L1 – Introduction and Production Functions I

Will Rafey

UCLA

January 4, 2022

Industrial organization

How do firms behave, and what are the consequences for welfare?

- Demand conditions (John's course!)
- Cost conditions – production
- Competition
- Externalities

What determines demand, costs, and competition?

- Investment, innovation
- Market structure (entry, mergers, acquisitions)
- Natural and institutional environments

What policies, market rules should we adopt?

- Regulation of concentrated industries
- Corrective taxation
- Market design

Theoretical IO: 1980s

Approach and broad lessons (see Tirole's textbook):

- Game-theoretic models of firm behavior
- Rich set of results explaining a wide range of phenomena
- Details of the specific market matter
  - Classic example: Investment could accommodate or deter entry depending on strategic effects (Fudenberg and Tirole, 1984)
  - Conundrum for public policy

Tirole's Nobel lecture (2014) consequently calls for a case-by-case or "**rule of reason**" approach:

> *"Economists must (1) develop a rigorous analysis of how markets work, accounting for*
>
> - *specificities of industries*
> - *what regulators do and do not know*
>
> *(2) participate in policy debate."*

Underlying theme: **heterogeneity** across industries

Empirical approach

Tirole (2014), cont'd

> *"One cannot underestimate the interaction between theory and empirics: empirical work needs theory, both to guide it and to make it useful for policy. Theory needs empirical work to strengthen the confidence in policy recommendations and also to suggest key omitted ingredients."*

Einav and Levin (JEP, 2010):

> *"[T]he applied work we often find most exciting relies on careful measurement based on data with good underlying variation, but then continues by framing the empirical exercise in terms of a coherent economic model. The model can then provide a way to think about the operation of the industry and potentially to draw conclusions about policy or general principles." (p. 159)*

Empirical approach

Theoretical literature offered:

- Several explanations for similar observations
- A basis for building empirical models
- An understanding of important mechanisms

Empirical revolution:

- Focus on narrowly-defined industries (heterogeneous effects)
- Close attention to theory
  1. deriving testable hypotheses
  2. interpreting the data (through the lens of a model)
- Greater emphasis on
  1. strategies (conduct)
  2. quantifying welfare, economic costs, profits (performance)
  3. counterfactual simulations (policy) and decompositions (channels)

Empirical approach

- Phrase a question in terms of a counterfactual
  - ✓ What would happen if the U.S. regulated carbon emissions in the concrete industry?
  - Only data available is from a world where carbon is not priced

- Approach: build the primitives of the model
  - Industry cost structure; emissions; demand
  - Interaction between incumbent firms and entrants; foreign imports

- Estimate a model: attention to data limitations and institutional details, e.g.
  - Entry and investment costs may not be observed
  - Undifferentiated final good

- Simulate a counterfactual world and analyze outcomes of interest
  - Factual world: forward-looking entry, investment; oligopolistic production
  - Counterfactual: introduce a permanent carbon pricing rule
  - ✓ Consumer welfare, profits, incumbent v. entrant surplus.

Primitives are held **fixed** $\implies$ tailor model to question of interest

Identification

We hope to identify and estimate empirical models of market behavior.

Estimation: procedure for recovering parameters from the data

Identification: articulation of how the data reveals the model parameters

History of identification

Marschak and Andrews 1944 (1/3)

Marschak and Andrews 1944, Econometrica, vol. 12 no. 3–4, pp. 133–134:

> **To describe and measure causation, the economist cannot perform experiments.** *That is, he cannot choose one variable as "dependent," and, while keeping the other, "independent," ones under control (i.e., while making them assume deliberately chosen sets of values), watch the values taken by the dependent, i.e., uncontrolled variable. The economist has no independent variables at his disposal because he has to take the values of all variables as they come, produced by a mechanism outside his control.*

History of identification

Marschak and Andrews 1944 (2/3)

[. . .]

For example, in agricultural experimentation preassigned quantities of fertilizers are added to the soil of various plots, and the yields compared. Substitute "firms" for "plots," and "labor, capital" for "nitrate phosphate." Can the economist measure the effect of changing amounts of labor and capital on the firm's output—the "production function"—in the same way in which the agricultural research worker measures the effect of changing amounts of fertilizers on the plot's yield? **He cannot** *because the manpower and capital used by each firm is determined by the firm, not by the economist. This determination is expressed by a system of functional relationships; the production function, in which the economist happens to be interested, is but one of them.*

History of identification

Marschak and Andrews 1944 (3/3)

Mann and Wald's (1943) . . . work has brought further clarification to the older debate on "pitfalls in the statistical construction of demand and supply curves" (Frisch 1933) and on "what do statistical demand curves show?". That discussion was started **at least as early as 1910**, *when Pigou (1910) had to recapitulate economic theory to explain the fact that the "statistical demand curve" for steel, i.e., the curve fitted (by Moore, 1914) to the data on price and consumption of steel was an increasing one; and to point out that* **it could in no sense claim to be a demand curve**. *. . . Bronfenbrenner (1944, 1939) and Reder (1943) gave the name of "interfirm production functions" to functions fitted to the data on the output, manpower, and capital of a number of individual firms.* **They raised doubt whether these functions can be identified with the production function as understood by economists**. *(p. 147)*

Identification

Estimation: procedure for recovering parameters from the data

Identification: articulation of how the data reveals the model parameters

Various identifying assumptions:

1. Exclusion/independence assumptions — e.g., instruments, experiments
2. Parametric — e.g., normally distributed variables
3. Functional form — linearity, separability, etc.
4. Model/behavioral restrictions – equilibrium, optimality, etc.

Discussion

Empirical models in industrial organization provide

- Rigorous theoretical basis of analysis
- Theoretical interpretation of parameters
- Ability to move forward without ideal data or experimental variation
- Well-articulated link between data, theory, questions and answers

We must remain wary of

- Unnecessarily complicated or opaque analytical techniques
- Murkier identification strategies
- Focus on narrow questions; e.g.,
    - "Are high R&D industries more profitable?" vs. "Does R&D in the computer industry deter entry?"

Overview of topics

I. Firms

1. Production functions
2. Tests for misallocation
3. Innovation
4. Entry

II. Designing Markets

1. Hybrid markets (electricity flows)
2. Fiat markets (pollution)
3. Incomplete markets (water, kidneys)
4. Externalities (information, land use)

✓ Objective: Bring you to the research frontier.

Logistics

Requirements:

- Assignments (60%)
  - three problem sets with empirical exercises and conceptual questions
  - discussion, group work encouraged!
- Replication exercise (20%)
  - you will replicate a recent paper using actual or simulated data
  - due before the last day of the quarter (March 18)
- Final exam (20%)
  - in-class on March 8
- Class participation (0%)
  - there will be 1–2 required papers per week
  - I recommend that you read the assigned papers and come to class prepared to discuss them
  - choose your favorite figure, table, equation, or sentence

L1. Production Functions

Will Rafey

UCLA

January 4, 2022

Motivation

Supply side of the firm is central to many key economic questions:

- Returns to scale/scope
- Price regulation, corrective taxation
- Learning by doing, R&D

Productivity analysis:

- Dispersion in productivity over time and space
- Effects of policies/regulations

Useful in other areas as well

- Innovation and technological spillovers
- Trade, e.g. effects of tariffs
- Environmental policy and market design

Plan

~~Introduction~~

Today, basics of identification of production functions:

1. primitives
2. data and measurement
3. specification
4. identification

Production Functions, Part I

1. Setup
2. Data issues
3. Specification issues
4. Identification

Classic setup

We will suppose firm i produces Q_i with some production function

$$Q_i = F(L_i, K_i, \omega_i, \varepsilon_i),$$

where L_i is labor, K_i is capital, ω_i is "productivity," and ε_i is measurement error.

- the firm observes ω_i, but we do not
- no one observes ε_i

Various reasons to learn about F

- F_L, F_K: effects of labor and capital on output
- to recover ω_i:
  - across firms: factor misallocation, market structure
  - over time: learning-by-doing, innovation, R&D, exit/entry

Classic setup

Remark. Often we are also interested in some controls/instruments Z_i, relevant to production.

$\hookrightarrow$ e.g., tariffs, regulation, time, weather

Must decide which aspects of Z_i

- enter directly, as an argument in F
- not enter directly in F

If the latter, Z_i usually assumed to affect ω_i, L_i, and/or K_i.

We will therefore write

$$Q_i = F(L_i, K_i, Z_i, \omega_i, \varepsilon_i),$$

to allow for the possibility that Z_i affects F directly.

Datasets

Firm and plant-level data.

Examples:

- Administrative data from regulated industries (e.g., electricity; Wolak 2003)
  - can be very high quality, though potential reporting incentives
- Census data
  - e.g., Annual Survey of Manufactures (ASM)
  - detailed balance sheet and additional information.
  - accessing data requires proposals, confidentiality
  - differs by country and often between censuses
- Proprietary datasets
  - AMADEUS provides firm-level (plant-level) data for long list of countries with annual balance sheet and ownership information.
- Compustat: income, costs, investments
  - large, traded (always multiproduct) firms
  - input data is rarely broken down by the product
  - lots of attrition due to mergers and acquisition
  - selection into being a listed firm

Measurement

Measurement issues for $Q_i = F(L_i, K_i, Z_i, \omega_i, \varepsilon_i)$:

- What is output?
  - aggregate over products; often use revenue divided by some price index
    - what if prices differ across firms?
    - revenue productivity and physical productivity: have diff welfare implications
- What is labor?
  - often measured as hours or number of full-time employees
  - what about education, experience, skills?
- What is capital?
  - have to aggregate equipment, machinery, land, buildings, etc.
  - how to deal with depreciation, variation in utilization rates, etc.
- What else should we control for?
  - time, firm age, tariffs, regulatory status, weather. . .

$\hookrightarrow$ The answers to the above four questions give us ω_i:

- (everything that is not in your chosen L_i, K_i, Z_i) + (what is misspecified by F) − (measurement error)

Specification

Cobb-Douglas. Using lowercase letters to denote natural logarithms,

$$q_i = \beta_0 + \beta_\ell \ell_i + \beta_k k_i + \beta_z' z_i + \omega_i + \varepsilon_i$$

Economic implications:

1. Shape of the function

- constant output elasticities
- fixed substitutability across inputs
- will imply constant returns to scale if $\beta_\ell + \beta_k \equiv 1$

Quality of C-D approximations depend on the range of values of ℓ_i and k_i in the data and in counterfactuals.

- e.g., if you have some exogenous Z_i that enter into F, this approximation may not be so restrictive if β_0, β_ℓ, and/or β_k are allowed to flexibly depend on Z_i.
- some alternatives: translog, quadratic, CES, nested CES

Specification, cont'd

Cobb-Douglas: $q_i = \beta_0 + \beta_\ell \ell_i + \beta_k k_i + \beta_z' z_i + \omega_i + \varepsilon_i$.

2. Form of productivity:

- ω_i is additively separable in log output
  - i.e., "Hicks-neutral"
  - C-D $\implies$ HN, but not the reverse.
- more generally, HN rules out unobservable differences in input quality
  - i.e., "factor-augmenting technical change"
  - if you want to study non-HN shocks, you cannot use C-D
  - however, often easier to allow more flexible substitutability/complementarity across inputs (e.g., translog, CES) than to abandon HN

Identification

The basic problem is that, if productivity is economically interesting to us, it is probably also interesting to the firm. But, econometrically, we would prefer the firm be uninterested in this.

Suppose we have resolved the data issues and now have data on

$$\{q_{it}, \ell_{it}, k_{it}\}$$

for a set of firms i over times $t \geq 0$.

Assume that

$$q_{it} = f(\ell_{it}, k_{it}, t, \omega_{it}) + \varepsilon_{it},$$

and we want f and $\{\omega_{it}\}_{i,t}$.

Economics suggests that ℓ_{it} and k_{it} probably depend on $(\omega_{is})_{s \geq 0}$ (**simultaneity**).

As a corollary, whether or not i shows up at all in the data at t ($k_{it} > 0$) also depends on $(\omega_{is})_{s \geq 0}$ (**selection**).

Example of simultaneity bias

Suppose labor is selected conditional on productivity.

- labor choice in Cobb-Douglas, with fixed wage w_{it} and output price p_{it}, is

$$\ell_{it} = \frac{1}{1-\beta_\ell}\left[\ln \frac{p_{it}\beta_\ell}{w_{it}} + \beta_0 + \beta_z' z_{it} + \omega_{it} + \beta_k k_{it}\right]$$

- OLS is clearly biased. ✓

Propagates through to capital:

$$\hat{\beta}_k = \beta_k + \frac{\sigma_{\ell\ell}\sigma_{k\omega} - \sigma_{\ell k}\sigma_{\ell\omega}}{\sigma_{\ell\ell}\sigma_{kk} - \sigma_{\ell k}^2}$$

However, sign of $\hat{\beta}_k - \beta_k$ unclear:

- if $\sigma_{k\omega} = \sigma_{k\ell} = 0$, no bias in $\hat{\beta}_k$
- if $\sigma_{k\omega} = 0$ and $\sigma_{\ell k} > 0$, downwards bias
- if $\sigma_{k\omega} < \sigma_{\ell\omega}$ and $\sigma_{\ell k} > 0$, likely to be downwards bias

More generally, direction of bias unclear for multifactor production functions.

Some solutions to simultaneity

Long history, as we recall from Marschak and Andrews (1944).

Three sources of identification:

1. experimental or quasi-experimental variation in input choices
    - e.g., variation in factor prices
2. statistical restrictions on productivity
    - e.g, fixed effects using panel data
3. economic restrictions on input choices
    - e.g., timing, optimality

Most "production function estimation" strategies rely on a combination of the three.

- best combination is **highly dependent** on the application of interest
- many of the issues in this literature stem from combinations gone awry

#1 Instrumental variables

A valid instrument is

- relevant shifter; i.e., correlated with ℓ_{it} or k_{it}
  - depends on the theory of factor demand in this industry
- excluded; i.e., uncorrelated with ω_{it} and ε_{it}
  - depends on source of productivity shocks in this industry

Virtues:

- consistency does not require one to specify the relationship between the inputs and instruments correctly
  - i.e., robust to various theories of factor demand
- can also help with measurement error
  - e.g., in capital (Collard-Wexler and de Loecker 2020)

Examples of instruments: local labor supply (Syverson 2004, §5(E)); input tariffs (de Loecker Goldberg Khandelwal Pavcnik 2016; Rubens 2020; Ruane and Peter 2020); rainfall, water-sharing rules (Rafey 2020); gov't supply quotas (Syverson et al 2020)

Example: Input price instruments

Input prices sometimes used as instruments (e.g., Raval 2019 RAND).

Exclusion rules out variation in input prices due to

- firm choices
  - e.g., locating on a downward-sloping input supply curve
- unmeasured differences in input quality or utilization
  - e.g., wage skill premiums or overtime
- shocks affecting productivity
  - e.g., geographic or temporal factors may also affect productivity
  - in practice, after we control for these shocks, can be **insufficient variation** in input prices

Output prices could also be used, but face similar issues.

NB. The "right" prices to use are the expected prices at the time the firm makes the input decision, rather than realized and possibly endogenous ones.

#2 Statistical restrictions

Fixed effects (Mundlak 1961). Assume

$$\omega_{it} = \omega_i + \nu_{it}$$

for all t, and that $\mathbb{E}[\nu_{it} + \varepsilon_{it} | \ell_{it}, k_{it}, z_{it}]$. Then you can estimate

$$q_{it} = \beta_0 + \beta_\ell \ell_{it} + \beta_k k_{it} + \beta_z' z_{it} + \omega_i + \varepsilon_{it}.$$

Questions:

- If ν_{it} does not affect input decisions, why is there variation in capital/labor choices over time?
    - in practice, signal-to-noise issues can lead to downwards bias in capital coefficients (Griliches and Mairesse 1998)
- And: what is the economic importance of a transient shock ν_{it} that the firm cannot respond to or learn from?

#2 Statistical restrictions, continued

Dynamic panel data solutions (Chamberlain 1983; Arellano and Bond 1991; Blundell and Bond 1999; Bond and Söderbom 2005)

- Take differences
- Use lagged values of capital and labor as instruments
- Did not, historically, appear to work well in practice because of weak instrument issues
- However, maybe worth revisiting (see, e.g., ACF pp. 2433–2435)

#3 Economic restrictions

First-order conditions. For example, suppose

- the firm takes wages w_{it} and the output price p_{it} as given
- and that its labor choice solves $\max_{L \geq 0} [p_{it} Q - w_{it} L]$.

The optimal labor choice is characterized by $p_{it} e^{\omega_{it}} \frac{\partial F}{\partial L} = w_{it}$. The LHS becomes

$$p_{it} e^{\omega_{it}} \frac{\partial F}{\partial L} \frac{Q_{it}}{Q_{it}} \frac{L_{it}}{L_{it}} = e^{\omega_{it}} \frac{\partial F}{\partial L} \frac{L_{it}}{e^{\omega_{it} + \varepsilon_{it}} F} \frac{p_{it} Q_{it}}{L_{it}}$$

and $e^{\omega_{it}}$ vanishes, so we obtain the factor share equation

$$\beta_\ell = \frac{\partial F}{\partial L} \frac{L}{F} = \frac{w_{it} L_{it}}{p_{it} Q_{it}} e^{\varepsilon_{it}} \tag{1}$$

which identifies β_ℓ for any ω_{it}.

The ratio $\frac{w_{it} L_{it}}{p_{it} Q_{it}}$ is the **expenditure** (or cost) **share**. Sometimes taken directly as the elasticity (e.g., Solow 1957, Syverson 2004). If F also has constant returns to scale ($\beta_k = 1 - \beta_\ell$), then these assumptions can identify all of F.

#3 Economic restrictions, cont'd

More generally, any optimality condition on a factor decision gives an additional moment we can use.

Suggests several ways forward:

- if the choice depends on ω, and in particular, if we can invert it for ω, then we can form a **control function**
- for example, from
  - investment (Olley and Pakes 1996)
  - materials (Levinsohn and Petrin 2003)

Other things you could do:

- identify an additional unobservable, e.g., labor-augmenting productivity, by inverting something like (1) (Doraszelski and Jaumandreu 2018)
- recover all of F nonparametrically, by integrating over moments like (1) (Gandhi Navarro Rivers 2016)
- identify markups, using a version of (1) that lets p depend on Q

Next week

Two required readings:

- Olley and Pakes 1996 ECTA
- Ackerberg Caves Fraser 2015 ECTA

Full citations in the syllabus.

No other homework.

See you next week!