Estimating Production Functions

Paul Schrimpf

UBC
Economics 565

January 9, 2018

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)

These slides based on:

- Aguirregabiria (2017) chapter 2
- Ackerberg et al. (2007) section 2
- Van Beveren (2012)

Section 2

Setup

Setup

- Cobb Douglas production

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

- In logs,

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

  with $\log A_{it} = \omega_{it} + \epsilon_{it}$, ω_{it} known to firm, ϵ_{it} not
- Problems:
  1. Simultaneity: if firm has information about $\log A_{it}$ when choosing inputs, then inputs correlated with $\log A_{it}$, e.g. price p, wage w, perfect information

$$L_{it} = \left(\frac{p}{w} \beta_l A_{it} K_{it}^{\beta_k} \right)^{\frac{1}{1-\beta_l}}$$

  2. Selection: firms with low productivity will exit sooner
  3. Others: measurement error, specification

Section 3

Simultaneity

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 some 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

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

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

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

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

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

- 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

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

      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 $\mathrm{E}\left[\hat{\xi}_{it}(\beta_k) k_{it}\right] = 0$

Critiques and extensions

- Levinsohn and Petrin (2003): investment often zero, so use other inputs instead of investment to form control function
- Ackerberg, Caves, and Frazer (2015): control function often collinear with l_{it} — for it not to be must be firm specific unobervables affecting l_{it} (but not investment / other input or else demand not invertible and cannot form control function)
- Gandhi, Navarro, and Rivers (2013): relax scalar unobservable in investment / other input demand
- Wooldridge (2009): more efficient joint estimation
- Maican (2006) and Doraszelski and Jaumandreu (2013): endogenous productivity

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 would 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

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

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

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

- 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

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

$$\text{s.t. } K_t = (1 - \delta_k) K_{t-1} + I_t$$

$$L_t = (1 - \delta_l) L_{t-1} + H_t$$

Implies

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

$$P_t \frac{\partial F_t}{\partial K_t} - P_t^K \frac{\partial G_t}{\partial K_t} = \lambda_t^K \left(1 - (1 - \delta_k)\psi \mathrm{E}\left[\frac{\lambda_{t+1}^K}{\lambda_t^K} | \mathcal{I}_t\right]\right)$$

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 $\mathrm{E}\left[\frac{\lambda_{t+1}^L}{\lambda_t^L} | \mathcal{I}_t\right]$
- 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))

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}

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

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

Section 5

Applications

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

Part II

Selected applications and extensions

Section 6

Ackerberg, Caves, and Frazer (2015)

Ackerberg, Caves, and Frazer (2015): contributions

- Document collinearity problem in OP and Levinsohn and Petrin (2003)
  - Need l_{it}, $f_{it}(k_{it}, i_{it})$ not collinear, i.e. something causes variation in l, but not k
- Propose alternative estimator
- Relates estimator to dynamic panel (Blundell and Bond, 2000) approach
- Illustrates estimator using Chilean data

[0]*These slides are based on the working paper version Ackerberg, Caves, and Frazer (2006).

Collinearity in OP 1

- OP assume $i_{it} = I_t(k_{it}, \omega_{it})$
- Symmetry, parsimony suggest $l_{it} = L_t(k_{it}, \omega_{it})$
- Then $l_{it} = L_t(k_{it}, I_t^{-1}(k_{it}, i_{it})) = g_t(k_{it}, i_{it})$

$$y_{it} = \beta_l l_{it} + f_t(k_{it}, i_{it}) + \epsilon_{it}$$

  l_{it} collinear with $f_t(k_{it}, i_{it})$
- Worse in Levinsohn and Petrin (2003)
  - Uses other input m_{it} to form control function

$$\begin{aligned} y_{it} =& \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \omega_{it} + \epsilon_{it} \\ m_{it} =& M_t(k_{it}, \omega_{it}) \end{aligned}$$

  - Even less reason to treat labor demand differently than other input demand

Collinearity in OP 2

- Collinearity still problem with parametric input demand
- Plausible models that do not solve collinearity
    - Input price data
        - Must include in control function to preserve scalar unobservable
        - Same logic above implies m and l are functions of both prices, so still collinear
    - Adjustmest costs in labor
        - Need to add l_{it-1} to control function
    - Change in timing assumptions
    - Measurement error in l (but not m)
        - Solves collinearity, but makes $\hat{\beta}_l$ inconsistent
- Potential model change that removes collinearity
    - Optimization error in l (but not m)
    - m chosen, l specific shock revealed, l chosen
    - OP only: l_{it} chosen at $t - 1/2$, $l_{it} = L_t(\omega_{it-1/2}, k_{it})$, i_{it} chosen at t

ACF estimator

- Idea: like capital, labor is harder to adjust than other inputs
- Model: l_{it} chosen at time $t - 1/2$, m_{it} at time t
    - Implies $m_t = M_t(k_{it}, l_{it}, \omega_{it})$
- Estimation:
    1. $y_{it} = \underbrace{\beta_k k_{it} + \beta_l l_{it} + f_t(m_{it}, k_{it}, l_{it})}_{\equiv \Phi_t(m_{it}, k_{it}, l_{it})} + \epsilon_{it}$ gives
    $$\hat{\omega}_{it}(\beta_k, \beta_l) = \hat{\Phi}_{it} - \beta_k k_{it} - \beta_l l_{it}$$
    2. Moments from timing and Markov process for ω_{it} assumptions:
    $$\omega_{it} = \mathrm{E}[\omega_{it} | \omega_{it-1}] + \xi_{it}$$
        - $\mathrm{E}[\xi_{it} | k_{it}] = 0$ as in OP
        - $\mathrm{E}[\xi_{it} | l_{it-1}] = 0$ from new timing assumption
        - $\hat{\xi}_{it}(\beta_k, \beta_l)$ as residual from nonparametric regression of $\hat{\omega}_{it}$ on $\hat{\omega}_{it-1}$
        - Can add moments based on $\mathrm{E}[\epsilon_{it} | \mathcal{I}_{it}] = 0$

Relation to dynamic panel estimators

- 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}

Empirical example

- Chilean plant level data
- Compare OLS, FE, LP, ACF, and dynamic panel estimators
- LP and ACF using three different inputs (materials, electricity, fuel) for control function
- Results:
    - 311=food, 321=textiles, 331=wood, 381=metal
    - Expected biases in OLS and FE
    - ACF and LP significantly different
    - ACF less sensitive to which input used for control function
    - Dynamic panel closer to ACF than LP, but still significant differences

TABLE 1

	Industry 311					
	Capital		Labor		Returns to Scale	
	Estimate	SE	Estimate	SE	Estimate	SE
OLS	0.336	0.025	1.080	0.042	1.416	0.026
FE	0.081	0.038	0.719	0.055	0.800	0.066
ACF – M	0.371	0.037	0.842	0.048	1.212	0.034
ACF – E	0.379	0.031	0.865	0.047	1.244	0.032
ACF – F	0.395	0.033	0.884	0.046	1.279	0.028
LP – M	0.455	0.038	0.676	0.037	1.131	0.035
LP – E	0.446	0.032	0.764	0.040	1.210	0.034
LP – F	0.410	0.032	0.942	0.040	1.352	0.036
DP	0.391	0.026	0.987	0.043	1.378	0.028

	Industry 321					
	Capital		Labor		Returns to Scale	
	Estimate	SE	Estimate	SE	Estimate	SE
OLS	0.256	0.035	0.953	0.056	1.210	0.034
FE	0.204	0.068	0.724	0.087	0.927	0.108
ACF – M	0.242	0.041	0.893	0.063	1.135	0.040
ACF – E	0.272	0.037	0.832	0.060	1.104	0.039
ACF – F	0.272	0.038	0.873	0.061	1.145	0.040
LP – M	0.320	0.037	0.775	0.059	1.094	0.049
LP – E	0.241	0.037	0.978	0.065	1.219	0.047
LP – F	0.254	0.039	1.008	0.062	1.262	0.048
DP	0.320	0.042	0.837	0.064	1.157	0.041

	Industry 331					
	Capital		Labor		Returns to Scale	
	Estimate	SE	Estimate	SE	Estimate	SE
OLS	0.236	0.047	1.038	0.074	1.274	0.052
FE	-0.028	0.103	0.897	0.095	0.869	0.136
ACF – M	0.196	0.064	0.923	0.085	1.119	0.076
ACF – E	0.195	0.065	0.897	0.088	1.092	0.073
ACF – F	0.212	0.062	0.915	0.086	1.127	0.075
LP – M	0.352	0.056	0.678	0.077	1.030	0.072
LP – E	0.305	0.059	0.786	0.086	1.090	0.075
LP – F	0.241	0.052	0.993	0.079	1.234	0.071
DP	0.252	0.054	0.998	0.073	1.249	0.061

Industry 381

TABLE 2

	Industry 311				Industry 321		
	M	E	F		M	E	F
ACF vs OLS				ACF vs OLS			
K	0.111	0.040	0.010	K	0.585	0.192	0.192
L	1.000	1.000	1.000	L	0.970	1.000	0.996
RTS	1.000	1.000	1.000	RTS	0.998	1.000	0.998
ACF vs LP				ACF vs LP			
K	1.000	1.000	0.707	K	0.982	0.052	0.070
L	0.000	0.000	0.899	L	0.048	0.998	1.000
RTS	0.000	0.061	0.990	RTS	0.198	1.000	1.000
ACF vs DP				ACF vs DP			
K	0.737	0.788	0.505	K	1.000	0.992	0.992
L	1.000	1.000	1.000	L	0.052	0.511	0.084
RTS	1.000	1.000	1.000	RTS	0.820	0.996	0.669

	Industry 331				Industry 381		
	M	E	F		M	E	F
ACF vs OLS				ACF vs OLS			
K	0.892	0.840	0.830	K	0.060	0.058	0.054
L	0.974	0.990	0.984	L	1.000	1.000	1.000
RTS	1.000	1.000	1.000	RTS	1.000	1.000	1.000
ACF vs LP				LP vs ACF			
K	1.000	1.000	0.860	K	0.996	0.980	0.683
L	0.000	0.024	0.876	L	0.000	0.072	0.910
RTS	0.056	0.431	0.984	RTS	0.002	0.323	0.984
ACF vs DP				ACF vs DP			
K	0.962	0.922	0.884	K	0.834	0.916	0.892
L	0.940	0.986	0.962	L	0.852	0.844	0.649
RTS	1.000	1.000	0.998	RTS	0.984	0.992	0.934

Note: Value is the % of bootstrap reps where ACF coeff is less than OLS, LP, or DP coef. A value either above 0.95 or below 0.05 indicates that coefficients are significantly different from each other.

Section 7

Gandhi, Navarro, and Rivers (2013)

Gandhi, Navarro, and Rivers (2013)

- Show that control function method is not nonparametrically identified when there are flexible inputs
- Propose alternate estimate that uses data on input shares and information from firm's first order condtiion
- Show that value-added and gross output production functions are incompatible
- Application to Colombia and Chile

Assumptions

1. Hicks neutral productivity $Y_{jt} = e^{\omega_{jt}+\epsilon_{jt}} F_t(L_{jt}, K_{jt}, M_{jt})$
2. ω_{jt} Markov, ϵ_{jt} i.i.d.
3. K_{jt} and L_{jt} determined at $t-1$, M_{jt} determined flexibly at t
   - K and L play same role in the model, so after this slide I will drop L
4. $M_{jt} = \mathbb{M}_t(L_{jt}, K_{jt}, \omega_{jt})$, monotone in ω_{jt}

Reduced form

- Let $h(\omega_{jt-1}) = \mathrm{E}[\omega_{jt}|\omega_{jt-1}]$, $\eta_{jt} = \omega_{jt} - h(\omega_{jt-1})$
- log output

$$
\begin{aligned}
y_{jt} =& f_t(k_{jt}, m_{jt}) + \omega_{jt} + \epsilon_{jt} \\
=& f_t(k_{jt}, m_{jt}) + \underbrace{h(\mathbb{M}_{t-1}^{-1}(k_{jt-1}, m_{jt-1}))}_{=h_{t-1}(k_{jt-1}, m_{jt-1})} + \eta_{jt} + \epsilon_{jt}
\end{aligned}
$$

- Assumptions imply

$$
\mathrm{E}[\eta_{jt}|\underbrace{k_{jt}, k_{jt-1}, m_{jt-1}, \ldots k_{j1}, m_{j1}}_{=\Gamma_{jt}}] = 0
$$

- Reduced form

$$
\mathrm{E}[y_{jt}|\Gamma_{jt}] = \mathrm{E}[f_t(k_{jt}, m_{jt})|\Gamma_{jt}] + h_{t-1}(k_{jt-1}, m_{jt-1}) \tag{1}
$$

- Identification: given observed $\mathrm{E}[y_{jt}|\Gamma_{jt}]$ is there a unique f_t, h_{t-1} that satisfies (3)?

Example: Cobb-Douglas 1

- Let $f_t(k, m) = \beta_k k + \beta_m m$
- Assume firm is takes prices as given
- First order condition for m gives

$$m = constant + \frac{\beta_k}{1-\beta_m} k + \frac{1}{1-\beta_m} \omega$$

- Put into reduced form

$$E[y_{jt}|\Gamma_{jt}] = C + \frac{\beta_k}{1-\beta_m} k_{jt} + \frac{\beta_m}{1-\beta_m} E[\omega_{jt}|\Gamma_{jt}] + h_{t-1}(k_{jt-1}, m_{jt-} \tag{2}$$

- ω Markov and $\omega_{jt-1} = \mathbb{M}_{t-1}^{-1}(k_{jt-1}, m_{jt-1})$ implies

$$E[\omega_{jt}|\Gamma_{jt}] = E[\omega_{jt}|\omega_{jt-1} = \mathbb{M}_{t-1}^{-1}(k_{jt-1}, m_{jt-1})] = \\ = h_{t-1}(k_{jt-1}, m_{jt-1})$$

Example: Cobb-Douglas 2

- Which leaves

$$\mathrm{E}[y_{jt}|\Gamma_{jt}] = constant + \frac{\beta_k}{1-\beta_m} k_{jt} + \frac{1}{1-\beta_m} h_{t-1}(k_{jt-1}, m_{jt-1}) \tag{3}$$

  from which β_k, β_m are not identified
- Rank condition fails, $\mathrm{E}[m_{jt}|\Gamma_{jt}]$ is colinear with $h_{t-1}(k_{jt-1}, m_{jt-1})$
- After conditioning on $k_{jt}, k_{jt-1}, m_{jt-1}$, only variation in m_{jt} is from η_{jt}, but this is uncorrelated with the instruments

Identification from first order conditions 1

- Since m flexible, it satisfies a simple static first order condition,

$$\rho_t = p_t \frac{\partial F_t}{\partial M} \mathrm{E}[e^{\epsilon_{jt}}] e^{\omega_{jt}}$$

$$\log \rho_t = \log p_t + \log \frac{\partial F_t}{\partial M}(k_{jt}, m_{jt}) + \log \mathrm{E}[e^{\epsilon_{jt}}] + \omega_{jt}$$

- Problem: prices often unobserved, endogenous ω
- Solution: difference from output equation to eliminate ω, rearrange so that it involves only the value of materials and the value of output (which are often observed)

$$\underbrace{s_{jt}}_{\equiv \log \frac{\rho_t M_{jt}}{p_t Y_{jt}}} = \log \underbrace{G_t(k_{jt}, m_{jt})}_{\equiv \left(M_t \frac{\partial F_t}{\partial M}\right)/F_t} + \log \underbrace{\mathrm{E}[e^{\epsilon_{jt}}]}_{\mathcal{E}} - \epsilon_{jt}$$

Identification from first order conditions 2

- Identifies elasticity up to scale, $G_t\mathcal{E}$ and ϵ_{jt} which identifie $\mathcal{E}$
- Integrating,

$$\int_{m_0}^{m_{jt}} G_t(k_{jt}, m)/m = f_t(k_{jt}, m_{jt}) + c_t(k_{jt})$$

identifies f up to location

- Output equation

$$y_{jt} = \int_{m_0}^{m_{jt}} \tilde{G}_t(k_{jt}, m)/m - c_t(k_{jt}) + \omega_{jt} + \epsilon_{jt}$$

$$-c_t(k_{jt}) + \omega_{jt} = \underbrace{y_{jt} - \int_{m_0}^{m_{jt}} \tilde{G}_t(k_{jt}, m)/m - \epsilon_{jt}}_{\equiv \mathcal{Y}_{jt}}$$

Identification from first order conditions 3

where the things on the right have already been identified

- Identify c_t from

$$\mathcal{Y}_{jt} = -c_t(k_{jt}) + \tilde{h}_t(\mathcal{Y}_{jt-1}, k_{jt-1}) + \eta_{jt}$$

Value added vs gross production

- Value added:

$$\begin{aligned} VA_{jt} =& p_t Y_{jt} - \rho_t M_{jt} \\ =& p_t F_t(K_{jt}, \mathbb{M}_t(K_{jt}, \omega_{jt}))e^{\omega_{jt}+\epsilon_{jt}} - \rho_t \mathbb{M}_t(K_{jt}, \omega_{jt}) \end{aligned}$$

- Envelope theorem implies $\text{elasticity}^{Y}_{e^\omega} \approx \text{elasticity}^{VA}_{e^\omega}(1 - \frac{\rho_t M_{jt}}{p_t Y_{jt}})$

Problems

- Production Hicks-neutral productivity does not imply value-added Hicks-neutral productivity
- Ex-post shocks ϵ_{jt} not accounted for in approximation

Empirical results

- Look at tables
- Value-added estimates imply much more productivity dispersion than gross (90-10) ratio of 4 vs 2

Section 8

Grieco and McDevitt (2017)

Grieco and McDevitt (2017)

`https://www.ftc.gov/sites/default/files/documents/public_events/fifth-annual-microeconomics-conference/grieco-p_0.pdf`

Model details

- Timing:
  1. Quality chosen $q_{it} = q(k_{it}, \ell_{it}, x_{it}, \omega_{i,t-b})$
  2. Production occurs, ω_{it} revealed to firm
  3. Hiring chosen $\ell_{i,t+1} - \ell_{it} = h_{it} = h(k_{it}, \ell_{it}, x_{it}, \omega_{it})$
- ω follows Markov process:

$$\mathrm{E}[\omega_{i,t-b}|\mathcal{I}_{i,t-b}] = \mathrm{E}[\omega_{i,t-b}|\omega_{i,t-1}] \;\&\; \mathrm{E}[\omega_{it}|\mathcal{I}_{i,t}] = \mathrm{E}[\omega_{it}|\omega_{i,t-b}$$

and $\omega_{it} = \mathrm{E}[\omega_{it}|\omega_{i,t-1}] + \eta_{it} = g(\omega_{i,t-1}) + \eta_{it}$

Moment conditions

- Control function assumption: hiring is a monotonic function of ω

$$h_{it} = h(k_{it}, \ell_{it}, x_{it}, \omega_{it})$$

so

$$\omega_{it} = h^{-1}(k_{it}, \ell_{it}, x_{it}, h_{it})$$

- Substitute into production function:

$$\begin{aligned} y_{it} =& \alpha_q q_{it} + \beta_k k_{it} + \beta_\ell \ell_{it} + h^{-1}(k_{it}, \ell_{it}, x_{it}, h_{it}) + \epsilon_{it} \\ y_{it} =& \alpha_q q_{it} + \Phi(k_{it}, \ell_{it}, x_{it}, h_{it}) + \epsilon_{it} \end{aligned}$$

- Evolution of ω

$$\begin{aligned} \omega_{it} =& y_{it} - \alpha_q q_{it} - \beta_k k_{it} - \beta_\ell \ell_{it} - \epsilon_{it} = g(\omega_{i,t}) + \eta_{it} \\ =& g(\Phi(k_{it-1}, \ell_{it-1}, x_{it-1}, h_{it-1}) - \beta_\ell \ell_{it-1} - \beta_k k_{it-1}) + \eta_{it-1} \end{aligned}$$

- Moment conditions:

$$\mathrm{E}[\epsilon_{it} | q_{it}, k_{it}, \ell_{it}, x_{it}, h_{it}] = 0$$
$$\mathrm{E}[\eta_{it} | k_{it}, \ell_{it}, x_{it}, k_{it-1}, \ell_{it-1}, x_{it-1}] = 0$$

Estimation

1. Estimate, α_q, Φ from

$$y_{it} = \alpha_q q_{it} + \Phi(k_{it}, \ell_{it}, x_{it}, h_{it}) + \epsilon_{it}$$

   by semiparametric regression

2. Estimate β_k, β_ℓ
   - Let $\omega(\beta)_{it} = P\hat{h}i(k_{it}, \ell_{it}, x_{it}, h_{it}) - \beta_k k_{it} - \beta_\ell \ell_{it}$
   - For each β estimate $g()$

$$\omega(\beta)_{it} = g(\omega(\beta)_{it-1}) + \eta_{it}(\beta)$$

     by nonparametric regression
   - Minimize empirical moment condition for η

$$\hat{\beta} = \arg\min(\frac{1}{NT}\sum_{it} k_{it}\eta_{it}(\beta))^2 + (\frac{1}{NT}\sum_{it} \ell_{it}\eta_{it}(\beta))^2$$

- Should hemoglobin level be controlled for when measuring quality?
    - Anemia (low hemoglobin) is risk-factor for infection
    - Anemia can be treated through diet, iron supplements (pills or IV), EPO, etc
        - Are dialysis facilities responsible for this treatment?
        - In 2006-2014 data average full-time dieticiens = 0.5, average part-time = 0.6
- Estimation details:

Step 1: Estimate α_q

$$y_{jt}\hat{E}[y|h_{jt}, i_{jt}, k_{jt}, \ell_{jt}, x_{jt}] = \alpha_q(qjt\hat{E}[q|h_{jt}, i_{jt}, k_{jt}, \ell_{jt}, x_{jt}]) + \epsilon_{jt}$$

- Drop observations with $h_{jt} = 0$ (not invertible)
- Okay here, because selecting on ω, and residual, ϵ_{jt} is uncorrelated with ω
- Problematic in last step? No, see footnote 49

Step 2: Estimate β_k, β_ℓ from

$$y_{jt} + \hat{\alpha}_q + \beta_k k_{jt} + \beta_\ell \ell_{jt} = g(\hat{\omega}_{jt-1}(\beta)) + \eta_{jt} + \epsilon_{jt}$$

  - Only have $hat\omega_{jt-1}(\beta)$ when $h_{jt-1} \neq 0$, okay because ϵ_{jt} and η_{jt} are uncorrelated with ω_{jt-1}, would be problem if using $\hat{\omega}_{jt}$
- Nothing about selection – number of centers, 4270, vs center-years, 18295, implies there must be entry and exit

- Would like to see some results related to productivity dispersion e.g.
  - Decompose variation in infection rate into: productivity variation, incentive variation, quality-quantity choices, and random shocks
  - Compare strengthening incentives vs closing least productive facilities as policies to increase quality

Section 9

Amiti and Konings (2007)

Overview

- Effect of reducing input and output tariffs on productivity
- Reducing output tariffs affects productivity by increasing competition
- Reducing input tariffs affects productivity through learning, variety, and quality effects
- Previous empirical work focused on output tariffs; might be estimating combined effect
- Input tariffs hard to measure; with Indonesian data on plant-level inputs can construct plant specific input tariff

Methodology

- Estimate TFP using Olley-Pakes
  - Output measure is revenue ⇒ may confound productivity and markups
- Estimate relation between TFP and tariffs

$$\log(TFP_{it}) = \gamma_0 + \alpha_i + \alpha_{tl(i)} + \gamma_1(\text{output tariff})_{tk(i)} + \\ + \gamma_2(\text{input tariff})_{tk(i)} + \epsilon_{it} \quad (4)$$

  - $k(i)$ = 5-digit (ISIC) industry of plant i
  - $l(i)$ = island of plant i
- Explore robustness to:
  - Different productivity measure
  - Specification of 4
  - Endogeneity of tariffs

Data and tariff measure

- Indonesian annual manufacturing census of 20+ employee plants 1991-2001, after cleaning $15,000$ firms per year
- Input tariffs:
  - Data on tariffs on goods, τ_{jt}, but also need to know inputs
  - 1998 only: have data on inputs, use to construct input weights at industry level, w_{jk}
  - Industry input tariff $= \sum_j w_{jk}\tau_{jt}$

Results

- Look at tables
- Input tariffs have larger effect than output, $\hat{\gamma}_1 \approx -0.07$, $\hat{\gamma}_2 \approx -0.44$
- Robust to:
    - Productivity measure
    - Tariff measure
    - Including/excluding Asian financial crisis
- Less robust to instrumenting for tariffs
    - Qualitatively similar, but larger coefficient estimates
- Explore channels for productivity change
    - Markups (maybe), product switching/addition (no), foreign ownership (no), exporters (no)

Section 10

Doraszelski and Jaumandreu (2013)

Overview

- Estimable model of endogenous productivity, which combines:
  - Knowledge capital model of R&D
  - OP & LP productivity estimation
- Application to Spanish manufacturers focusing on R&D
  - Large uncertainty (20%-60% or productivity unpredictable)
  - Complementarities and increasing returns
  - Return to R&D larger than return to physical capital investment

Model (simplified) 1

- Cobb-Douglas production:

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

- Controlled Markov process for productivity, $p(\omega_{it+1}|\omega_{it}, r_{it})$,

$$\omega_{it} = g(\omega_{it-1}, r_{it-1}) + \xi_{it}$$

- Labor flexible and non-dynamic
- Value function

$$V(k_t, \omega_t, u_t) = \max_{i,r} \Pi(k_t, \omega_t) - C_i(i, u_t) - C_r(r, u_t) + \\ + \frac{1}{1+\rho} \mathrm{E}\left[V(k_{t+1}, \omega_{t+1}, u_{t+1}) | k_t, \omega_t, i, r, u_t\right]$$

Model (simplified) 2

- u scalar or vector valued shock
- u not explicitly part of model, but identification discussion (especially p10 and footnote 6) implicitly adds it
- u independent of? k, l? across time?

- Control function incorporating Cobb-Douglas assumption (and perfect competition):

$$\omega_{it} = h(l_{it}, k_{it}, w_{it} - p_{it}; \beta) = \lambda_0 + (1 - \beta_l) l_{it} - \beta_k k_{it} + (w_{it} - p_{it})$$

- Form moments based on

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + g\left(h(l_{it-1}, k_{it-1}, w_{it-1} - p_{it-1}; \beta), r_{it-1}\right) + \xi_{it} + \epsilon$$

- No collinearity because:
  - Parametric h
  - Variation in k, r due to u
- Estimated model adds

Model (simplified) 3

  - Material input instead of labor for control function
  - h based on imperfect competition
- Comparison to OP, LP, ACF

Results

- Look at tables and figures
- Large uncertainty (20%-60% or productivity unpredictable)
- Complementarities and increasing returns
- Return to R&D larger than return to physical capital

Ackerberg, D., K. Caves, and G. Frazer. 2006. "Structural identification of production functions." URL http://mpra.ub.uni-muenchen.de/38349/.

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. 2017. "Empirical Industrial Organization: Models, Methods, and Applications." URL http://www.individual.utoronto.ca/vaguirre/courses/eco2901/teaching_io_toronto.html.

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.

Doraszelski, Ulrich and Jordi Jaumandreu. 2013. "R&D and Productivity: Estimating Endogenous Productivity." *The Review of Economic Studies* 80 (4):1338–1383. URL http://restud.oxfordjournals.org/content/80/4/1338.abstract.

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.

Gandhi, A., S. Navarro, and D. Rivers. 2013. “On the Identification of Production Functions: How Heterogeneous is Productivity?” URL https://sites.google.com/site/econsalvador/Research/production_9_25_13_FULL.pdf?attredirects=0.

Grieco, Paul LE and Ryan C McDevitt. 2017. “Productivity and Quality in Health Care: Evidence from the Dialysis Industry.” *Review of Economic Studies (forthcoming)* URL http://www.restud.com/wp-content/uploads/2016/08/MS17933manuscript.pdf.

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.

Maican, F.G. 2006. "Productivity dynamics, r&d, and competitive pressure." *ECONOMIC STUDIES DEPARTMENT OF ECONOMICS SCHOOL OF BUSINESS, ECONOMICS AND LAW UNIVERSITY OF GOTHENBURG* .

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.

Wooldridge, Jeffrey M. 2009. "On estimating firm-level production functions using proxy variables to control for unobservables." *Economics Letters* 104 (3):112 – 114. URL http://www.sciencedirect.com/science/article/pii/S0165176509001487.