airlines' Decisions

- An airline network $\mathbf{x}_{it}$ determines the set of routes that the airline provides, $L(\mathbf{x}_{it})$.
- Every period, active airlines in a route compete in prices
- Price competition determines variable profits for each airline.
- Every period (quarter), each airline decides its network for next period. There is *time-to-build*.
- We represent this decision as $\mathbf{a}_{it} \equiv \{a_{imt} : m = 1, 2, ..., M\}$, though $a_{imt} \equiv x_{imt+1}$.

Model: Profit Function

- The airline's total profit function is:

$$\begin{aligned} \Pi_{it} &= \sum_{r \in L(\mathbf{x}_{it})} (p_{irt} - c_{irt}) q_{irt} \\ &- \sum_{m=1}^{M} a_{imt} \left(FC_{imt} + (1 - x_{imt}) \; EC_{imt} \right) \end{aligned}$$

- $(p_{irt} - c_{irt})q_{irt}$ = Variable profit in route r.
- FC_{imt} and EC_{imt} are fixed cost and entry cost

Model Network effects in demand and costs

- An important feature of the model is that demand, variable costs, fixed costs, and entry costs depend on the scale of operation (number of connections) of the airline in the origin and destination airports of the city-pair.

- For instance,

$$\begin{aligned} FC_{imt} &= \gamma_1^{FC} + \gamma_2^{FC}\ HUB_{imt} + \gamma_3^{FC} DIST_m + \gamma_{4i}^{FC} + \gamma_{5c}^{FC} \\ EC_{imt} &= \eta_1^{EC} + \eta_2^{EC}\ HUB_{imt} + \eta_3^{EC} DIST_m + \eta_{4i}^{EC} + \eta_{5c}^{EC} \end{aligned}$$

- This implies that markets are interconnected through these hub-size effects. Entry-exit in a market has implications of profits in other markets.

Dynamic Game / Strategy Functions

- Airlines maximize intertemporal profits, are forward-looking, and take into account the implications of their entry-exit decisions on future profits and on the expected future reaction of competitors.
- Airlines' strategies depend only on payoff-relevant state variables, i.e., Markov perfect equilibrium assumption.
- An airline's payoff-relevant information at quarter t is $\{\mathbf{x}_t, \mathbf{z}_t, \varepsilon_{it}\}$.
- Let $\sigma \equiv \{\sigma_i(\mathbf{x}_t, \mathbf{z}_t, \varepsilon_{it}) : i = 1, 2, ..., N\}$ be a set of strategy functions, one for each airline.
- A MPE is a set of strategy functions such that each airline's strategy maximizes the value of the airline for each possible state and taking as given other airlines' strategies.

Dynamic Game: Reducing the dimensionality

- Given the number of cities and airlines in our empirical analysis, the number of possible industry networks is $|X| = 2^{NM} \simeq 10^{10,000}$.
- We consider **two types of simplifying assumptions** that reduce the dimension of the dynamic game and make its solution and estimation manageable.
- 1. An **airline's choice of network is decentralized** in terms of the separate decisions of local managers.
- 2. The state variables of the model can be aggregated in a vector of **inclusive-values** that belongs to a space with a much smaller dimension than the original state space.

Decentralizing the Airline's Choice of Network

- Each airline has M local managers, one for each city-pair.
- A local manager decides whether to operate or not non-stop flights in his local-market: i.e., he chooses a_{imt}.
- Let R_{imt} be the sum of airline i's variable profits over all the routes that include city-pair m as a segment.
- **ASSUMPTION:** *Local managers maximize the expected and discounted value of*

$$\Pi_{imt} \equiv R_{imt} - a_{imt} \left(FC_{imt} + (1 - x_{imt}) EC_{imt} \right).$$

- IMPORTANT: A local manager internalizes the effects of his own entry-exit decision in many other routes. Entry deterrence.

Inclusive-Values

- Decentralization of the decision simplifies the computation of players' best responses, but the state space of the decision problem of a local manager is still huge.

- Notice that the profit of a local manager depends only on the state variables:

$$\mathbf{x}_{imt}^{*} \equiv \left(x_{imt}, \;\; R_{imt}, \;\; HUB_{imt} \right)$$

- **ASSUMPTION:** The *vector* $\mathbf{x}_{imt}^{*}$ *follows a controlled first-order Markov Process:*

$$\Pr\left(\mathbf{x}_{im,t+1}^{*} \mid \mathbf{x}_{imt}^{*}, a_{imt}, \mathbf{x}_{t}, \mathbf{z}_{t} \right) = \Pr\left(\mathbf{x}_{im,t+1}^{*} \mid \mathbf{x}_{imt}^{*}, a_{imt} \right)$$

Dynamic Game: Reducing the dimensionality

- A MPE of this game can be describe as a vector of probability functions, one for each local-manager:

$$P_{im}(\mathbf{x}_{imt}^{*}) : i = 1, 2, ..., N; \;\; m = 1, 2, ..., M$$

- $P_{im}(\mathbf{x}_{imt}^{*})$ is the probability that local-manager (i, m) decides to be active in city-pair m given the state $\mathbf{x}_{imt}^{*}$.
- An equilibrium exits.
- The model typically has multiple equilibria.

Data

- Airline Origin and Destination Survey (DB1B) collected by the Office of Airline Information of the BTS.
- Period 2004-Q1 to 2004-Q4.
- $C = 55$ largest metropolitan areas. $N = 22$ airlines.
- City Pairs: $M = (55 * 54)/2 = 1,485$.

Airlines: Passengers and Markets

	Airline (Code)	# Passengers (in thousands)	# City-Pairs (max = 1,485)
1.	Southwest (WN)	25,026	373
2.	American (AA)[3]	20,064	233
3.	United (UA)[4]	15,851	199
4.	Delta (DL)[5]	14,402	198
5.	Continental (CO)[6]	10,084	142
6.	Northwest (NW)[7]	9,517	183
7.	US Airways (US)	7,515	150
8.	America West (HP)[8]	6,745	113
9.	Alaska (AS)	3,886	32
10.	ATA (TZ)	2,608	33
11.	JetBlue (B6)	2,458	22

Distribution of City-Pairs by # Airlines with non-stop flights	
Markets with 0 airlines	35.44%
Markets with 1 airline	29.06%
Markets with 2 airlines	17.44%
Markets with 3 airlines	9.84%
Markets with 4 or more airlines	8.22%

Number of Monopoly Markets by Airline	
Southwest	157
Northwest	69
Delta	56
American	28
Continental	24
United	17

Entry and Exit	
	All Quarters
Distribution of Markets by Number of New Entrants	
Markets with 0 Entrants	84.66%
Markets with 1 Entrant	13.37%
Markets with 2 Entrants	1.69%
Markets with 3 Entrants	0.27%
Distribution of Markets by Number of Exits	
Markets with 0 Exits	86.51%
Markets with 1 Exit	11.82%
Markets with 2 Exits	1.35%
Markets with more 3 or 4 Exits	0.32%

Estimation of the Structural Model

- Our estimation approach proceeds in three stages.

1. **Estimation of demand system.** IV estimation (a la BLP) where the IV's are the competitors' hub-sizes.
2. **Estimation of marginal cost functions.**
3. **Estimation of dynamic game of entry-exit.** Nested Pseudo Likelihood (NPL) method.

Estimation of Dynamic Game of Entry-Exit		
Data: 1,485 markets $\times$ 22 airlines $\times$ 3 quarters $=$ 98,010 observations		
	Estimate (in thousand $)	**(Std. Error)**
Fixed Costs (quarterly):		
Fixed cost (average)	119.15	(5.233)
Effect of hub-size on FC	-1.02	(0.185)
Effect of distance on FC	4.04	(0.317)
Entry Costs:		
Entry cost (average)	249.56	(6.504)
Effect of hub-size on EC	-9.26	(0.140)

Counterfactual Experiments

		Zero Hub-Size Effects in:			
Carrier	Observed	var. profits	fixed costs	entry costs	No entry-deter
Southwest	18.2	17.3	15.6	8.9	16.0
American	42.0	39.1	36.5	17.6	29.8
United	45.7	42.5	39.3	17.8	32.0
Delta	48.0	43.7	34.0	18.7	**25.0**
Continental	68.3	62.1	58.0	27.3	43.0
Northwest	49.2	44.3	36.9	18.7	**26.6**
US Airways	45.3	41.7	39.0	18.1	34.4

Summary of empirical results

1. Hub-size effects on **demand, variable costs and fixed operating costs** are significant but can **explain very little of the propensity to hub-spoke networks**.
2. **Hub-size effects on Sunk Entry Costs are large**. This is the most important factor to explain hub-spoke networks.
3. **Strategic factors: hub-spoke network as a strategy to deter entry** is the second most important factor for some of the largest carriers (Northwest and Delta).
4. **Sunk Entry Costs are positively with Entry Deterrence.** Airlines with larger entry costs tend to have higher propensity to use hub-and-spoke networks to deter entry of competitors.

Ongoing Research

- Though the estimated provides a very good fit to the Quarterly data for 2004, it has limitations to explain some important features for the evolution of airline networks for the period 1990-2008.

- For instance, some airlines experienced very abrupt changes in their network structure: e.g., closings of large hubs, almost 'instantaneous birth' of new hubs.

- Explaining the transition of some airlines from almost pure hub-and-spoke networks to point-to-point networks is also challenging.

- Economic interpretation of the negative effect of hub-size on entry costs: is it due to technological reasons, OR it has to do with contracts between airports and airlines ???