IO1 Pset 1

Dylan Baker

October 2024

1 Problem 1

Suppose value added is a general function of TFP, labor, and capital,i.e., $Y = f(A, L, K)$. Assuming perfect competition and constant returns to scale, derive the following expression for logged TFP: $a = y - \alpha l - (1-\alpha)k$, where lower-case letters denote logs of their level counterparts and α the wage bill as a share of total revenues.

$$\begin{aligned}
Y &= f(A, L, K) \\
\Rightarrow dY &= \frac{\partial f}{\partial A} dA + \frac{\partial f}{\partial L} dL + \frac{\partial f}{\partial K} dK && \text{Total differential} \\
\Rightarrow \frac{dY}{Y} &= \frac{\partial f}{\partial A}\frac{dA}{Y} + \frac{\partial f}{\partial L}\frac{dL}{Y} + \frac{\partial f}{\partial K}\frac{dK}{Y} && \text{Divide by } Y \\
&= \frac{A}{Y}\frac{\partial f}{\partial A}\frac{dA}{A} + \frac{L}{Y}\frac{\partial f}{\partial L}\frac{dL}{L} + \frac{K}{Y}\frac{\partial f}{\partial K}\frac{dK}{K} && (1)
\end{aligned}$$

Note that a cost maximizing firm will set the marginal product equal to the ratio of the real input prices. That is,

$$\frac{\partial F}{\partial L} = \frac{w}{P}$$
$$\frac{\partial F}{\partial K} = \frac{r}{P}$$

Additionally,

$$d \ln Y = \frac{dY}{Y}$$

and the analogous equalities hold for A, L, and K.

Thus, we can return to (1) to get:

$$
\begin{aligned}
\frac{dY}{Y} =& \frac{A}{Y}\frac{\partial f}{\partial A}\frac{dA}{A} + \frac{L}{Y}\frac{\partial f}{\partial L}\frac{dL}{L} + \frac{K}{Y}\frac{\partial f}{\partial K}\frac{dK}{K} \\
=& \frac{A}{Y}\frac{\partial f}{\partial A}\frac{dA}{A} + \frac{L}{Y}\frac{w}{P}\frac{dL}{L} + \frac{K}{Y}\frac{r}{P}\frac{dK}{K} \\
\Rightarrow d\ln Y =& \frac{A}{Y}\frac{\partial f}{\partial A}d\ln A + \frac{L}{Y}\frac{w}{P}d\ln L + \frac{K}{Y}\frac{r}{P}d\ln K \\
=& d\ln A + \frac{L}{Y}\frac{w}{P}d\ln L + \frac{K}{Y}\frac{r}{P}d\ln K && \text{since } \frac{\partial f}{\partial A} = \frac{Y}{A} \\
\Rightarrow dy =& da + \alpha dl + (1-\alpha)dk \\
\Rightarrow da =& dy - \alpha dl - (1-\alpha)dk && \text{rearrange} \\
\Rightarrow a =& y - \alpha l - (1-\alpha)k
\end{aligned}
$$

2 Problem 2

Do the following Monte Carlo exercise to convince yourself of the possible endogeneity problems in production function estimation. (Matlab, Stata's matrix programming language Mata, or even Excel can be used for this, though eventually you're going to have to move the data into a program that runs regressions.)

Construct a balanced panel dataset of 100 firms, each operating in 50 periods (for a total of 5000 observations). You will have to create data for firms' value added (output), TFP, labor and capital inputs, and real wage rates.

The firm's production function is, $y_{it} = a_{it} + 0.7l_{it} + 0.3k_{it}$, where all variables are expressed in logs.

Logged TFP is described by the following statistical process: $a_{it} = \gamma_i + \omega_{it}$. Here γ_i is a firm fixed effect and ω_{it} follows an AR(1) process: $\omega_{it} = \rho\omega_{it-1} + \varepsilon_{it}$. Assume $\gamma_i \sim$ N$(0, 0.25)$-that is, has a standard deviation of 0.5 -across firms, $\rho = 0.8$, and ε_{it} is i.i.d. N$(0, 0.01)$ across firms and time. The firm observes γ_i and ω_{it-1} at the beginning of period t and then chooses its inputs. After that, ε_{it} is revealed.

Assume capital is exogenous to the firm and randomly determined (obviously nonsensical, but it makes things easy). Specifically, k_{it} is i.i.d. N$(0, 0.01)$ across firms and time. Likewise, the log real wage rate facing the firm is also i.i.d., with $w_{it} \sim$ N$(0, 0.25)$.

a. After constructing these exogenous variables, figure out what the firm's labor inputs are. To do so, assume the firm is a price taker in its output and labor markets. Use the expression for the profit-maximizing labor level to derive the labor demand equation, and construct the firms' labor input values according to this equation.

b. Construct the firms' output levels from their TFP, capital, and labor values.

c. Estimate the production function by regressing y on k and l using OLS. Are the coefficients biased? In which direction, if so? Explain why.

d. Now estimate the production function using firm fixed effects. Is there a bias, and how does this compare to (c)? Explain. (Just for fun, try this using random effects too.)

e. Is there anything in the data you created that would be a suitable instrument? Explain why, if so, and then estimate the production function using IV. How do the results change?

2.1 Part A

After constructing these exogenous variables, figure out what the firm's labor inputs are. To do so, assume the firm is a price taker in its output and labor markets. Use the expression for the profit-maximizing labor level to derive the labor demand equation, and construct the firms' labor input values according to this equation.

A given firm, in a given period, is solving:

$$\max_{L_{i,t}} P\mathbb{E}[A_{i,t} \mid \omega_{i,t-1}] L_{i,t}^{\alpha} K_{i,t}^{1-\alpha} - W L_{i,t}$$

which gives us:

$$\begin{aligned}\frac{\partial \pi}{\partial L} &= P\mathbb{E}[A_{i,t} \mid \omega_{i,t-1}] \alpha L_{i,t}^{\alpha-1} K_{i,t}^{1-\alpha} - W = 0 \\ \Rightarrow L_{i,t}^{*} &= \left(\frac{\alpha \mathbb{E}[A_{i,t} \mid \omega_{i,t-1}]}{W/P} \right)^{\frac{1}{1-\alpha}} K_{i,t}\end{aligned}$$

Switching to logs and plugging in $\alpha = 0.7$ we get:

$$l_{it}^{*} = k_{it} + \frac{10}{3} \left(\ln(0.7) + \gamma_i + \rho \omega_{it-1} - w_{it} \right)$$

2.2 Part B

Construct the firms' output levels from their TFP, capital, and labor values.

See code.

2.3 Part C & D

Part C: Estimate the production function by regressing y on k and l using OLS. Are the coefficients biased? In which direction, if so? Explain why.

Part D: Now estimate the production function using firm fixed effects. Is there a bias, and how does this compare to (c)? Explain. (Just for fun, try this using random effects too.)

Dependent Variable:	y	
Model:	(1)	(2)
Variables		
Constant	0.1929***	
	(0.0056)	
l	0.8408***	0.7125***
	(0.0022)	(0.0014)
k	0.1710***	0.2992***
	(0.0503)	(0.0187)
Fixed-effects		
firm_id		Yes
Fit statistics		
Observations	5,000	5,000
R^2	0.96675	0.99472
Within R^2		0.98627

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The estimate for l is biased up in each case, though most so for the baseline OLS estimate. The estimate for k is biased down in each case, though, again, most so for the baseline OLS estimate. The random effects version is between the two regressions presented in the table here.

Here, we are suffering from omitted variable bias, since l_{it} is correlated with the error term, since it's a function of ω_{it-1} and γ_i. By including fixed effects, we are able to partially address this by accounting for the permanent piece of firm productivity.

2.4 Part E

Is there anything in the data you created that would be a suitable instrument? Explain why, if so, and then estimate the production function using IV. How do the results change?

We can use log real wage as an instrument. If we do so, we get the below results:

Dependent Variable:	y
Model:	(1)
Variables	
Constant	0.0439***
	(0.0081)
l	0.6987***
	(0.0041)
k	0.3078***
	(0.0681)
Fit statistics	
Observations	5,000
R^2	0.93924
Adjusted R^2	0.93921

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The estimates are still slightly off, but they are quite close.

3 Problem 3

A simple model of equilibrium industry productivity and output heterogeneity.

Suppose that firms in an industry have a variable profit function $\pi(\phi_i, \mu, \sigma)$, where ϕ_i is the firm's productivity level, and σ is the elasticity of substitution between the output of industry firms. μ is a value that indexes the intensity of industry competition (like s in the Ericson-Pakes model); assume that it is an increasing function of both the average productivity level and mass of industry producers. Assume that $\pi_\phi(\cdot) > 0$ and $\pi_\mu(\cdot) < 0$; i.e., being more productive increases profits, and stiffer competition decreases profits. There is also a fixed cost of production, f, so total operating profits are $\Pi(\phi_i, \mu, \sigma, f) = \pi(\phi_i, \mu, \sigma) - f$.

Entry is a simple two stage game. In the first stage, a very large number of ex-ante identical potential entrants decide whether to enter into the industry. If they choose to do so, they pay a sunk entry cost, s, and learn their own productivity draw (taken from a common distribution $G(\phi)$ over $[\phi, \phi']$) as well as those of all other firms who paid the entry cost. Upon learning these productivity draws, they simultaneously decide whether or not to produce in the second stage, and earn operating profits as specified above.

Two conditions hold in equilibrium. No firm produces at a loss (they can always choose not to produce and earn zero profits). Second, there is free entry: in equilibrium, the expected value of entry is equal to s.

a. Use the first equilibrium condition and the firm's operating profit function to argue that there is a unique cutoff productivity level, ϕ^*, where firms with productivity draws $\phi_i < \phi^*$ do not produce. Assume that $G(\phi)$ has a large enough domain so that $\pi\left(\varphi^l, \mu, \sigma\right) < f$ and $\pi(\varphi^u, \mu, \sigma) > f$ for all μ and σ. Use the unique cutoff result to write a zero-marginal-profit equation that embodies the first equilibrium condition.

b. Derive, using this equilibrium equation, the condition under which $d\phi^*/df > 0$; that is, higher fixed production costs raise the efficiency bar for profitable operation. Give the intuition behind the condition and this result. (Note that the average productivity level in the industry is a function of ϕ^*.)

c. Write down an equation that embodies the second (free-entry) equilibrium condition. (Remember that the expected value of entry is equal to the probability of successful entry multiplied by average profits conditional upon successful entry.) Use it to show that $d\phi^*/ds < 0$; i.e., high entry costs imply it is easier for inefficient firms to operate profitably. (Hint: use the zero-cutoff-profit equation to substitute in for the fixed cost in the free-entry equation before doing the comparative statics.) Why does an increase in the lump-sum sunk cost have an effect opposite that of an increase in the lump-sum fixed production cost?

d. Show the condition on the profit function required so that $d\phi^*/d\sigma > 0$, i.e., an increase in substitutability makes it more difficult for inefficient producers to survive. Give the intuition for the derived condition. Do you find it intuitively plausible?

e. How would you empirically test these implications?

3.1 Part A

Use the first equilibrium condition and the firm's operating profit function to argue that there is a unique cutoff productivity level, ϕ^*, where firms with productivity draws $\phi_i < \phi^*$ do not produce. Assume that $G(\phi)$ has a large enough domain so that $\pi\left(\varphi^l, \mu, \sigma\right) < f$ and $\pi\left(\varphi^u, \mu, \sigma\right) > f$ for all μ and σ. Use the unique cutoff result to write a zero-marginal-profit equation that embodies the first equilibrium condition.

A firm will produce if

$$\pi\left(\phi_i, \mu, \sigma\right) > f$$

Given that for any fixed μ and σ, there is some $\underline{\phi}$ such that $\pi\left(\underline{\phi}, \mu, \sigma\right) < f$ and some $\overline{\phi}$ such that $\pi\left(\overline{\phi}, \mu, \sigma\right) > f$. With this established, the desired result follows from the fact that $\pi_\phi(\cdot) > 0$.[1]

3.2 Part B

Derive, using this equilibrium equation, the condition under which $d\phi^*/df > 0$; that is, higher fixed production costs raise the efficiency bar for profitable operation. Give the intuition behind the condition and this result. (Note that the average productivity level in the industry is a function of ϕ^*.)

Use total differentiation:

[1] I'm taking π to be continuous in ϕ.

$$\left(\frac{d\pi\left(\phi^*,\mu\left(\phi^*\right)\right)}{d\phi^*}+\frac{d\pi\left(\phi^*,\mu\left(\phi^*\right)\right)}{d\mu\left(\phi^*\right)}\frac{d\mu\left(\phi^*\right)}{d\phi^*}\right)d\phi^*-df=0$$
$$\Leftrightarrow\frac{d\phi^*}{df}=\frac{1}{\frac{d\pi\left(\phi^*,\mu\left(\phi^*\right)\right)}{d\phi^*}+\frac{d\pi\left(\phi^*,\mu\left(\phi^*\right)\right)}{d\mu\left(\phi^*\right)}\frac{d\mu\left(\phi^*\right)}{d\phi^*}}$$

Given

$$\frac{d\pi\left(\phi^*,\mu\left(\phi^*\right)\right)}{d\mu\left(\phi^*\right)}<0$$
$$\frac{d\pi\left(\phi^*,\mu\left(\phi^*\right)\right)}{d\phi^*}>0$$

we know that

$$\frac{d\phi^*}{df}>0\Rightarrow\frac{d\mu\left(\phi^*\right)}{d\phi^*}>0$$

and

$$\frac{d\pi\left(\phi^*,\mu\left(\phi^*\right)\right)}{d\phi^*}\frac{d\mu\left(\phi^*\right)}{d\phi^*}>-\frac{d\pi\left(\phi^*,\mu\left(\phi^*\right)\right)}{d\mu\left(\phi^*\right)}$$

The intuition behind this condition is that we are looking at the difference in the rising profit of a productive firm against the falling profit of a firm in a more competitive industry.

3.3 Part C

Write down an equation that embodies the second (free-entry) equilibrium condition. (Remember that the expected value of entry is equal to the probability of successful entry multiplied by average profits conditional upon successful entry.) Use it to show that $d\phi^*/ds<0$; i.e., high entry costs imply it is easier for inefficient firms to operate profitably. (Hint: use the zero-cutoff-profit equation to substitute in for the fixed cost in the free-entry equation before doing the comparative statics.) Why does an increase in the lump-sum sunk cost have an effect opposite that of an increase in the lump-sum fixed production cost?

Using g to be the density of G, the free entry condition is given by:

$$\int_{\phi^*}^{\phi^u}(\pi(\phi,\mu,\sigma)-f)g(\phi)d\phi-s=0$$

Then we can write:

$$\left[-\left[\pi\left(\phi^*,\mu,\sigma\right)-f\right]g\left(\phi^*\right)+\int_{\phi^*}^{\phi^*}\pi_\mu(\phi)\frac{\partial\mu}{\partial\phi^*}g(\phi)d\phi\right]d\phi^*-ds=0 \quad \text{total differentiation}$$

$$\Rightarrow\left\{\int_{\phi^*}^{\phi^u}\pi_\mu(\phi)\frac{\partial\mu}{\partial\phi^*}g(\phi)d\phi\right\}d\phi^*=ds \qquad \text{since } \pi\left(\phi^*,\mu,\sigma\right)-f=0$$

$$\Rightarrow\frac{d\phi^*}{ds}=\frac{1}{\left\{\int_{\phi^*}^{\phi^u}\pi_\mu(\phi)\frac{\partial\mu}{\partial\phi^*}g(\phi)d\phi\right\}}$$

Noting $\pi_\mu(\cdot)<0$ and $\frac{\partial\mu}{\partial\phi^*}>0$, $\frac{d\phi^*}{ds}<0$.

An increase in f corresponds to a relative increase in the difficulty of production for less productive firms. If s is increase, meanwhile, it is relatively easier for less productive firms to produce, since the fixed cost of discovering productivity is raised.

3.4 Part D

Show the condition on the profit function required so that $d\phi^*/d\sigma>0$, i.e., an increase in substitutability makes it more difficult for inefficient producers to survive. Give the intuition for the derived condition. Do you find it intuitively plausible?

We have

$$\int_{\phi^*}^{\phi^u}\left(\pi(\phi,\mu,\sigma)-\pi\left(\phi^*,\mu,\sigma\right)\right)g(\phi)d\phi-s=0$$

from

$$\pi\left(\phi^*,\mu,\sigma\right)=f$$

Using total differentiation and the zero-marginal-profit equation, we get:

$$\left\{\int_{\phi^*}^{\phi^\mu}\left[\pi_\mu(\phi)\frac{\partial\mu}{\partial\phi^*}-\pi_\phi\left(\phi^*\right)-\pi_\mu\left(\phi^*\right)\frac{\partial\mu}{\partial\phi^*}\right]g(\phi)d\phi\right\}d\phi^*+\left\{\int_{\phi^*}^{\phi^u}\left[\pi_\sigma(\phi)-\pi_\sigma\left(\phi^*\right)\right]g(\phi)d\phi\right\}d\sigma=0$$

Thus,

$$\frac{d\phi^*}{d\sigma}=-\frac{\int_{\phi^*}^{\phi^u}\left[\pi_\sigma(\phi)-\pi_\sigma\left(\phi^*\right)\right]g(\phi)d\phi}{\int_{\phi^*}^{\phi^\mu}\left[\pi_\mu(\phi)\frac{\partial\mu}{\partial\phi^*}-\pi_\phi\left(\phi^*\right)-\pi_\mu\left(\phi^*\right)\frac{\partial\mu}{\partial\phi^*}\right]g(\phi)d\phi}$$

The denominator is positive by earlier parts and $\frac{d\phi^*}{d\sigma} > 0$. Thus,

$$\int_{\phi^*}^{\phi^u} \left[\pi_\sigma(\phi) - \pi_\sigma\left(\phi^*\right)\right] g(\phi) d\phi < 0$$

If σ changes and elicits a positive effect on ϕ such that $\phi > \phi^*$, then it becomes increasingly challenging for less productive firms to survive, moving the cutoff.

3.5 Part E

How would you empirically test these implications?

We are looking for exogenous variation in f, s, and σ. These are challenging given the clear endogeneity issues accompanying them. In principle, we'd be interested in something like valid instruments or natural experiments lending themselves to design-based identification strategies. However, this is challenging in its own right.