Production Functions

Bryan S. Graham, UC - Berkeley & NBER

March 30, 2023

Let O_t, K_t and L_t respectively denote a firm's period $t = 1, \ldots, T$ output, capital stock and employed labor. Output is produced with the value-added Cobb-Douglas technology

$$O_t = K_t^{\alpha} L_t^{\beta} \exp\left(U_t + \varepsilon_t\right),$$

where total factor productivity (TFP), $A_t \overset{def}{\equiv} \exp\left(U_t + \varepsilon_t\right)$, consists of two components, a forecastable component, U_t, which is revealed to the firm at the beginning-of-period t, and an unforecastable component, ε_t, which is revealed at the end-of-period t when output is "realized". Our goal is to recover the capital and labor output elasticities, α and β, as well as to learn the process which governs the evolution of U_t.

Griliches & Mairesse (1998) provides an introduction to the challenges of production function identification and estimation. The basic method outlined in this note is an adaption of an approach due to Olley & Pakes (1996); it is also closely related to ideas feature in the dynamic panel data literature (e.g., Blundell & Bond, 2000).

Each period the firm makes a sequence of decisions under an information set which grows with time. The sequence of these decisions, and the nature of the information set used to make them, will play a central role in the development of our ideas. In each period it is useful to imagine the firm taking the following actions.

1. At the beginning of the period, each firm chooses its current labor input, L_t, to maximize expected current profits. Let P_t and W_t respectively denote the output price and wage rate faced by the firm. The observed labor input is thus

$$L_t = \arg\max_{l_t} \mathbb{E}\left[P_t K_t^{\alpha} l_t^{\beta} \exp\left(U_t + \varepsilon_t\right) - W_t l_t \middle| \mathcal{I}_t \right], \tag{1}$$

where $\mathcal{I}_t$ denotes the firm's beginning-of-period t information set. Here $P_t O_t = P_t K_t^{\alpha} l_t^{\beta} \exp\left(U_t + \varepsilon_t\right)$ equals total revenue, while $W_t L_t$ is the firm's wage bill or total variable input costs. The firm's beginning-period information set includes the histories $W_1^t = (W_1, \ldots, W_t)'$, $P_1^t = (P_1, \ldots, P_t)'$, $K_1^t = (K_1, \ldots, K_t)'$, $L_1^{t-1} = (L_1, \ldots, L_{t-1})'$, $Y_1^{t-1} = (Y_1, \ldots, Y_{t-1})'$, $U_1^{t-1} = (U_1, \ldots, U_{t-1})'$ and $I_1^t = (I_1, \ldots, I_t)$. Here I_t denotes a firm's period t capital expenditures or investment. Firms do not know U_t or ε_t when making their hiring decisions. This last assumption is important for the results which follow.

2. After choosing their labor input, firms choose their total capital expenditures, I_t. This decision determines their period $t+1$ capital stock according to the perpetual inventory rule
$$K_{t+1} = (1-\delta) K_t + I_t,$$
where δ is the capital depreciation rate. While firm's are able to freely adjust their labor input each period, capital is a dynamic input. It's level is determined one period in advance. Capital investment decisions have consequence for a firm which persist across many periods. We assume that U_t is observed prior to the capital investment decision, but after the labor hiring decision. This assumption is somewhat artificial, but also crucial.

3. Finally the firm produces their period t output. At this stage the firm also learns ε_t; although this information plays no role in future decisions since we assume ε_t is a pure white noise idiosyncratic production shock.

Will consider each of these steps in detail.

Step 1: the firm's hiring decision

Begin by considering the firm's employment decision. The first order condition associated with (1) yields
$$\beta P_t K_t^{\alpha} L_t^{\beta-1} \mathbb{E}\left[\exp\left(U_t + \varepsilon_t\right) \middle| \mathcal{I}_t\right] - W_t = 0, \tag{2}$$
We will assume that the persistent component of productivity, U_t, follows a first order linear Markov process:
$$U_t = \lambda_t + \rho U_{t-1} + V_t \tag{3}$$
with V_t, the period t "productivity innovation", Gaussian
$$V_t | \mathcal{I}_t \sim \mathcal{N}\left(0, \sigma^2\right). \tag{4}$$

The white noise output shock is also Gaussian

$$\varepsilon_t | \mathcal{I}_t \sim \mathcal{N}\left(0, \tau^2\right) \tag{5}$$

and orthogonal to V_t (i.e., $\mathbb{C}\left(V_t, \varepsilon_t | \mathcal{I}_t\right) = 0$).

An important implication of (3) and (4) is that

$$U_t | \mathcal{I}_t \overset{D}{=} U_t | U_{t-1}.$$

In words: the only element of the beginning-of-period information set, $\mathcal{I}_t$, that is helpful for predicting current productivity, U_t, is productivity last period, U_{t-1}. This is a strong assumption. For example, it rules out purposeful firm behaviors, such as research and development (R&D), that alter the trajectory of TFP. We will likely return to this observation in a review sheet question.

The Gaussian assumption is not essential to our development, but will allow for a cleaner analysis.[1] Recall that if $Z \sim \mathcal{N}\left(\mu, \sigma^2\right)$, then $\exp(Z)$ follows a log-normal distribution with $\mathbb{E}\left[\exp(Z)\right] = \exp\left(\mu + \frac{\sigma^2}{2}\right)$. Using this fact we can calculate a firm's beginning-of-period expectation about is current productivity:

$$\mathbb{E}\left[A_t | \mathcal{I}_t\right] = \mathbb{E}\left[\exp\left(U_t + \varepsilon_t\right) | \mathcal{I}_t\right] = \exp\left(\lambda_t + \rho U_{t-1} + \frac{\sigma^2 + \tau^2}{2}\right). \tag{6}$$

Taking logs of (2) and substituting in the right-hand-side of (6) yields a (log-) labor input demand equation of

$$\begin{aligned}
\ln L_t &= \frac{\ln(\beta)}{1-\beta} + \frac{\alpha}{1-\beta}\ln K_t + \frac{1}{1-\beta}\ln\left(\frac{P_t}{W_t}\right) + \frac{1}{1-\beta}\ln\left(\mathbb{E}\left[\exp\left(U_t + \varepsilon_t\right) | \mathcal{I}_t\right]\right) \\
&= \frac{\ln(\beta)}{1-\beta} + \frac{\alpha}{1-\beta}\ln K_t + \frac{1}{1-\beta}\ln\left(\frac{P_t}{W_t}\right) + \frac{\lambda_t}{1-\beta} + \frac{\rho}{1-\beta}U_{t-1} + \frac{1}{1-\beta}\frac{\sigma^2+\tau^2}{2} \\
&= \frac{\ln(\beta) + \sigma^2 + \tau^2 + \lambda_t}{1-\beta} + \frac{\alpha}{1-\beta}\ln K_t + \frac{1}{1-\beta}\ln\left(\frac{P_t}{W_t}\right) + \frac{\rho}{1-\beta}U_{t-1} \\
R_t &= \zeta_t + \frac{\alpha}{1-\beta}S_t + \frac{1}{1-\beta}\ln\left(\frac{P_t}{W_t}\right) + \frac{\rho}{1-\beta}U_{t-1},
\end{aligned} \tag{7}$$

where $R_t \overset{def}{\equiv} \ln L_t$ and $S_t \overset{def}{\equiv} \ln K_t$.

Firms hire more labor when their capital stock is large, their output price is high (and/or their the wage rate is low), and when they believe their current period productivity will be

[1] We could also relax the linearity part of the Markov assumption.

high.

Step 2: the firm's investment decision

After firms choose their labor level, they set their capital expenditure level for the current period. This decision doesn't impact their current output because investment today only influences capital next period; current capital stocks are fixed one period in advance.

The investment decision is a forward-looking one: firms try to choose an investment level which will make their capital stock optimal for next (and future) periods. The analysis of dynamic optimization problems is a fascinating area, for our purposes we will just appeal to a result in Olley & Pakes (1996). They show that the firm's optimal investment rule takes the form

$$I_t = \nu\left(K_t, U_t; W_t, P_t\right)$$

for some (unknown) function $\nu\left(k_t, u_t; w_t, p_t\right)$

If we assume that all firms face the same output prices and wages, then we can write

$$I_t = \nu_t\left(K_t, U_t\right).$$

Olley & Pakes (1996) further show that $\nu_t\left(k_t, u_t\right)$ is strictly increasing (and hence invertible) in its second argument. Since productivity is persistent, firms with high levels of productivity today anticipate having high productivity in the future. Consequently they make larger capital expenditures since firms with high levels of productivity prefer large capital stocks. Labor levels do not enter the firms investment rule, since the firm is free to costlessly re-optimize the amount of labor they employ each period. It is interesting to think about how making labor adjustment costly would change our analysis.

The assumption that all firms face the same output price and wage is very strong. It is made because we rarely observe such variables at the firm level. Typically the methods outlined is this note are applied to a panel of firms within the same narrow industry (e.g., semiconductor firms). In such a setting it is still a strong assumption, but not a priori unreasonable.

Let $Y_t \overset{def}{\equiv} \ln O_t$; taking logs of the production function yields

$$\begin{aligned}
Y_t &= \beta R_t + \alpha S_t + U_t + \varepsilon_t \\
&= \beta R_t + \alpha S_t + \nu_t^{-1}\left(K_t, I_t\right) + \varepsilon_t \\
&= \beta R_t + h_t\left(S_t, I_t\right) + \varepsilon_t,
\end{aligned}$$

where we use the substitution $U_t = \nu_t^{-1}(K_t, I_t)$ in the second equality and define $h_t(S_t, I_t) \stackrel{def}{\equiv} \alpha S_t + \nu_t^{-1}(K_t, I_t)$ to get the third equality.

Next observe that $\mathbb{C}(V_t, \varepsilon_t | \mathcal{I}_t) = 0$ implies that $\mathbb{C}(U_t - \lambda_t - \rho U_{t-1}, \varepsilon_t | \mathcal{I}_t) = 0$ and hence orthogonality of ε_t and U_t. Investment is a function of K_t – a component of $\mathcal{I}_t$ – and U_t. From these observations and (5) we have the mean independence restriction

$$\mathbb{E}[\varepsilon_t | R_t, S_t, I_t] = 0.$$

and hence

$$\mathbb{E}[Y_t | R_t, S_t, I_t] = \beta R_t + h_t(S_t, I_t). \tag{8}$$

We can use (8) to identify the labor output elasticity, β. To understand why note that

$$Y_t - \mathbb{E}[Y_t | S_t, I_t] = \beta(R_t - \mathbb{E}[R_t | S_t, I_t]) + \varepsilon_t.$$

If we look at two firms with identical period t capital stocks and identical capital expenditure levels, then we can conclude that they have identical values for U_t, the persistent component of productivity. Output across firms varies due to labor, capital and total factor productivity (TFP). If we hold the last two "fixed", then we can attribute variation in output levels across the two firms as due to their varying labor levels. Intuitively a firm's current capital stock and investment decisions "reveals" the persistent component of productivity, U_t, to the econometrician.

In thinking about this thought experiment is important to verify that two firms identical in K_t and U_t might actually employ different levels of labor. Here our timing assumptions about firm decisions and information revelation play a key role.[2] A firm's labor input is given by (7). Because firm's choose labor before they know U_t, their choice varies with U_{t-1} (their best predictor under the Markov assumption for current period productivity). Consequently firms with identical capital stocks and identical realizations of U_t, *but different realizations of* U_{t-1} will employ different amounts of labor. This is the variation we use to identify β.

Equation (8) defines what is called a partially linear regression model (PLM). This is a canonical and well-studied semiparametric model. Robinson (1988) and Chamberlain (1992) are classic references, while I especially like the paper by Robins et al. (1992). We will take a particularly simple approach, one that is convenient and appropriate for our needs, but would ideally be replaced by something less sensitive to modeling assumptions in actual empirical work.

[2] Ackerberg et al. (2015) provide an extensive discussion of the role of information revelation assumptions and decision timing in the "Olley-Pakes" setting.

Specifically we will assume that the unknown function $h_t(S_t, I_t)$ in (8) is well approximated by a second order polynomial in S_t and I_t :

$$\begin{aligned} h_t(S_t, I_t) &\simeq \pi_{0t} + \pi_{1t}S_t + \pi_{2t}I_t + \pi_{3t}S_tI_t + \pi_{4t}S_t^2 + \pi_{5t}I_t^2 \\ &= T_t'\pi_t, \end{aligned}$$

with $T_t \overset{def}{\equiv} (1, S_t, I_t, S_tI_t, S_t^2, I_t^2)$ and $\pi_t = (\pi_{0t}, \pi_{1t}, \pi_{2t}, \pi_{3t}, \pi_{4t}, \pi_{5t})$'. The coefficient vector π_t is indexed by t since the $h_t(s_t, i_t)$ function may vary with t due to changing input-output prices and/or other non-stationarities in the firm's decision problem. We could, of course, use a more complex approximation (e.g., a third order polynomial), however the simple quadratic approximation often works well in practice (e.g., de Loecker & Warzynski, 2012).

Using the quadratic approximation we get

$$\begin{aligned} \mathbb{E}[Y_t | R_t, S_t, I_t] =& \beta R_t + h_t(S_t, I_t) \\ \simeq& \beta R_t + \pi_{0t} + \pi_{1t}S_t + \pi_{2t}I_t + \pi_{3t}S_tI_t + \pi_{4t}S_t^2 + \pi_{5t}I_t^2 \\ =& \beta R_t + T_t'\pi_t, \end{aligned}$$

which implies that β is consistently estimated by the coefficient on R_t in the least squares fit of Y_t onto R_t and T_t.

Consider the two period setting with $t = 0, 1$. Imagine we have computed the labor output elasticity $\tilde{\beta}$ by the least squares fit of Y_0 onto R_0 and T_0 (this also gives us the coefficient vector $\tilde{\pi}_0$ such that $T_0'\tilde{\pi}_0 = \tilde{h}(S_0, I_0)$). For any period t – using the log production function, the linear Markov process for U_t, and the invertibility of the firm's optimal capital expenditures rule – we have

$$\begin{aligned} Y_t =& \beta R_t + \alpha S_t + U_t + \varepsilon_t \\ =& \beta R_t + \alpha S_t + \lambda_t + \rho U_{t-1} + V_t + \varepsilon_t \\ =& \beta R_t + \alpha S_t + \lambda_t + \rho h_{t-1}(S_{t-1}, I_{t-1}) + V_t + \varepsilon_t \\ \simeq& \alpha\beta R_t + \alpha S_t + \lambda_t + \rho \left(T_{t-1}'\pi_{t-1}\right) + V_t + \varepsilon_t. \end{aligned}$$

Note that S_t and T_{t-1} are all contained in the firm's beginning-of-period information set, $\mathcal{I}_t$. This implies that

$$\mathbb{E}[V_t + \varepsilon_t | S_t, T_{t-1}] = 0.$$

Hence the least squares fit of the "labor adjusted" log output, $Y_t - \beta R_t$ onto S_t, a constant and the linear combination $(T_{t-1}'\pi_{t-1})$ will provide consistent estimates of, respectively, α,

λ_t and ρ. This OLS fit is not feasible, but a feasible approximation is available. Returning to our two period example, we simply compute the least squares fit of $Y_1 - \tilde{\beta} R_1$ onto S_1, a constant, and the linear combination $(T_0'\tilde{\pi}_0)$.

Recovering the distribution of total factor productivity

Let θ be the vector containing all the model parameters. With a consistent estimate of θ in hand it is straightforward to estimate features of the distribution of productivity across firms. For example period t TFP for firm i can be estimated by

$$\hat{A}_{it} = \frac{O_{it}}{K_{it}^{\hat{\alpha}} L_{it}^{\tilde{\beta}}}.$$

One can then estimate quantiles of A_t by the sample quantiles of $\hat{A}_{1t}$,$\hat{A}_{2t}, \ldots, \hat{A}_{Nt}$. The difference between the 90th and 10th percentiles of A_t is a common measure of TFP dispersion across firms (see Syverson, 2011).

Bootstrap inference

Let $Z = (Y_0, Y_1, R_0, R_1, S_0, S_1, I_0, I_1)'$ and define the vector valued functions

$$\psi_1 (Z; \beta, \pi_0) = \begin{pmatrix} R_0 \\ T_0 \end{pmatrix} (Y_0 - \beta R_0 - T_0'\pi_0)$$

and

$$\psi_2 (Z; \beta, \pi_0, \alpha, \lambda_1, \rho) = \begin{pmatrix} 1 \\ S_1 \\ T_0'\pi_0 \end{pmatrix} (Y_1 - \beta R_1 - \lambda_1 - \alpha S_1 - \rho (T_0'\pi_0)).$$

Observe that $\mathbb{E}[\psi_1 (Z; \beta, \pi_0)] = 0$ and $\mathbb{E}[\psi_2 (Z; \beta, \pi_0, \alpha, \lambda_1, \rho)] = 0$ at the population parameters. Next note that our two step-procedure coincides with the solutions to

$$\frac{1}{N} \sum_{i=1}^{N} \psi_1 \left(Z_i; \tilde{\beta}, \tilde{\pi}_0 \right) = 0 \tag{9}$$

and, in the second step, fixing β and π_0 at their first step values, to

$$\frac{1}{N}\sum_{i=1}^{N}\psi_2\left(Z;\hat{\alpha},\tilde{\pi}_0,\tilde{\beta},\hat{\lambda}_1,\hat{\rho}\right)=0. \tag{10}$$

This casts our procedure as a special case of sequential generalized method-of-moments (GMM). See Wooldridge (2010) for a textbook introduction to GMM. We could use this characterization to derive an explicit form for the limit distribution of $\hat{\theta}$. This is beyond the scope of the course. Instead we will use a simple percentile bootstrap.

Our sample consists of the N independent and identically distributed realizations $Z_1, Z_2, \ldots, Z_N$.

For $s = 1, \ldots, S$ bootstrap repetitions:

1. Construct the bootstrap sample $Z_1^{(s)}, Z_2^{(s)}, \ldots, Z_N^{(s)}$ by drawing observations uniformly from our sample $Z_1, Z_2, \ldots, Z_N$ *with replacement*;

2. Using the s^{th} bootstrap sample repeat the Olley & Pakes (1996) estimation procedure, producing $\hat{\theta}^{(s)}$. We can also construct other statistics, say, median estimated TFP: $\hat{\eta}^{(s)}$.

We repeat steps 1 and 2 $s = 1 \ldots S$ times ($S = 1,000$ would be a common choice).

Let $\hat{\beta}^{0.025}$ and $\hat{\beta}^{0.0975}$ be the 0.025 and 0.975 sample quantiles of the bootstrap estimates $\hat{\beta}^{(1)}, \hat{\beta}^{(2)}, \ldots, \hat{\beta}^{(S)}$. We use $\left[\hat{\beta}^{0.025}, \hat{\beta}^{0.975}\right]$ as an approximate 0.95 confidence interval for β. The idea is that our bootstrap procedure approximates the the sampling distribution of β. See Efron & Hastie (2016) for additional discussion.

References

Ackerberg, D. A., Caves, K., & Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6), 2411 – 2451.

Blundell, R. & Bond, S. (2000). Gmm estimation with persistent panel data: an application to production functions. *Econometric Reviews*, 19(3), 321 – 340.

Chamberlain, G. (1992). Efficiency bounds for semiparametric regression. *Econometrica*, 60(3), 567 – 596.

de Loecker, J. & Warzynski, F. (2012). Markups and firm-level export status. *American Economic Review*, 102(6), 2437 – 2471.

Efron, B. & Hastie, T. (2016). *Computer Age Statistical Inference*. Cambridge: Cambridge University Press.

Griliches, Z. & Mairesse, J. (1998). *Econometrics and Economic Theory in the 20th Century*, chapter Production functions: the search for identification, (pp. 169 – 203). Cambridge University Press: Cambridge.

Olley, G. S. & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6), 1263 – 1297.

Robins, J. M., Mark, S. D., & Newey, W. K. (1992). Estimating exposure effects by modeling the expectation of exposure conditional on confounders. *Biometrics*, 48(2), 479 – 495.

Robinson, P. M. (1988). Root-n-consistent semiparametric regression. *Econometrica*, 56(4), 931 – 954.

Syverson, C. (2011). What determines productivity. *Journal of Economic Literature*, 49(2), 326 – 365.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press, 2nd edition.