UG Empirical Industrial Organization

Lecture 1: Introduction to the Course

Anubhav Jha[1]
Ashoka University

March 11, 2024

[1] This course is based on Victor Aguirregabiria's Empirical IO book. The slides (in my first year of teaching) are extremely close to his slides. These will change in the future iterations of this course.

Introduction / Logistics: Outline

- Basic information (meetings, tutorials, evaluation, etc)
- Course topics and objectives
- Introduction to Empirical IO

Basic Information

Class meetings and tutorials

- Lecture: Tue & Thu 3:00 PM - 4:20 PM at AC-04-LR-004
- Office hours: Fri 11:30 AM - 12:30 PM at zoom/google-meet

Course Prerequisites

- Microeconomic Theory: General Eqm. Theory and/or Game Theory.
- Quantitative Methods: Statistics and Econometric courses.

Evaluation

- Problem Sets (2): 40%
- Presentations: 25%
- Midterm Exam: 25%
- In-class attendance 5% + Student presentation participation 5%

Course Topics

1. Measuring Productivity: Estimation of Production Functions.
2. Measuring Consumer Preferences: Demand Estimation.
3. Firm Competition in Prices and Quantities.
4. (**May-be**) Empirical Models of Market Entry.

- In addition to above, lectures on how to do empirical analysis on R

Course Materials

- Will be provided to you
- Need R for PSets.

Course Objectives

- Understand the main features of empirical models:
  - demand;
  - production function;
  - price and quantity competition;
  - market entry.
- Learn how to use market real world data to estimate the parameters of these empirical models, and interpret the economic implications of these estimations.
- Acquire enough programming experience using R and practical experience using actual data such that you can work in a research project in empirical IO.

Introduction to Empirical IO: Outline

- What is Empirical IO?
- Empirical Questions in IO.
- The Basic Structure of EIO Models.
- An Example

What is Empirical IO?

- **IO studies:**
  1. How markets work
  2. How firms compete or collude with each other
  3. How these interactions determine profits & consumer welfare.

- **In Empirical IO, we use data on consumers' and firms' decisions to:**
  1. Measure consumer demand, firm productivity and profits
  2. To understand firms' strategies
  3. To analyze how government regulations affect market competition and social welfare.

Econometrics and Empirical IO

- **Econometrics and data analysis are fundamental tools for the modern economist of the 21st century.**
- In this course, we review and apply basic econometric models and methods:
  - linear regression model;
  - instrumental variables estimation;
  - discrete choice models.
- You will gain practical experience working with economic data, and making use of R.

Basic Structure of Empirical IO Models

- To study competition in an industry, IO researchers propose and estimate models of demand and supply.
- **Behavior:**
  - Models of consumer and firm behavior where consumers are utility maximizers and firms are profit maximizers.
- **Unobservables:**
  - Model recognizes that we, as researchers, have less information about demand & costs than firms in the industry.
  - Some variables in the model observable to firms are unobservable to us.
- **Parameters:**
  - The parameters of the model have a clear economic interpretation in terms of consumer preferences, production technology, or institutional constraints.

Principle of Revealed Preference

- We observe consumers' and firms' choices.
- We assume that they make these decisions to maximize:
  - Utility (consumers).
  - Profits (firms).
- Under this condition, their observed choices reveal information to us about:
  - Their utilities (consumers).
  - Their profits (firms).
- **Examples:**
  - Consumer purchase of a particular model of car.
  - Firm opening a store in a particular location.
  - Firm choice of price, etc.

Typical IO Models

- **Model of consumer behavior (Demand)**
  - Product differentiation?
- **Model for firms' costs**
  - Economies of scale; Economies of scope? Entry costs? Investment costs?
- **Equilibrium model of static competition**
  - Price (Bertrand), Quantity (Cournot).
- **Equilibrium model of market entry.**
- **Equilibrium model of dynamic competition**
  - Investment, advertising, quality, product characteristics, stores, etc.

An Example: (0) Empirical Question

- We always start with an empirical question.
- **US cement industry:** Evaluation of the effects in this industry of a new Air Clean Act.
- **Key regulatory change:**
  - The new law restricts the amount of emissions a cement plant can make.
  - Complying with the new law requires the adoption of a green technology that involves larger fixed costs than the old/non-green technology.
- **Research question:**
  - What is the effect of this new policy on prices, quantities, number of plants, profits, and consumer welfare?

An Example: (1) Key Characteristics of the Industry

- **Product differentiation?**
- **Size fixed costs**
- **Slope of Marginal Costs?**
- **How easy to increase capacity?**
- **Local industry structure?**
- **Market Structure: Perfect? Monopoly?**

An Example: (2) Key Characteristics of the Industry

- **Homogeneous product**
  - We abstract from spatial differentiation.
- **Substantial fixed costs**
  - Operating a plant (cement furnace) involves high fixed costs.
- **Marginal cost dynamics**
  - Marginal cost increases when output approaches full capacity.
- **Capacity investment**
  - It is an important strategic variable.
- **Local industry structure**
  - High transportation costs per dollar value make the industry very local.
  - Many "isolated" local markets exist.
- **Oligopolistic industry**
  - Small number of firms compete in each local market.

An Example: (3) Data

- The specification of the model depends crucially on the data that is available for the researcher.
- **Market structure:**
  - M local markets (e.g., towns) observed over T consecutive quarters.
  - Markets are indexed by m and quarters by t.
- **For every market-quarter observation, the dataset contains information on:**
  - Number of plants operating in the market (N_{mt}).
  - Aggregate amount of output produced by all the plants (Q_{mt}).
  - Market price (P_{mt}).
  - Market characteristics that may affect demand and/or costs:
    - Population.
    - Average income.
    - Input prices (X_{mt}).
- **Data Representation:**
  - Data $= \{P_{mt}, Q_{mt}, N_{mt}, X_{mt} : m = 1, 2, ..., M; t = 1, 2, ..., T\}$

An Example: (4) Specification of Demand

- **Functional Forms of Demand:**
  - **Linear in prices and parameters:**
    - $P_{mt} = \beta_0 + \beta_X X^D_{mt} - \beta_1 Q_{mt} + \epsilon^D_{mt}$
  - **Log-linear (isoelastic):**
    - $\ln P_{mt} = \beta_0 + \beta_X \ln X^D_{mt} - \beta_1 \ln Q_{mt} + \epsilon^D_{mt}$
- **Interpretation of Terms:**
  - β_0 are parameters.
  - X^D_{mt} are observable variables affecting demand (e.g., population, average income).
  - ϵ^D_{mt} is a random variable with zero mean that is unobservable to researchers and affects demand.

An Example: (5) Specification of Costs

- **The cost of a firm in this industry is:**
  - $C(q_{imt}) = VC_{mt}(q) + FC_{mt}$
- **Variable Cost (Quadratic Form):**
  - $VC_{mt}(q) = \left(\gamma_{MC1} + \gamma_{MCX} X_{mt}^{MC} + \epsilon_{mt}^{MC}\right) q + \frac{\gamma_{MC2}}{2} q^2$
- **Marginal Cost:**
  - $MC_{mt}(q) = \gamma_{MC1} + \gamma_{MCX} X_{mt}^{MC} + \gamma_{MC2} q + \epsilon_{mt}^{MC}$
- **Fixed Cost:**
  - $FC_{mt} = \gamma_{FC1} + \gamma_{FCX} X_{mt}^{FC} + \epsilon_{mt}^{FC}$
- **Unobservables:**
  - ϵ_{mt}^{MC} and ϵ_{mt}^{FC} are unobservable to the researcher.

For simplicity, I will ignore constant terms γ_1, β_0, γ_{FC1}. Mathematically their treatment in this setting will be the same as the treatment of $\gamma_{MCX} X_{mt}^{MC}$, $\gamma_{FCX} X_{mt}^{FC}$, and $\beta_X X_{mt}^{D}$.

An Example: (6) Re-specifying

- **Linear in prices and parameters:**
  - $P_{mt} = \beta_X X^D_{mt} - \beta_1 Q_{mt} + \epsilon^D_{mt}$
- **Variable Cost (Quadratic Form):**
  - $VC_{mt}(q) = \left(\gamma_{MCX} X^{MC}_{mt} + \epsilon^{MC}_{mt}\right) q + \frac{\gamma_{MC2}}{2} q^2$
- **Marginal Cost:**
  - $MC_{mt}(q) = \gamma_{MCX} X^{MC}_{mt} + \gamma_{MC2} q + \epsilon^{MC}_{mt}$
- **Fixed Cost:**
  - $FC_{mt} = \gamma_{FCX} X^{FC}_{mt} + \epsilon^{FC}_{mt}$
- **Unobservables:**
  - ϵ^D_{mt}, ϵ^{MC}_{mt} and ϵ^{FC}_{mt} are unobservable to the researcher.

An Example: (7) Cournot Competition

- Suppose that firms active in a local market compete with each other à la Cournot.
- **Profit function of a firm:**
  - Let q be the firm's own output and $\tilde{Q}$ be the output of competitors.
  - The profit function is given by:

$$\Pi_{mt}(q, \tilde{Q}) = P_{mt}(q, \tilde{Q})q - VC_{mt}(q) - FC_{mt}$$

- **Best response output is characterized by the condition of optimality:**
  - First-order condition:

$$P_{mt} + \frac{\partial P_{mt}(q + \tilde{Q})}{\partial q} q = MC_{mt}(q)$$

- **With our linear specification of demand and costs, this condition simplifies to:**
  -

$$P_{mt} - \beta_1 q_{mt} = \gamma_{MCX} X_{mt}^{MC} + \gamma_{MC2} q_{mt} + \epsilon_{mt}^{MC}$$

An Example: (8) Model of Market Entry

- Let $\Pi_{mt}(N)$ be the Cournot equilibrium profit per firm with N active plants.
  - It is a strictly decreasing function.
- **Equilibrium entry condition:**
  - Every active firm and every potential entrant is maximizing profits.
- **Conditions for market equilibrium:**
  - Active firms should be making non-negative profits:

$$\Pi_{mt}(N_{mt}) \geq 0.$$

  - Potential entrants are not leaving positive profits on the table:

$$\Pi_{mt}(N_{mt} + 1) < 0.$$

- **Unique equilibrium value of N:**
  - The unique value of N satisfies both equilibrium conditions:

$$\Pi_{mt}(N) \geq 0 \quad \text{and} \quad \Pi_{mt}(N+1) < 0.$$

An Example: (9) Model of Market Entry

- Suppose that we approximate the equilibrium conditions with:
  - $\Pi_{mt}(N) = 0$, which is equivalent to:
  $$P_{mt} q_{mt} - VC(q_{mt}) = FC_{mt}.$$
- **Derivation for the model with linear demand and quadratic variable cost function:**
  - First-order condition (F.O.C.) for Cournot equilibrium:
  $$P_{mt} - \beta_1 q_{mt} - MC_{mt} = 0,$$
  - Which can be rewritten as:
  $$P_{mt} = \beta_1 q_{mt} + MC_{mt}.$$
- **Substituting into the profit condition:**
  - $$P_{mt} q_{mt} - VC(q_{mt}) = (\beta_1 q_{mt} + MC_{mt}) q_{mt} - VC(q_{mt}).$$
- **Using the variable and marginal cost functions:**
  - $$VC(q_{mt}) = \left(\gamma_{MCX} X_{mt}^{MC} + \epsilon_{mt}^{MC}\right) q + \frac{\gamma_{MC2}}{2} q^2$$
  $$MC_{mt}(q) = \gamma_{MCX} X_{mt}^{MC} + \gamma_{MC2} q + \epsilon_{mt}^{MC}$$
  - solve!

An Example: (10) Market Entry

- We obtain:

$$\left(\beta_1 + \frac{\gamma_{MC2}}{2}\right) q_{mt}^2 = FC_{mt}.$$

An Example: (11) Structural Equations

- For every market-year (m, t), the model can be described as a system of three equations with three endogenous variables: N, P, and Q.
- **Demand Equation:**
  - $$P_{mt} = \beta_X X_{mt}^D - \beta_1 Q_{mt} + \epsilon_{mt}^D$$
- **Cournot Competition Equation:**
  - $$P_{mt} - \beta_1 \frac{Q_{mt}}{N_{mt}} = \gamma_{MCX} X_{mt}^{MC} + \gamma_{MC2} \frac{Q_{mt}}{N_{mt}} + \epsilon_{mt}^{MC}$$
- **Entry Equation:**
  - $$\left(\beta_1 + \frac{\gamma_{MC2}}{2}\right)\left(\frac{Q_{mt}}{N_{mt}}\right)^2 = \gamma_{FCX} X_{mt}^{FC} + \epsilon_{mt}^{FC}$$
- This system of equations is denoted as the **structural equations** of the model.

An Example: (12) Market Equilibrium

- For any value of the exogenous parameters, this model has an equilibrium and it is unique.
- **The proof is simple:**
  - The entry equation determines output per firm:

$$\frac{Q_{mt}}{N_{mt}} = ?$$

    where $\lfloor x \rfloor$ is the greatest integer smaller than x.
  - Given output per firm, the Cournot equation determines price:

$$P_{mt} = ?$$

  - Finally, given price, the demand equation determines:
    - Total output $Q_{mt} = ?$.
    - Number of plants $N_{mt} = ?$.
- The solution of this system, which relates each endogenous variable with only exogenous variables, is called the **reduced form equations**.

An Example: (13) Market Equilibrium

- For any value of the exogenous parameters, this model has an equilibrium and it is unique.
- **The proof is simple:**
  - The entry equation determines output per firm:

$$\frac{Q_{mt}}{N_{mt}} = \left\lfloor \sqrt{\frac{\gamma_{FCX} X_{mt}^{FC} + \epsilon_{mt}^{FC}}{\left(\beta_1 + \frac{\gamma_{MC2}}{2}\right)}} \right\rfloor = q_{mt}^*$$

    where $\lfloor x \rfloor$ is the greatest integer smaller than x.
  - Given output per firm, the Cournot equation determines price:

$$P_{mt} = (\beta_1 + \gamma_{MC2})\, q_{mt}^* + \gamma_{MCX} X_{mt}^{MC} + \epsilon_{mt}^{MC}$$

  - Finally, given price, the demand equation determines:
    - Total output $Q_{mt} = \frac{1}{\beta_{mt}} \left(\beta_X - P_{mt} + \epsilon_{mt}^D\right)$.
    - Number of plants $N_{mt} = \frac{1}{q_{mt}^* \cdot \beta_1} \left(\beta_X - P_{mt} + \epsilon_{mt}^D\right)$
- The solution of this system, which relates each endogenous variable with only exogenous variables, is called the **reduced form equations**.
- These are non-linear in parameters and errors. Cannot be estimated using OLS nor 2SLS, need MLE or GMM.

An Example: (14) Estimation

- As researchers, we are interested in using the model and data to estimate the model parameters:
  - β's
  - γ_{MC}'s
  - γ_{FC}'s
- **Econometric Model:**
  - **Demand equation:**

$$P_{mt} = \beta_X X_{mt}^D - \beta_1 Q_{mt} + \epsilon_{mt}^D$$

  - **Cournot competition equation:**

$$P_{mt} - \beta_1 \frac{Q_{mt}}{N_{mt}} = \gamma_{MCX} X_{mt}^{MC} + \gamma_{MC2} \frac{Q_{mt}}{N_{mt}} + \epsilon_{mt}^{MC}$$

  - **Entry equation:**

$$\left(\beta_1 + \frac{\gamma_{MC2}}{2}\right) \left(\frac{Q_{mt}}{N_{mt}}\right)^2 = \gamma_{FCX} X_{mt}^{FC} + \epsilon_{mt}^{FC}$$

- **Estimation method:**
  - Each of these equations is linear in parameters.
  - They can be estimated using a Linear Regression Model.

An Example: (15) Endogeneity

- **Endogeneity problem:**
  - A key econometric issue in estimating these equations.
  - Some regressors (e.g., price, output, or number of plants) are correlated with the error term of the regression.
  - The error term represents unobservables ϵ's.
- **Consequences of ignoring endogeneity:**
  - Estimating the parameters using Ordinary Least Squares (OLS) can lead to serious bias.
  - This bias can distort our estimates and affect the validity of our empirical conclusions.
- **Solution: Instrumental Variables (IV) Estimation**
  - A potential approach to addressing endogeneity.
  - IV estimation helps obtain unbiased parameter estimates.

An Example: (16) Answer to Empirical Question

- **Evaluating the impact of the policy change:**
  - Once we have estimated the parameters of the model before and after the policy change, we can answer our empirical question.
- **Key outcomes analyzed:**
  - Changes in output.
  - Changes in price.
  - Number of plants.
  - Production costs.
  - Firm profits.
  - Consumer surplus.
- **Counterfactual analysis:**
  - What are the equilibrium values of these variables in every market-year (m, t) after the policy change?
  - Assumption: The cost parameters remain the same as before the policy change.