Productivity Learning

Chunxiao Jing *

February 2014

*credit to https://www.aeaweb.org/content/file?id=3015

1 Introduction

- Production function solves:
  - What is the level of returns to scale?
  - How do input coefficients on capital and labor change over time?
  - How does adoption of a new technology affect production?(Shock)
  - How much heterogeneity is there in measured productivity across firms, and what explains it?(IO)
  - How does the allocation of firm inputs relate to productivity?(allocation)
- Production function also solves:
  - dynamic models of industry evolution
  - evaluation of firm conduct(e.g.collusion)
- Here, we will work with two input C-D production function:

  $$Y_i = e^{\beta_0} K_i^{\beta_1} L_i^{\beta_2} e^{\epsilon_i}$$

  where i indexes firms, K_i is units of labor, L_i is units of output. $(\beta_0, \beta_1, \beta_2)$ are parameters and ϵ_i captures unobservables that affects output (e.g. weather, soil quality, management quality)
- Take natural logs:

  $$y_i = \beta_0 + \beta_1 k_i + \beta_2 l_i + \epsilon_i$$
- The function can be extended to additional inputs: R&D, dummies representing discrete technologies, different types of labor/capital, intermediate inputs

2 Endogeneity Issues

- Problem:
  1. Inputsk_i,l_i are typically choices variable of firms
  2. Firms choose inputs to maximize profits, and hence will often depend on unobservables ϵ_1.

3. The dependence depends on what the firm knows aboutϵ_i when they make these input choices.
4. Example: Suppose a firm operating in perfectly competitive output and input markets(with respective prices p_i, r_i,w_i), and perfectly observes ϵ_i before optimally choosing inputs.
   Profit maximization problem is :

$$\max_{K_i,L_i} p_i e^{\beta_0} K_i^{\beta_1} L_i^{\beta_2} e^{\epsilon_i} - r_i K_i - w_i L_i$$

- FOC of the optimal function implies that K_i and L_i (k_i and l_i) will depend on ϵ_i.
  Intuition: ϵ_i which can be seen as productivity positively affects marginal product of inputs.

- Endogeneity issue: one cannot estimate

$$y_i = \beta_0 + \beta_1 k_i + \beta_2 l_i + \epsilon_i$$

  using OLS because of the correlation. The coefficient would be positively biased.

- In some more complicated models(e.g. non-perfectly competitive output or input markets, ϵ_i only partially observed), except the special case where the firm has not knowledge of ϵ_i when choosing inputs (in this case, both firms and economists don't know the productivity, and thus the productivity is exogenous)

- k_i is less correlated with ϵ_i than l_i is: k_i is a "less variable" input than l_i because one might expect the firm to be less knowledge about ϵ_i when choosing k_i. This is because k_i is decided by firms in previous period, in which firm has little knowledge of ϵ_i.

- WLOG,let's think about ϵ_i above having 2 components (ω_i and ϵ_i), so that we can rewrite the equation as:

$$y_i = \beta_0 + \beta_1 k_i + \beta_2 l_i + \omega_i + \epsilon_i$$

  where:

  - ω_i is an unobservable that is predictable (or partially predictable) to the firm when it makes its input decisions

- ϵ_i is an unobservable that the firm has no information about when making input decisions. ϵ_i could also represent measurement error in output.
- in this formulation, ω_i causes the endogeneity problem, which is called the "productivity shock".

3 Traditional Solutions

There are two traditional solutions to endogeneity problems : instrumental variables and fixed effects model.

3.1 Instrumental Variables

- Basic idea: Find "instruments" that are
  - correlated with the endogenous inputs k_i and l_i
  - not directly determine y_i
  - not correlated with ω_i and ϵ_i
- Consider input and output prices w_i, r_i and p_i.
  - These prices will affect firms' optimal choices of k_i and l_i
  - These prices are excluded from the production function as they do not directly determin output y_i conditional on the inputs
- The problem is when will be the case that w_i, r_i and p_i are not correlated with the productivity shock ω_i.
- One key issue is the form of competition in input and output markets:
  1. p_i will be negatively correlated with ω_i, invalidating p_i as an instrument:
     If output markets are imperfectly competitive (i.e. firms face downward sloping demand curves), then a higher ω_i will increase a firm's output supply ,driving p_i down.

2. w_i and/or r_i may be positively correlated with ω_i, invalidating them as instruments:
   If input markets are imperfectly competitive(i.e. firms face upward sloping supply curves), then a higher ω_i will increase a firm's input demand, driving w_i and/or r_i up.

- To use these instruments, firms must be operated in perfectly competitive input or output markets. Typically, this is more believable for input markets than for output markets.

- Problem: IV solutions haven't been broadly used in practice because:

  1. one needs data on w_i and r_i.
  2. there is often very little variation in w_i and r_i across firms (often there is a real question of whether observed variation operate in different input markets?)
  3. one often wonders whether observed variation in e.g $w-i$ actually represents firms facing different input prices, or whether it represents things like variation in unobserved labor quality (i.e. the firm with the higher w_i is employing workers of higher quality.)
     If latter, then w_i is not a valid instrument.

- While there might be "true" variation in input prices across time, this is usually not helpful, because if one has data across time, one often wants to allow the production function to change across time (i.e. $f_t(k_{it}, l_{it})$), e.g.

$$y_{it} = \beta_{0t} + \beta_1 k_{it} + \beta_2 l_{it} + \omega_{it} + \epsilon_{it}$$

- To sum up, if one can find a market where there is convincing exogenous input price variation, IV approach is probably more convincing than the approaches in the rest of this note.

- Notes:

  1. Randomized experiments: either directly manipulating inputs, or manipulating input prices.
  2. In this note, an implicit "homogeneous treatment effects" assumption is made. A heterogeneous treatment effects model would be

$$y_i = \beta_0 + \beta_{1i} k_i + \beta_{2i} l_i + \omega_i + \epsilon_i$$

This affects the interpretation of IV estimators, e.g. Heckman and Robb (1985),Angrist and Imbens(1994, Ecta)

3. If there are unobserved firm choice variables in ω_i, it becomes quite hard to find valid instruments, even with the above assumption.

3.2 Fixed Effects

- The approach relies on having panel data on firms across time, i.e.

$$y_{it} = \beta_0 + \beta_1 k_{it} + \beta_2 l_{it} + \omega_{it} + \epsilon_{it}$$

- assume:

1. ϵ_{it} is independent across t(intuition: ϵ_{it} is not predictable by the firm when choosing k_{it} and l_{it})
2. the productivity shock is constant over time (fixed effect assumption), i.e.

$$w_{it} = w_i$$

- Then one can either mean difference

$$y_{it} - \bar{y}_i = \beta_1(k_{it} - \bar{k}_i) + \beta_2(l_{it} - \bar{l}_i) + (\epsilon_{it} - \bar{\epsilon}_i)$$

or first difference

$$y_{it} - y_{\bar{it-1}} = \beta_1(k_{it} - k_{\bar{it-1}}) + \beta_2(l_{it} - l_{\bar{it-1}}) + (\epsilon_{it} - \epsilon_{\bar{it-1}})$$

Since the problematic unobservable ω_{it} have been difference out of these expressions (recall that we have assume that the $\epsilon_{it} areuncorrelatedwithinputchoices$), these equation can be estimated with OLS.

- Problems:

1. ω_{it}=ω_i is a strong assumption
2. These estimators often produce strange estimates. In particular, they often generate very small (or even negative) capital coefficients. Perhaps this is due to measurement error in capital.

- Other notes:

- Mean Difference Approach vs First Difference Approach:
  1. The mean difference approach requires all the input choices to be uncorrelated with all the ϵ_{it} (strict exogeneity).
  2. The first difference approach only requires current and lagged inputs to be uncorrelated with current and lagged ϵ_{it}. Using k_{it-1} and l_{it-1} (or other lags) as instruments for (k_{it}-k_{it-1}) and (l_{it}-l_{it-1}), one can allow current inputs to be arbitrarily correlated with past ϵ_{it} (sequential exogeneity)
- Panel data approach can be extended to richer error structures (Arellano and Bond (1991, ReStud), Arellano and Bover (1995, JoE), Blundell and Bond (1998, JoE; 2000, ER), Arellano and Honore (2001, Handbook)), e.g.

$$\omega_{it} = \rho\omega_{it-1} + \xi_{it}$$

  or

$$\omega_{it} = \alpha_i + \lambda_{it}$$

  where

$$\lambda_{it} = \rho\lambda_{it-1} + \xi_{it}$$

3.3 First Order Conditions

- The approach to estimating production function is based on information in the FOCs of optimizing firms.
- For a firm operating in perfectly competitive input and output markets, static cost minimization implies that

$$\frac{\partial Y}{\partial L}\frac{L}{y} = \frac{wL}{pY}$$

  i.e. the output elasticity w.r.t an input must equal its (cost) share in revenue.
- In a C-D context, these output elasticities are the production function coefficients β_1 and β_2, so observations on these revenue shares across firms could provide estimates of the coefficients.

- Note that r can often be assumed known (interest rate or sth. equivalent) and often one directly observes wL and pY (rather than L and Y - i.e. labor input and output are measured in terms of dollar units (that are implicitly assumed to be comparable across firms))

- Problem:

  1. This assumes static cost minimization - i.e. it assumes away dynamics, adjustment costs, etc. At the very least, we often think about the capital input being subject to a dynamic accumulation process, e.g. $K_{it} = \delta K_{it-1} + i_{it-1}$
  2. There are additional terms when firms are not operating in perfectly competitive markets, e.g. when firms face downward sloping demand curve
  $$\frac{\partial Y}{\partial L}\frac{L}{Y} = \mu\frac{wL}{pY}$$
  $$\frac{\partial Y}{\partial K}\frac{L}{K} = \mu\frac{rK}{pY}$$
  where $\mu = \frac{p}{mc}$, i.e. percentage markup. Note that profit maximization implies $\frac{p}{mc} = \frac{\epsilon}{1+\epsilon}$, where ϵ is the elasticity of demand. So, for example, one could still identify production coefficients using this method if the elasticity of demand was known (Hsieh and Klenow (2009, QJE)). Or one might be able to identify both with additional restrictions, e.g. CRS (related to Hall (1988, JPE))