Estimating Production Functions

Introduction

Paul Schrimpf

UBC
Economics 567

January 13, 2022

Section 1

Introduction

Why estimate production functions?

- Primitive component of economic model
- Gives estimate of firm productivity – useful for understanding economic growth
    - Stylized facts to inform theory, e.g. Foster, Haltiwanger, and Krizan (2001)
    - Effect of deregulation, e.g. Olley and Pakes (1996)
    - Growth within old firms vs from entry of new firms, e.g. Foster, Haltiwanger, and Krizan (2006)
    - Effect of trade liberalization, e.g. Amiti and Konings (2007)
    - Effect of FDI Javorcik (2004)

General references:

- Aguirregabiria (2019) chapter 3
- Ackerberg et al. (2007) section 2
- Van Beveren (2012)

Section 2

Setup

Setup

- Cobb Douglas production

$$Y_{it} = A_{it} K_{it}^{\beta_k} L_{it}^{\beta_\ell}$$

- In logs,

$$y_{it} = \beta_k k_{it} + \beta_\ell l_{it} + \omega_{it} + \epsilon_{it}$$

  with $\log A_{it} = \omega_{it} + \epsilon_{it}$, ω_{it} known to firm, ϵ_{it} not

Empirical Challenges

1. Simultaneity: if firm has information about $\log A_{it}$ when choosing inputs, then inputs correlated with $\log A_{it}$
2. Selection: firms with low productivity will exit sooner
3. Others: measurement error, specification

Section 3

Simultaneity

Simultaneity

- Firm's inputs will be correlated with firm's knowledge of productivity
- E.g. output price p, wage w, choosing L given K

$$\max_L p\mathrm{E}[A]K_k^{\beta}L_{\ell}^{\beta} - wL$$

implies

$$L = \left(\frac{p}{w}\beta_\ell \mathrm{E}[A]K^{\beta_k}\right)^{\frac{1}{1-\beta_\ell}}$$

or in logs,

$$l = \frac{1}{1-\beta_\ell}\log\left(\frac{p}{w}\beta_\ell\right) + \frac{\beta_k}{1-\beta_\ell}k + \frac{1}{1-\beta_\ell}\log\left(\mathrm{E}[A]\right)$$

- l correlated with productivity through correlation of $\log A$ and $\log(\mathrm{E}[A])$
- Exercise: calculate bias of $\hat{\beta}_\ell^{OLS}$ (with some further assumptions)

Simultaneity solutions

1. IV
2. Panel data
3. Control functions

Instrumental variables

- Instrument must be
  - Correlated with k and l
  - Uncorrelated with $\omega + \epsilon$
- Possible instrument: input prices
  - Correlated with k, l through first-order condition
  - Uncorrelated with ω if input market competitive
- Other possible instruments: output prices (more often endogenous), input supply or output demand shifter (hard to find)

Problems with input prices as IV

- Not available in many data sets
- Average input price of firm could reflect quality as well as price differences
- Need variation across observations
    - If firms use homogeneous inputs, and operate in the same output and input markets, we should not expect to find any significant cross-sectional variation in input prices
    - If firms have different input markets, maybe variation in input prices, but different prices could be due to different average productivity across input markets
    - Variation across time is potentially endogenous because could be driven by time series variation in average productivity

Fixed effects

- Have panel data, so should consider fixed effects
- FE consistent if:
  1. $\omega_{it} = \eta_i + \delta_t + \omega^*_{it}$
  2. ω^*_{it} uncorrelated with l_{it} and k_{it}, e.g. ω^*_{it} only known to firm after choosing inputs
  3. ω^*_{it} not serially correlated and is strictly exogenous
- Problems:
  - Fixed productivity a strong assumption
  - Estimates often small in practice
  - Worsens measurement error problems

$$\text{Bias}(\hat{\beta}_k^{FE}) \approx -\frac{\beta_k \text{Var}(\Delta\epsilon)}{\text{Var}(\Delta k) + \text{Var}(\Delta\epsilon)}$$

Dynamic panel: motivation 1

- General idea: relax fixed effects assumption, but still exploit panel
- Collinearity problem: Cobb-Douglas production, flexible labor and capital implies log labor and log capital are linear functions of prices and productivity (Bond and Söderbom (2005))
- If observed labor and capital are not collinear then there must be something unobserved that varies across firms (e.g. prices), but that could invalidate monotonicity assumption of control function

Dynamic panel: moment conditions

- See Blundell and Bond (2000)
- Assume $\omega_{it} = \gamma_t + \eta_i + v_{it}$ with $v_{it} = \rho v_{i,t-1} + e_{it}$, so

$$y_{it} = \beta_\ell l_{it} + \beta_k k_{it} + \gamma_t + \eta_i + v_{it} + \epsilon_{it}$$

subtract $\rho y_{i,t-1}$ and rearrange to get

$$\begin{aligned} y_{it} =& \rho y_{i,t-1} + \beta_\ell (l_{it} - \rho l_{i,t-1}) + \beta_k (k_{it} - \rho k_{i,t-1}) + \\ &+ \gamma_t - \rho \gamma_{t-1} + \underbrace{\eta_i(1-\rho)}_{=\eta_i^*} + \underbrace{e_{it} + \epsilon_{it} - \rho \epsilon_{i,t-1}}_{=w_{it}} \end{aligned}$$

- Moment conditions:
  - Difference: $\mathrm{E}[x_{i,t-s} \Delta w_{it}] = 0$ where $x = (l, k, y)$
  - Level: $\mathrm{E}\left[\Delta x_{i,t-s}(\eta_i^* + w_{it})\right] = 0$

Dynamic panel: economic model 1

- Adjustment costs

$$
\begin{aligned}
V(K_{t-1}, L_{t-1}) = \max_{I_t, K_t, H_t, L_t} & P_t F_t(K_t, L_t) - P_t^K \left(I_t + G_t(I_t, K_{t-1})\right) - \\
& - W_t \left(L_t + C_t(H_t, L_{t-1})\right) + \\
& \psi \mathrm{E}\left[V(K_t, L_t) | \mathcal{I}_t\right] \\
\text{s.t. } K_t = & (1 - \delta_k) K_{t-1} + I_t \\
L_t = & (1 - \delta_l) L_{t-1} + H_t
\end{aligned}
$$

Implies

$$
\begin{aligned}
P_t \frac{\partial F_t}{\partial L_t} - W_t \frac{\partial C_t}{\partial L_t} = & W_t + \lambda_t^L \left(1 - (1 - \delta_l) \psi \mathrm{E}\left[\frac{\lambda_{t+1}^L}{\lambda_t^L} | \mathcal{I}_t\right]\right) \\
P_t \frac{\partial F_t}{\partial K_t} - P_t^K \frac{\partial G_t}{\partial K_t} = & \lambda_t^K \left(1 - (1 - \delta_k) \psi \mathrm{E}\left[\frac{\lambda_{t+1}^K}{\lambda_t^K} | \mathcal{I}_t\right]\right)
\end{aligned}
$$

Dynamic panel: economic model 2

- Current productivity shifts $\frac{\partial F_t}{\partial L_t}$ and (if correlated with future) the shadow value of future labor $\mathrm{E}\left[\frac{\lambda^L_{t+1}}{\lambda^L_t} | \mathcal{I}_t\right]$
- Past labor correlated with current because of adjustment costs

Dynamic panel data: problems

- Problems:
  - Sometimes imprecise (especially if only use difference moment conditions)
  - Differencing worsens measurement error
  - Weak instrument issues if only use difference moment conditions but levels stronger (see Blundell and Bond (2000))
    - Level moments require stronger stationarity assumption - η_i uncorrelated with Δx_{it}

Control functions

- From Olley and Pakes (1996) (OP)
- **Control function**: function of data conditional on which endogeneity problem solved
    - E.g. usual 2SLS $y = x\beta + \epsilon$, $x = z\pi + v$, control function is to estimate residual of reduced form, $\hat{v}$ and then regress y on x and $\hat{v}$. $\hat{v}$ is the control function
- Main idea: model choice of inputs to find a control function

OP assumptions

$$y_{it} = \beta_k k_{it} + \beta_\ell l_{it} + \omega_{it} + \epsilon_{it}$$

1. ω_{it} follows exogenous first order Markov process,
$$p(\omega_{it+1} | \mathcal{I}_{it}) = p(\omega_{it+1} | \omega_{it})$$
2. Capital at t determined by investment at time $t-1$,
$$k_t = (1-\delta)k_{it-1} + i_{it-1}$$
3. Investment is a function of ω and other observed variables
$$i_{it} = I_t(k_{it}, \omega_{it}),$$
and is strictly increasing in ω_{it}
4. Labor variable and non-dynamic, i.e. chosen each t, current choice has no effect on future (can be relaxed)

OP estimation of β_ℓ

- Invertible investment implies $\omega_{it} = I_t^{-1}(k_{it}, i_{it})$

$$\begin{aligned} y_{it} =& \beta_k k_{it} + \beta_\ell l_{it} + I_t^{-1}(k_{it}, l_{it}) + \epsilon_{it} \\ =& \beta_\ell l_{it} + f_t(k_{it}, i_{it}) + \epsilon_{it} \end{aligned}$$

- Partially linear model
  - Estimate by e.g. regress y_{it} on l_{it} and series functions of t, k_{it}, i_{it}
  - Gives $\hat{\beta}_l$, $\hat{f}_{it} = \hat{f}_t(k_{it}, i_{it})$

OP estimation of β_k

- Note: $\hat{f}_t(k_{it}, i_{it}) = \hat{\omega}_{it} + \beta_k k_{it}$
- By assumptions, $\omega_{it} = \mathrm{E}[\omega_{it}|\omega_{it-1}] + \xi_{it} = g(\omega_{it-1}) + \xi_{it}$ with $\mathrm{E}[\xi_{it}|k_{it}] = 0$
- Use $\mathrm{E}[\xi_{it}|k_{it}] = 0$ as moment to estimate β_k.
  - OP: write production function as

$$\begin{aligned} y_{it} - \beta_\ell l_{it} = & \beta_k k_{it} + g(\omega_{it-1}) + \xi_{it} + \epsilon_{it} \\ = & \beta_k k_{it} + g\left(f_{it-1} - \beta_k k_{it-1}\right) + \\ & + \xi_{it} + \epsilon_{it} \end{aligned}$$

    Use $\hat{\beta}_l$ and $\hat{f}_{it}$ in equation above and estimate $\hat{\beta}_k$ by e.g. semi-parametric nonlinear least squares
  - Ackerberg, Caves, and Frazer (2015): use $\mathrm{E}\left[\hat{\xi}_{it}(\beta_k) k_{it}\right] = 0$

Dynamic panel vs control function

- Both derive moment conditions from assumptions about timing and information set of firm
- Dealing with ω
  - Dynamic panel: AR(1) assumption allows quasi-differencing
  - Control function: makes ω estimable function of observables
- Dynamic panel allows fixed effects, does not make assumptions about input demand
- Control function allows more flexible process for ω_{it}

Applications

- Olley and Pakes (1996): productivity in telecom after deregulation
- Söderbom, Teal, and Harding (2006): productivity and exit of African manufacturing firms, uses IV
- Levinsohn and Petrin (2003): compare estimation methods using Chilean data
- Javorcik (2004): FDI and productivity, uses OP
- Amiti and Konings (2007): trade liberalization in Indonesia, uses OP
- Aw, Chen, and Roberts (2001): productivity differentials and firm turnover in Taiwan
- Kortum and Lerner (2000): venture capital and innovation

Section 4

Selection

Selection

- Let $d_{it} = 1$ if firm in sample.
  - Standard conditions imply $d = 1\{\omega \geq \omega^*(k)\}$
- Messes up moment conditions
  - All estimators based on $\mathrm{E}[\omega_{it}\text{Something}] = 0$, observed data really use $\mathrm{E}[\omega_{it}\text{Something}|d_{it} = 1]$
  - E.g. OLS okay if $\mathrm{E}[\omega_{it}|l_{it}, k_{it}] = 0$, but even then,

$$\begin{aligned}\mathrm{E}[\omega_{it}|l_{it}, k_{it}, d_{it} = 1] =& \mathrm{E}[\omega_{it}|l_{it}, k_{it}, \omega_{it} \geq \omega^*(k_{it})] \\ =& \lambda(k_{it}) \neq 0\end{aligned}$$

- Selection bias negative, larger for capital than labor

Selection in OP model

- Estimate β_ℓ as above
- Write $d_{it} = 1\{\xi_{it} \leq \omega^*(k_{it}) - \rho(f_{i,t-1} - \beta_k k_{it-1}) = h(k_{it}, f_{it-1}, k_{it-1})\}$
- Propensity score $P_{it} \equiv \mathrm{E}[d_{it} | k_{it}, f_{it-1}, k_{it-1}]$
- Similar to before estimate β_k, from

$$
\begin{aligned}
y_{it} - \beta_\ell l_{it} = & \beta_k k_{it} + \tilde{g}\left(f_{it-1} - \beta_k k_{it-1}, P_{it}\right) + \\
& + \xi_{it} + \epsilon_{it}
\end{aligned}
$$

Ackerberg, D., C. Lanier Benkard, S. Berry, and A. Pakes. 2007. "Econometric tools for analyzing market outcomes." *Handbook of econometrics* 6:4171–4276. URL http://www.sciencedirect.com/science/article/pii/S1573441207060631. Ungated URL http://people.stern.nyu.edu/acollard/Tools.pdf.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica* 83 (6):2411–2451. URL http://dx.doi.org/10.3982/ECTA13408.

Aguirregabiria, Victor. 2019. "Empirical Industrial Organization: Models, Methods, and Applications." URL http://aguirregabiria.net/wpapers/book_dynamic_io.pdf.

Amiti, Mary and Jozef Konings. 2007. “Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia.” *The American Economic Review* 97 (5):pp. 1611–1638. URL http://www.jstor.org/stable/30034578.

Aw, Bee Yan, Xiaomin Chen, and Mark J. Roberts. 2001. “Firm-level evidence on productivity differentials and turnover in Taiwanese manufacturing.” *Journal of Development Economics* 66 (1):51 – 86. URL http://www.sciencedirect.com/science/article/pii/S0304387801001559.

Blundell, R. and S. Bond. 2000. “GMM estimation with persistent panel data: an application to production functions.” *Econometric Reviews* 19 (3):321–340. URL http://www.tandfonline.com/doi/pdf/10.1080/07474930008800475.

Bond, Steve and Måns Söderbom. 2005. "Adjustment costs and the identification of Cobb Douglas production functions." IFS Working Papers W05/04, Institute for Fiscal Studies. URL http://ideas.repec.org/p/ifs/ifsewp/05-04.html.

Foster, L., J.C. Haltiwanger, and C.J. Krizan. 2001. "Aggregate productivity growth. Lessons from microeconomic evidence." In *New developments in productivity analysis*. University of Chicago Press, 303–372. URL http://www.nber.org/chapters/c10129.pdf.

Foster, Lucia, John Haltiwanger, and C. J. Krizan. 2006. "Market Selection, Reallocation, and Restructuring in the U.S. Retail Trade Sector in the 1990s." *The Review of Economics and Statistics* 88 (4):pp. 748–758. URL http://www.jstor.org/stable/40043032.

Javorcik, Beata Smarzynska. 2004. "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages." *The American Economic Review* 94 (3):pp. 605–627. URL http://www.jstor.org/stable/3592945.

Kortum, Samuel and Josh Lerner. 2000. "Assessing the Contribution of Venture Capital to Innovation." *The RAND Journal of Economics* 31 (4):pp. 674–692. URL http://www.jstor.org/stable/2696354.

Levinsohn, James and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *The Review of Economic Studies* 70 (2):pp. 317–341. URL http://www.jstor.org/stable/3648636.

Olley, G.S. and A. Pakes. 1996. "The dynamics of productivity in the telecommunications equipment industry." *Econometrica* 64 (6):1263–1297. URL http://www.jstor.org/stable/2171831.

Söderbom, Måns, Francis Teal, and Alan Harding. 2006. "The Determinants of Survival among African Manufacturing Firms." *Economic Development and Cultural Change* 54 (3):pp. 533–555. URL http://www.jstor.org/stable/10.1086/500030.

Van Beveren, I. 2012. "Total factor productivity estimation: a practical review." *Journal of Economic Surveys* URL http://onlinelibrary.wiley.com/doi/10.1111/j.1467-6419.2010.00631.x/full.