1 Introduction

- Estimation of games has become an important topic for research in IO.
- Here we are going to focus on discrete choice simultaneous move games (more popular than sequential move).

- Basic structure:
  - **Choices:** A finite number of firms compete in the market and simultaneously decide whether to entry or to stay out of the market, whether to adopt or not adopt a given technology, etc. taken as given other firms’ best responses.
  - **Information:** Firms have complete or incomplete information about other firms’ payoffs.
  - **Time horizon:** Static or dynamic.
  - **Equilibrium:** Define the type of the equilibrium used to solve the game (Nash, Bayesian, Markovian).
  - **Data:** Researchers observe actions taken by the firms, market characteristics and, possibly, firms’ characteristics and assume that firms decisions are taken based on the theoretical model previously designed.
  - **Goal:** Theoretical model (choices, information structure, time horizon) provides us with structure that is used to identify and estimate firms’ payoffs; use the estimated payoffs to solve the model (find equilibrium decisions) and make counterfactuals.

- Main advantages compared to reduced form stuff:
  1. Estimates have a clear economic interpretation.
  2. Essentially, we can use the model to predict the effects of changes in the market (mergers and acquisitions, mandatory technological changes, privatization, etc.) on market outcomes (welfare).
- Main disadvantages compared to reduced form stuff:

  1. Computational burden: Estimation algorithms (in general) are not canned in Stata; estimation can be very time consuming.

  2. Identification (and estimation) heavily relies on theoretical assumptions.
2 Some Literature

- Structural versus reduced form:
  - Reduced form “manifesto”: Angrist and Pischke (JEP, 2010).
  - Structural “counter attack”: Levin and Einav (JEP, 2010).

- Relevant papers discussed today:
  - Estimation of static games of complete information: Bresnahan and Reiss (1991, JoE) and Bresnahan and Reiss (1991, JPE).
  - Estimation of static games of incomplete information: Seim (Rand, 2006), Bajari, Hong, Krainer and Nekipelov (JBES, 2010) and Pesendorfer and Schmidt-Dengler (REStud, 2008).

- For recent surveys you can check:
  - Static games: Berry and Reiss (Handbook of IO, 2007).
3 Part I: Static Games of Complete Information

- We are going to start with a very simple structure, developed initially in Bresnahan and Reiss (1990), and address the main issue involved in the estimation of discrete games, namely, the existence of multiple equilibria. The games analyzed here are simple 2x2 games (two players and two actions). Although very stylized (which reduces their applicability) these simple games are useful to understand the spirit of problem.

- We analyze the solution proposed by BR (1990) to identify the parameters of the model - identification here is very linked to the existence of multiple equilibria.

- In the end we generalize the model to include a finite number of players.

3.1 Model and Identification

- Bresnahan and Reiss (1990) start with a discrete game of complete information. In the discrete game, actions are discrete, i.e., for example, in this game, players choose either to enter in the market or not.

- The idea is the following: Researchers observe equilibrium actions (for example if the firm entered or not in the market, if the firm adopted or not a given technology, etc.) but do not observe players’ payoffs. BR (1990) develops some structure that allows us to recover players’ payoffs from observed (equilibrium) actions.

3.1.1 Theoretical model

- Let’s consider the following $2 \times 2$ simultaneous entry game.

- There are two players indexed by $i \in \{1, 2\}$.

- Players decide simultaneously to choose either $a_i = 1$ or $a_i = 0$ (enter or stay out of the market, respectively).

- Payoffs are described by:

$$u_i(a_i, a_{-i}, x_i, \varepsilon_i; \Theta) = \begin{cases} \alpha \left( \sum_{j \neq i} a_j \right) + \beta x_i - \varepsilon_i \end{cases} a_i$$  \hspace{1cm} (1)

- Firms observe $\varepsilon$’s; econometricians do not observe $\varepsilon$’s. For the econometrician, $\varepsilon_i$’s are iid random draws from a given distribution with continuous and unbounded support.

- $x_i$ is a given vector of observed variables and $\Theta = (\alpha, \beta)$ are parameters to be estimated.
Write a given profile strategy as \( a = (a_1, a_2) \), with \( a_1, a_2 \in \{0, 1\} \) and the set of all possible strategies \( A = \times_{i=1,2} a_i = \{0, 1\} \times \{0, 1\} \) or, simply, \( A = \{(0, 0); (0, 1); (1, 0); (1, 1)\} \).

Write player \( j \)'s best response to player \( i \)'s actions \( i \neq j \) as \( B_j(a_i) \) and \( B(a) = B_1(a_2) \times B_2(a_1) \).

Remember that the profile of strategies \( a^* \) is a Nash Equilibrium in pure strategies if \( a^* \in B(a^*) \).

The conditions for the Nash Equilibrium in pure strategies are:

- \((0, 0)\) is a N.E. iff \((0, 0) \in B ((0, 0)) = B_1 (0) \times B_2 (0)\). This happens if \( B_1 (0) = 0 \) and \( B_2 (0) = 0 \) or \( \beta x_1 - \varepsilon_1 < 0 \) for \( i = 1, 2 \).

- \((0, 1)\) is a N.E. iff \((0, 1) \in B ((0, 1)) = B_1 (1) \times B_2 (0)\). This happens if \( B_1 (1) = 0 \) and \( B_2 (0) = 1 \) or \( \alpha + \beta x_1 - \varepsilon_1 < 0 \) and \( \beta x_2 - \varepsilon_2 \geq 0 \).

- \((1, 0)\) is a N.E. iff \((1, 0) \in B ((1, 0)) = B_1 (0) \times B_2 (1)\). This happens if \( B_1 (0) = 1 \) and \( B_2 (1) = 0 \) or \( \beta x_1 - \varepsilon_1 \geq 0 \) and \( \alpha + \beta x_2 - \varepsilon_2 < 0 \).

- \((1, 1)\) is a N.E. iff \((1, 1) \in B ((1, 1)) = B_1 (1) \times B_2 (1)\). This happens if \( B_1 (1) = 1 \) and \( B_2 (1) = 1 \) or \( \alpha + \beta x_1 - \varepsilon_1 \geq 0 \) for \( i = 1, 2 \).

### 3.1.2 Econometric model and identification issues

- Suppose that the econometrician observes a random sample of equilibrium actions (outcomes) and covariates \( \{a^*_m, x_{1m}, x_{2m}\}_{m=1}^M \), where \( m = 1, 2, ..., M \) is a given number of markets (in some applications instead of repetition of this game across markets the econometrician observes a repetition across time or across time and markets). Econometrician’s objective is to use the structure above and the observed data to estimate \( \Theta = \{\alpha, \beta\} \). The \( \beta \)'s can be understood as the effect of demand and/or cost shifters on players’ payoffs and the \( \alpha \)'s can be understood as the effect of competition (or the entry of a new firm) on incumbent’s profits.

- BR (1990) develop an econometric framework to estimate the primitives of this games by noting that under certain conditions there exists a function mapping payoffs into the set of equilibria. Let’s call this function as \( f (\cdot) : \{x_{1m}, x_{2m}, \varepsilon_{1m}, \varepsilon_{2m}; \Theta\} \rightarrow A^*_m \), where \( A^*_m \) is the set of possible equilibrium actions in market \( m \). There are however some relevant issues...

- The first is that for certain values of parameters and covariates, the game above can have more than one equilibrium; the second is that for another set of payoff values, the game will not have any equilibrium.

  - To see this, firstly fix \( \{x_{1m}, x_{2m}; \beta\} \) and \( \alpha < 0 \). Now if we vary \( (\varepsilon_{1m}, \varepsilon_{2m}) \) over the unbounded support of this vector we can observe...
different equilibrium configurations. Multiple equilibrium arises because for a certain range of \((\varepsilon_{1m}, \varepsilon_{2m})\) we have two possible equilibrium configurations. Have a look on the picture below.

Figure 1: Simultaneous Move Entry Game: Case 1 \((\alpha < 0)\)

- That is, for \(\{x_{1m}, x_{2m}; \beta\}\) and \(\alpha < 0\) for certain values of \((\varepsilon_{1m}, \varepsilon_{2m})\) - see the orange box in the middle of the figure - we can observe either \((0, 1)\) or \((1, 0)\) as the possible equilibrium configuration. As we are going to see this pattern creates identification problems.
- Now, fix \(\{x_{1m}, x_{2m}; \beta\}\) and \(\alpha > 0\). We have the same problem. Have a look on the following figure:
Therefore, when $\alpha > 0$, if $(\varepsilon_{1m}, \varepsilon_{2m})$ is in the orange box we can have one equilibrium in pure strategies where both firms enter or an equilibrium where neither firm 1, nor firm 2 enters.

- **Question:** What are the implications of multiple equilibria for the identification of $\Theta$?

- **Let’s start with the case $\alpha < 0$.**

  - The existence of a multiple equilibrium region creates a well known identification problem - i.e. a problem to identify $\Theta$ given that the econometrician observes $\{a^*_m, x_{1m}, x_{2m}\}_{m=1}^M$. To see this define $p(a^*_m|x_{1m}, x_{2m}; \Theta)$ as the probability that equilibrium $a^*_m \in A^*_m$ is played in market $m$ given the set of covariates for the market and the parameters of interest. Notice now that $p(\cdot|x_{1m}, x_{2m}; \Theta)$ is a well defined probability function if $\sum_{a^* \in A^*} p(a^*|x_{1m}, x_{2m}; \Theta) = 1$ for all vectors of covariates and parameters. If we find this probability function we can use the NE conditions defined above to write down a likelihood function for the equilibrium actions and estimate $\Theta$.

  - Can we always define such a probability function in the presence of multiple equilibria? The answer for this question is that if the game has multiple equilibria $p(a^*|x_{1m}, x_{2m}; \Theta)$ cannot be defined and it is not possible to identify $\Theta$. To see this write:
\[ p (a_m^* = (0, 0)|x_1, x_2; \Theta) = \]
\[ p (\varepsilon_1 > x_1, \varepsilon_2 > x_2; \Theta) \]
\[ p (a_m^* = (1, 1)|x_1, x_2; \Theta) = \]
\[ p (\varepsilon_1 \leq \alpha + \beta x_1, \varepsilon_2 \leq \alpha + \beta x_2; \Theta) \]

- Using independence and normality of the \( \varepsilon \), it is easy to get a closed form for these expressions. Notice that to define the probability that the equilibrium \((0, 0)\) is played we are integrating the area “no entrants” in the figure above; we are doing the same for \((1, 1)\).

- The multiple equilibrium problem appears when we want to use the same procedure to define the probabilities of \((1, 0)\) and \((0, 1)\). Because the orange area in the figure is counted twice (one to define the probability of \((1, 0)\) and the other to define the probability of \((0, 1)\)) then we have that \( \sum_{a^* \in A} p (a^*|x_1, x_2; \Theta) > 1 \) and we cannot define the likelihood function for the equilibrium actions.

- The same happens when \( \alpha > 0 \).

### 3.1.3 An Identified Econometric Model

- In principle, as we observed, this simple game cannot be estimated or, in other words, we cannot identify \( \Theta \) using pure strategies in a game of complete information.

- BR (1990) suggests however a clever solution for the identification problem. The solution is based on two identification assumption:

1. \( \alpha < 0 \) is the relevant case. This is true for most models of competition and simply says that entrants reduce the profits of incumbents.

2. When \( \alpha < 0 \) we still have multiplicity of equilibrium. BR (1990) suggest that we can modify the equilibrium set of actions, \( A_m^* \). By treating non unique equilibria as equivalent they transform the problem from one that predicts strategies to one that predicts the equilibrium number of firms operating in the market.

- To see how (1) and (2) solve the identification problem let’s suppose that \( \alpha < 0 \) and that non unique equilibria are equivalent. Then write:

\[ p (N_m = 0|x_1, x_2; \Theta) = p (a_m^* = (0, 0)|x_1, x_2; \Theta) \]
\[ p (N_m = 2|x_1, x_2; \Theta) = p (a_m^* = (1, 1)|x_1, x_2; \Theta) \]

- With this we can write:
\[ p(N_m = 1|x_1m, x_2m; \Theta) = \\
1 - p(N_m = 0|x_1m, x_2m; \Theta) - p(N_m = 2|x_1m, x_2m; \Theta) \]

- And finally we have a likelihood function to estimate \( \Theta \):

\[
p(N_1 = n_1, ..., N_M = n_M | x; \Theta) = \\
\prod_{m=1}^{M} \sum_{n_m=1}^{2} p(N_m = n_m | x_1m, x_2m; \Theta) I(N_m = n_m)
\]

- With this likelihood you can find \( \hat{\Theta} \) s.t. \( \alpha < 0 \).

### 3.1.4 Some Conclusions

1. Usually, parameters of static games of complete information cannot be identified without some assumptions.

2. According to BR (1990) identification of these games must be done on a case by case basis.

3. The identification based on the model for \( N \) proposed by BR (1990) is just one solution for the multiple equilibria problem. There are more recent and more general developments - but in general they harder to implement.

- In what follows we generalize the model to include a finite number of players.

### 3.2 A More General Entry Model

- Let’s consider now a game with four players and 2 actions. There are four players indexed by \( i \in \{1, 2, 3, 4\} \).

- Players decide simultaneously to choose either \( a_i = 1 \) or \( a_i = 0 \) (enter or stay out of the market, respectively).

- Payoffs are described by equation (1).

- The \( \varepsilon_i \)'s are iid random draws from a given distribution with continuous and unbounded support, \( x_i \) is a given vector of observed variables and \( \Theta = (\alpha, \beta) \) are parameters to be estimated.

- Write a given profile strategy as \( a = (a_1, a_2, a_3, a_4) \), with \( a_1, a_2, a_3, a_4 \in \{0, 1\} \) and the set of all possible strategies \( A = \times_{i=1}^{4} a_i \). This means that \( A \) has \( 2^N = 2^4 = 16 \) elements.

- Finding the equilibria of this game for all values of parameters, observables and unobservables becomes a little bit cumbersome. Check it out:
3.2.1 A general entry model with symmetric players

This payoff function is similar to the one we analyzed above, except by

BR (1991) start with a model similar to that showed above.

There is however a much simpler way to proceed. This leads us to the BR

sically imposes symmetry across players operating in each market. This

the fact that now shocks and covariates are not player specific. This ba-

\[ \begin{align*}
- (0,0,0,0) \text{ is a N.E. iff } (0,0,0,0) & \in B ((0,0,0,0)) = B_1 (0,0,0) \times B_2 (0,0,0) \times B_3 (0,0,0) \times B_4 (0,0,0). \text{ This happens if } B_i (0,0,0) = 0 \\
- (1,0,0,0) \text{ is a N.E. iff } (1,0,0,0) & \in B ((1,0,0,0)) = B_1 (0,0,0) \times B_2 (1,0,0) \times B_3 (1,0,0) \times B_4 (1,0,0). \text{ This happens if } B_i (0,0,0) = 1 \\
- (0,1,0,0) \text{ is a N.E. iff } (0,1,0,0) & \in B ((0,1,0,0)) = B_1 (1,0,0) \times B_2 (0,0,0) \times B_3 (0,1,0) \times B_4 (0,1,0). \text{ This happens if } B_i (0,0,0) = 1 \\
- (0,0,1,0) \text{ is a N.E. iff } (0,0,1,0) & \in B ((0,0,1,0)) = B_1 (0,1,0) \times B_2 (0,1,0) \times B_3 (0,0,0) \times B_4 (0,0,1). \text{ This happens if } B_i (0,0,0) = 1 \\
- (0,0,0,1) \text{ is a N.E. iff } (0,0,0,1) & \in B ((0,0,0,1)) = B_1 (0,0,1) \times B_2 (0,0,1) \times B_3 (0,0,1) \times B_4 (0,0,0). \text{ This happens if } B_i (0,0,0) = 1 \\
\end{align*} \]

You can keep on and check the conditions for the remaining 11 possible equilibria. Once again you can see that multiple equilibria or no equilibrium in pure strategies will be a problem and we need to impose restrictions in the parameters and redefine the set of equilibrium actions to solve the model. After this you can construct the probability of each equilibrium, write down a likelihood function and estimate the model using market-time data.

- There is however a much simpler way to proceed. This leads us to the BR (1991) model.

3.2.1 A general entry model with symmetric players

- BR (1991) start with a model similar to that showed above.

- Suppose that you have data from \( m = 1, 2, ... , M \) markets. In each market there are \( N_m > 0 \) potential competitors.

- In market \( m \), player \( i \)'s utility is given by:

\[ u_{im}(a_{im}, a_{-im}, x_m, \varepsilon_m; \Theta) = \left\{ \alpha \left( \sum_{j \neq i} a_j \right) + \beta x_m - \varepsilon_m \right\} a_i \]

- This payoff function is similar to the one we analyzed above, except by the fact that now shocks and covariates are not player specific. This basically imposes symmetry across players operating in each market. This
assumption simplifies a lot the construction of our econometric model. BR (1991) avoid the multiple equilibria problem by focusing on the equilibrium number of players (instead of writing down a model for equilibrium actions).

• To see how symmetry helps let’s find the conditions that determine the equilibrium number of players in each market \( m \) (just for simplicity we dropped the market subscript):

- \( N = 0 \) is an equilibrium if \( \beta x - \varepsilon < 0 \).
- \( N = 1 \) is an equilibrium if \( \beta x - \varepsilon \geq 0 \) and \( \alpha + \beta x - \varepsilon < 0 \), or simply if \( \alpha + \beta x < \varepsilon \leq \beta x \).
- \( N = 2 \) is an equilibrium if \( \alpha + \beta x - \varepsilon \geq 0 \) and \( 2\alpha + \beta x - \varepsilon < 0 \), or simply if \( 2\alpha + \beta x < \varepsilon \leq (n-1)\alpha + \beta x \).
- And, generically, \( N = n \) is an equilibrium if \( (n-1)\alpha + \beta x - \varepsilon \geq 0 \) and \( n\alpha + \beta x < \varepsilon \leq (n-1)\alpha + \beta x \).

• Notice that the symmetry assumption really simplifies the problem. Go back to last section and see that the equilibrium conditions for \( a = (1,0,0,0) \), \( a = (0,1,0,0) \), \( a = (0,0,1,0) \), \( a = (0,0,0,1) \), which determines \( N = 1 \), are equal across all these 4 equilibria. The same happens when any two players play \( a_i = 1 \) and so on.

• Besides the symmetry, notice that in equilibrium we must have \( \alpha < 0 \) or, equivalently, that the entry a new competitor reduces the profits of the incumbent(s). This is true for a large class of oligopoly models. If \( \alpha > 0 \), on the other hand, there is not equilibrium in pure strategies.

• To write down a likelihood function to estimate \( \Theta = (\alpha, \beta) \) suppose that you observe \( \{N_m, x_m\}_{m=1}^{M} \). Assume also that \( \varepsilon_m \sim N(0,1) \) iid across markets. Then use the equilibrium conditions above and write:

\[
p(N_m = n_m|x_m; \Theta) = P(n_m\alpha + \beta x_m < \varepsilon \leq (n_m - 1)\alpha + \beta x_m)
= \Phi((n_m - 1)\alpha + \beta x_m) - \Phi(n_m\alpha + \beta x_m)
\]

• Here \( \Phi(\cdot) \) is the CDF of a standard normal variable. This holds for all \( N_m \geq 1 \). For \( N_m = 0 \) then we have \( p(N_m = n_m|x_m; \Theta) = 1 - \Phi(\beta x_m) \).

Observing data from several markets:

\[
p(N_1 = n_1,...,N_M = n_M|x; \Theta) = \\
\prod_{m=1}^{M} \sum_{n_m=0}^{\bar{N}_m} p(N_m = n_m|x_m; \Theta) I(N_m = n_m)
\]
• This is simply the likelihood function of an Ordered Logit. An econometric package like STATA can be used to estimate $\Theta = (\alpha, \beta)$. Have in mind that you must restrict $\alpha < 0$. This assumption guarantees that an equilibrium in pure strategies exists in every market.

3.2.2 Some Conclusions

• BR (1991) is attractive because of its simplicity. There are however some strong assumptions behind the model, namely:

1. Players in the same market are symmetric.
2. The parameter that measures “competition”, $\alpha$, must be necessarily negative - that is, the application is restricted to models of competition, where new entrants reduce the profits of incumbents.
3. Equilibrium in pure strategies - we are ruling out other potential equilibria (for example, in mixed strategies, joint profit maximization, etc.).
4 Part II: Static Games of Incomplete Information

• In models of incomplete information there are components in player $i$'s payoff function that are observed only by $i$ herself but not by the other players and the econometrician.
  – This unobserved part is modeled as a profitability shock that player $i$ draw from a known distribution before taking her decision.
  – All other players (and the econometrician) do not observed $i$’s draw but they know that this draw comes from a given distribution.
  – Under incomplete information players are assumed to play a Bayes-Nash equilibrium.

4.1 Model

• Let’s consider the following $2 \times 2$ simultaneous entry game.
  – Seim (Rand, 2006) has a more complete model including entry/exit and locational choices.
  – Simpler models are studied in Bajari, Hong, Krainer and Nekipelov (JBES, 2010) and in the introduction of Pesendorfer and Schmidt-Dengler (REStud, 2008).
  – Here we illustrate the idea using the simple example in the introduction of Pesendorfer and Schmidt-Dengler (REStud, 2008).

• Each firm $i$ has two possible choices: Be active or not active, $a^t_i \in A_i = \{0,1\}$, where 0 corresponds to “not active” and 1 to “active”.

• Firm $i$’s state space in period $t$ has four elements, denoting the actions made by both firms in period $t-1$, $s^t = a^{t-1} \in \{(0,0); (0,1); (1,0); (1,1)\}$.

• State variables are common knowledge.

• Firm $i$’s period payoffs are described as follows:
  $$\Pi(a^t, s^t, \xi^t_i) =$$
  $I(a^t_i = 1) \cdot [\pi_0 + \pi_1 a^t_{i-1} - \xi^t_i] + I(a^t_i = 1) \cdot I(a^{t-1}_i = 0) \cdot F,$
  – Where the vector of parameters, $\Theta = (\pi_0, \pi_1, F)$, describes respectively firm $i$’s monopoly profits, duopoly profits and entry costs;
  – $I(\cdot)$ is an indicator function and $\xi^t_i$ denotes firm $i$’s iid profitability with distribution $N(0,1)$. We denote its cdf by $\Phi(\cdot)$. The profitability shock is firm $i$’s private information. The distribution of $\xi^t_i$ is common knowledge.
4.2 Equilibrium

• We are looking for Bayes-Nash equilibria in pure strategies.

• Suppose that the observed vector of states is \( s^t = (0, 0) \).

  – If firm \( i \) plays \( a^t_i = 1 \) it gets:

  \[
  E_{a^t_i} \left[ \Pi(a^t_i, s^t, \xi^t_i) | a^t_i = 1, s^t = (0, 0), \xi^t_i \right]
  = [\pi_0 + \pi_1 - \xi^t_i + F] \sigma_i(a^t_{-i} = 1 | (0, 0)) + [\pi_0 - \xi^t_i + F] \sigma_i(a^t_{-i} = 0 | (0, 0))
  = \pi_0 + \pi_1 \sigma_i(a^t_{-i} = 1 | (0, 0)) - \xi^t_i + F,
  \]

  where, \( \sigma_i(a^t_{-i} = 1 | (0, 0)) \) denotes \( i \)'s beliefs on the probability of \(-i\) playing \( a^t_{-i} = 1 \) in state \((0, 0)\).

  – If firm \( i \) plays \( a^t_i = 0 \) it gets zero.

  – Therefore, firm \( i \) plays \( a^t_i = 1 \) if \( \pi_0 + \pi_1 \sigma_i(a^t_{-i} = 1 | (0, 0)) - \xi^t_i + F \geq 0 \) or \( \pi_0 + \pi_1 \sigma_i(a^t_{-i} = 1 | (0, 0)) + F \geq \xi^t_i \). Using the fact that \( \xi^t_i \) is iid \( N(0, 1) \) then \( \sigma_i(a^t_i = 1 | (0, 0)) = \Phi \left( \pi_0 + \pi_1 \sigma_i(a^t_{-i} = 1 | (0, 0)) + F \right) \).

• By stacking these best response for all states and all players one can solve for the equilibrium probabilities \( \sigma_i(a^t_{-i} = 1 | s^t) \) and \( \sigma_{-i}(a^t_i = 1 | s^t) \).

  – By Brower’s fixed point this equilibrium exists.

  – Due to non linearity of best responses the equilibrium, however, is not unique.

4.3 Estimation ideas: NFPX and CCP

• There are two basic ways of estimating \( \Theta = (\pi_0, \pi_1, F) \):


  – Two step Conditional Choice Probability methods (CCP): Developed in Hotz and Miller (RESTud, 1993) and used in Bajari, Hong, Krainer and Nekipelov (JBES, 2010).

• NFPX is much more complicated than CCP. By this reason CCP methods have dominated the literature in recent years. We are going to give you an overview of NFPX and describe with more details the CCP idea.
4.3.1 NFXP

- Start with a guess for the vector of \((\pi_0, \pi_1, F)\) and solve for the equilibrium probabilities using the system of best responses defined above. Notice that with multiple equilibria it is possible that changes in the vector of parameters lead to different equilibria.

- Given a time series of observed states and actions, \(\{a_t, s_t\}_{t=1}^T\) find \((\pi_0, \pi_1, F)\) that maximizes the likelihood \(\prod_{t=1}^T p(a_t|s_t; \Theta)\) using the fact that \(p_i(a_t^i = 1|0, 0) = \Phi(\pi_0 + \pi_1 \sigma_i(a_{t-1}^i = 1|0, 0) + F)\) and the like for all other possible states.

- Use the new parameters to solve the model and then go to the likelihood again.

- Solve this iteratively until parameters converge.

4.3.2 CCP

- Assume that the data (actions conditional on states) is generated from a single equilibrium.

- Go to the data and estimate \(\sigma_i(a_{t-1}^i = 1|s^t)\) and \(\sigma_{-i}(a_t^i = 1|s^t)\) using a logit/probit or any non-parametric thing. Call your estimates as \(\hat{\sigma}_i(a_{t-1}^i = 1|s^t)\) and \(\hat{\sigma}_{-i}(a_t^i = 1|s^t)\).

- Notice that by getting the equilibrium probabilities directly from the data you do not have to solve the model (avoid the first step of NFXP).

  - Finding the fixed point in the system of best responses for every iteration can be extremely time consuming for more realistic models.

- From here you can proceed either by plugging the “hat” probabilities into the best response and using MLE (second stage of NFXP) or a moment estimator to find the parameters of interest.

- A popular moment estimator was proposed in Pesendorfer and Schmidt-Dengler (2008). The estimator is called ALSE (Asymptotic Least Squares). Here is how it works:

  - Write down the best responses for all players and states.

  - Stacking these best responses one can get a system of “moments” \(\hat{p} - \Psi(\hat{\sigma}; \Theta) = 0\). In the context of our example this system has 8 equations (four states for 2 players).

  - \(\hat{\Theta}_{ALSE}\) solves \(\min_{\Theta} (\hat{p} - \Psi(\hat{\sigma}; \Theta))'W(\hat{p} - \Psi(\hat{\sigma}; \Theta))'\), where \(W\) is a matrix of weights.

  - An efficient matrix of weights (that minimizes the variance of \(\hat{\Theta}_{ALSE}\)) can be derived.
While Pesendorfer and Schmidt-Dengler (2008) proposes an ALSE in the probability space, Sanches, Silva and Srisuma (2013) proposes an ALSE in the payoff space. This is done by inverting $\Phi(\cdot)$ in the best responses.

- They prove that both estimators are asymptotically equivalent.
- Advantage of our procedure is that under the linearity of payoffs when $W = I$ ALSE becomes OLS. We derive an efficient estimator that is a GLS estimator.
- OLS/GLS have closed form solutions - no need of numerical methods.
5 Part III: Dynamic Games of Incomplete Information

- Usually entry/investment decisions involve substantial sunk costs that must be paid upfront.
- In these cases it is perhaps more realistic to think that firms pay (or do not pay) these costs based on future payoffs.
- Next:
  - Canonical dynamic model.
  - Estimation procedures.

5.1 Model

- **Time and markets.** Time is discrete, \( t = 1, 2, ..., \infty \). There is one market denoted by \( m \).
- **Players.** The set of players in market \( m \) is \( N = \{1, 2, ..., N\} \). We denote each player in market \( m \) by \( i \in N \).
- **Actions.** A player’s action in market \( m \), period \( t \) is denoted by \( a^t_i \in \{0, 1, ..., K\} \). The \( 1 \times N \) vector \( a^t \in A = \times_{i \in N} a^t_i \) denotes the action profile in market \( m \), period \( t \). We sometimes use \( a^{t-1}_i \in A_{-i} = \times_{j \neq i, j \in N} a^t_j \) to denote the actions of all players but player \( i \). The cardinality of the action space in market \( m \) is \( N_a = (K+1)^N \).
- **State space.** The state space is discrete and finite. The state variables for player \( i \in N \) is composed by a vector \( s^t_i \in S_i = \{1, 2, ..., L\} \) of exogenous variables. The state variables are publicly known to the players and to the econometrician. The vector of all players’ state variables is \( s^t = (s^t_1, s^t_2, ..., s^t_N) \) such that \( s^t \in S = \times_{i \in N} S_i \). The cardinality of the state space \( S \) is \( N_s = L^N \).
- **Shocks.** In each period players draw a vector of profitability shocks. We use \( \xi^t_i \) to denote the \( (K+1) \times 1 \) vector \( (\xi^t_{i0}, \xi^t_{i1}, ..., \xi^t_{iK}) \) of profitability shocks. The profitability shock is iid across individuals, time and actions. This is the only source of asymmetric information in the model. We denote the cumulative distribution function of \( \xi^t_i \) by \( G(\cdot) \).
- **Payoffs.** Player \( i \)'s period payoff in market \( m \) is given by \( \Pi_i(a^t, s^t, \xi^t_i) = \pi_i(a^t, s^t) + \sum_{k=0}^{K} \xi^t_{ik} \cdot I(a^t_i = k) \), where \( \pi_i(a^t, s^t) \) denotes player’s deterministic profits and \( I(\cdot) \) is an indicator function that assumes 1 if the condition \( (\cdot) \) is satisfied and 0 otherwise.
• **Transitions.** The vector \( s^{t+1} \) evolves according to the conditional cumulative density function \( p(s^{t+1}|a^t, s^t) \), described by next period distribution of possible values for the vector \( s^{t+1} \) conditional on each \( (a^t, s^t) \). We sometimes use \( p \) to denote the \( N_a \cdot N_s \cdot N_o \times 1 \) vector of transitions, \( p(s^{t+1}|a^t, s^t) \), for every possible future state \( s^{t+1} \in S \) given all \( (a^t \in A, s^t \in S) \).

• **Sequence of decisions.** The sequence of events in this game is the following:

1. States are observed by all the players.
2. Each player draws the private profitability shock \( \xi_i^t \).
3. Actions are simultaneously chosen. Players maximize their payoffs given beliefs on competitor’s actions. The total payoff of a player is given by the discounted sum of player’s period payoffs. The discount rate is given by \( \beta < 1 \) and is the same for all players.
4. After actions are chosen the law of motion for \( s^{t+1} \) determines the distribution of states in the next period; the problem restarts.

### 5.2 Equilibrium characterization

• We restrict our attention to pure *Markovian strategies*. This means that players’ actions are fully determined by the current vector of state variables. Intuitively, whenever a player observes the same vector of states it will take the same actions and the history of the game until period \( t \) does not influence player’s decisions.

• Player \( i \)'s best response function solves the following Bellman equation:

\[
\max_{a_i \in \{0,1,...,K\}} \left\{ \sum_{s_{i-1}} \sigma_i(a_{i-1}^t|s^t) \Pi_i(a_i^t=k, a_{i-1}^t, s^t, \xi_i^t) + \beta \sum_{s_{i+1}} \Pi_i(s_{i+1}^t|s^t, a_i^t, p) E_{\xi_i} V_i(\sigma_i, p) \right\}.
\]

(2)

Here \( \Pi_i(\cdot) \) is player’s period payoff; the function \( \sigma_i(a_{i-1}^t|s^t) \) accounts for \( i \)'s beliefs on other players’ actions given current states; \( \sigma_i \) is a \( N_a \cdot N_s \times 1 \) vector of beliefs, \( \sigma_i(a_i^t|s^t) \), for all and \( a_i \in A \) and \( s^t \in S \); \( z_k(s^{t+1}|s^t; \sigma_i, p) \) is a \( 1 \times N_s \) vector containing the transitions \( \sigma_i(a_{i-1}^t|s^t)p(s^{t+1}|a_i^t=k, a_{i-1}^t, s^t) \) and \( E_{\xi_i} V_i(\sigma_i, p) \) is a \( N_s \times 1 \) vector with the expected continuation value for the player, \( E_{\xi_i} V_i(s^{t+1}; \sigma_i, p, \pi) \), for all \( s^{t+1} \in S \).

• The value function conditional on \( a_i^t = k \in \{0,1,...,K\} \) being played in period \( t \) is then defined as:

\[
V_i^k(s^t; \sigma_i, p) = \sum_{a_{i-1}^t} \sigma_i(a_{i-1}^t|s^t) \pi_i(a_i^t=k, a_{i-1}^t, s^t) + \beta \sum_{s_{i+1}} \Pi_i(s_{i+1}^t|s^t; \sigma_i, p) E_{\xi_i} V_i(\sigma_i, p) + \xi_i^t.
\]

(3)
\[ V_i^k(s^t; \sigma_i, p) = \hat{V}_i^k(s^t; \sigma_i, p) + \xi_{ik}^t, \text{ where } \hat{V}_i^k(s^t; \sigma_i, p) \text{ comprises all} \]

the terms in (3) except the profitability shock.

- The solution to problem (2) implies that player \(i\)'s probability of playing action \(k\) when states are \(s^t\) satisfies the following equilibrium restrictions:

\[
P_i(a_i^t = k|s^t; \sigma_i, p) = \text{Prob}\left( \hat{V}_i^k(s^t; \sigma_i, p) - \hat{V}_i^{k'}(s^t; \sigma_i, p) \geq \xi_{ik}^t - \xi_{ik'}^t, \forall k' \neq k \right) \quad (4)
\]

- that holds for all \(k, k' \in \{0, 1, ..., K\}\), all \(s^t \in S\) all \(i \in N\).

- The solution to this problem is a vector of player \(i\)'s optimal actions when he faces each possible configuration for the state vector \(s^t\) and has consistent beliefs about other players actions in the same states of the world.

- By stacking up the equilibrium restrictions derived above for every action except action \(k = 0\) of every player and every possible state one can form a system of \(N \cdot K \cdot N_s \times 1\) equations. This system is used to find the \(N \cdot K \cdot N_s \times 1\) vector of players' beliefs.

- A formal proof of the existence of this vector can be found in Pesendorfer and Schmidt-Dengler (2008). Uniqueness of this equilibrium is not, however, guaranteed. This is a common feature of games.

5.3 Estimation

- Firstly have to fix a distribution for \(\xi_{ik}^t\) - this distribution cannot be identified from data (Pesendorfer and Schmidt-Dengler (2008) and Magnac and Thesmar (Econometrica, 2002)).

- Second have to fix a discount factor - Sanches, Silva and Srisuma (2014) discuss conditions for the identification of \(\beta\), but usually this thing is taken as given.

- Let's assume that \(\xi_{ik}^t\) is iid EV(0,1). In this case the probability in (1) becomes:

\[
P_i(a_i^t = k|s^t; \sigma_i, p) = \frac{\exp(\hat{V}_i^k(s^t; \sigma_i, p))}{\sum_{k' = 0}^K \exp(\hat{V}_i^{k'}(s^t; \sigma_i, p))}.
\]

- This holds for any \(k \in \{0, 1, ..., K\}\). Dividing both sides by \(P(a_i^t = k|s^t; \sigma_i, p)\) and taking logs, for any \(k \in \{1, ..., K\}\) the equilibrium restriction becomes:

\[
q_{k0}(s^t; \sigma_i, p) = \hat{V}_i^k(s^t; \sigma_i, p) - \hat{V}_i^0(s^t; \sigma_i, p),
\]
where \( q_{k0} (s^t; \sigma_i, p) = ln \{ P(a_i^t = k|s^t; \sigma_i, p) \} - ln \{ P(a_i^t = 0|s^t; \sigma_i, p) \}. \)

\( \xi_{tik} \) iid EV(0,1) implies

\[
E [\xi_{tik} | a_i^t = k, s^t; \sigma_i, p] = \gamma - ln \{ P(a_i^t = k|s^t; \sigma_i, p) \},
\]

where \( \gamma \) is the Euler constant - see Hotz and Miller (1993).

• Some ways to proceed from here:

  - NFXP of Rust (1987) very complicated here because we have to solve for the value functions and then for tequilibrium probabilities and then go to MLE. Solution for the value function and entry probabilities at each iteration is hell.
  - CCP idea of Hotz and Miller (1993): Estimate equilibrium probabilities from data (logits, probits, multinomial logits, non parametric, etc.). Use probabilities to solve for the value functions and then use MLE, ALSE, etc. to estimate the parameters of the model using the theoretical probabilities derived above to find the parameters of interest.

  • Here we focus on CCPs - NFXP not used in practice (problems with multiple equilibria, computational burden).

5.3.1 Value functions

• Estimate the CCPs and transitions from data.

• Two ways to proceed: Forward Simulation or \textit{ex-ante} value function.

• Hotz, Miller, Sanders and Smith (1994): Forward simulation of value functions.

  - Start with a guess for the parameters.
  - Use the CCPs and the transitions to simulate \( S \) paths of future profits \( T \) periods ahead: Take the average across paths.

• \textit{Ex-ante} value functions:

\[
E_{\xi} V_i (\sigma_i, p) = \Delta_i \left( \bar{\pi}_i + \tilde{E}_{\xi} \right)
\]

where \( \Delta_i = [I_{N_s} - \beta Z_i]^{-1}; \)

- \( \bar{\pi}_i \) is a \( N_s \times 1 \) vector stacking \( \sum_{s_{t+1}} \sigma_i(a^t|s^t)\pi_i(a^t, s^t), \) for every state;

- \( \tilde{E}_{\xi} \) is a \( N_s \times 1 \) vector of expected shocks \( \tilde{E}_{\xi} (s^t; \sigma_i, p) = \sum_{k=0}^{K} \sigma_i(a_i^t = k|s^t; \sigma_i, p)E [\xi_{tik} | a_i^t = k, s^t] \) for every state;

- \( I_{N_s} \) is a \( N_s \times N_s \) identity matrix;

- \( Z_i \) is a \( N_s \times N_s \) matrix stacking the \( 1 \times N_s \) vector \( z (s^{t+1}|s^t; \sigma_i, p) \) containing the transitions \( \sigma_i(a^t|s^t)p(s^{t+1}|a^t, s^t) \) for every state.
• Start with a guess for the parameters. Plug CCPs and transitions into the formula above and solve for $E_\xi V_i(\sigma_i, p)$.

• Which one works better?
  – Forward-simulation: Have to truncate the simulation for a given $T$. This will introduce some bias in the estimation. Moreover, you have to use numerical simulation (write down a code, etc.).
  – Ex-ante: No bias (do not have to truncate anything) but have to invert a (usually) large matrix ($\Delta_i$). If this matrix is too large (the state space is too large) this thing either will not invert at all or will take some time.

• Sanches, Silva and Srisuma (2014) derive a procedure to estimate the model without solving for the value functions.

5.3.2 Parameters

• Plug the value functions into the theoretical probabilities and use MLE, ALSE, whatever to estimate the parameters.

• Sanches, Silva and Srisuma (2013): When payoffs are linear in the parameters ALS is OLS.
  – No need of numerical methods.
  – Useful when the model has a large number of parameters to be estimated.

• Focus on the linear-in-the-parameters payoff function (Sanches, Silva and Srisuma (2013)).

• Show that theoretical probabilities can be written as a linear function of the parameters: $y_k(s^t; \sigma_i, p) - D_k(s^t; \sigma_i, p)\Theta = 0$.

• We introduce a sequence of $H$ auxiliary parameters containing estimates for transitions, $p$, and beliefs, $\sigma_i$, for all $i \in N$. Call it as $(\hat{\sigma}(T), \hat{p}(T))$.

• We assume that the sequence $(\hat{\sigma}(T), \hat{p}(T))$ exists, converges in probability to $(\sigma, p)$ and is asymptotically normally distributed, that is, $(\hat{\sigma}(T), \hat{p}(T)) \overset{p}{\rightarrow} T \uparrow \infty (\sigma, p)$ and $\sqrt{T} ((\hat{\sigma}(T), \hat{p}(T)) - (\sigma, p)) \overset{d}{\rightarrow} T \uparrow \infty N(0, \Omega)$, where $\Omega$ is a positive definite $H \times H$ matrix.
  – This assumption follows Pesendorfer and Schmidt-Dengler (2008).

  Pesendorfer and Schmidt-Dengler (2008) discuss estimators for $(\sigma, p)$ that satisfy E3.
To keep the exposition neater we abuse notation and drop the $T$ subscript from the hat variables. From now on keep in mind that all the hat variables are indexed on $T$. We define $\hat{y}_{ikt} = y_k(s^t; \hat{\sigma}_i, \hat{p})$ and $\hat{D}_{ikt} = D_k(s^t; \hat{\sigma}_i, \hat{p})$.

By summing and subtracting the term $\hat{y}_{ikt} - \hat{D}_{ikt} \Theta'$ from the equilibrium restriction we write $\hat{y}_{ikt} - \hat{D}_{ikt} \Theta' - \hat{u}_{ikt} = 0$ or, equivalently:

$$\hat{y}_{ikt} = \hat{D}_{ikt} \Theta' + \hat{u}_{ikt}, \quad (5)$$

where 

$$u_{ikt} = \left( \hat{y}_{ikt} - \hat{D}_{ikt} \Theta' \right) - \left( y_{ikt} - D_{ikt} \Theta' \right), \quad y_{ikt} = y_k(s^t; \sigma_i, p)$$

and $D_{ikt} = D_k(s^t; \sigma_i, p)$.

In matrix form, stacking (5) for all players, states and actions except action $k = 0$ we write:

$$\hat{y} = \hat{D} \Theta' + \hat{u}, \quad (6)$$

where the variable $\hat{y}$ is a $N \cdot K \cdot N_s \times 1$ vector stacking $\hat{y}_{ikt}$ for all individuals, actions and states, $\hat{D}$ is a $N \cdot K \cdot N_s \times N_p$ matrix stacking $\hat{D}_{ikt}$ for all individuals, actions and states and $\hat{u}$ is a $N \cdot K \cdot N_s \times 1$ vector stacking $\hat{u}_{ikt}$ for all individuals, actions and states. The asymptotic properties of $\hat{u}$ are derived below.

Show that $\sqrt{T} \hat{u} \xrightarrow{d}{T \to \infty} N(0, \Lambda)$.

Then $\hat{\Theta}'_{GLS} = \left(D'\Lambda^{-1}D\right)^{-1} \left(D'\Lambda^{-1}\hat{y}\right)$ is consistent.

Also $\sqrt{T} \left( \hat{\Theta}_{GLS} - \Theta \right) \xrightarrow{d}{T \to \infty} N(0, \Xi_{\Theta})$, where $\Xi_{\Theta} = \left(D'\Lambda^{-1}D\right)^{-1}$ is the asymptotic variance of $\hat{\Theta}_{GLS}$.

Or you can forget about $\Lambda^{-1}$ and apply OLS directly on (6). Consistent but not efficient but does a good job when $T$ - number of observations used to estimate CCPs and transitions - is large.

22