Industry Dynamics and Productivity III

Paul T. Scott
Toulouse School of Economics
ptscott@gmail.com

Empirical IO
Spring 2014

Brief Digression on Cost Function Estimation

CD example

- If we have a Cobb-Douglas production function:

$$Q = \exp\left(\omega + \varepsilon\right) L^{\beta_l} K^{\beta_k}$$

- Then the cost function (assuming cost minimization) is

$$c = \mu + \frac{\beta_l}{\beta_l + \beta_k} w + \frac{\beta_k}{\beta_l + \beta_k} r + \frac{1}{\beta_l + \beta_k} q - \frac{1}{\beta_l + \beta_k} \left(\omega + \varepsilon\right)$$

  where w is the wage, r is the rental rate, and μ is a constant that depends on the β's

Cost function estimation

- We can estimate

$$c = \mu + \gamma_l w + \gamma_r r + \gamma_q q - (\omega^* + \varepsilon^*)$$

where $\omega^* = \gamma_q \omega$ and $\varepsilon^* = \gamma_q \varepsilon$.

- And the parameters of the production function could be recovered:

$$\beta_l = w_l/\gamma_q, \qquad \beta_k = \gamma_r/\gamma_q.$$

Insights from Nerlove (1963)

- Note that cost-minimization theory tells us that $\gamma_l + \gamma_k = 1$. Nerlove suggests measuring all prices relative to one price (say, w), eliminating one of the redundant terms. In logs,

$$
\begin{aligned}
c - w &= \mu + \gamma_w w + \gamma_r r + \gamma_q q - \gamma_q (\omega + \varepsilon) - w \\
&= \mu + \gamma_w w + \gamma_r r + + \gamma_q q - \gamma_q (\omega + \varepsilon) - \gamma_w w - \gamma_r w \\
& \quad \mu + \gamma_r (r - w) + \gamma_q q - \gamma_q (\omega + \varepsilon)
\end{aligned}
$$

- Note that we only need variation in either r or w in order to estimate the model.

- Another insight from Nerlove: we could do without observing r as long as it is constant across firms. With just a cross section of firms, the term with r would be absorbed by the constant. With a panel, we might control for temporal variation in r with time fixed effects.

Concerns

- The cost function is based on static cost minimization, so using it only makes sense if all inputs are variable.
- With productions functions, we worry about the endogneity of inputs. With cost functions, we worry about the endogeneity of output. In Nerlove's application (electricity producers), Q was arguably exogneous because of regulation. In other settings, we would need some instrument for Q (e.g., demand shifters).

"Markups and Firm-Level Export Status"
De Loecker and Warzynski (2012)

Overview

- Demonstrates how production function can be used to make inferences about markups
- Applied question: how do markups of exporters differ from non-exporters, and how does a firm's productivity change when it becomes an exporter.
- Findings:
  - Exporters have higher markups than importers
  - Markups increase when a firm becomes an exporter
  - Note similarity to De Loecker (2011), but focus is now on exporter status rather than trade liberalization

Sketch of main idea I

- Definition of markup: $\mu = P/MC$
- Let P_{it}^{v} represent the price of input v and let P_{it} represent the price of output.
- Production function:

$$Q_{it} = Q_{it}\left(X_{it}^{1}, \ldots, X_{it}^{V}, K_{it}, \omega_{it}\right)$$

  where $v = 1, 2, \ldots, V$ indexes variable inputs.
- Assumption: variable inputs are set each period to minimize costs.

Sketch of main idea II

- Lagrangian for cost minimization problem:

$$L\left(X_{it}^{1},\ldots,X_{it}^{V},K_{it},\lambda_{it}\right)=\sum_{v=1}^{V}P_{it}^{v}X_{it}^{v}+r_{it}K_{it}+\lambda_{it}\left(Q_{it}-Q_{it}\left(\cdot\right)\right)$$

- First-order condition:

$$P_{it}^{v}-\lambda_{it}\frac{\partial Q_{it}\left(\cdot\right)}{\partial X_{it}^{v}}=0,$$

where λ_{it} is the marginal cost of production at production level Q_{it}.

Sketch of main idea III

- First-order condition:

$$P_{it}^{v} - \lambda_{it} \frac{\partial Q_{it}(\cdot)}{\partial X_{it}^{v}} = 0.$$

- Multiplying by X_{it}^{v}/Q_{it}:

$$\frac{\partial Q_{it}(\cdot)}{\partial X_{it}^{v}} \frac{X_{it}^{v}}{Q_{it}} = \frac{1}{\lambda} \frac{P_{it}^{v} X_{it}^{v}}{Q_{it}}.$$

- With $\mu_{it} \equiv P_{it}/\lambda_{it}$,

$$\frac{\partial Q_{it}(\cdot)}{\partial X_{it}^{v}} \frac{X_{it}^{v}}{Q_{it}} = \mu_{it} \frac{P_{it}^{v} X_{it}^{v}}{P_{it} Q_{it}}$$

  where we have multiplied and divided by P_{it} on the RHS.

The markup formula

This leads to a simple expression:

$$\mu_{it} = \theta^{v}_{it} \left(\alpha^{v}_{it}\right)^{-1}$$

where θ^{v}_{it} is the output elasticity with respect to input v, and α^{v}_{it} is expenditures on input v as a share of revenues.

- On its own, this formula is nothing new
- What's new about DLW is how flexible they are about estimating θ^{v}_{it} and how they base their inferences about markups on careful production function estimation.

The demand-based approach

- Recall the formula for monopoly pricing:

$$\frac{p}{mc} = \frac{1}{1 + \mathcal{E}_D^{-1}}$$

  where $\mathcal{E}_D^{-1}$ is the inverse elasticity of demand.

- In more complicated settings (e.g., differentiated products), we can still solve for markups as a function of demand elasticities.

- Demand-based approach has been the standard, but notice the many assumptions involved:
  - Typically static Nash-Bertrand competition (or at least some imperfect competition game where we can easily solve for the equilibrium)
  - Instruments to identify demand
  - Functional form assumptions on demand system, model of consumer heterogeneity

CD: example

- Assume labor is a flexible input.
- With Cobb-Douglas production function,

$$Q_{it} = \exp(\omega_{it}) L^{\beta_L} K^{\beta_K},$$

output elasticity of labor is just a constant:

$$\theta_{it}^L = \frac{\partial Q_{it}}{\partial L_{it}} \frac{L_{it}}{Q_{it}} = \beta_L.$$

- Markup:

$$\mu_{it} = \frac{\beta_L}{\alpha_{it}^L}$$

CD: concerns

Cobb-Douglas markup:

$$\mu_{it} = \frac{\beta_L}{\alpha_{it}^L}$$

Some things we might worry about:

- Bias in estimating β_L without appropriate econometric strategy (always a concern in production function estimation)
- Cobb-Douglas is very restrictive, imposing output elasticity which does not depend on Q nor the relative levels of inputs. Variation in expenditure shares will be only source of variation in markups.
- If we assume variation of input share is independent of output elasticity, then any variation in productivity which affects the input share is being treated as variation in markups.

Translog production function

- DLW's main results are based on a translog production function:

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{lk} l_{it} k_{it} + \omega_{it} + \varepsilon_{it}.$$

- Translog output elasticities:

$$\hat{\theta}_{it}^L = \hat{\beta}_l + 2\hat{\beta}_{ll} l_{it} + \hat{\beta}_{lk} k_{it},$$

so translog production is flexible enough to allow for a first-order approximation to how output elasticities vary with input use.

Empirical framework

- Consistent with production function estimation literature, they assume Hicks-neutral productivity shocks:

$$Q_{it} = F\left(X_{it}^{1}, \ldots, X_{it}^{V}, K_{it}; \beta\right) \exp\left(\omega_{it}\right).$$

- Also allow for some measurement error in production:

$$y_{it} = \ln Q_{it} + \varepsilon_{it}$$

$$y_{it} = f\left(x_{it}, k_{it}; \beta\right) + \omega_{it} + \varepsilon_{it}$$

The control function

- Following Levinsohn and Petrin, use materials to proxy for productivity

$$m_{it} = m_t \left(k_{it}, \omega_{it}, \mathbf{z}_{it} \right)$$

where $\mathbf{z}_{it}$ are controls.

- Note: a big claim of the paper is estimating "markups without specifying how firms compete in the product market"
- But here, $\mathbf{z}_{it}$ must control for everything which shifts input demand choices or else there will be variation in productivity they're not controlling for (and hence some of the variation in their inferred markups may actually come from variation in productivity)
- In the appendix, they explain that $\mathbf{z}_{it}$ includes input prices, lagged inputs (meant to capture variation in input prices), and exporter status.

Physical output vs. sales

- Note that the theory is developed in terms of outputs, but DLW only have sales (as usual).
- For a price-taking firm, there's no problem rewriting the formula in terms of sales:

$$\frac{\partial R_{it}(\cdot)}{\partial X_{it}^{v}} \frac{X_{it}^{v}}{R_{it}} = \frac{P_t \partial Q_{it}(\cdot)}{\partial X_{it}^{v}} \frac{X_{it}^{v}}{P_t Q_{it}} = \mu_{it} \frac{P_{it}^{v} X_{it}^{v}}{P_{it} Q_{it}}$$

  because $\frac{\partial R_{it}(\cdot)}{\partial X_{it}^{v}} = \frac{P_t \partial Q_{it}(\cdot)}{\partial X_{it}^{v}}$.
- However, if the firm has market power,

$$\frac{\partial R_{it}(\cdot)}{\partial X_{it}^{v}} = \frac{\partial Q_{it}(\cdot)}{\partial X_{it}^{v}} \left(P_{it} + \frac{\partial P_{it}}{\partial Q_{it}} \right).$$

TABLE 2—ESTIMATED MARKUPS

Methodology	Markup
Hall[a]	1.03 (0.004)
Klette[a]	1.12 (0.020)
Specification	
I (Cobb-Douglas)	1.17
II (I w/ endog. productivity)	1.10
III (I w/ additional moments)	1.23
IV (Translog)	1.28
V (II w/ export input)	1.23
VI (Gross Output: labor)	1.26
VI (Gross Output: materials)	1.22
VII[a] (**I** w/ single markup)	1.16 (0.006)
VIII[a] (First difference)	1.11 (0.007)

[a]Markups are estimated jointly with the production function (as discussed in Section III), and we report the standard errors in parentheses. The standard deviation around the markups in specifications **I–VI** is about 0.5.

Table 3—Markups and Export Status I: Cross-Section

	Methodology	Export Premium
	Hall	0.0155 (0.010)
	Klette	0.0500 (0.090)
Specification		
	I (Cobb-Douglas)	0.1633 (0.017)
	II (I w/ endog. productivity)	0.1608 (0.017)
	IV (Translog)	0.1304 (0.014)
	V (II w/ export input)	0.1829 (0.017)
	VIII (First difference)	0.1263 (0.013)

Notes: Estimates are obtained after running equation (21) where the different specifications refer to the different markup estimates, and we convert the percentage markup difference into levels as discussed above. The standard errors under specifications **I–V** are obtained from a nonlinear combination of the relevant parameter estimates. All regressions include labor, capital, and full year and industry dummies as controls. Standard errors are in parentheses.

"Structural Identification of Production Functions"
Ackerberg, Caves, and Frazer (2006)

Overview

- ACF argue that Olley and Pakes's (1996) and Levinsohn and Petrin's (2003) approach suffer from collinearity issues.
- They propose a new approach which involves modified assumptions on the timing of input decisions and moves the identification of all coefficients of the production function to the second stage of the estimation.

Overview

- ACF argue that Olley and Pakes's (1996) and Levinsohn and Petrin's (2003) approach suffer from collinearity issues.
- They propose modifying the timing assumptions on input choice to avoid the collinearity

LP's first stage

- Levinsohn and Petrin's first-stage regression:

$$y_{it} = \beta_l l_{it} + f_t^{-1}\left(m_{it}, k_{it}\right) + \varepsilon_{it}.$$

- LP's approach was based on the premise that materials inputs are a variable input and therefore a function of state variables:

$$m_{it} = m_t\left(\omega_{it}, k_{it}\right),$$

- They also assume that labor is a variable input (or else we would not be able to exclude it from the inversion), so

$$l_{it} = l_t\left(\omega_{it}, k_{it}\right).$$

LP's identification problem

- This means we can write:

$$y_{it} = \beta_l l_t \left(f_t^{-1} (m_{it}, k_{it}), k_{it} \right) + f_t^{-1} (m_{it}, k_{it}) + \varepsilon_{it},$$

  and since we're being nonparametric about f_t^{-1},
  it should absorb $\beta_l l_t \left(f_t^{-1} (m_{it}, k_{it}), k_{it} \right)$.

- There should be no variation in l_{it} left over to identify β_l.

Collinearity with CD production

- Cobb-Douglas production with inverted productivity:

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + f_t^{-1}(m_{it}, k_{it}) + \epsilon_{it}.$$

- FOC for materials:

$$\beta_m K_{it}^{\beta_k} L_{it}^{\beta_l} M_{it}^{\beta_m - 1} e^{\omega_{it}} = \frac{p_m}{p_y}.$$

- Solving for ω:

$$\omega_{it} = \ln\left(\frac{1}{\beta_m}\right) + \ln\left(\frac{p_m}{p_y}\right) - \beta_k k_{it} - \beta_l l_{it} + (1 - \beta_m) m_{it}$$

- Plugging this into the production function, the $\beta_l l_{it}$ terms cancel:

$$y_{it} = \ln\left(\frac{1}{\beta_m}\right) + \ln\left(\frac{p_m}{p_y}\right) + m_{it} + \epsilon_{it}.$$

Collinearity in practice and in principle

- It could be the case that l_{it} takes different values in the data for the same values of (m_{it}, k_{it}). ACF's argument is about collinearity in principle, given the assumptions of LP.
- Some potential sources of independent variation:
  (*Which one works?*)
  - unobserved variation in firm-specific input prices.
  - measurement error in l_{it} or m_{it}
  - optimization error in l_{it} or m_{it}

Collinearity in practice and in principle

- It could be the case that l_{it} takes different values in the data for the same values of (m_{it}, k_{it}). ACF's argument is about collinearity in principle, given the assumptions of LP.
- Some potential sources of independent variation:
  (*Which one works?)*
  - unobserved variation in firm-specific input prices.
  - measurement error in l_{it} or m_{it}
  - optimization error in l_{it} or m_{it}
- While optimization error in l_{it} works econometrically, it's not the most appealing assumption economically.

Another failed solution

- Note that the whole problem comes about because labor and materials are set simultaneously. This means one way to break the collinearity is to assume they are set with respect to different information sets.
- Let's try to break the informational equivalence with timing assumptions. Suppose:
  - m_{it} is set at time t
  - l_{it} is set at time $t-b$ with $0<b<1$
  - ω has Markovian in between subperiods:

$$\begin{aligned} p\left(\omega_{i,t-b}|I_{i,t-1}\right) &= p\left(\omega_{i,t-b}|\omega_{it-1}\right) \\ p\left(\omega_{it}|I_{i,t-b}\right) &= p\left(\omega_{i,t}|\omega_{i,t-b}\right) \end{aligned}$$

- But this doesn't work! And neither does having m_{it} set first. *(Why?)*

An implausible solution

- Let's try again:
  - l_{it} is set at time t
  - m_{it} is set at time $t - b$ with $0 < b < 1$
  - we have a more complicated structure of productivity shocks:

$$y_{it} = \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + \omega_{i,t-b} + \eta_{it},$$

$$p\left(\omega_{i,t-b} | I_{i,t-1}\right) = p\left(\omega_{i,t-b} | \omega_{i,t-1}\right) \ ,$$

  - and there is some unobservable shock to labor prices which is realized between $t - b$ and t. This shock must be i.i.d.

- l_{it} has its own shock to respond to, creating independent variation, and the productivity inversion still works because the new shock is not a state variable.

- This works, but as ACF argue, it's rather ad-hoc and difficult to motivate.

Collinearity in Olley Pakes

- Olley Pakes's control function has the same collinearity issue, but ACF argue it can be avoided with assumptions which "might be a reasonable approximation to the true underlying process."

- Assume that l_{it} is set at $t - b$ with $0 < b < 1$. ω has a Markovian between subperiods. Then:

$$l_{it} = l_t\left(\omega_{i,t-b}, k_{it}\right),$$

  so we have variation in l_{it} which is independent of (ω_{it}, k_{it}).

- Note that even though l_{it} is set before investment i_{it}, investment won't depend on l_{it} because it is a static input. So the productivity inversion is unchanged.

- These timing assumptions cannot save LP, but they work well with OP.

ACF's alternative procedure I

- Consider value added production function:

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

- ACF's procedure is based on the same timing assumption that "saves" OP: labor chosen at $t - b$, slightly earlier than when materials are chosen at t.
- Point of first stage is just to get expected output:

$$y_{it} = \Phi_t \left(m_{it}, k_{it}, l_{it} \right) + \epsilon_{it}$$

where

$$\Phi_t \left(m_{it}, k_{it}, l_{it} \right) = \beta_k k_{it} + \beta_l l_{it} + f_{it}^{-1} \left(m_{it}, k_{it}, l_{it} \right)$$

... first stage no longer recovers β_l.

ACF's alternative procedure II

- After the first stage, we have $\hat{\Phi}_{it}$, expected output.
- We can construct a measure of productivity given coeffiencts:

$$\hat{\omega}_{it}\left(\beta_k, \beta_l\right) = \hat{\Phi}_{it} - \beta_k k_{it} - \beta_l l_{it}$$

- Then, non-parametrically regressing $\hat{\omega}_{it}\left(\beta_k, \beta_l\right)$ on $\hat{\omega}_{i,t-1}\left(\beta_k, \beta_l\right)$, we can construct the innovations:

$$\hat{\xi}_{it}\left(\beta_k, \beta_l\right) = \hat{\omega}_{it}\left(\beta_k, \beta_l\right) - E\left(\hat{\omega}_{it}\left(\beta_k, \beta_l\right) | \hat{\omega}_{i,t-1}\left(\beta_k, \beta_l\right)\right)$$

ACF's alternative procedure III

- Estimation relies on the following moments:

$$T^{-1}N^{-1}\sum_t \sum_i \hat{\xi}_{it}\left(\beta_k, \beta_l\right)\begin{pmatrix} k_{it} \\ l_{i,t-1} \end{pmatrix}$$

- In the second stage, these two moments are used to estimate both β_k and β_l.
- In ACF's framework, l_{it} isn't a function of ω_{it} but of $\omega_{i,t-b}$. However, labor will still be correlated with part of the innovation in productivity, so we still need to use lagged labor in the moments.
- The moment with lagged labor is very much in the spirit of OP and LP, and they actually used it as an overidentifying restriction.

Comments

- The approach also works with an investment proxy,.
- Wooldridge (2009) proposes estimating the first and second stages together. This makes computation of standard errors easier (standard GMM formulas rather than boostrap), and it improves efficiency.

"On the Identification of Production Functions:
How Heterogeneous is Productivity?"
Gandhi, Navarro, and Rivers (2006)

Overview

- A bit like ACF's critique of OP and LP (but more formal), GNR argue that past approaches suffer from non-identification due to flexible inputs.
- They argue that ACF's "solution" merely moves the non-identification problem.
- Instead, they suggest using first-order conditions from profit maximization for identification.
- Their identification argument is in the context of gross output production functions, and then they attack the use of value added production functions.

Setup

- For the general intuition for the non-identification result, consider a gross output production function:

$$\begin{aligned} y_{it} &= f\left(l_{it}, k_{it}, m_{it}\right) + \omega_{it} + \varepsilon_{it} \\ y_{it} &= f\left(l_{it}, k_{it}, m_{it}\right) + E\left(\omega_{it} | l_{i,t-1}, k_{i,t-1}, m_{i,t-1}\right) + \xi_{it} + \varepsilon_{it} \end{aligned}$$

- This is an equation in $\left(l_{it}, k_{it}, m_{it}, l_{i,t-1}, k_{i,t-1}, m_{i,t-1}\right)$. Theory tells us that there is no instrument for m_{it} which will be a source of exogenous variation.

Intuition for identification issue

- Letting h_t be the productivity inversion and $g_t(\omega_{i,t-1}) = E(\omega_{i,t}|\omega_{i,t-1})$,

$$
\begin{aligned}
y_{it} &= f(l_{it}, k_{it}, m_{it}) + g_t(h_t(l_{i,t-1}, k_{i,t-1}, m_{i,t-1})) + \xi_{it} + \varepsilon_{it} \\
m_{it} &= m_t(l_{it}, k_{it}, g_t(h_t(l_{i,t-1}, k_{i,t-1}, m_{i,t-1})) + \xi_{it})
\end{aligned}
$$

- There is no instrument for m_{it} which is both valid and relevant:
  - variation in (l_{it}, k_{it}) won't help identify effect of m_{it} if we're being flexible about the functional form of f;
  - variation in $(l_{i,t-1}, k_{i,t-1}, m_{i,t-1})$ won't help if we're being flexible about the functional form of h_t.
  - variation in ξ_{it} is not exogenous.
- A Cobb-Douglas example in their appendix suggests that functional form assumptions shouldn't help.

Using FOCs from profit maximization I

- FOC for optimal choice of intermediate inputs:

$$p_{yt} F_{M,t}\left(L_{it}, K_{it}, M_{it}\right) e^{\omega_{it}} \mathcal{E} = p_{mt}$$

where $F_{M,t} = \frac{\partial F_t}{\partial M}$ and $\mathcal{E} = E\left(e^{\varepsilon_{it}}\right)$.

- Note: this is a *static* profit maximization assumption.
- We can form a system of equations with the production function:

$$\begin{aligned} \ln p_{mt} &= \ln p_{yt} + \ln F_{M,t}\left(L_{it}, K_{it}, M_{it}\right) + \ln \mathcal{E} + \omega_{it} \\ y_{it} &= f_t\left(L_{it}, K_{it}, M_{it}\right) + \omega_{it} + \varepsilon_{it}. \end{aligned}$$

Using FOCs from profit maximization II

- We can form a system of equations with the production function:

$$\begin{aligned} \ln p_{mt} &= \ln p_{yt} + \ln F_{M,t}\left(L_{it}, K_{it}, M_{it}\right) + \ln \mathcal{E} + \omega_{it} \\ y_{it} &= f_t\left(L_{it}, K_{it}, M_{it}\right) + \omega_{it} + \varepsilon_{it}. \end{aligned}$$

- Differencing and adding m_{it} to both sides:

$$\ln \frac{m_{it} p_{mt}}{y_{it} p_{yt}} = \ln G_t\left(L_{it}, K_{it}, M_{it}\right) + \ln \mathcal{E} - \varepsilon_{it}$$

where G_t is the elasticity of output with respect to M_{it}:

$$G_t = \frac{F_{M,t}\left(L_{it}, K_{it}, M_{it}\right) M_{it}}{F_t\left(L_{it}, K_{it}, M_{it}\right)}$$

Using FOCs from profit maximization III

- The basic idea is that the input expenditure share,

$$\ln s_{it} = \ln \frac{m_{it} p_{mt}}{y_{it} p_{yt}},$$

gives us information about how the production function depends on m_{it}.

- Notice that for a Cobb-Douglas production function,

$$\ln s_{it} = \beta_m + \varepsilon_{it},$$

and β_m is the coefficient not identified by the ACF approach.