Cost and Production[1]

Fall 2017

[1] These lectures notes are adapted from earlier lecture notes originally used by Jan de Loecker. All errors are mine.

Overview Lectures:

1. Introduction to production/cost analysis
2. Estimating Production Functions I (Olley Pakes)
3. Estimating Production Functions II (OP Extensions)
4. Examples of applications

Introduction

- The focus is on the estimation of production and cost functions as tools to analyze firm performance. In this respect it is an important part of IO, but got divided into subfield and into other fields.
- All of the models will rely on symmetric information, and we will discuss the primitive on the supply side: productivity.

Cost and Production: Intro

- **Reasons for interest**. Both functions are a crucial component of a general framework.
    1. Needed for any analysis of static equilibrium, or assumptions on what it is.
    2. Its form and properties will determine which model to use (e.g. learning by doing).
- Evaluating efficiency of industry output allocation and its relationship to changes in the operating environment or policy: cost complementarities and mergers, deregulation and efficient allocation, IRS and size, etc.
- Productivity analysis includes the analysis of both **efficiency of production** and **technological change**
    1. Analyzing effect of deregulation, tariff changes on productivity
    2. Externalities like R&D on productivity
    3. Technological change

Cost versus Production

- Cost data are typically proprietary (remember breakthrough demand systems with FOC on price setting) or hard to rely on due to various accounting standards. Explains little work on direct estimation of cost functions. Often the costs are backed out from the implications of a behavioral model like Nash in prices in a merger analysis.
- Ever growing literature on the use of production functions and comes from
  1. access to firm (plant) level panel data through various government and private agencies.
  2. need to evaluate efficiency impact of the major changes in the operating environment of firms. These will include privatization, deregulation, trade liberalization, health care, infrastructure in developing countries, etc.
- Therefore we will spend most of our time on methods developed to estimate production functions, but always remember that we can go to cost from there by adding behavioral model (simply $max\pi$).

Data Sources

- Computstat. Large traded (always multiproduct) firms. Input data is seldom broken down down by the product and lots of attrition due to mergers and acquisition, in addition to a selection on being a listed firm! Research question will determine whether this data is useful.
- Regulated industries, in particular electric utilities and water provide cost data but we can easily question reliability due to incentives to report higher costs. A nice overview and discussion is Wolak (2003).
- LRD and other census data allow for detailed balance sheet and additional information. Problem of accessing data through procedure of proposals and confidentiality.

- Alternatives to LRD and Computstat with middle ground in terms of access and data availability: AMADEUS provides firm-level (plant-level) data for long list of countries with annual balance sheet and ownership information.
- Special datasets on particular industries and or countries (via eg IMF, Worldbank, etc.).
- Fixed costs getting to know and cleaning the data, **measurement, definition, aggregation** are present when dealing with big micro datasets, let alone say finding out relevant institutional details of each market (industry).

Estimation Issues

- The availability of new datasets (micro data) lead to strong emphasis on substantive and technical issues.
  1. **Substantive Issues**. Micro data allows to distinguish the impact of the change on efficiency of the output allocation among firms, from productivity of individual producers and their correlates.
  2. **Technical Issues**. Data contains panel of firms in a given industry where we face i) large serially correlated differences in *productivity* among plants, ii) lots of entry and exit and this leads to econometric concerns.

Econometric issues

- Econometric Concerns are **simultaneity bias**, **selection bias** and **Omitted price bias**
  1. Simultaneity bias due to endogeneity of inputs. Firms with positive productivity shock will grow and need to buy more inputs (static versus dynamic inputs).
  2. Selection bias due to attrition in data where firms with (significant) negative productivity shocks exit the market. Especially when we study periods of drastic changes in the operating environment. This counts for both biases we will study
  3. Omitted Price variable bias, we will turn to this issue of price heterogeneity later in the class. Essentially comes from observing a version of sales or valued added in the data and not quantity produced (Q). Until we get to the explicit treatment of this let us call it sales generating production functions and sales per input.

Estimation: An Overview

- Let us start with a simple Cobb-Douglas production function with an input neutral efficiency term ω_i and an *iid* shock u_i

$$Q_{it} = L_{it}^{\alpha} K_{it}^{\beta} exp(\omega_{it} + u_{it}) \tag{1}$$

- where we keep in mind *two empirical facts*
  1. Large variance in productivity ω across firms
  2. Productivity of a given firm is highly correlated over time $E(\omega_{it}) = f(\omega_{it-1})$
- Comment on R^2 in time and cross section, and interpretation.

- We therefore expect inputs to depend on ω and therefore we need a model that determines it and how the inputs respond to those determinants. The property of estimators (α, β) depend on $E(X\omega)$ and therefore need to know how inputs respond to
    1. true technology and management differenes
    2. measurement error in inputs or output!
    3. variance in external factors (strikes, machine breakdowns, weather, etc.)
- Factors response depends on how easy to adjust(in order M, L, then K) [we will revisit this issue in depth]

Measurement Problems

- Major source is aggregation over products to input and output. Even if more disaggregated is available, we face problem of how to use it.
- How do you aggregate say labor to the firm level? This requires an assumption on the cost function, that is are there economies of scope, or cost synergies?
- Are the productivity shocks occuring at the level of the product or the firm, or is it an interaction. Important set of assumptions are needed to use more disaggregated data.

Measurement error in variables

- We almost never measure **output** produced Q but instead we observe sales divided by some price index, where the latter is mostly applied to all firms within a given industry. This is in particular problematic due to wide range of products that are being produced by a single firm and across firms within a sector (Bernard, Redding, Schott, 2005).
- **Labor** is usually in man-hours or simply number of full time employees. You would really like to control for type of labor, education, experience and specific skills when needed.
- **Materials** are often not thought to be too problematic but the same problems arise as in output. We rely on a deflated measure of materials, which requires all firms to pay same price. This does not allow for bulk discounts or quality differences in inputs.

Measurement ctd.

- **Capital**. We need to aggregate investment over various categories of capital (equipment, machinery, land and buildings, etc.) and correct for the appropriate depreciation. There are basically two ways of measuring capital: either directly via the book value (not free from problems) or through investment sequence (Perpetual Inventory Method) which requires assumption on initial capital stock. Measurement error large and implications for stylized facts - Collard-Wexler and De Loecker (2016).
- The application at hand will determine the need for explicit treatment of any of these problems. However, the first problem will never go away which lead to reinterpretation to sales per input measures for productivity. Except for a few studies this differences has been ignored and welfare interpretation are quite different. For instance how deregulation or tariff change impact sales per input or output per input are two different questions with different implications on the evaluation of the policy [see Module 2: Market power].

Standard Estimation Techniques and Problems

- I will go over the various techniques the literature has proposed (almost an historical overview, but not in all fields). I will also turn to the dual problem (cost function). We will think of the following simple production function

$$y_{it} = \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + \omega_{it} + u_{it} \tag{2}$$

1. Standard OLS regression of the production function
2. IV techniques
3. Panel data techniques: Fixed effects
4. Extended Panel data: dynamic panel Arrellano-Bond,and the like
5. Proxy - structural estimators (in depth).

OLS estimation

- This clearly leads to biased estimates of the production function as the inputs are choosen endogenously. Obviously, the bias will be more severe for those inputs that respond quickly to a productivity shock and thus freely adjustable (here materials, labor).
- This has lead to the use of value added production functions where we eliminate the need to estimate the coefficient on materials, however, we pin it down to be one and rely on Leontief.

IV techniques

- What are possible instruments Z_{it} in this case such that $E(lz) \neq 0$ and $E(z\omega) = 0$?
- Lagged values of inputs, but this needs assumptions on the time series properties of ω and there is a real question of robustness here (dynamic panel techniques).
- input prices (w, r)? Is there sufficient variance in them? Usually not much in the cross section for a specific industry - unless for regional differences but then we usually think of putting in region dummies to control for all kinds of unobserved regional specific differences. This now leads to all the variance coming from the time series and arguably common across firms and mixes up other common factors. Keep the variation in input prices in mind when we turn to structural models to identify coefficients, they will be ruled out!
- If the variance in input prices is exogenous - unrelated to productivity - say due to the fact you are not measuring quality aspects of inputs (think about education or experience) than quality of input is correlated with say wage and this is problematic.
- All of this lead to step away from instruments (think about cost function implications).

Panel data techniques

- The standard notion in panel data approaches is that the unobservable can be decomposed into a time invariant productivity shock and a shock.

$$\omega_{it} = \omega_i + \nu_{it} \tag{3}$$

- where $E_{t-1}(\nu_{it}) = 0$ and now assume that inputs are chosen at the beginning of the period. This would allow for correction of the time invariant productivity shock. This approach is very popular throughout applied economics, but often not very useful if interest is exactly in recovering ω_{it}.
- It is usefuly to rewrite the production function as the within OLS regression

$$\Delta y_{it} = \alpha_l \Delta l_{it} + \alpha_k \Delta k_{it} + \Delta \nu_{it} \tag{4}$$

- In the context of production functions it is particularly problematic
  1. **Noise-to-signal-ration** in inputs (in particular capital) is large when looking only at changes over time within a firm, and our timing models are at best poor approximations to reality [timing of inputs will be revisited]. We mostly do poor on capital coefficient as capital is a (quasi) fixed input in production and will magnify noise leading to a downward bias (to zero) of α_k, which you will find in almost every dataset.
  2. It assumes that productivity is time invariant, and we are typically interested in verifying its change due to change in environment.
  3. Problem of short panels in T and all asymptotics of the estimators are in N, but we kill that dimension.
-

Dynamic Panel Data

- This techniques are essentially a combination of the FE approach combined with IV approach, where the instruments are lagged inputs. In addition they allow for serial correlated (AR(1) process) productivity shocks.

$$y_{it} = \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + \gamma y_{it-1} + \omega_i + u_{it} \tag{5}$$

- Problem of fixed effect correlated with lagged dependent variable. Methods take first difference approach and rely on IV on (l, k) where lagged values are usually considered. Note there are tons of moments in the data to rely on for estimation and they will typicall perform a test of overidentifying restrictions. But, which ones work?
- I will leave it to you to read the Arellano and Bond, Blundell and Bond papers. The same issues mentioned above are still prevalent, especially the consequences of essentially taking a within transformation.

Issues with dynamic panel techniques for Productivity Functions

- Think about the selection problem in these estimators for a while, what is introduced by construction? These methods do not provide a correction for this. Sample size heavily reduced, by at least 3-4 time periods.
- Again, problem of close to zero estimates for capital with these methods.
- What if you want allow for a partly endogenous process for productivity? $\omega_{it} = g(\omega_{it-1}, action_{t-1})$.
- Ackerberg et al (2006) does show a high degree of similarity in the GMM approach of proxy estimators (later).

Cost Functions: Duality

- We can generate the cost function for our production function by considering profit maximization. In our case we get the following cost function ($MRTS = w/r$) in logs

$$c_{it} = c_0 + \frac{1}{\alpha+\beta} q_{it} + \frac{\alpha}{\alpha+\beta} w_{it} + \frac{\beta}{\alpha+\beta} r_{it} + \frac{1}{\alpha+\beta} \omega_{it} + u^*_{it} \qquad (6)$$

- The same issues are clearly present. Now it is about the $E(q_{it}\omega_{it})$, which requires treatment.
- Also note the impact of assumping CRS, data requirements on (C, Q, r)!
- There are some interesting applications studying regulated industries relying on Q being exogenous and dictated by regulator. Example Nerlove (1963) who studied Returns to scale in electricity supply. This paper was about an important matter and had the right data. Still a correlation though price setting of regulator based on ω. Subsequent literature on this moved to cost efficient allocation and market power.

Requirements

- Requirements to use the dual:
  1. all inputs can be varied costlessly, problem for quasi and fixed inputs. Think about the equilibrium conditions (FOC need an adjustment cost).
  2. Demand shifters to instrument for Q or explicity orthogonal assumption
  3. good variation in input prices
- In order to rely on duality we need competitive input markets and very specific input demand functions. Given this we might add information to problem and jointly estimate the input demand functions with the cost funcion. However, its leads to overdetermination and we need more, e.g. measurement error on inputs.

Factor shares

- If you are taking cost function approach you might as well run factor share through framework.
- Relying on observing costs: wage bill, material expenditures and user cost of capital (or CRS assumption) allows you to back out productivity under perfect competition in both input and output market.
- We can relax output market with the Hall(.) approach (revisit in market power lecture).

Functional forms

- Cobb-Douglas is predominant functional form in applied work with obvious restrictions on substitutions among inputs, i.e. elasticity of one. And all of the technical change is neutral to inputs. Think about skill-biased technological change.
- Returns to scale are constant across firms and over time. This has implications on optimal scale and size of plants in an industry.
- Multi-product firms are not allowed to experience economies of scope or benefit from cost synergies. This *the* argument for mergers on the part of firms. This is fundamentally a question whether the cost function is additive separable in the costs of producing different outputs (products).
- In applied work the data will constrain the functional form we can work with and allows identification.
- Some articles on the reading list discuss these issues.

Subaddivity

- This is curcial in the assesment of regulation and antitrust (natural monopolies and cost synergies leading to lower prices for consumers).
- Note that this concept is defined for a given output and could be true at one level and not at another.
- For a single product the cost function $C(.)$ is subadditive at a particular Q if and only if the following holds for all vectors $(q_1, q_2, ...)$ s.t. $\sum_i (q_i) = Q$.
$$C(Q) < \sum_i C(q_i) \tag{7}$$
- For a single good production, we just need a decreasing average cost curve until Q.
- For multi-product firms just treat Q as a vector now. There is no definition for production functions with multi-products. E.g. for 2 products: $C(Q_1, Q_2) < C(Q_1, 0) + C(0, Q_2)$
- Read Evans and Heckman (1983) for an example of such a test.

Multi-product producers: missing analysis

- Old literature on multi-product cost functions (see above), but very few analysis or implications thereof in production function literature.
- Hard problem of input aggregation and no such thing as a production function, but a *transformation*, i.e. $q_{ijt} = \boldsymbol{x}_{it}\beta + \boldsymbol{q}\gamma_{i,-j,t}\gamma + \omega_{it}$.
- Multi-product firms are large part of production anywhere, so more analysis is needed.
- Think about $\#$ 1 line of defense for merging parties: **creating synergies**, how do we evaluate this in a convincing way?
- Some issues worth keeping in mind: synergies in physical or cost space?; product or firm specific productivity?; cost or production approach?

Selection bias in estimation of PF

- Remember our discussion on attrition
- It is clear that we need a model of entry/exit in addition to input demand as a function of productivity.
- correct model for exit will depend on the data at hand (small firms, large firms)
- Remember the importance of this control especially when we study a drastic change in the environment of firms.

Need for underlying model

- All of the above shows the need for a model of firms thinking of inputs (L,M, K or investment) and market participation.
- In addition we need to model heterogeneity of productivity across producers and serial correlation of these differences (persistence).
- Theory should help us tackle this problem in a general way and will imply different datasets and questions to have different solutions. This is not an off-the-shelf approach!
- Next we introduce a flexible empirical framework to allow for exactly:
    1. heterogeneity in efficiency,
    2. persistence (unobserved) productivity,
    3. endogenous inputs and survival choices
    4. potential to integrate with price heterogeneity (last class)

Estimating Production Functions

Semi-structural approaches: Olley-Pakes, Levinsohn-Petrin and Ackerberg-Caves-Frazier.

Estimating Production Functions

- Back to basics before *plunging in*.
- Overview of recent, structural, approach to production function estimation: Olley-Pakes, Levinsohn-Petrin and Ackerberg-Caves-Frazier.
- For now, let us assume we have correct units, either by assumption or by measurement.

Factor share approach

- Often we are interested in recovering a measure of productivity (ω, TFP) as a residual from the production function:

$$y_{it} = \alpha_l l_{it} + \alpha_k k_{it} + \omega_{it} + \epsilon_{it} \tag{8}$$

- Under the classical setup of perfectly competitive output and input markets, constant returns to scale and static labor choices, we **do not** need to estimate anything – i.e. the so-called factor share approach delivers:

$$\alpha_l = \frac{WL}{PQ}$$
$$\alpha_k = 1 - \alpha_l$$

- Keep this in mind as we go into structural approaches: it has to be about departures from either CRS and or input flexibility. See Asker, Collard-Wexler and De Loecker (2014, *JPE*) for an application to WBES.

- In this class I will go over the Olley and Pakes (1996) article in detail and refer to the various tables. This article will teach us how to correct for the selection bias and the simultaneity bias while relying on a model of industry dynamics.
- The underlying stucture of this model will turn out to be very important if you want to expand, modify and adjust it to settings with specific questions such as R&D investment, export status, dynamic inputs, product differentiation, etc.

Simultaneity, Selection and Reallocation

- **Telecommunication equipment example**. Some facts in this industry: Major restructuring of the industry since late 1960's due to two related changes.
    1. Technological change which brought many new products, both for delivering phone services (digital switches, fiber optics, etc.)and for using the phone lines (fax, modem, etc.).
    2. Deregulation of the industry.
- Prior History. Monopoly in services and essential monopoly in equipment because of Equipment procurement practices of Western Electric (AT&T subs) near monopoly in equipment used to route and send phone calls. Monopoly for all equipment that needed the public network by making it illegal to connect without consent of AT&T.
- This implied that potential entrants had to bring in their own network, which was not useful without service (and even illegal in many states).

Deregulation

- Carter phone decision (1968) and FCC sets up registration and certification program and the entry of new products (modems, fax machines, etc.) lead to new companies entering (Table 1).
- Biq equipment remained in the hands of Western Electric. Follows further deregulation and establishes *none could own an equipment manufacturer*. Also forcing ATT to lease out their lines to other long-distance carriers (competitive effects).
- Immediate impact on Bell system purchases. Table 2.
- Industry remains under change, recent changes in phone services through cable providers (Timewarner, Verizon, etc.).

TABLE I

Characteristics of the Data

Year	Plants	Firms	Shipments (billions 1982 $)	Employment
1963	133	104	5.865	136899
1967	164	131	8.179	162402
1972	302	240	11.173	192248
1977	405	333	13.468	192259
1982	473	375	20.319	222058
1987	584	481	22.413	184178

Data and analysis

- The data comes from Census and combination of annual and census data.
- Raw data shows importance of entry. Table 3. Here the unbalanced panel aspect is huge. If we would not allow for entrants, we would eliminate around 79 percent of plants with around 30 percent shipments in 1987.
- Importance of exit in Table 4. Not allowing for exit would imply 60 percent of plants and 40 percent of shipments.
- Large changes in the size of continuing establishments. There was clearly a dramatic change in the operating environment of firms and different firms reacted differently to this change.
- **Analysis**. The goal of this paper is to analyze the change in industry productivity that accompanied regulation. Therefore we need production function coefficients.

TABLE II

BELL COMPANY EQUIPMENT PROCUREMENT (PERCENT PURCHASED FROM WESTERN ELECTRIC)

1982	1983	1984	1985	1986[E]
92.0	80.0	71.8	64.2	57.6

[E] Estimated for 1986.
Source: NTIA (1988, p. 336, and discussion pp. 335–337).

TABLE III

ENTRANTS ACTIVE IN 1987

	Number	Share of Number Active in 1987 (%)	Share of 1987 Shipments (%)	Share of 1987 Employment (%)
Plants: New since 1972	463	79.0	32.8	36.0
Firms: New since 1972	419	87.0	30.0	41.4
Plants: New since 1982	306	52.0	12.0	13.5
Firms: New since 1982	299	60.1	19.4	27.5

TABLE IV

INCUMBENTS EXITING BY 1987

	Number	Share of Number Active in Base Year (%)	Share of Shipments in Base Year (%)	Share of Employment in Base Year (%)
Plants active in 1972 but not in 1987	181	60.0	40.2	39.0
Firms active in 1972 but not in 1987	169	70.0	13.8	12.1
Plants active in 1982 but not in 1987	195	41.2	26.0	24.1
Firms active in 1982 but not in 1987	184	49.1	17.3	16.1

Productivity Analysis

- Analyzing productivity requires the estimation of a production function for this industry (given data constraints here sales generating production functions).
- Large serially correlated differences in active plants to begin with.
- Deregulation and technological change impacted the implication of this productivity distribution and induced entry/exit and changes in size of continuing.
- The error term in the production function therefore changed over time in particular way: as a determinant of the exit decision (generates selection problem) and as a deteminant of the input decisions (simultaneity problem).
- This paper provides us with a model to correct for both in a dynamic consistent manner. It needs to be dynamic (to allow for exit) and rich enough to allow for differences among firms, idiosyncratic sources of uncertainty and entry.

Underlying model of industry dynamics

- The dynamic model is based on a Hopenhayn-Rogerson and Ericson and Pakes framework. The estimation procedure requires an insight into these models (Details not covered but urge you to read them).
- **Definitions**
  1. Firm specific state variables: (k_t, ω_t, a_t)
  2. Market structure: value for these triple for all active firms
  3. Profits map own state variables and market structure into dollars.

behavior of agents

- Beginning of each period agent decides whether to continue operation and if so choose variable factors and investment (to increase next period's capital stock), where t denotes market structure.

$$V_t(k_t, \omega_t, a_t) \equiv \max(\Phi, \pi_t(.)) \tag{9}$$

$$\pi_t(.) = \pi_t(k_t, \omega_t, a_t) + \sup_{i_t \geq 0}(-c(i_t) + \beta E(V_{t+1}(k_{t+1}, a_{t+1}, \omega_{t+1}) | J_t)) \tag{10}$$

- Note that the expectation is over the market structure and the firm's own state in the future. For the model to make sense the probability distribution generating those market structures must be consistent with the behavior of multiple agents solving problems like this (Markov perfect equilibrium game structure).

state variables

$$K_{t+1} = K_t(1-\delta) + I_t, a_{t+1} = a_t + 1 \tag{11}$$

$$F_\omega \equiv F(.|\omega), \omega \in \Omega \tag{12}$$

$$\omega_{it+1} = g(\omega_{it}) + \xi_{it+1} \tag{13}$$

MPE in payoff relevant variables

- The equilibrium wil deliver an exit and investment rule. They have the following properties.
- **Exit rule** is $\chi_t(k_t, a_t, \omega_t)$
  1. $\chi = 1 \Leftrightarrow \omega_t \geq \underline{\omega}_t(k_t, a_t)$
  2. Note that the function is indexed by t
  3. Note that $\underline{\omega}_t(.)$ falls in k, alternatively ?
- **Investment rule** is $i_t(k_t, a_t, \omega_t)$
  1. Note that $i_t(.)$ increasing in ω if $i > 0$
  2. Note that this function also is indexed by t.

Model building options

- Make enough assumptions to be able to compute policy functions (χ_t, i_t) directly. Not in this approach.
- Here without precise functional forms and therefore non-parametric.
- We face tradeoffs of keeping a very robust form of profits and entry/exit rule, versus, using more assumptions and increased computational burden to answer more questions, but might enhance efficiency of estimates.

Estimation of OP model

- They start out with a homogeneous good producer with a Cobb-Douglas (value added) production function (in logs)[2]

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \omega_{it} + \eta_{it} \tag{14}$$

- Here η is defined to be part of the error that is not known when input decision are made, whereas ω is crucially observed by the firm (not by the econometrician).
- Bias in coefficient using OLS: simultaneity especially in labor (no material here!).
- Selection bias due to $E(\omega_t|\omega_t > \underline{\omega}(k_t), \omega_{t-1}, \chi_t = 1)$. Conditional on ω_{t-1} this function is decreasing in k_{t-1}. I.e. the value function is increasing in both, so if k is higher we will continue with a lower ω. This would lead to a negative capital coefficient bias.

[2] The model has age as an input as well, I will drop this and just keep it in investment proxy to show that one can include other observed variables as states to further control.

Estimating labor: stage 1

- OP work on condition that we can invert investment policy function that is solution to the dynamic problem outlined before. There is actually a bit of work to proof this (see Pakes, 1994).

$$i_t = i_t(k_t, \omega_t, a_t) \tag{15}$$

$$\omega_t = h_t(i_t, k_t, a_t) \tag{16}$$

- where we now have a model that proxies (controls for) productivity and note the time subscript, i.e. market structure. Now we can substitute this into the production function and collect terms on capital, investment into non parametric function $\phi_t(.)$.

$$y_{it} = \beta_l l_{it} + \phi_t(i_{it}, k_{it}, a_{it}) + \eta_{it} \tag{17}$$

  where $\phi_t(i_t, k_t, a_t) \equiv \beta_0 + \beta_k k_{it} + h_t(i_{it}, k_{it}, a_{it})$

- We simply estimate this as a partial linear model (Robinson, 1988) and get estimate for β_l and ϕ_{it}. Practically, just use polynomial expansion in state variables or use LWLS.

Selection control: stage 2

- We now want explicitly control for non random exit of firms, i.e. firms with lower productivity (conditional on capital stock) have higher probability to exit the market. Or

$$Pr(\chi_{t+1} = 1|\underline{\omega}_{t+1}(.), J_t) = Pr(\omega_{t+1} > \underline{\omega}_{t+1}(.)|.) \tag{18}$$

$$= 1 - F(\underline{\omega}_{t+1}(.)|\omega_t) \equiv P_t(i_t, a_t, k_t) = P_{it} \tag{19}$$

- This is again a non parametric function of the exit probability in the state variables (generated by the model) and we can estimate this using kernel estimation techniques or probit in polynomial of state variables. This will generate estimate for P_{it}.

Estimating fixed input coefficients β_k: stage 3

- We now have the following estimates $(b_l, \widehat{\phi}_{it}, \widehat{P}_{it})$ and let us consider one period ahead, where we realize that this is only observed for surviving firms (Exit rule!). This implies that we know that

$$E(y_{t+1} - \beta_l l_{t+1} | a_{t+1}, k_{t+1}, \chi_{t+1} = 1) = \beta_0 + \beta_k k_{t+1} + E(\omega_{t+1} | \omega_t, \chi_{t+1} = 1) \tag{20}$$

- where now we will use the non parametric evolution of the productivity process, i.e. an AR(1) for ω would be a special case. We have that

$$E(\omega_{it+1} | \omega_t, \chi_{t+1} = 1) = \tag{21}$$

$$\int_{\underline{\omega}_{t+1}} \omega_{t+1} \frac{F(d\omega_{t+1} | \omega_t)}{\int_{\omega_{t+1}} F(d\omega_{t+1} | \omega_t)} \tag{22}$$

$$\equiv g(\underline{\omega}_{t+1}, \omega_t) = g(P_{it}, \phi_{it} - \beta_k k_{it}) \tag{23}$$

3rd stage

- Now using the Markovian assumption and its implication for expected productivity into the production function one period ahead, subtracting the know variation in labor (or whatever the input is you are estimating in the first stage).

$$y_{t+1} - b_l l_{t+1} = c + \beta_k k_{t+1} + g(\widehat{P}_{it}, \widehat{\phi}_{it} - \beta_k k_{it} - \beta_0) + \xi_{it+1} + \eta_{it+1} \quad (24)$$

- where crucially ξ is the innovation in the Markov process of productivity, this is exactly what forces us to use a first stage to estimate labor since $\omega_{it} = g(\omega_{it-1}) + \xi_{it}$ and induces correlation between labor and productivity, when labor is optimally chosen each period.
- Labor can respond to ξ and therefore l is a function of it, that is why we need to subtract it out.
- Questions:
  1. Think about what if labor has adjustment cost and is a dynamic input.
  2. What if doing RD impacts future productivity (probabilistic vs deterministic)?

Implementing 3rd stage

- We will estimate this equation using NLLS while using a series approximation, say order 4 (typically you expand until no change).
- Recall that from the first stage we have an estimate for $\phi_{it} = y_{it} - b_l l_t$, which no longer includes η due to estimation. Under the model's structure measurement error is purged.
- Estimating capital coefficient (or any coefficient of an input that has dynamics!) on the following

$$y_{t+1} - b_l l_{t+1} = \beta_0 + \beta_k k_{t+1} + \sum_{r=0}^{4-s} \sum_{s=0}^{4} \widehat{P}_{it}^{s} (\widehat{\phi}_{it} - \beta_k k_{it})^r + \xi_{it+1} + \eta_{it+1} \quad (25)$$

- Apply NNLS with OLS starting values to search for β_k, in stata *nl* for instance.

Simplification: linear productivity process

- Useful exercise is to consider a Martingale in productivity and recover parameters with two OLS equations, and see where identification comes from.

Alternative estimation: GMM

- Alternatively we can estimate the capital coefficient with GMM and rely on only one moment, while adding and testing for overidentifying restrictions.
- Introduce method here as will be useful for later (LP, ACF, DL)
- From first stage we have an estimate of productivity given parameter β_k

$$\omega_{it}(\beta_k) = \phi_{it} - \beta_k k_{it} \tag{26}$$

- Relying on productivity evolution process, $\omega_{it+1} = g(\omega_{it}) + \xi_{it+1}$, we can recover $\xi_{it+1}(\beta_k)$ by non parametrically regressing ω_{it+1} on ω_{it}.
- Now we have $\xi_{it+1}(\beta_k)$ and can form moments to identify β_k

$$E\left\{\xi_{it+1}(\beta_k)\begin{pmatrix} k_{it+1} \\ k_{it} \end{pmatrix}\right\} = 0 \tag{27}$$

TABLE VI

ALTERNATIVE ESTIMATES OF PRODUCTION FUNCTION PARAMETERS[a] (STANDARD ERRORS IN PARENTHESES)

Sample:	Balanced Panel		Full Sample[c,d]						
								Nonparametric F_ω	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimation Procedure	Total	Within	Total	Within	OLS	Only P	Only h	Series	Kernel
Labor	.851	.728	.693	.629	.628			.608	
	(.039)	(.049)	(.019)	(.026)	(.020)			(.027)	
Capital	.173	.067	.304	.150	.219	.355	.339	.342	.355
	(.034)	(.049)	(.018)	(.026)	(.018)	(.02)	(.03)	(.035)	(.058)
Age	.002	−.006	−.0046	−.008	−.001	−.003	.000	−.001	.010
	(.003)	(.016)	(.0026)	(.017)	(.002)	(.002)	(.004)	(.004)	(.013)
Time	.024	.042	.016	.026	.012	.034	.011	.044	.020
	(.006)	(.017)	(.004)	(.017)	(.004)	(.005)	(.01)	(.019)	(.046)
Investment	—	—	—	—	.13	—	—	—	—
					(.01)				
Other Variables	—	—	—	—	—	Powers of P	Powers of h	Full Polynomial in P and h	Kernel in P and h
# Obs.[b]	896	896	2592	2592	2592	1758	1758	1758	1758

Estimation Results

- Comparing various samples and methods to test our intuition and sign of the bias predicted by theory.
- Balanced versus unbalanced sample: big effect on capital coefficient! Important takeaway.
- FE regression confirms our prior and this is a typical result for Prod Functions. Implications for returns to scale!
- Moving to unbalanced panel corrects part of the selection problem but leaves simultaneity untouched.
- Introduction of control function $\phi(i, k, a)$ has expected effect and reduces labor coefficient.
- Series and kernel, estimates go in the right direction and have implications for RTS and productivity estimates.

Results Ctd.

- Standard errors are somewhat bigger but still relatively tight. Note, you run bootstraps over entire procedure.
- One interesting robustness check on the inversion. Is investment enough to control for heterogeneity (essentially about scalar unobservable). If not then there is an error in the third stage equal to $(b_l - \beta_l)l_{t+1}$. We can check this by including l_t in last stage and test whether γ_l in that regression is different from one.

$$y_{t+1} - b_l l_{t+1} = c + \beta_k k_{t+1} + \gamma_l l_t + g(.) + \xi_{it+1} + \eta_{it+1} \quad (28)$$

- Furthermore, do robustness over sample periods, groups of firms, etc.

Implications for productivity

- Now we can measure productivity either in logs or levels as

$$p_{it} \equiv \exp\left(y_{it} - b_l l_{it} - b_k k_{it}\right) \tag{29}$$

- Aggregate productivity is constructed as a weighted average over firms using market share as weights.
- Implications are different when relying on balanced panel and different estimates.
- Rely on estimates to analyze aggregate productivity growth sources: individual firms are getting better or a reallocation of market share towards more productive firms? This is where the micro data kicks in.

Reallocation

- Evidence is shown that productivity increase was due to either average productivity increase or reallocation of fixed factors towards more productive firms. This comes from analysis on cost efficiency before and after regulation. Results indicate that after regulation, not a better cost-efficient allocation *conditional on output produced and fixed factors (Capital, productivity)*. Therefore the next step.
- Reallocation towards more productive firms versus technical change (at the firm level). Let us consider the aggregate productivity index and now decompose it into average productivity and a covariance term between market share and productivity (look at this term over time to find out whether there was any reallocation).

$$p_t \equiv \sum_i s_{it} p_{it} \tag{30}$$

$$p_t = \overline{p_t} + \sum (s_{it} - \overline{s}_t)(p_{it} - \overline{p}_t) \tag{31}$$

TABLE XI

DECOMPOSITION OF PRODUCTIVITY[a]
(EQUATION (16))

Year	p_t	$\bar{p}_t$	$\Sigma_i \Delta s_{it} \Delta p_{it}$	$\rho(p_t, k_t)$
1974	1.00	0.90	0.01	−0.07
1975	0.72	0.66	0.06	−0.11
1976	0.77	0.69	0.07	−0.12
1977	0.75	0.72	0.03	−0.09
1978	0.92	0.80	0.12	−0.05
1979	0.95	0.84	0.12	−0.05
1980	1.12	0.84	0.28	−0.02
1981	1.11	0.76	0.35	0.02
1982	1.08	0.77	0.31	−0.01
1983	0.84	0.76	0.08	−0.07
1984	0.90	0.83	0.07	−0.09
1985	0.99	0.72	0.26	0.02
1986	0.92	0.72	0.20	0.03
1987	0.97	0.66	0.32	0.10

[a] See text for details.

reallocation ctd

- Little increase in average productivity, but allocation improves dramatically after regulation.
- Conclude. Reallocation of capital and shift in production to more productive plants after divesture underlies the increase in productivity that followed regulation.
- Caveats. Read this section carefully! Especially on the use of scalar unobservable controlled by simple investment. That is only rely on investment to control for unobserved heterogeneity across producers.

Reallocation and IO topics

- Identifying precise mechanism underlying the reallocation of resources is hard and very few papers here.
- What is the role of market structure and market power in the process?
- Dynamics: entry and exit requires a different decomposition, i.e. cross sectional covariance over time is obtained by summing over different set of firms due entry/exit process. See Melitz and Polanec (2012).
- Collard-Wexler and De Loecker (2015) focus on technology and competition using detailed panel data covering 40 years and directly measuring technology by producer, in the US Steel industry.

Follow-up work

- This method became standard in IO, trade, development, etc. However, there are several technical issues that were raised and addressed in the literature.
- Investment data is hard to come by, or simply lumpy in nature. For estimation procedure we need $i > 0$ in order to rely on inversion result. (Levinsohn and Petrin, 2003)
- What if all inputs have some adjustment costs, can we still estimate the coefficients? (Ackerberg, Caves and Frazier, 2006). Towards a more general treatment on identification.
- What if we believe there to be more than one source of heterogeneity, can we still invert? (Ackerberg and Pakes, 2006).
- Other important decision variables in a given research question will lead to inclusion of another state or control variable. (see applications Van Biesebroeck, 2005 and De Loecker, 2007).
- Sales/Value added used for quantity with reinterpretation. Model structure can be accommodated to handle revenue explicitly and incorporate demand shocks (De Loecker, 2011).

Focus on identification:

- Identification of parameters in OP/LP framework
- Identification of parameters under alternatives or extensions such as:

1. Endogenous productivity process: R&D, LBE, LBD, etc.
2. Functional form of production function
3. Relaxing scalar assumption of unobservable