Lecture 1: Firm Productivity & Trade-Industry Link

A Selective Primer

Abhiroop Mukhopadhyay

IGC-ISI Summer School

16TH JULY 2015

Firm Productivity

- Linked to estimation of Production Functions
- For example: Cobb Douglas Production Function for i^{th} firm at time t :

$$Y_{it} = A_{it} \cdot L_{it}^{\beta} \cdot K_{it}^{\gamma} \cdot M^{\delta}$$

- Y : Measure of Output
- L : Labour, K : Capital, M: Intermediate Input: Raw Material/Energy
- A : Total Factor Productivity (TFP): Increases all factor's marginal product simultaneously

Output or Value Added

- Often data is available for real value added (VA): that is, Value of real output (Y) minus cost of Intermediate Inputs (M).
- Use of VA instead of Y in production function estimation assumes that the production function is additive separable in primary inputs (L, K) and intermediate inputs (Materials, Intermediate energy inputs).
- If you use VA, you cannot look at the coefficients of intermediate inputs.

Output or Value Added

- Additive Separability may be a strong assumption because:
  - Choice of primary inputs is in general not independent of intermediate input prices
  - Marginal rate of substitution between primary inputs will vary in the amount of intermediate inputs.
- So use data on physical output whenever possible or deflated plant sales (though using aggregate product prices for deflation can lead to some bias when products across firms vary in quality).

Empirical Model: Cobb Douglas

- $$Y_{it} = A_{it} \cdot L_{it}^{\beta} \cdot K_{it}^{\gamma} \cdot M^{\delta}$$
- Transform by taking logs.

  $$y_{it} = \beta \cdot l_{it} + \gamma \cdot k_{it} + \delta m_{it} + u_{it}$$

  where lower cases reflect logs of the variables.
- If coefficients of l and k and m are consistently estimated, then u is consistently estimated.
- Notice that $u = \log A$

OLS Regression

- Suppose we run an Ordinary Least Square Regression (OLS) on all the plant level data where our regressors are l and k and where the production function is as given before. Therefore

$$y_{it} = \beta \cdot l_{it} + \gamma \cdot k_{it} + u_{it}$$

- Assumption for consistency of estimators of β and γ :

$$Cov(l, u) = 0; \ Cov(k, u) = 0$$

Problem 1:

- The choice of k and l depend on the unobserved m.
- So classic omitted variable problem: Any residual (TFP) from the OLS regression is inconsistent.
- Standard Solution: Use Instrumental Variables (i.e. variables that affect l and k choice but are not m.)
- A lot of early studies took this approach but instruments are difficult to find.

Problem 2

- Suppose we have data for m.
- At least a part of TFP will be observed by the firm at a point in time early enough so as to allow the firm to change the factor input decision.
  $\Rightarrow$
- The error term (TFP) of the production function is expected to influence the choice of factor inputs.
- To see this, let us split the error term u into two parts so that:

$$y_{it} = \beta \cdot l_{it} + \gamma \cdot k_{it} + \delta m_{it} + \omega_{it} + e_{it}$$

- So ω_{it} is observed by the firm early enough to affect input choice. But since ω_{it} is unobserved to the economist, the conditions for consistency are not met.

Standard Remedy: Fixed Effects Estimation

- If the part of TFP that influences firm behaviour (ω_{it}) is specific to the firm and is invariant over time, then $\omega_{it} = \omega_i$
- Then differencing (or demeaning) the observations over time for each firm is going to remove ω_i and estimators will be consistent.

$$\begin{aligned} y_{it+1} - y_{it} &= \beta \cdot (l_{it+1} - l_{it}) + \gamma \cdot (k_{it+1} - k_{it}) + \delta(m_{it+1} - m_{it}) \\ &\quad + (e_{it+1} - e_{it}) \end{aligned}$$

Problem with Remedy

- A Fixed Effects Estimation method uses only within firm variation over time, which tends to be lower than the between firm variation (cross sectional variation). This means you are losing a lot of variation in the dataset. Usually these make standard errors high. Therefore you are likely to see many insignificant coefficients.
- Assumption that $\omega_{it} = \omega_i$ implies that the part of TPF that influences firm behaviour is fixed over time: May not be true; in which case we are back to the problem of inconsistency

Problem 3

- Many firm (plant) level data sets contain missing values associated with firms dropping out of the sample.
- Standard practice when using panel data: keep firms who are always in the panel: *"Creating a balanced panel"*
- But what if the exit decision of firms that are not in the sample biases estimates? (More on this later)

A Standard Approach to Address these Problems

- Olley Pakes (1996)
- A Modification to OP : Levinhson Petrin (1996)

Olley Pakes

- Dynamic Model of Firm Behaviour
- To tackle Simultaneity of Input Choice: Model must make explicit what is known at the time inputs are decided
- To deal with Firms exiting the Sample: Model must generate an exit rule.

Olley Pakes: The Model

- At the time a firm makes decisions what is known to it?:
- Information about itself (Firm State Variables)
  - Capital Stock k_t
  - An index of the firm's efficiency: ω_t :This is modelled as a random variable (a productivity shock) which is revealed to the firm at the start of every period. However, firms with higher ω today are more likely to have a higher expected ω in the future.
- Information about Input Prices (and some probability over future input prices): Assumed Same across all firms.

Olley Pakes: Decisions

- .At the beginning of a period a firm makes 3 decisions
  -Exit or Continue in Operation
- If it exits it received a selling value of the firm Φ (and never reappears)
- If it continues in operation: it chooses investment (i) and variable input (l)

Olley Pakes: Biases in OLS

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + e_{it}$$

- Endogeneity of Variable Input Choices: You expect more productive firms to hire more labour Therefore $Corr(l_{it}, \omega_{it}) > 0$: Coefficient of l :Upward bias
- Self Selection induced by plant shut down
  - Survival will depend in part on TFP (ω) revealed every period. Higher the ω,better the chance.
  - Firms which get a bad productivity shock in any period can withstand that if they have a larger capital stock. This is because they expect larger future returns.
- When we are looking at the effect of capital on output, we will underestimate it since higher capital stock allows firms with lower ω to survive.

Olley Pakes: Algorithm

- The trick is to be able to take care of the unknown ω_{it}.
- Use the fact that

$$i_{it} = i_t(\omega_{it}, k_{it})$$

- If $i_t > 0$, it can be shown that the function can be inverted, so that:

$$\omega_{it} = h_t(i_t.k_t)$$

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + h_t(i_{it}, k_{it}) + e_{it}$$

- Lets say we assume some polynomial approximation for the unknown function h
- But now how to interpret the coefficient of k.Since it will not only be β_k but also the coefficient of linear term in the unknown function h.

Olley Pakes: Algorithm

- The way out is: Define some polynomial function (3rd/4th order) that approximates $.\beta_0 + \beta_k k_{it} + h_t(i_{it}, k_{it})$. call it $\phi(i_{it}.k_{it})$
- So now

$$y_{it} = \beta_l l_{it} + \phi_{it} + e_{it}$$

- We can estimate the coefficient of l consistently, since ϕ takes into account ω.

Olley Pakes: Algorithm

- Suppose that there is no attrition. Then

$$V = y_{it} - \beta_l l_{it} = \beta_k k_{it} + \omega_t + e_{it}.$$

- Since ω_t follow a time series process, we assume first order. So $\omega_t = g(\omega_{t-1})$
- But now, we can substitute for ω_{t-1} by $h_{t-1}(i_{it-1}, k_{it-1})$
- So we get $\omega_t = g(h_{t-1}(i_{it-1}, k_{it-1}))$. But this is just $g(\phi_{t-1} - .\beta_0 + \beta_k k_{it-1})$
- Note that now we have

$$V = y_{it} - \beta_l l_{it} = \beta_k k_{it} + g(\phi_{t-1} - .\beta_0 + \beta_k k_{it-1}) + e_{it}.$$

- We don't know g.Again take an approximation. Estimate to get β_k.

Olley Pakes: Algorithm

- With Attrition: Estimate a probit model of survival. Predict probability of suvival $\hat{p_t}$.A little more arithmetic and taking expectations condition on survival show that:

$$V = y_{it} - \beta_l l_{it} = \beta_k k_{it} + g(\phi_{t-1} - .\beta_0 + \beta_k k_{it-1}, \hat{p_{t-1}}) + e_{it}.$$

- with the $\hat{p_t}$., we are correcting for attrition when we run the model on firms that survive.
- This is in spirit the same as a heckman type correction. As a matter of fact, if there is only attrition and no correlated productivity differences, then its an approximation to a heckman correction. [If you dont know heckman correction, ignore this line!]

Levinhson Petrin: Algorithm

- One of problems with OP is that it requires the inversion and this is true only when $i > 0$
- In real life data, huge number of firms for many years report 0 investment. Then OP is not useful
- LP idea: the same idea as OP except:

$$\omega_t = h(m_t, k_t)$$

where m is intermediate input: materials.

- This induces some endogeneity issues but allows the inversion more since intermediate inputs are always non zero.

Plant and Firm Level Analysis of Trade: Some Background

- In the literature, 3 static predictions of trade theory

- Protection can change firm's pricing behavior
- When trade policies affect prices, they generally also change the set of active producers (market share reallocation) and/or their output levels (scale). These affect productivity.
- Changes in intensity of foreign competition and/or firm's opportunities to export can affect their technical efficiency.

Pricing

- Static Optimization

$$\frac{p}{c} = \frac{\eta}{\eta - 1}$$

- The Right hand side is elasticity of demand. Trade liberalization increases η. Mark up therefore falls.
- Many ways to motivate this:

-Demand Elasticity for domestic goods rises as relative price of foreign goods fall (if they are substitutes)
- Removal of import quota has the same effect.
-Increase in product varieties increases demand elasticity of domestic firms.

Price Cost Margin: Empirical

- Usually no data on prices and marginal cost (esp the latter)
- Let PCM be the price cost margin.

$$PCM = \frac{\Pi_{it}}{p_{it}q_{it}} + \frac{(r_t + \delta)\,k_{it}}{p_{it}q_{it}}$$

- Controlling for ratio of capital stocks to sales, variables that measure intensity of foreign competition should contribute nothing to explain PCM if $\Pi_{it} = 0$ (perfect competition)
- On the other hand, if $\Pi_{it} > 0$ and trade liberalization increases demand elasticity, then Π_{it} will fall.
- Therefore in the regression

$$PCM_{it} = \beta_o + \beta_1(\frac{k_{it}}{p_{it}q_{it}}) + \beta_2\,\mathrm{Im}_{it} + ... + \varepsilon_{it}$$

we expect coefficient of *Im* (proxy for intensity of import competition) to be negative.

Price Cost Margin: Alternate

- If $q_{it} = A_{it} h(v_{it})$ where $v_{it} = (v_{it}^1, v_{it}^2, ..., v_{it}^J)$ is the vector of J factor inputs, then output growth can be decomposed as:

$$d\ln(q_i) = \left(\frac{\eta}{\eta - 1}\right) \sum \left(\frac{v_i^j w_j}{p_i q_i}\right) d\ln(v_i^j) + d\ln(A_i)$$

- Regression of output growth on the share weighted rate of input growth, treating $d\ln(A_i)$ as the mean productivity growth plus noise. We get the *PCM* as the slope coefficient.
- let the slope coefficient vary through time (pre reform/ post reform): one can test if trade policy affects mark up.
- Problem: This requires instruments that are correlated with input growth but not with the transitory part of productivity.

Firm Size Distribution/Market Share and its effect on productivity

- When demand elasticities rise with liberalization, price cost mark ups are squeezed. This should induce exit until the remaining firms can make up on volume what they lost on margin.
- Let B be industry productivity. It can be shown that

$$\frac{dB}{B} = \Delta Eff_{Scale\ Economies} + \Delta Eff_{Mkt\ Sh\ \mathrm{Re}\,all} + \Delta Eff_{intra\ firm\ TFP}$$

- There is evidence that the first two components have a big effect.

Intra Firm Productivity Gains

- Fall in price of Imported Inputs prices
- Product Variety Changes: Increase in menu of available inputs (We discuss this more in detail)
- Technology Diffusion through imports (we won't take about it more in this lecture)

Using Theory to Generate Testable Implications

- In the second lecture, we will explore an empirical fact: As input tariffs fall, the number of product varieties in an Industry rise.
- Usual empirical work can provide a correlation, sometimes even causality: But without a mechanism, it is difficult to interpret results.
- In this part of the lecture, we will provide an example of a theoretical structure that motivates empirical work on the empirical fact mentioned above

Model

- Specify a Cobb Douglas Production Function:

$$Y_q = AL^{\alpha_{Lq}} S^{\alpha_{Sq}} \prod_{i=1}^{I} X_i^{\alpha_{iq}}$$

where Y_q denotes Output of product q, A: TFP , L :Labour S :Other non tradeable Inputs (electricity, water). $X_i = fn(X_{iD}, X_{iF})$ are domestic and imported inputs. Let the coefficients add up to 1.

- There is a fixed cost for production of final good.

Cost per Unit/ CES Aggregator

- The Minimum cost of manufacturing one unit of output is:

$$C_q = A^{-1}\left[\prod P_i^{\alpha_{iq}}\right]\left(P_L^{\alpha_{Lq}} P_S^{\alpha_{Sq}}\right) \bullet Coeffs$$

- Each input sector i has a domestic and an imported component that are combined according to a CES aggregator:

$$X_i = \left(X_{iD}^{\frac{\gamma_i - 1}{\gamma_i}} + X_{iF}^{\frac{\gamma_i - 1}{\gamma_i}}\right)^{\frac{\gamma_i}{\gamma_i - 1}}$$

- γ_i is the elasticity of substitution between the two input bundles.

Price of Input Bundle

- The over all price index of input industry i is the weighted average of the price index for the domestic (Π_{iD}) and foreign input bundles (Π_{iF})
$$P_i = (\Pi_{iD})^{\omega_{iD}} (\Pi_{iF})^{\omega_{iF}}$$

Import Input Bundle

- The imported input industry X_{IF} is itself a CES aggregator of imported varieties:

$$X_{iF} = \left[\sum_{v \in I_{iF}} \alpha_{iv}^{\sigma_i} x_{iv}^{\frac{\sigma_i - 1}{\sigma_i}} \right]^{\frac{\sigma_i}{\sigma_i - 1}}$$

- Cost of purchasing one unit of foreign input bundle:

$$c = \left[\sum_{v \in I_{iF}} \alpha_{iv} p_{iv}^{\sigma_i - 1} \right]^{\frac{1}{\sigma_i - 1}}$$

Price Index over a constant set of Imported Varieties

- Feenstra (1994) showed that relative to period 1 the price index in period 2 if the set of imported varieties was constant was:

$$P_{iF}^{Conv} = \frac{c_2}{c_1} = \prod_{v \in I_{iF}} \left(\frac{p_{iv}^2}{p_{iv}^1} \right)^{w_{iF}}$$

Price Index with new varieties

- Feenstra (1994):

$$\Pi_{iF} = P_{iF}^{Conv} \Lambda_{iF}$$

Here Λ_{iF} is the variety index. : measures the expenditure of varieties that are available in both periods relative to the expenditure on the varieties that are available in the current period.

- The more important the new varieties are, the lower the Λ_{iF}. and smaller the exact price index will be relative to the conventional index P_{iF}^{Conv}
- The more substitutable the varieties are, the lower the difference between exact and conventional price.

Cost After Substitution of Price of Foreign Input

- Recall

$$C_q = A^{-1}\left[\prod P_i^{\alpha_{iq}}\right]\left(P_L^{\alpha_{Lq}} P_S^{\alpha_{Sq}}\right) \bullet Coeffs$$

- Substitution of P_{iF} and taking logs yields:

$$\begin{aligned}\ln C_q &= \left\{\sum \alpha_{iq}\omega_{iF} \ln P_{iF}^{Conv} + \alpha_{Lq} \ln P_L + \alpha_{sq} \ln P_S\right\} \\ &+ \left\{\sum \alpha_{iq}\omega_{iF} \ln \Lambda_{iF}\right\} + v\end{aligned}$$