Estimating Static Discrete Games

Paul B. Ellickson[1]

[1]Simon School of Business Administration
University of Rochester

August 2013

Introduction

- Marketing concerns understanding, predicting, and influencing various agent's choices & behaviors
  - Consumers: what & how much to buy, where & when to shop, whom to emulate & interact with
  - Firms: what to sell & how much to charge, promotion, placement, positioning, when to introduce new products
- These decisions aren't made in a vacuum, but depend on the actions of others
  - Consumers are influenced by their peers (social interactions, social media)
  - Firms are constrained by the reactions of their rivals
- Accounting for strategic *interactions* makes these strategic *games*
- These two sessions are an introduction to estimating such games

Introduction (cont'd)

- The structure of the game depends on what's being modelled
  - Decisions may be discrete or continuous
    - What car to sell vs. what price to charge
  - Payoffs may be observed or latent
    - Do you have data on P, Q and C, or just choices?
  - Information may be complete or incomplete
    - Do players observe everything, or is there uncertainty?
  - It may be a one-shot game, or it may continue for many periods
    - Do today's choices impact tomorrow's payoffs?
- I'll focus on static discrete games, but consider both information structures.
  - Slides follow Ellickson & Misra "Estimating Discrete Games" (2011)

Working Example: Entry by Discount Stores

- Let's start with a concrete example
  - Consider an entry game between Walmart and Kmart.
- Entry is a discrete choice
  - Let's assume they compete in local markets (e.g. small towns) where they can build at most one store (big assumption!)
  - Let's also assume that these decisions are made simultaneously, once and for all
  - We have data on choices, but not payoffs
- We can revisit some of these assumptions later..

Working Example: Entry by Discount Stores

- Strategic entry is a static discrete game (think long run equilibrium..)
  - K & W choose either *enter* or *don't enter*.
    - The smallest markets can't support any stores.
    - Larger markets can support one.
    - Largest markets can support both.
- We are interested understanding who enters which markets and why.
- We might then
  - Recover structural parameters (determine what drives profit)
  - Solve for counterfactual outcomes
    - Introduce a new product
    - Start a social media campaign

Notation

- Profit of firm $i = \{K, W\}$ in market m is $\pi_{im}(\theta; y_{-im}, X_m, Z_m)$ where
  - y_{im} is the action (enter/don't enter) of firm i,
  - y_{-im} is the action of its rival,
  - X_m is a vector of market characteristics,
  - $Z_m = (Z_{Km}, Z_{Wm})$ contains firm characteristics, and
  - θ is a finite-dimensional parameter vector.
- Note that these are *latent* profits, like *utility* in DC models
  - If these were consumers, we'd work with utilities
- Let's choose a simple functional form

$$\pi_{im} = \alpha_i' X_m + \beta_i' Z_{im} + \delta_i y_{-im} + \varepsilon_{im}$$

where ε_{im} is a component of profits the firm sees but we don't.

Information Structure

- We need to decide what the players *do* and *do not* observe.
  - If we assume that both firms see X_m, Z_m **and** $(\varepsilon_{Km}, \varepsilon_{Wm})$, this becomes a game of *complete information*.
  - If, instead, we assume that firms do not see some components of these profit shifters, this becomes a game of *incomplete information*.
- Let's start with the complete information case, following Bresnahan and Reiss (1990, 1991) and Berry (1992) who pioneered the empirical games literature.
- Then we'll discuss the incomplete information case (Rust, 1994), which segues nicely into the treatment of dynamics (session 2).

Nash Equilibrium (Simultaneous Moves)

- In equilibrium, firms maximize profits, taking rivals actions as given.
- A Nash equilibrium is characterized by

$$\begin{aligned} y_{Km} &= 1\left[\alpha_K' X_m + \beta_K' Z_{Km} + \delta_K y_{Wm} + \varepsilon_{Km} \geq 0\right] \\ y_{Wm} &= 1\left[\alpha_W' X_m + \beta_W' Z_{Wm} + \delta_W y_{Km} + \varepsilon_{Wm} \geq 0\right] \end{aligned}$$

  which captures each firm's non-negative profit condition.
- Note that these could just as easily be utilities in a system of social interactions.
- An equilibrium is a configuration that satisfies both equations.

Nash Equilibrium (Simultaneous Moves)

$$\begin{aligned} y_{Km} &= 1\left[\alpha_K' X_m + \beta_K' Z_{Km} + \delta_K y_{Wm} + \varepsilon_{Km} \geq 0\right] \\ y_{Wm} &= 1\left[\alpha_W' X_m + \beta_W' Z_{Wm} + \delta_W y_{Km} + \varepsilon_{Wm} \geq 0\right] \end{aligned}$$

- This is now a binary simultaneous equation system.
- This structure distinguishes a discrete *game* from a standard discrete *choice* problem.
  - The outcomes are determined via equilibrium conditions
  - The RHS of each equation contains a dummy endogenous variable
- To proceed, we must confront this endogeneity problem

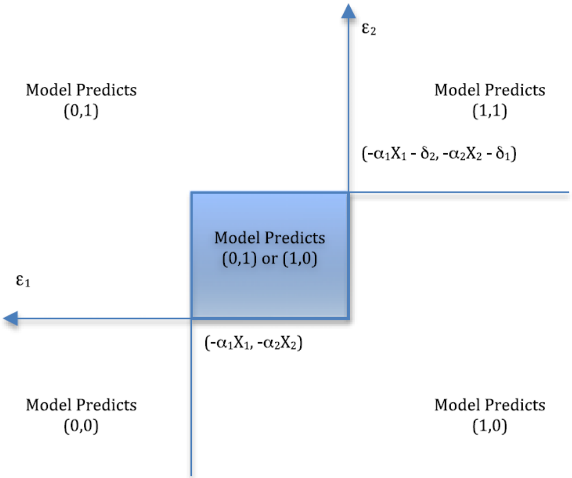
Multiple Equilibrium

- Why not solve the system for its reduced form & match it to what we see in the data?
- Unfortunately, in many cases, the reduced form won't be unique
  - This is a direct result of this being a game $\rightarrow$ games may admit more than one equilibria!
- For example, if the δ's are < 0 (competition reduces profits), multiple equilibria arise in the region of ε space for which

$$-\left(\alpha_i' X + \beta_i' Z_i\right) \leq \varepsilon_i \leq -\left(\alpha_i' X + \beta_i' Z_i\right) - \delta_{3-i} \text{ for } i = 1, 2$$

  - Let's look at a picture from Bresnahan & Reiss...
- Thus, for a given set of parameters there may be more than one possible vector of equilibrium outcomes (y).

Multiple Equilibrium



Incompleteness due to multiplicity

- The center (shaded) box is the problem region
  - The model does not yield a unique prediction
- Multiplicity makes the econometric model **incomplete**
  - A **complete** econometric response model asserts that a random variable y is a *function* of a random pair (x, ε) where x is observable and ε is not. (Manski, 1988)
  - An **incomplete** econometric model is one where the relationship from (x, ε) to y is a *correspondence*. (Tamer, 2003)
    - Other examples: selection, censoring
- Key issue: incompleteness makes it difficult to define (simple) probability statements about players' actions.
  - To proceed, we must complete the model somehow (or forgo simple probability statements).

Restoring econometric completeness

- What are our options for completing the model?
- There are four main approaches in the literature
  1. Aggregate to a prediction which is unique (e.g. the number of entrants)
  2. Impose additional assumptions to guarantee uniqueness (e.g. sequential moves)
  3. Specify an equilibrium selection rule (e.g. most profitable)
  4. Employ a method that can handle non-uniqueness (e.g. set inference)
- I'll discuss the first 3
  - Option 4 is a frontier technique (beyond our scope)
  - To start, see Tamer (2003)
- The original papers to read are Bresnahan & Reiss (1990, 1991), Berry (1992), and Bjorn & Vuong (1985)

Implementation

- Let's go through option 1 in detail, and then sketch out options 2 & 3.
- To fix ideas, consider Walmart & Kmart again, letting profits be just

$$\pi_{im} = \alpha' X_m - \delta y_{-im} + \varepsilon_{im}$$

- X_m might include things like population, income & retail sales.
- We are ignoring the firm characteristics (Z's) in this case.
  - This raises problems for (non-parametric) identification
    - Ideally, we'd like *exclusion* restrictions: covariates that shift around each firm's profits separately from its rivals
    - If not, we are relying on *functional form*
- The ε_{im}'s are i.i.d. shocks, distributed $N(0, 1)$ perhaps.
  - We can relax this later..

Likelihood Function (Aggregation)

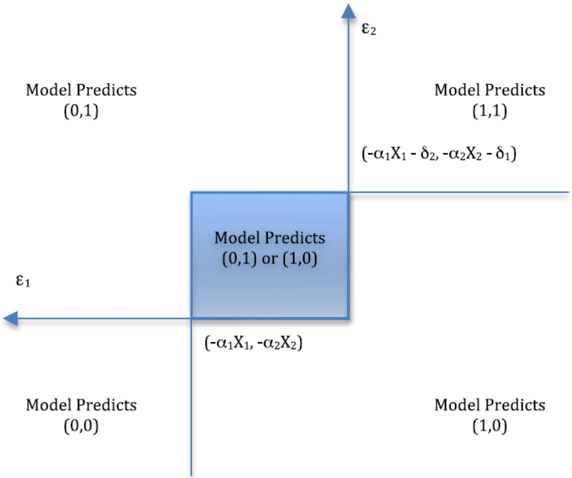
- For option 1, we aggregate to a unique prediction.
  - We predict *how many* firms enter, not *who* enters.
- Given this structure, the likelihood of observing n_m firms in a given market m can be computed in closed form

$$
\begin{aligned}
\Pr(n_m = 2) &= \prod_i \Pr\left(\alpha' X_m - \delta y_{-im} + \varepsilon_{im} \geq 0\right) \\
\Pr(n_m = 0) &= \prod_i \Pr\left(\alpha' X_m - \delta y_{-im} + \varepsilon_{im} < 0\right) \\
\Pr(n_m = 1) &= 1 - \Pr(n_m = 2) - \Pr(n_m = 0)
\end{aligned}
$$

- The sample log-likelihood is then

$$
\ln \mathcal{L} = \sum_{m=1}^{M} \sum_{l=0}^{2} \mathcal{I}(n_m = l) \ln \Pr(n_m = l).
$$

Multiple Equilibrium



Limitations

- Aggregation clearly involves the loss of some information (Tamer, 2003)
- It can also require strong assumptions to generalize (e.g. to many players and/or player types).
- Not clear how to apply it to mixed strategies or the incomplete information case.
- Also, trickier to see where identification comes from.
- However, it's closest in spirit to the set inference approach...

Alternatives (Options 2 & 3)

- Berry (1992) completes the model by assuming sequential entry.
  - Assigns all "contested" outcomes to a single firm (e.g. Walmart)
  - Has the advantage of being scalable (Berry allows up to 26 entrants) via simulation
  - Mazzeo (2002) is a nice example
- The third option is to provide a selection rule: a way to "select" amongst many equilibria
  - Proposed by Bjorn & Vuong (1985), extended/formalized by Bajari, Hong & Ryan (2010).
  - Simple example: assign probability π and $1 - \pi$ to the two monopoly outcomes and estimate π as part of overall likelihood (mixture).
  - Issues: finding all equilibria, somewhat ad hoc..
- EM (MS, 2011) provide sample code for options 1 and 2...

Extensions

- Heterogeneity & Bayesian approaches
  - 'Full information' structure facilitates both
    - Hartmann (2010) & Narayanan (2013)
- "Post-entry" data
  - Can bring in data on prices, quantities, etc.
  - Key challenge: accounting for selection
    - Mazzeo (2002), Singh & Zhu (2008), Zhu et al. (2009)
- Multiple discreteness/networks
  - High dimensional structure yields small probabilities (& steep computational burden)
  - Can exploit 'profit inequalities' instead
    - Jia(2008), Pakes et al. (2005), Ellickson et al. (2013)

Incomplete Information Setting

- So far, we've assumed firms know everything about everyone
  - Realistic if the market is in *long run* equilibrium
- A second class of models instead assumes that firms have some private information
  - They can no longer perfectly predict each other's actions
    - They must form *expectations* over what their rivals will do
  - They may then have "regret" (if they predict wrong)
- Whether this is a more or less reasonable assumption than complete information is a matter of debate, but it does ease the computational burden considerably
  - Might test using Grieco (2013)

Incomplete Information Setting

- Incomplete information allows us to recast decision problem as a collection of 'games against nature'
  - Similar role to conditional independence in DDC setting...
  - ...or IPVs in auction setting
- This simplifies estimation and provides an additional option for dealing with multiplicity: two-step estimation
- It also readily extends to dynamic games, where the uncertainty is more intuitive
- Let's see how the information assumption changes the set up...

Incomplete Information Setting

- Under incomplete information, player's no longer observe everything about their rivals.
- To fix ideas, let's assume that each player observes its own ε_i, but only knows the *distribution* of ε_j for its rivals.
  - Suppose **we** also know the distribution, but don't see actual draws.
- Each firm now forms expectations about its rivals' behavior, choosing the action that maximizes *expected* profits *given those beliefs*.
  - The equilibrium concept is now Bayes Nash.
- This yields a new system of equations

$$\begin{aligned} y_{Km} &= 1\left[\alpha_K' X_m + \beta_K' Z_{Km} + \delta_K p_W + \varepsilon_{Km} \geq 0\right] \\ y_{Wm} &= 1\left[\alpha_W' X_m + \beta_W' Z_{Wm} + \delta_W p_K + \varepsilon_{Wm} \geq 0\right] \end{aligned}$$

where $p_i \equiv E_i\left(y_{-i}\right)$ captures firm i's beliefs.

Bayesian Nash equilibrium (BNE)

- The BNE satisfies the following set of *equalities*

$$
\begin{aligned}
p_K &= \Phi_K(\alpha_K' X_m + \beta_K' Z_{Km} + \delta_K p_W) \\
p_W &= \Phi_W(\alpha_W' X_m + \beta_W' Z_{Km} + \delta_W p_K)
\end{aligned}
$$

  where the form of Φ depends on the distribution of ε.
  - For example, if ε's are $N(0,1)$, it's the standard Normal CDF.
- The $\Phi(\cdot)$'s are *best response probability functions*, mapping expected profits (conditional on beliefs p) into (ex ante) choice probabilities.
  - An equilibrium is a *fixed point* of these equations
    - Existence follows directly from Brouwer's FPT
  - Once again, it will likely be non-unique
    - Simple example: coordination game

Multiplicity Again

- Incompleteness arise here as well (there can still be more than one outcome associated with a given set of parameters)
  - Options 2 and 3 still apply
    - See, e.g., Einav (2010) or Sweeting (2009)
  - However, there is now a new "solution"
    - Use two-step estimation to 'condition on the equilibrium that's played in the data'
- To see how it works, let's start with a simpler case that's closer to what we've already seen...

Nested Fixed Point Approach

- Let's consider a 'full-solution' approach like we've seen so far
  - Suppose we *know* there is only one equilibrium!
  - The fixed point representation provides a direct method of solving for it: successive approximation
    - Or you can make it a constraint (MPEC)...
    - ...see Su & Judd (2012)
  - This fixed point calculation yields "reduced form" choice probabilities (CCPs) which you can then match to the data
- This estimation approach is called nested fixed point (NFXP)
  - The fixed point (equilibrium) calculation is nested inside the likelihood
  - Developed for DDC problems by Rust (1987)
  - Applied to discrete games by Seim (2006)

Implementation (NFXP)

- **NFXP approach:** Consider the same profit function as before

$$\pi_{im} = \alpha' X_m - \delta y_{-im} + \varepsilon_{im}$$

  and include the same covariates.
- The ε's are still *iid* $N(0,1)$, but private information now.
- The estimation routine requires solving the fixed point problem

$$p^*_{im} = \Phi(\alpha' X_m - \delta p^*_{-im})$$

  for a given guess of the parameter vector.
- The resulting probabilities feed into a binomial log likelihood

$$\ln \mathcal{L} = \sum_{m=1}^{M} \sum_{i \in \{W,K\}} y_{im} \ln(p^*_{im}) + (1 - y_{im}) \ln(1 - p^*_{im})$$

  which is then maximized to obtain parameter estimates.

Two-step Estimation

- A problem with the NFXP approach is that it *assumes* uniqueness
  - If there's more than one equilibrium, it's mis-specified
- To complete the model: either assume an order of entry (ensuring uniqueness) or find all the equilibria and impose a selection rule.
  - However, there is now another option
- Suppose we have *estimates* of each player's beliefs $\hat{p}_i \equiv \hat{E}_i\left(y_{-i}\right)$
  - Can either get (non-parametrically) from data or (potentially) from an auxiliary survey...
- Idea comes from Hotz & Miller (1993) and Rust (1994)

Two-step Estimation

- Simply plug these estimates into the RHS of

$$p^*_{im} = \Phi(\alpha' X_m - \delta p^*_{-im})$$

  revealing a standard discrete choice problem!
    - This could be done in Excel!!!
- We have also "solved" the multiplicity problem by selecting the equilibrium that was *played in the data*.
    - Assuming there is only one...
    - ...most realistic if we see the same market over time...
- If we have collected 'expectations' data (e.g. a survey), we can even relax the (implicit) rational expectations assumption...

Small sample noise (a hidden curse of dimensionality)

- Back to the rational expectations case...
- In principal, the first stage must be done non-parametrically
  - Why? The "descriptive" CCPs are not economic primitives!
    - We cannot impose structure on them
  - Also makes Bayesian approaches more challenging here...
- In practice, nonparametric models will be noisy (or infeasible) and parametric models will be mis-specified (& thus inconsistent).

Nested Pseudo Likelihood

- A clever fix is to use an iterative approach: treat the fitted probabilities from the second stage as new first stage beliefs
  - Continue (iterate!) until the probabilities no longer change
- This is called the Nested Pseudo Likelihood approach
  - Developed by Aguirregabiria & Mira (2002, 2007) in DDC context
  - It reduces small sample bias and eliminates the need for a consistent first stage
    - Can also allow unobserved heterogeneity...
    - ...but it's not guaranteed to converge
- EM provide code for three incomplete information approaches

Extensions

- Unobserved heterogeneity
  - Simplest with NPL or NFXP (imposes structure)
  - Ellickson & Misra (2008), Sweeting (2009), Orhun (2013)
- Bayesian methods
  - Misra (2013)
- Post-entry data (e.g. revenues, prices)
  - Draganska et al. (2009), Ellickson & Misra (2013)

Looking forward to dynamics...

- The (two-step) incomplete information setting provides a natural segue to dynamic games.
    - Dynamic games pose a doubly nested fixed point problem!
        - A dynamic programming problem coupled with the overall game
- Earlier, by estimating beliefs in a first stage, we reduced a complex strategic interaction to a collection of games against nature.
    - This bypassed the fixed point problem associated with the game
- A similar trick can be used to handle the 'future': invert CCPs to recover (differenced) choice specific value functions
    - This eliminates the other fixed point problem (from DP problem)
- There are several methods for doing so, which Sanjog will discuss next!