

Simultaneity and Other "Simple" Problems : Lecture VI

Charles B. Moss

September 5, 2018

1 Simultaneity and Estimation of the Production Function

- Irving Hock - Econometrica
- Lawrence Klein - A Textbook of Econometrics

2 Two Models of Simultaneity

- Empirical Setup
- Indirect Least Squares Solution
- Seemingly Unrelated Regression Formulation

3 Zeros in the Cobb-Douglas

Simultaneity and Estimation of the Production Function

- The above discussion (and estimates) makes the experimental plot design assumption regarding the data.
 - I essentially assumed that the data are being generated from some sort of experimental design so that the errors are truly random.
 - If the data are actually the result of farm level decisions, the data are endogenous.

Irving Hock - Econometrica

- Hock, Irving. 1958. Simultaneous Equation Bias in the Context of the Cobb-Douglas Production Function. *Econometrica* 26(4), 566-78.
- The basic firm-level model is that we have an empirical model under the assumption of:
 - A Cobb-Douglas production function, and
 - Competition.

$$\begin{aligned}
 X_0 &= K_0 \prod_{q=1}^Q X_q^{a_q} \\
 P_0 \frac{\partial X_0}{\partial X_q} &= P_0 \left(a_q \frac{X_0}{X_q} \right) = P_q \\
 a_q \left(\frac{X_0 P_0}{X_q P_q} \right) &= a_q \frac{Y_0}{Y_q} = 1
 \end{aligned} \tag{1}$$

Lawrence Klein - A Textbook of Econometrics

- Klein demonstrates that the best linear unbiased estimate of a_q is

$$\hat{a}_q = \prod_{i=1}^I \left(\frac{Y_{qi}}{Y_{0i}} \right)^{\frac{1}{I}} \quad (2)$$

- In this approach the “average” firm is defined to be the optimal firm.
- As an alternative

$$P_0 \left(a_0 \frac{X_0}{X_q} \right) = R_q P_q \quad (3)$$

where R_q is “some constant and the investigator wishes to test whether it is equal to one.”

- “The firm sets the value of the marginal product equal to the price augmented by the effect of any restrictions that exist.”
- “In this formulation, R_q can of course vary among firms; but, for a sample of firms, the investigator would be interested in testing whether the average R_q is equal to one.”

Two models of Simultaneity

- Model 1: Production disturbance not transmitted to the “independent” variables
 - “If the disturbance in the production equation affects only the output and is not transmitted to the other variables in the system, then there is no simultaneous equation bias. Single equation estimates are consistent.”
 - For example if inputs are fixed or are predetermined .
- Model 2: Production disturbance transmitted to the “independent” variables.
 - “Simultaneous equation bias arises when disturbances in the production relations affect the observed values of all variables, and, as a result, single equation estimates are not consistent.”

Empirical Setup

- Starting in level space

$$X_0 = K_0 \prod_{q=1}^Q X_q^{a_q} U$$
$$X_q = \frac{a_q}{R_q} \frac{P_0}{P_q} X_0 V_q, \quad q = 1, 2, \dots, Q$$
(4)

- Taking the logarithms of this expression

$$x_0 = k_0 + \sum_{q=1}^Q a_q x_q + u$$
$$x_q = k_q + x_0 + v_q \quad q = 1, 2, \dots, Q$$
$$x_0 = \ln(X_0) \quad x_q = \ln(X_q)$$
$$k_q = \ln \left[\frac{a_q P_0}{R_q P_q} \right]$$
(5)

Indirect Least Square Solution

- Kmenta, J. 1964. Some Properties of Alternative Estimates of the Cobb-Douglas Production Function. *Econometrica* 32(1/2), 183-188.
- Simplifying the general system of equations

$$\begin{aligned}x_{0i} &= k_0 + a_1 x_{1i} + a_2 x_{2i} + \nu_{0i} \\x_{1i} &= k_1 + x_{0i} + \nu_{1i} \\x_{2i} &= k_2 + x_{0i} + \nu_{2i}\end{aligned}\tag{6}$$

- Transforming the estimation problem to

$$x_{0i} = b_0 + b_1 (x_{1i} - x_{0i}) + b_2 (x_{2i} - x_{0i}) + e_i \quad (7)$$

yields estimates of b_1 and b_2 that are consistent. Note by the definitions

$$\begin{aligned} x_{1i} &= k_1 + x_{0i} + \nu_{1i} \Rightarrow x_{1i} - x_{0i} = k_1 + \nu_{1i} \\ x_{2i} &= k_2 + x_{0i} + \nu_{2i} \Rightarrow x_{2i} - x_{0i} = k_2 + \nu_{2i} \end{aligned} \quad (8)$$

- Working through the mechanics

$$\begin{aligned}
 x_{0i} &= b_0 + b_1 x_{1i} - b_1 x_{0i} + b_2 x_{2i} - b_2 x_{0i} + e_i \\
 (1 + b_1 + b_2) x_{0i} &= b_0 + b_1 x_{1i} + b_2 x_{2i} + e_i \\
 x_{0i} &= \frac{b_0}{1 + b_1 + b_2} + \frac{b_1}{1 + b_1 + b_2} x_{1i} + \frac{b_2}{1 + b_1 + b_2} x_{2i} + \frac{1}{1 + b_1 + b_2} e_i \\
 \tilde{a}_r &= \frac{\hat{b}_r}{1 + \hat{b}_1 + \hat{b}_2} \quad r = 1, 2
 \end{aligned}
 \tag{9}$$

Seemingly Unrelated Regression Formulation

- Building on the simple Cobb-Douglas function

$$\begin{aligned}
 \pi &= pAx_1^\alpha x_2^\beta x_3^\gamma - w_1x_1 - w_2x_2 - w_3x_3 \\
 \frac{\partial \pi}{\partial x_1} &= p\alpha \frac{Y}{x_1} - w_1 = 0 \Rightarrow \alpha \frac{Y}{x_1} = \frac{w_1}{p} \Rightarrow \alpha Yp = x_1w_1 \Rightarrow \alpha R = S_1 \\
 \frac{\partial \pi}{\partial x_2} &= p\beta \frac{Y}{x_2} - w_2 = 0 \Rightarrow \beta \frac{Y}{x_2} = \frac{w_2}{p} \Rightarrow \beta Yp = x_2w_2 \Rightarrow \beta R = S_2 \\
 \frac{\partial \pi}{\partial x_3} &= p\gamma \frac{Y}{x_3} - w_3 = 0 \Rightarrow \gamma \frac{Y}{x_3} = \frac{w_3}{p} \Rightarrow \gamma Yp = x_3w_3 \Rightarrow \gamma R = S_3
 \end{aligned}
 \tag{10}$$

- Three related empirical equations

$$\begin{aligned}
 \ln(Y) &= a_0 + a_1 \ln(x_1) + a_2 \ln(x_2) + \ln(x_3) + \nu_1 \\
 S_1 &= k_1 + a_1 R + \nu_2 \\
 S_2 &= k_2 + a_2 R + \nu_3 \\
 S_3 &= k_3 + a_3 R + \nu_4
 \end{aligned} \tag{11}$$

- System of Equations

$$\begin{bmatrix} \ln(Y_1) \\ S_{11} \\ S_{21} \\ S_{31} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & \ln(x_{11}) & \ln(x_{21}) & \ln(x_{31}) \\ 0 & 1 & 0 & 0 & R_1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & R_1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & R_q \end{bmatrix} \begin{bmatrix} a_0 \\ k_1 \\ k_2 \\ k_3 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} \tag{12}$$

$$\begin{bmatrix} \ln(Y_1) \\ S_{11} \\ S_{21} \\ S_{31} \\ \ln(Y_2) \\ S_{12} \\ S_{22} \\ S_{32} \\ \ln(Y_3) \\ S_{13} \\ S_{23} \\ S_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & \ln(x_{11}) & \ln(x_{21}) & \ln(x_{31}) \\ 0 & 1 & 0 & 0 & R_1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & R_1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & R_1 \\ 1 & 0 & 0 & 0 & \ln(x_{12}) & \ln(x_{22}) & \ln(x_{32}) \\ 0 & 1 & 0 & 0 & R_2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0R_2 \\ 1 & 0 & 0 & 0 & \ln(x_{13}) & \ln(x_{23}) & \ln(x_{33}) \\ 0 & 1 & 0 & 0 & R_3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & R_2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & R_3 \end{bmatrix} \begin{bmatrix} k_0 \\ k_1 \\ k_2 \\ k_3 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} + \begin{bmatrix} \nu_{11} \\ \nu_{21} \\ \nu_{31} \\ \nu_{41} \\ \nu_{12} \\ \nu_{22} \\ \nu_{32} \\ \nu_{42} \\ \nu_{13} \\ \nu_{23} \\ \nu_{33} \\ \nu_{34} \end{bmatrix} \quad (13)$$

$$\begin{aligned} y &= x\beta + \nu \Rightarrow \hat{\beta} = (x'x)^{-1} x'y \\ \tilde{\beta} &= (x'\hat{V}^{-1}x)^{-1} x'\hat{V}^{-1}y \end{aligned} \quad (14)$$

$$\hat{V} = \frac{1}{3} \begin{bmatrix} \nu_{11} & \nu_{12} & \nu_{13} \\ \nu_{21} & \nu_{22} & \nu_{23} \\ \nu_{31} & \nu_{32} & \nu_{33} \end{bmatrix} \begin{bmatrix} \nu_{11} & \nu_{21} & \nu_{31} \\ \nu_{12} & \nu_{22} & \nu_{32} \\ \nu_{13} & \nu_{23} & \nu_{33} \end{bmatrix} \quad (15)$$

Zeros in the Cobb-Douglas Functional Form

- Moss, Charles B. 2000. Estimation of the Cobb-Douglas with Zero Input Levels: Bootstrapping and Substitution. *Applied Economics Letters* 7(10), 677-679.
- Zeros raise several difficulties in estimating the Cobb-Douglas production function.
 - On the theoretical side, the presence of a zero-level input is that it violates weak necessity of inputs.
 - On the empirical side, how do you take the natural logarithm of zero.

- The existence of zeros is the result of measurement error.
 - Agronomically, production of a crop is impossible without some level of each fertilizer.
 - Thus, production occurs based on some true level of each nutrient available to the plant

$$x_i^* = x_i + \epsilon_i \quad (16)$$

- In fact we can think of the soil as a sponge that contains a variety of nutrients that we can augment by applying fertilizer. The actual level of fertilizer used by the crop could then be a function of what we add, the weather (i.e., if adequate moisture is not present the crop does not use the full potential), etc.
- Second, a production function that does not admit zero input could represent a misspecification.

- In this paper, I consider two techniques for adjusting the zero observations.
 - First, I redraw from the sample averaging the result until the pseudo sample contains no zeros.
 - Second, I substitute a small non-zero number for those observations that contain zeros (i.e., 0.1, 0.01, 0.001).
- The goodness of fit for each procedure is then compared using a Strobel measure of information

$$I = \sum_{i=1}^N s_i \ln \left(\frac{s_i}{\tilde{s}_i} \right) \quad (17)$$

Where s_i is the theoretically appropriate budget share for each input and $\tilde{s}_i$ is the budget share estimated using each empirical approximation of zero.