1 Syllabus and Course Overview.

- Welcome to Economics 8681.

- Intended for Ph.D. students interested in conducting research in empirical IO.

- Focus on research frontier.

- Main goal: Facilitate student research!

We will discuss the following classes of models (estimation and applications)

1. Static games of Imperfect Competition.

2. Dynamic games of Imperfect competition.

3. Auctions.

4. IO and Mechanism Design

First two readings:

1. Bajari, Hong, Krainer and Nekipelov, "Estimating Static Models of Strategic Interactions

2. Katja Seim, "An empirical model of firm entry with endogenous product-type choices", 2006, RAND Journal of Economics 37(3)

1.1 Requirements.

- This class will be rather demanding.

- However, I hope that you will find the benefits will exceed the cost.

- Philosophy– last chance to be exposed to the frontier in class!

1. Read the papers **before** class.

2. There will be 3 or 4 intense empirical exercises.

-We will impliment the estimators studied in class.

-Familiarity with some programming language (e.g. Matlab, Fortran, C, Gauss. . . .)

1.2 Grading.

- Students will be graded on both items

- There will be no in class exams.

- Students are encouraged to work together by discussing the papers and in implementing the estimators.

- Course information can be found on my personal web page.

- Office hours are on Friday and by appointment.

- Registered students are encouraged to drop by– don't be shy!

- Don't be passive. Ask questions– this is an important part of being a good researcher.

- Good reseachers love to discuss economics.

2 Helpful Hints for Graduate Students.

- Objective of most graduate students– Get a Good Job/ write well cited research!

- If you want to get a job at a research university, the following two attributes will determine your level of success:

1. The quality of your job market paper.

2. Your ability to communicate in an articulate manner with other people.

- In economics departments, there is relatively more weight on the first item.

- In business schools, more weight on the second item.

Remark. If your first language is not English and you wish to consider the B-School or even consulting market, you should exploit available ESL resources.

-Weak oral communication skills $\Longrightarrow$ trouble with the MBA's.

-Oral communication skills also a consideration at Econ. Departments.

Things that don't matter very much in your placement:

1. Your grades.

2. Being an outstanding TA or RA.

3. Politics.

- The New Economics PhD market is a big matching market (lots of candidates with lots of schools).

- All the department can do is get other people to read your file.

- Typically, most faculty are generous in recommending students.

- However, after someone else gets your file, our ability to influence the outcome at the margin is small.

- This is true even if you believe your advisor is a "big shot".

- Economists are an independent bunch who make up their own minds about the quality of alternative candidates.

- The market is quite efficient and it is relatively rare that students are massively under or over placed.

- If they are, the market will reallocate them shortly.

2.1 Characteristics of High Quality Job Market Papers in Empirical Microeconomics?

1. The authors asks an interesting and novel research question.

-The data and questions are new and of interest to your audience.

-Rehashing the same questions with not very different results from other researchers is not a good strategy.

2. The analysis is technically well executed.

-You should attempt to use state of the art methods.

-Making at least a small technical innovation in some part of the paper is a good signal to future employers (particularly at top departments).

-They want to make sure that you are capable of teaching graduate students.

-If you don't try hard technically on some part of your paper, don't expect your employers to give you the benefit of the doubt that you are technically up to speed.

-Clean identification.

3. The paper is well written and presented well.

-Remember to write your paper to a general audience of economists, not just specialists.

-Don't skip steps of logic.

-This is much, much harder than most graduate students think.

-Budget at least 6 months for writing and polishing your results.

2.2 Common Pitfalls.

1. Research project starts off by making a small extension to an existing, well known paper.

-Typically, if easy, but important extensions could be made the author would have done it already.

2. Technique motivates the empirical question (instead of vice versa).

-This is appropriate for a theorist, since producing purely technical papers is the domain of theorists.

-Pure econometric theory, or economic theory, is an extremely competitive field (much more so than applied work).

-Outcomes in this field have a high variance– not for the risk averse!

-When an applied economist comes up with the technique before the theory, the result is frequently that neither the theoretical innovation nor the application are very compelling.

3. Fail to produce a first draft early enough.

-The most successful candidates have a solid first draft a year before they go onto the market.

-The profession puts a high premium on polished work.

-Polishing a paper your first time out requires a considerable amount of time.

2.3 Helpful Hints

1. Start out simply.

-Start your project with a good question.

-Before you estimate a complicated model, examine the data using simpler techniques (e.g. a table or a regression).

-Research is a process of polishing.

2. Get started right away.

-Rough in your first draft as soon as possible with all the major moving parts.

-Gradually polish and extend your results.

-Most people squander their third year and don't make forward progress on research.

-Few people have time left to spare before the job market.

3. Attend a reading group and a seminar.

-A reading group allows you to keep up on recent research.

-If you read a paper a week for 3 years, you will know 100 papers (at least) by the time your graduate.

-A seminar allows you to see leading researchers present and discuss their work.

-Giving a clear, persuasive seminar is important in getting a job. Learn from researchers with more experience.

4. Start working with an advisor.

-You should meet with an advisor on a regular basis (weekly or bi-weekly).

-If you aren't making forward progress, you might want to see if you can find a better fit.

3 Overview of topics.

- The course will primarily focus on structural models.

- This is due in part to the instructor's research interests.

- Also, this is a natural way to organize the course.

Why use structural models if people aren't rational?

- We are all aware, either from industry, or personal experience, of examples where people or firms act irrationally.

- Recent work in theory suggests that Nash equilibrium can be the result of irrational behavior.

- Evolution- survival of the fittest leads to equilibrium.

- Learning- backward looking behavior can frequently lead to equilibrium.

- If agents experience feedback causes them to improve their behavior (e.g. learning or equilibrium), the result will be equilibrium if the economy doesn't degenerate into infinite cycles.

- You can also estimate structural models where people are less than rational.

- This has a mixed track record.

- In many methods for structural estimation, a first step requires estimation of a "reduced form".

- The structural estimation boils down to a change of variables, from the reduced form to the structural parameters.

- In many cases, this change of variables is close to just identified.

- The conflict between structural and reduced form methods is not particularly intense when viewed from this perspective.

- Structural estimation just involves a relabeling of the parameters.

- IO economists increasingly spend time simulating computationally sophisticated oligopoly models.

- Like macro, computational methods are used to study models.

- Estimation is a reaonable way to choose benchmark parameter values.

- The simulations are often interesting as pieces of applied theory.

4 Static Games

- A game is a generalization of a standard discrete choice model (e.g. logit or probit)

- Payoffs depend on exogenous covariates and preference shocks

- In a game, payoffs also depend on actions of other players

- Observed behavior is assumed to be an equilibrium to the game

- Objective: Estimate agents utilities and equilibrium selection mechanism

- Early work:Vuong and Bjorn (1984), Bresnahan and Reis (1990,1991) and Berry (1992).

- More recent work: Mazzeo (2002), Tamer (2003), Sweeting (2004), Ackerberg and Gowrisankaran (2006), Aradillas-Lopez (2005), Ryan and Tucker (2006), Pakes, Porter, Ho and Ishii (2005), Ho(2005), Ishii (2005), Ciliberto and Tamer (2007)

- Dynamic Games: Aguirregabiria and Mira (2007), Pesendorfer and Schmidt-Dengler (2007),Pakes, Ovtrovsky and Berry (2007) and Bajari, Benkard and Levin (2007).

Research Goals

1. Point estimation for static games

2. Computationally simple

3. Nonparametric identification and estimation

4. Semiparametric estimator

5 Outline

1. Simple entry example

2. Static Games

3. Nonparametric identification

4. Nonparametric/semiparametric estimator

5.1 Entry Example.

- Data on a cross section of markets.
- Entry by Walmart and/or Target.
- Markets $t = 1, ..., T$ and firms $i = 1, 2$
- Let $a_{i,t} = 1$ denote entry and $a_{i,t} = 0$ denote no entry.

- Economic theory suggests that profits depends on demand, costs and number of competitors

- POP_t is population of market t (demand)

- $DIST_{it}$ is distance from headquarters (costs)

- a_{-i} indicates entry by competitors

- The profit of firm i is:

$$u_{it} = \alpha \cdot POP_t + \beta \cdot DIST_{it} + \delta \cdot a_{-i,t} + \varepsilon_{it} \text{ if } a_{i,t} = 1$$

$$u_{it} = 0 \text{ if } a_{i,t} = 0$$

- ε_{it} is private information

- $\sigma_i(a_{i,t} = 1)$ is probability that i enters market t.

- In a Bayes-Nash equilibrium, agent i makes best response to $\sigma_{-i}(a_{-i,t} = 1)$

- Therefore, i's decision rule is:

$$a_i = 1 \iff \alpha \cdot Pop_t + \beta \cdot DIST_{it} + \delta \cdot \sigma_{-i}(a_{-it} = 1) + \varepsilon_{it} > 0$$

- If error terms are extreme value, then

$$\sigma_i(a_i = 1) = \frac{\exp(\alpha \cdot POP_t + \beta \cdot DIST_{it} + \delta \cdot \sigma_{-i}(a_{-i} = 1))}{1 + \exp(\alpha \cdot POP_t + \beta \cdot DIST_{it} + \delta \cdot \sigma_{-i}(a_{-i} = 1))}$$

- Equilibrium- two equations in two unknowns ($\sigma_1(a_1 = 1)$ and $\sigma_2(a_2 = 1)$):

$$\sigma_1(a_1 = 1) = \frac{\exp(\alpha \cdot POP_t + \beta \cdot DIST_{1t} + \delta \cdot \sigma_2(a_2 = 1))}{1 + \exp(\alpha \cdot POP_t + \beta \cdot DIST_{1t} + \delta \cdot \sigma_2(a_2 = 1))}$$

$$\sigma_2(a_2 = 1) = \frac{\exp(\alpha \cdot POP_t + \beta \cdot DIST_{2t} + \delta \cdot \sigma_1(a_1 = 1))}{1 + \exp(\alpha \cdot POP_t + \beta \cdot DIST_{2t} + \delta \cdot \sigma_1(a_1 = 1))}$$

Two-Step Estimator

- First, estimate $\widehat{\sigma}_i(a_i = 1|POP_j, DIST_{1j}, DIST_{2j})$ using a "flexible" method.

- This is the frequency that entry is observed empirically.

- Standard problem.

- We are recovering an agent's equilibrium beliefs by using the sample analogue.

- Given this first stage estimate, agent i's decision rule is estimated as :

$$a_i = 1 \Longleftrightarrow \alpha \cdot Pop_t + \beta \cdot DIST_{it} + \delta \cdot \widehat{\sigma}_{-i}(a_{-it} = 1) + \varepsilon_{it} > 0$$

- The probability that i choose to enter is

$$\sigma_i(a_i = 1) = \frac{\exp(\alpha \cdot POP_t + \beta \cdot DIST_{it} + \delta \cdot \widehat{\sigma}_{-i}(a_{-i} = 1))}{1 + \exp(\alpha \cdot POP_t + \beta \cdot DIST_{it} + \delta \cdot \widehat{\sigma}_{-i}(a_{-i} = 1))}$$

- This is the familiar conditional logit model!

- In second step, let $L(\alpha, \beta, \delta)$ denote the pseudo likelihood function defined as:

$$L(\alpha, \beta, \delta) = \prod_{t=1}^{T} \prod_{i=1}^{2} \left(\frac{\exp(POP_t \cdot \alpha + DIST_{it} \cdot \beta + \widehat{\sigma_{-i}}(a_{-i}=1) \cdot \delta)}{1+\exp(POP_t \cdot \alpha + DIST_{it} \cdot \beta + \widehat{\sigma_{-i}}(a_{-i}=1) \cdot \delta)} \right)^{1\{a_{i,t}=1\}}$$

$$\left(1 - \frac{\exp(POP_t \cdot \alpha + DIST_{it} \cdot \beta + \widehat{\sigma_{-i}}(a_{-i}=1) \cdot \delta)}{1+\exp(POP_t \cdot \alpha + DIST_{it} \cdot \beta + \widehat{\sigma_{-i}}(a_{-i}=1) \cdot \delta)} \right)^{1\{a_{i,t}=0\}}$$

- Maximize psuedo-likelihood to estimate α, β, δ.

- Bottom line: this is the logit model with $\widehat{\sigma}_i(a_i = 1|POP_j, DIST_{1j}, DIST_{2j})$ as an additional independent variable.

- Simple generalization of widely used model.

- Computationally simple and accurate

- Easy to include unobserved heterogeneity.

- Generalize to richer models, including dynamic models.

6 General Model

- Players, $i = 1, ..., n$ and actions $a_i \in \{0, 1\}$.

- Let $A = \{0, 1\}^n$ and $a = (a_1, ..., a_n)$.

- Let $s \in S$ denote a vector of state variables.

- Two strategy assumption can be generalized.

- State s is common knowledge and observed by econometrician.

- Preference shocks ϵ_i private information, extreme value.

- Period utility for i:

$$u_i(a, s, \epsilon_i) = \Pi_i(a_i, a_{-i}, s) + \epsilon_i$$

- $\Pi_i(a_i, a_{-i}, s)$ mean utility

- Utility similar to standard discrete choice model (e.g. conditional logit).

- Normalize $\Pi_i(a_i = 0, a_{-i}, s) = 0$

- E.g. profit from not entering is zero.

- "Outside good" assumption

- Define the *choice specific value function* as

$$\Pi_i(a_i = 1, s) = \sum_{a_{-i}} \sigma_{-i}(a_{-i}|s)\Pi_i(a_i = 1, a_{-i}, s)$$

- This is the expected utility from choosing $a_i = 1$ (excluding the preference shock)

- The optimal decision rule satisfies:

$$a_i = 1 \Longleftrightarrow \Pi_i(a_i = 1, s) + \varepsilon_i > 0$$

- Since ϵ_i is extreme value

$$\sigma_i(a_i = 1|s) = \frac{\exp(\Pi_i(a_i=1,s))}{1+\exp(\Pi_i(a_i=1,s))}$$

- Analogous to our simple example.

7 Identification.

- The model is *identified* if we can reverse engineer $\Pi_i(a_i, a_{-i}, s)$ that uniquely rationalize $\sigma_i(a|s)$.

- $\Pi_i(a_i, a_{-i}, s)$ is a nonparametric function of (a_i, a_{-i}, s)

- The error terms $\epsilon_i(a_i)$ are distributed i.i.d. extreme value across actions a_i and agents i.

- We cannot nonparametrically identify both error terms and $\Pi_i(a_i, a_{-i}, s)$

- This is true in single agent models as well.

- We first do the "Hotz-Miller" inversion, i.e.

$$\sigma_i(a_i = 1|s) = \frac{\exp(\Pi_i(a_i = 1, s))}{1 + \exp(\Pi_i(a_i = 1, s))} \Rightarrow$$
$$\Pi_i(a_i = 1, s) = \log(\sigma_i(a_i = 1|s)) - \log(\sigma_i(a_i = 0|s))$$

- We can reverse engineer the choice specific value function $\Pi_i(a_i = 1, s)$ from choice probabilities $\sigma_i(a_i = 1|s)$

- Identification requires inversion of the following system:

$$\Pi_i(a_i = 1, s) = \sum_{a_{-i}} \sigma_{-i}(a_{-i}|s)\Pi_i(a_i = 1, a_{-i}, s), \forall i = 1, \ldots, n$$

- For a fixed s, there are $n \times 2^{n-1}$ unknowns corresponding to the $\Pi_i(a_i = 1, a_{-i}, s)$

- However, there are only n equations

- In general, this system cannot be inverted and the model is underidentified!

- One way to identify the system is to impose exclusion restrictions.

- Partition $s = (s_i, s_{-i})$, and suppose $\Pi_i(a_i, a_{-i}, s) = \Pi_i(a_i, a_{-i}, s_i)$ depends only on s_i.

- Ex. profit of firm i excludes distance of other agents.

- Then

$$\Pi_i(a_i, s_{-i}, s_i) = \textstyle\sum_{a_{-i}} \sigma_{-i}(a_{-i}|s_{-i}, s_i)\Pi_i(a_i, a_{-i}, s_i).$$

- By varying s_{-i}, we increase the number of equations but not the number of unknowns

- Sufficient condition: $\exists\ 2^{n-1}$ points in the support of the conditional distribution of s_{-i} given s_i.

8 Nonparametric Estimation

- Nonparametric estimation in 3 steps.

- Empirical analogue of identification argument

Step 1: Estimation of Choice Probabilities.

- There are $t = 1, \ldots, T$ repetitions of the game with actions and states $(a_{i,t}, s_{i,t})$, $i = 1, ..., n$.

- In the first step we form an estimate $\widehat{\sigma}_i(a_i|s)$ of $\sigma_i(a_i|s)$ using flexible method.

- E.g. sieve logit (see Newey (1990) and Ai and Chen (2003))

- Let $z_k(s)$ denote the vector of terms in a k^{th} order polynomial

- $\sigma_i(a_i = 1|s, \beta) = \frac{\exp(z_k(s)'\beta)}{1+\exp(z_k(s)'\beta)}$

- Let $k \rightarrow \infty$ as sample size $T \rightarrow \infty$, but not too fast, i.e. $\frac{k}{T} \rightarrow 0$

- Other basis functions (e.g. splines, orthogonal polynomials) are also possible

Second Step: Inversion

- Perform the empirical analogue of the Hotz-Miller inversion.

  $\hat{\Pi}_i(a_i = 1, s_t) = \log(\widehat{\sigma}_i(a_i = 1|s_t)) - \log(\widehat{\sigma}_i(a_i = 0|s_t))$

- This gives an estimate of the choice specific value function.

Third Step: Recovering The Structural Parameters

- In the third step, we "invert" our system of equations to estimate $\Pi_i(a_i, a_{-i}, s_i)$.

- Choose $\Pi_i(a_i, a_{-i}, s_i)$ to solve the following weighted least squares problem:

$$\frac{1}{T}\sum_{t=1}^{T}\left(\begin{array}{c}\hat{\Pi}_i(a_i, s_t)- \\ \sum_{a_{-i}}\widehat{\sigma}_{-i}(a_{-i}|s_{-it}, s_{it})\Pi_i(a_i, a_{-i}, s_i).\end{array}\right)^2 w(t, s_i)$$

- Local linear regression.

- Note that $\hat{\Pi}_i(a_i, s_t)$ and $\widehat{\sigma}_{-i}(a_{-i}|s_{-it}, s_{it})$ are from first two steps.

- The weights $w(t, s_i)$ are kernel weights:

$$K\left(\frac{s_{it} - s_i}{h}\right)$$

- The kernel measures the distance between s_i and s_{it}

- As we send the bandwidth h to zero, we get a consistent estimate of $\Pi_i(a_i, a_{-i}, s_i)$

8.1 A Semiparametric Estimator

- Nonparametric estimators suffer from a curse of dimensionality.

- We might wish to use a parametric model of $\Pi_i(a_i, a_{-i}, s_i)$

- Approximate $\Pi_i(a_i, a_{-i}, s_i)$ by a set of basis functions Φ_i (e.g. linear index, polynomial, etc...)

$$\Pi_i(a_i, a_{-i}, s_i) = \Phi_i(a_i, a_{-i}, s)'\theta$$

- Replace $\Pi_i(a_i, a_{-i}, s_i)$ in the above with $\Phi_i(a_i, a_{-i}, s)'\theta$

$$\sum_{t=1}^{T} \left(\begin{array}{c} \hat{\Pi}_i(a_i, s_t) - \\ \sum_{a_{-i}} \widehat{\sigma}_{-i}(a_{-i}|s_{-it}, s_{it})\Phi_i(a_i, a_{-i}, s)'\theta. \end{array} \right)^2$$

- The estimator $\widehat{\theta}$ is least squares!

- Alternatively, we can do a psuedo-mle procedure as in our simple example:

$$\begin{aligned}
\sigma_i(a_i &= 1|s) = \frac{\exp(\Pi_i(a_i = 1, s))}{1 + \exp(\Pi_i(a_i = 1, s))} \\
&= \frac{\exp(\sum_{a_{-i}} \widehat{\sigma}_{-i}(a_{-i}|s)\Pi_i(a_i = 1, a_{-i}, s))}{1 + \exp(\sum_{a_{-i}} \widehat{\sigma}_{-i}(a_{-i}|s)\Pi_i(a_i = 1, a_{-i}, s))} \\
&= \frac{\exp(\sum_{a_{-i}} \widehat{\sigma}_{-i}(a_{-i}|s)\Phi_i(a_i, a_{-i}, s)'\theta)}{1 + \exp(\sum_{a_{-i}} \widehat{\sigma}_{-i}(a_{-i}|s)\Phi_i(a_i, a_{-i}, s)'\theta)}
\end{aligned}$$

- The log-likelihood function is:

$$L(\theta) = \frac{1}{T}\sum_{i,t} log\left(\frac{\exp(\sum_{a_{-i}} \widehat{\sigma}_{-i}(a_{-i}|s)\Phi_i(a_i,a_{-i},s)'\theta)}{1+\exp(\sum_{a_{-i}} \widehat{\sigma}_{-i}(a_{-i}|s)\Phi_i(a_i,a_{-i},s)'\theta)}\right) 1\left\{a_{i,t} =\right.$$

$$log\left(\frac{1}{1+\exp(\sum_{a_{-i}} \widehat{\sigma}_{-i}(a_{-i}|s)\Phi_i(a_i,a_{-i},s)'\theta)}\right) 1\left\{a_{i,t} = 0\right\}$$

- This is just the logit model with a generated regressor $\widehat{\sigma}_{-i}(a_{-i}|s)$

- The hard work in Bajari, Hong, Krainer and Nekipelov (2007) is the asymptotic theory.

- We show that $\widehat{\theta}$ converges at a $T^{1/2}$ rate and has normal asymptotics.

- The asymptotic distribution is invariant to the choice of method used to estimate the first stage.

- The bootstrap can be used to compute standard errors

9 Examples

- Seim

- location choices in video retailing

- Estimate demand and markups using BLP

- Discrete choices are locations

- Andrew Sweeting

- Radio stations benefit from have the commercials at the same time

- Discrete choice is when to have a commercial during drive time

- Spillover from havinng commerials at the same time

- Ackerberg and Gowrisankaran

- Discrete Choice is choice of payment system

- Spillover from network effect

- Frequency of choices made by

- Bajari, Hong, Krainer and Nekipelov

- stock analyst ratings during tech bubble

- strong buy, buy, hold, sell are ratings choices (ordered logit)

- conflicts of interest, publicly available information and ratings of other analysts determine ratings

10 Conclusions

- Simple model of strategic interaction

- Nonparametric identification

- Simple Nonparametric/Semiparametric Estimators

- Identification and estimation of dynamic games follows similar principals