

Empirical Industrial Organization (ECO 310)  
Winter 2021 – Sections 0101, 9101, 0201, 9201  
Victor Aguirregabiria

Solution to Problem Set #1

---

INSTRUCTIONS. Please, follow the following instructions for the submission of your completed problem set.

1. Write your answers electronically in a word processor.
2. For the answers that involve coding in STATA, include in the document the code in STATA that you have used to obtain your empirical results.
3. Convert the document to PDF format.
4. Submit your problem set (in PDF) online via Quercus.
5. You should submit your completed problem set before Friday, February 5th, at 11:59pm.
6. Problem sets should be written individually.

The total number of marks is 200.

---

QUESTION 1. [60 points]. Consider an industry for an homogeneous product with the following characteristics.

- The production function uses labor, capital, and TFP, and it has a Cobb-Douglas form.
- All the firms use the same amount of capital (the same fixed equipment).
- Firms sell their output in the same product market at the same price.
- Firms are in different geographic locations with different labor markets.
- Firms are price takers in labor and product markets and maximize profits.

Q1.1. (10 points) (a) Write the expression for the Cobb-Douglas production function, and for this PF in logarithms. (b) Write the expression for a firm's profit function. (c) Obtain the first order conditions for the maximization of profits with respect to the amount of labor. (d) Obtain the equation that describes a firm's labor demand. (e) Write the labor demand equation in logarithms: in terms of the logarithms of output and labor.

ANSWER.

(a) Let  $Y$ ,  $L$ , and  $K$  be the physical amounts of output, labor, and capital, respectively. Then, the expression for the production function is:

$$Y = AL^{\alpha_L} K^{\alpha_K}$$

Taking logarithms in both sides of the equation, we have the following PF in logs:

$$y = \omega + \alpha_L \ell + \alpha_K k$$

where  $y = \log(Y)$ ,  $\omega = \log(A)$ ,  $\ell = \log(L)$ , and  $k = \log(K)$ .

(b) A firm's profit is equal to revenue minus costs. Revenue is equal to the unit price  $P$  times amount the amount of output in physical units,  $Y$ . And the total cost is equal to the cost of labor,  $W_L L$ , plus the cost of capital,  $W_K K$ , where  $W_L$  and  $W_K$  are the prices of labor and capital inputs, respectively. Therefore, the profit function is:

$$\Pi = P Y - W_L L - W_K K$$

Taking into account the production function,  $Y = AL^{\alpha_L} K^{\alpha_K}$ , we have that the profit as a function of the amounts of inputs is:

$$\Pi = P A L^{\alpha_L} K^{\alpha_K} - W_L L - W_K K$$

(c) The first order condition of profit maximization with respect to the amount of labor is:  $\frac{d\Pi}{dL} = 0$ . Using the previous expression for the profit function, we have:

$$\frac{d\Pi}{dL} = P \alpha_L A L^{(\alpha_L-1)} K^{\alpha_K} - W_L = 0$$

(d) A firm's labor demand give us the amount of labor that a firm wants to employ given the amount of output that it wants to produce and the price of labor. That is, a function  $L = f(Y, W_L)$  that comes from profit maximization. Consider the marginal condition of optimality in Q1(c). Using the form of the Cobb-Douglas PF, we have that  $A L^{(\alpha_L-1)} K^{\alpha_K} = \frac{Y}{L}$ . Therefore, we can write this condition of optimality as:

$$P \alpha_L \frac{Y}{L} = W_L$$

Solving for  $L$ , we have:

$$L = \alpha_L \frac{Y}{R}$$

where  $R$  is the real wage,  $R = \frac{W_L}{P}$ .

(e) Taking logarithms in both sides of the labor demand equation, we have:

$$\ell = \log(\alpha_L) + y - r$$

where  $r = \log(R)$ .

**Q1.2. (10 points)** (a) Write the simultaneous equations model that consists of the production function and the demand equation in logarithms. (b) Solve this system of two equations and two unknowns to obtain the equilibrium amounts of output and labor as functions of the exogenous variables.

ANSWER.

(a) The simultaneous equations model (the structural equations), in logarithms, are:

$$\begin{cases} y &= \omega + \alpha_L \ell + \alpha_K k \\ \ell &= \log(\alpha_L) + y - r \end{cases}$$

(b) The endogenous unknowns are  $y$  and  $\ell$ . The rest of the variables ( $\omega$ ,  $k$ , and  $r$ ) and parameters ( $\alpha_L$  and  $\alpha_K$ ) are exogenous. Solving the labor demand equation into the production function, we have:

$$y = \omega + \alpha_L (\log(\alpha_L) + y - r) + \alpha_K k$$

Solving for log-output, we get the following reduced form equation for  $y$ :

$$y = \frac{1}{1 - \alpha_L} (\omega + \alpha_L \log(\alpha_L) - \alpha_L r + \alpha_K k)$$

Plugging this equation into the labor demand, we get the following reduced form equation for log-labor:

$$\begin{aligned} \ell &= \log(\alpha_L) + \frac{1}{1 - \alpha_L} (\omega + \alpha_L \log(\alpha_L) - \alpha_L r + \alpha_K k) - r \\ &= \frac{1}{1 - \alpha_L} (\omega + \log(\alpha_L) - r + \alpha_K k) \end{aligned}$$

**Q1.3. (10 points)** Suppose that the industry is such that firms are heterogeneous in their log-TFP (represented as  $\omega_i$ ) and in their log-real wage (represented as  $r_i$ ). Let  $\sigma_\omega^2$ ,  $\sigma_r^2$ , and  $\sigma_{\omega r}$  be the variance of  $\omega_i$ , the variance of  $r_i$ , and the covariance between  $\omega_i$  and  $r_i$ , respectively. Use the equilibrium equations in Q1.2 to obtain the expression for: (a) the covariance between log-output and log-labor as a function of  $\sigma_\omega^2$ ,  $\sigma_r^2$ , and  $\sigma_{\omega r}$ ; and (b) the variance of log-labor as a function of  $\sigma_\omega^2$ ,  $\sigma_r^2$ , and  $\sigma_{\omega r}$ .

ANSWER. First, we include the firm subindex  $i$  in those variables that have sample variation across firms (that is,  $\omega_i$ ,  $r_i$ ,  $\ell_i$ , and  $y_i$ ). We do not include this subindex for the variables that are constant across firms (that is,  $\alpha_L$ ,  $\alpha_K$ , and  $k$ ). Therefore, we have the following expression for the reduced form equations:

$$\begin{cases} y_i &= \frac{1}{1 - \alpha_L} (\omega_i + \alpha_L \log(\alpha_L) - \alpha_L r_i + \alpha_K k) \\ \ell_i &= \frac{1}{1 - \alpha_L} (\omega_i + \log(\alpha_L) - r_i + \alpha_K k) \end{cases}$$

By definition of variance and covariance, we have that  $\sigma_\omega^2 = \mathbb{E}([\omega_i - \bar{\omega}]^2)$ ,  $\sigma_r^2 = \mathbb{E}([r_i - \bar{r}]^2)$ , and  $\sigma_{\omega r} = \mathbb{E}([\omega_i - \bar{\omega}] [r_i - \bar{r}])$ , where  $\mathbb{E}(\cdot)$  is the expectation operator, and  $\bar{\omega}$  and  $\bar{r}$  are the mean values of  $\omega_i$  and  $r_i$ , respectively: that is,  $\bar{\omega} = \mathbb{E}(\omega_i)$  and  $\bar{r} = \mathbb{E}(r_i)$ . Similarly, we have that  $\text{Var}(\ell_i) = \mathbb{E}([\ell_i - \bar{\ell}]^2)$  and  $\text{Cov}(y_i, \ell_i) = \mathbb{E}([y_i - \bar{y}][\ell_i - \bar{\ell}])$ . Looking at these definitions of  $\text{Var}(\ell_i)$  and  $\text{Cov}(y_i, \ell_i)$ , it is clear that we need to obtain log-output and log-labor in deviations with respect to their means. Then, we obtain means in both sides of the reduced form equations. We get:

$$\begin{cases} \bar{y} &= \frac{1}{1 - \alpha_L} (\bar{\omega} + \alpha_L \log(\alpha_L) - \alpha_L \bar{r} + \alpha_K k) \\ \bar{\ell} &= \frac{1}{1 - \alpha_L} (\bar{\omega} + \log(\alpha_L) - \bar{r} + \alpha_K k) \end{cases}$$

If we subtract the equations in means to the original equations, we obtain the following reduced equations in deviations with respect to means:

$$\begin{cases} y_i - \bar{y} &= \frac{1}{1 - \alpha_L} (\omega_i - \bar{\omega} - \alpha_L [r_i - \bar{r}]) \\ \ell_i - \bar{\ell} &= \frac{1}{1 - \alpha_L} (\omega_i - \bar{\omega} - [r_i - \bar{r}]) \end{cases}$$

Note that the constant elements in the original equations (that is,  $\alpha_L \log(\alpha_L)$ ,  $\log(\alpha_L)$ , and  $\alpha_K k$ ) do not appear in the equations in deviations with respect to means. It is notational convenient using  $\tilde{y}_i$ ,  $\tilde{\ell}_i$ ,  $\tilde{\omega}_i$ , and  $\tilde{r}_i$  to represent the variables in deviations,  $(y_i - \bar{y})$ ,  $(\ell_i - \bar{\ell})$ ,  $(\omega_i - \bar{\omega})$ , and  $(r_i - \bar{r})$ , respectively. Using this notation the reduced form equations in deviations with respect to means are:

$$\begin{cases} \tilde{y}_i &= \frac{1}{1 - \alpha_L} (\tilde{\omega}_i - \alpha_L \tilde{r}_i) \\ \tilde{\ell}_i &= \frac{1}{1 - \alpha_L} (\tilde{\omega}_i - \tilde{r}_i) \end{cases}$$

(a) By definition, the covariance between log-output and log-labor is:

$$\text{Cov}(y_i, \ell_i) = \mathbb{E}([y_i - \bar{y}][\ell_i - \bar{\ell}]) = \mathbb{E}(\tilde{y}_i \tilde{\ell}_i)$$

Plugging into  $\mathbb{E}(\tilde{y}_i \tilde{\ell}_i)$  the reduced form equations in deviations, we have that:

$$Cov(y_i, \ell_i) = \mathbb{E} \left( \frac{1}{1 - \alpha_L} (\tilde{\omega}_i - \alpha_L \tilde{r}_i) \frac{1}{1 - \alpha_L} (\tilde{\omega}_i - \tilde{r}_i) \right)$$

Operating in this expression using the properties of the expectation operator  $\mathbb{E}(\cdot)$ , we have:

$$\begin{aligned} Cov(y_i, \ell_i) &= \left( \frac{1}{1 - \alpha_L} \right)^2 \mathbb{E}((\tilde{\omega}_i - \alpha_L \tilde{r}_i) (\tilde{\omega}_i - \tilde{r}_i)) \\ &= \left( \frac{1}{1 - \alpha_L} \right)^2 \mathbb{E}(\tilde{\omega}_i \tilde{\omega}_i - \tilde{\omega}_i \tilde{r}_i - \alpha_L \tilde{\omega}_i \tilde{r}_i + \alpha_L \tilde{r}_i \tilde{r}_i) \\ &= \left( \frac{1}{1 - \alpha_L} \right)^2 (\sigma_\omega^2 - [1 + \alpha_L] \sigma_{\omega r} + \alpha_L \sigma_r^2) \end{aligned}$$

where, for the last equality, we have taken into account that by definition of variance and covariance,  $\mathbb{E}(\tilde{\omega}_i \tilde{\omega}_i) = \sigma_\omega^2$ ,  $\mathbb{E}(\tilde{\omega}_i \tilde{r}_i) = \sigma_{\omega r}$ , and  $\mathbb{E}(\tilde{r}_i \tilde{r}_i) = \sigma_r^2$ .

(b) By definition, the variance of log-labor is:

$$Var(\ell_i) = \mathbb{E}([\ell_i - \bar{\ell}]^2) = \mathbb{E}(\tilde{\ell}_i \tilde{\ell}_i)$$

Plugging into  $\mathbb{E}(\tilde{\ell}_i \tilde{\ell}_i)$  the reduced form equations in deviations, we have that:

$$Var(\ell_i) = \mathbb{E} \left( \frac{1}{1 - \alpha_L} (\tilde{\omega}_i - \tilde{r}_i) \frac{1}{1 - \alpha_L} (\tilde{\omega}_i - \tilde{r}_i) \right)$$

Operating in this expression using the properties of the expectation operator  $\mathbb{E}(\cdot)$ , we have:

$$\begin{aligned} Var(\ell_i) &= \left( \frac{1}{1 - \alpha_L} \right)^2 \mathbb{E}((\tilde{\omega}_i - \tilde{r}_i) (\tilde{\omega}_i - \tilde{r}_i)) \\ &= \left( \frac{1}{1 - \alpha_L} \right)^2 \mathbb{E}(\tilde{\omega}_i \tilde{\omega}_i - \tilde{\omega}_i \tilde{r}_i - \tilde{\omega}_i \tilde{r}_i + \tilde{r}_i \tilde{r}_i) \\ &= \left( \frac{1}{1 - \alpha_L} \right)^2 (\sigma_\omega^2 - 2\sigma_{\omega r} + \sigma_r^2) \end{aligned}$$

where, for the last equality, we have taken into account that by definition of variance and covariance,  $\mathbb{E}(\tilde{\omega}_i \tilde{\omega}_i) = \sigma_\omega^2$ ,  $\mathbb{E}(\tilde{\omega}_i \tilde{r}_i) = \sigma_{\omega r}$ , and  $\mathbb{E}(\tilde{r}_i \tilde{r}_i) = \sigma_r^2$ .

**Q1.4. (10 points) Consider the OLS estimation of the labor intensity parameter  $\alpha_L$  in the production function. (a) Write the expression for this OLS**

estimator as a function of the data on log-output and log-labor. By the Law of Large Numbers, when the number of observations (firms)  $N$  is large, the sample variances and covariances in the expression for the OLS estimator converge to their population counterparts. (b) Given the results in Q1.3, obtain the expression for the OLS estimator (when the sample size is large) as a function of the parameters  $\alpha_L$ ,  $\sigma_\omega^2$ ,  $\sigma_r^2$ , and  $\sigma_{\omega r}$ .

ANSWER.

(a) The regression equation for the production function in logs is:

$$y_i = \omega_i + \alpha_L \ell_i + \alpha_K k$$

In order to clearly distinguish between the intercept (constant term) in this linear regression and the zero mean error term, it is convenient to take into account that  $\omega_i = \bar{\omega} + \tilde{\omega}_i$ , where  $\bar{\omega}$  is the mean of  $\omega_i$ , and  $\tilde{\omega}_i$  is its deviation with respect to the mean, such that  $\tilde{\omega}_i$  has mean zero. Therefore, we can re-write this regression equation as follows:

$$y_i = \alpha_0 + \alpha_L \ell_i + \tilde{\omega}_i$$

where:  $\alpha_0$  is the intercept parameter, that according to the model is equal to  $\bar{\omega} + \alpha_K k$ ; and  $\tilde{\omega}_i$  is the zero mean error term.

The OLS estimator of the slope parameter  $\alpha_L$  in this regression model is:

$$\hat{\alpha}_L = \frac{\sum_{i=1}^N (\ell_i - \bar{\ell}) (y_i - \bar{y})}{\sum_{i=1}^N (\ell_i - \bar{\ell})^2}$$

(b) Note that:

$$\hat{\alpha}_L = \frac{\frac{1}{N} \sum_{i=1}^N (\ell_i - \bar{\ell}) (y_i - \bar{y})}{\frac{1}{N} \sum_{i=1}^N (\ell_i - \bar{\ell})^2}$$

By the Law of Large Numbers, given  $N$  random draws of a random variable  $x_i$ ,  $\frac{1}{N} \sum_{i=1}^N x_i$  converges to the population mean  $\mathbb{E}(x_i)$  as  $N$  is large. For this reason, the sample covariance  $\frac{1}{N} \sum_{i=1}^N (\ell_i - \bar{\ell}) (y_i - \bar{y})$  converges to  $\mathbb{E}(\tilde{\ell}_i \tilde{y}_i)$ , the sample variance  $\frac{1}{N} \sum_{i=1}^N (\ell_i - \bar{\ell})^2$  converges to  $\mathbb{E}(\tilde{\ell}_i \tilde{\ell}_i)$ , and the OLS estimator converges to:

$$\hat{\alpha}_L = \frac{\mathbb{E}(\tilde{\ell}_i \tilde{y}_i)}{\mathbb{E}(\tilde{\ell}_i \tilde{\ell}_i)}$$

Using the expressions for  $\mathbb{E}(\tilde{\ell}_i \tilde{y}_i)$  and  $\mathbb{E}(\tilde{\ell}_i \tilde{\ell}_i)$  in terms of  $\alpha_L$ ,  $\sigma_\omega^2$ ,  $\sigma_r^2$ , and  $\sigma_{\omega r}$  that we have obtained in Q1.3, we have that:

$$\begin{aligned}\hat{\alpha}_L &= \frac{\left(\frac{1}{1-\alpha_L}\right)^2 (\sigma_\omega^2 - [1+\alpha_L]\sigma_{\omega r} + \alpha_L \sigma_r^2)}{\left(\frac{1}{1-\alpha_L}\right)^2 (\sigma_\omega^2 - 2\sigma_{\omega r} + \sigma_r^2)} \\ &= \frac{\sigma_\omega^2 - [1+\alpha_L]\sigma_{\omega r} + \alpha_L \sigma_r^2}{\sigma_\omega^2 - 2\sigma_{\omega r} + \sigma_r^2}\end{aligned}$$

**Q1.5. (20 points)** Using the expression that you have derived in Q1.4 obtain the expression for the bias of the OLS estimator under the following different scenarios for the industry. (a) **No heterogeneity in real wages:**  $\sigma_r^2 = 0$ . (b) **No heterogeneity in TFP:**  $\sigma_\omega^2 = 0$ . (c) **Heterogeneity in real wages and TFP but not correlation between them:**  $\sigma_r^2 > 0$ ,  $\sigma_\omega^2 > 0$ , and  $\sigma_{\omega r} = 0$ . Try to obtain the sign of the bias in each of these three cases. Explain the results.

ANSWER.

(a) When  $\sigma_r^2 = 0$ , the log-real-wage is constant. This implies that  $\sigma_{\omega r} = 0$  because the covariance of a random variable with a constant is always zero. Plugging these values in the formula for the OLS estimator  $\hat{\alpha}_L$  (as  $N$  is large) that we have obtained in Q1.4, we have:

$$\hat{\alpha}_L = \frac{\sigma_\omega^2 - [1+\alpha_L]0 + \alpha_L 0}{\sigma_\omega^2 - 2*0 + 0} = \frac{\sigma_\omega^2}{\sigma_\omega^2} = 1$$

And the bias of the OLS estimator (the difference between the estimator and the true value of the parameter) is:

$$\text{Bias of OLS} = \hat{\alpha}_L - \alpha_L = 1 - \alpha_L$$

Since  $\alpha_L$  is typically greater than zero and not greater than one, this bias is positive. The OLS estimator over-estimates the true causal effect of labor on output. Only if the true value  $\alpha_L$  is equal to one, the OLS estimator is unbiased. In general, this bias can be large.

When all the exogenous variation comes from log-TFP and there is not variation in input prices, the observations  $(\ell_i, y_i)$  are located in a straight line with slope equal to 1. This straight line is in fact the labor demand. In this case, the OLS estimator gives an unbiased estimate of the labor demand, but not of the production function.

(b) When  $\sigma_\omega^2 = 0$ , the log-TFP is the same for all the firms. This implies that  $\sigma_{\omega r} = 0$  because the covariance of a random variable with a constant is always zero. Plugging these values in the formula for the OLS estimator  $\hat{\alpha}_L$  (as  $N$  is large) that we have obtained in Q1.4, we have:

$$\hat{\alpha}_L = \frac{0 - [1+\alpha_L]0 + \alpha_L \sigma_r^2}{0 - 2*0 + \sigma_r^2} = \alpha_L$$

And the bias of the OLS estimator is:

$$\text{Bias of OLS} = \hat{\alpha}_L - \alpha_L = 0$$

That is, in this case, the OLS is an unbiased estimator of the production function parameter  $\alpha_L$ .

When all the exogenous variation comes from log-real-wage and there is not variation in TFP, the observations  $(\ell_i, y_i)$  are located in a straight line with slope equal to  $\alpha_L$ . This straight line is the product function in logs. Of course, this case is very unrealistic because it implies that there is not any error term in the production function, and there is a deterministic relationship between output and labor. This deterministic relationship will be rejected any real dataset.

(c) When  $\sigma_r^2 > 0$ ,  $\sigma_\omega^2 > 0$ , and  $\sigma_{\omega r} = 0$ , we have:

$$\hat{\alpha}_L = \frac{\sigma_\omega^2 - [1 + \alpha_L] 0 + \alpha_L \sigma_r^2}{\sigma_\omega^2 - 2 * 0 + \sigma_r^2} = \frac{\sigma_\omega^2 + \alpha_L \sigma_r^2}{\sigma_\omega^2 + \sigma_r^2}$$

And the bias of the OLS estimator is:

$$\text{Bias of OLS} = \hat{\alpha}_L - \alpha_L = \frac{\sigma_\omega^2 + \alpha_L \sigma_r^2}{\sigma_\omega^2 + \sigma_r^2} - \alpha_L = \frac{\sigma_\omega^2(1 - \alpha_L)}{\sigma_\omega^2 + \sigma_r^2}$$

For the interpretation of this bias, it is convenient to represent it as follows:

$$\text{Bias of OLS} = \lambda (1 - \alpha_L)$$

where  $\lambda = \frac{\sigma_\omega^2}{\sigma_\omega^2 + \sigma_r^2}$  is the ratio between the variance of log-TFP and the total variance of log-TFP and log-real-wage. It is clear that  $\lambda$  is always a parameter between 0 and 1,  $\lambda \in [0, 1]$ . It is equal to zero only if there is not variation in log-TFP across firms (if  $\sigma_\omega^2 = 0$ ). And it is equal to 1 only if there is not variation in log-real-wage across firms (if  $\sigma_r^2 = 0$ ). Based on this formula we can make the following statements about the bias of the OLS estimator.

(i) The bias is positive. Since  $\lambda \geq 0$ , and  $0 \leq \alpha_L \leq 1$ , we have that  $\lambda (1 - \alpha_L) \geq 0$ . That is, in general, the OLS estimator over-estimates the true causal effect of labor on output.

(ii) The bias is not greater than  $1 - \alpha_L$ . Since  $\lambda \leq 1$ , the worst case scenario in terms of the magnitude of the bias is when  $\lambda = 1$ , that corresponds to  $\sigma_r^2 = 0$ .

(iii) The bias of the OLS estimator increases monotonically with  $\lambda$ . The larger is the variance of log-TFP relative to the variance of log-real-wage, the larger is the bias of the OLS estimator. That is, in an industry where firms are heterogeneous in their log-TFPs but very homogeneous in the wages that they pay, the bias can be particularly large.

(iv) If  $\sigma_\omega^2$  and  $\sigma_r^2$  are of similar magnitude,  $\sigma_\omega^2 \simeq \sigma_r^2$ , we have that  $\lambda = 1/2$  and the bias is  $(1 - \alpha_L)/2$ . This bias can be substantial. For instance, if the true value of  $\alpha_L$  is 0.5 (a value that is realistic for many industries), the bias of the OLS is 0.25. That is, we are over-estimating the true value by 50%.

**QUESTION 2.** [40 points]. Consider an industry for an homogeneous product. Firms use capital and labor to produce output according to a Cobb-Douglas technology with parameters  $\alpha_L$  and  $\alpha_K$  and Total Factor Productivity (TFP)  $A$ . Suppose that firms are price takers in the input markets for labor and capital. Let  $W_L$  and  $W_K$  be the price of labor and capital, respectively.

**Q2.1.** (5 points) (a) Derive the expression for the cost function of a firm  $C(Y)$  as a function of output  $Y$ , the technological parameters  $\alpha_L$  and  $\alpha_K$ , TFP, and input prices. Explain in detail the different steps in your derivation.

ANSWER. The cost function  $C(Y)$  is defined as the minimum cost of producing the amount of output  $Y$ :

$$C(Y) = \left[ \begin{array}{l} \min_{\{L,K\}} W_L L + W_K K \\ \text{subject to: } Y = A L^{\alpha_L} K^{\alpha_K} \end{array} \right]$$

Or using a Lagrange representation:

$$C(Y) = \min_{\{\lambda, L, K\}} W_L L + W_K K + \lambda [Y - A L^{\alpha_L} K^{\alpha_K}]$$

where  $\lambda$  is the Lagrange multiplier of the restriction. The marginal conditions of optimality with respect to labor and capital are:

$$\left\{ \begin{array}{l} W_L - \lambda \alpha_L \frac{Y}{L} = 0 \\ W_K - \lambda \alpha_K \frac{Y}{K} = 0 \end{array} \right.$$

We can also write these equations as:

$$\left\{ \begin{array}{l} L = \lambda Y \frac{\alpha_L}{W_L} \\ K = \lambda Y \frac{\alpha_K}{W_K} \end{array} \right.$$

The marginal condition of optimality with respect to  $\lambda$  is:

$$Y - A L^{\alpha_L} K^{\alpha_K} = 0$$

Therefore, we have the following system of three equations with three endogenous unknowns ( $L, K, \lambda$ ):

$$\left\{ \begin{array}{l} L = \lambda Y \frac{\alpha_L}{W_L} \\ K = \lambda Y \frac{\alpha_K}{W_K} \\ Y = A L^{\alpha_L} K^{\alpha_K} \end{array} \right.$$

To solve this system, we solve the first two equations into the third one. We get:

$$Y = A \left( \lambda Y \frac{\alpha_L}{W_L} \right)^{\alpha_L} \left( \lambda Y \frac{\alpha_K}{W_K} \right)^{\alpha_K}$$

And solve for the endogenous  $\lambda$  in terms of the exogenous variables to get

$$\lambda = A^{\frac{-1}{\alpha}} Y^{\frac{1-\alpha}{\alpha}} \left(\frac{W_L}{\alpha_L}\right)^{\frac{\alpha_L}{\alpha}} \left(\frac{W_K}{\alpha_K}\right)^{\frac{\alpha_K}{\alpha}}$$

with  $\alpha = \alpha_L + \alpha_K$ . Plugging this equation into the first order conditions for labor and capital, we have:

$$\begin{cases} W_L L = \lambda Y \alpha_L = \alpha_L \left(\frac{Y}{A}\right)^{\frac{1}{\alpha}} \left(\frac{W_L}{\alpha_L}\right)^{\frac{\alpha_L}{\alpha}} \left(\frac{W_K}{\alpha_K}\right)^{\frac{\alpha_K}{\alpha}} \\ W_K K = \lambda Y \alpha_K = \alpha_K \left(\frac{Y}{A}\right)^{\frac{1}{\alpha}} \left(\frac{W_L}{\alpha_L}\right)^{\frac{\alpha_L}{\alpha}} \left(\frac{W_K}{\alpha_K}\right)^{\frac{\alpha_K}{\alpha}} \end{cases}$$

Finally, by definition, the cost function  $C(Y)$  is equal to the cost  $W_L L + W_K K$  when inputs are chosen to minimize the cost of producing  $Y$  unit of output. Then, adding up the previous expressions for  $W_L L$  and  $W_K K$  we get:

$$C(Y) = \alpha \left(\frac{Y}{A}\right)^{\frac{1}{\alpha}} \left(\frac{W_L}{\alpha_L}\right)^{\frac{\alpha_L}{\alpha}} \left(\frac{W_K}{\alpha_K}\right)^{\frac{\alpha_K}{\alpha}}$$

**Q2.2. (5 points) For this question Q2.2, suppose that  $\alpha_L = 0.6$  and  $\alpha_K = 0.3$ . Obtain the values of the following elasticities. Explain your answer. (a) Elasticity of cost with respect to output. (b) Elasticity of cost with respect to TFP. (c) Elasticity of cost with respect to the price of labor. (d) Elasticity of cost with respect to the price of capital. (e) Elasticity of output with respect to labor. (f) Elasticity of output with respect to TFP.**

ANSWER. To obtain the elasticities, it is very convenient to take logarithms in both sides of the cost function to represent all the variables in logs. Remember that if we have a log-linear relationship between two variables, say  $a$  and  $b$ , such that  $\log(a) = \alpha + \beta \log(b)$ , the parameter  $\beta$  represents the elasticity of variable  $a$  with respect to variable  $b$ . To prove this, note that by differentiating the equation we have that:  $d(\log(a)) = d(\alpha + \beta \log(b))$ , that taking into account the derivative of the logarithm function implies  $\frac{1}{a} da = \beta \frac{1}{b} db$ , and

solving for  $\beta$  we get  $\beta = \frac{da}{db} \frac{b}{a}$ , that by definition is the elasticity of  $a$  with respect to  $b$ .

Applying logarithms to both sides of the cost function, we have:

$$\begin{aligned} \log(C) &= \log(\alpha) + \frac{1}{\alpha} \log(Y) - \frac{1}{\alpha} \log(A) \\ &\quad + \frac{\alpha_L}{\alpha} \log(W_L) - \frac{\alpha_L}{\alpha} \log(\alpha_L) + \frac{\alpha_K}{\alpha} \log(W_K) - \frac{\alpha_K}{\alpha} \log(\alpha_K) \end{aligned}$$

Similarly, applying logs to the production function.

$$\log(Y) = \log(A) + \alpha_L \log(L) + \alpha_K \log(K)$$

- (a) Elasticity of cost with respect to output  $= \frac{1}{\alpha} = \frac{1}{0.6 + 0.3} = 1.11$ .
- (b) Elasticity of cost with respect to TFP  $= \frac{-1}{\alpha} = \frac{-1}{0.6 + 0.3} = -1.11$ .
- (c) Elasticity of cost with respect to the price of labor  $= \frac{\alpha_L}{\alpha} = \frac{0.6}{0.6 + 0.3} = 0.66$ .
- (d) Elasticity of cost with respect to the price of capital  $= \frac{\alpha_K}{\alpha} = \frac{0.3}{0.6 + 0.3} = 0.33$ .
- (e) Elasticity of output with respect to labor  $= \alpha_L = 0.60$ .
- (f) Elasticity of output with respect to TFP  $= 1.00$

**Suppose that the output market in this industry is competitive: firms are price takers. The demand function is linear with the following form:  $P = 100 - Q$ , where  $P$  and  $Q$  are the industry price and total output, respectively. Suppose that  $\alpha_L = \alpha_K = 1/4$ , and the value of input prices are  $W_L = W_K = 1/4$ . Suppose also that each firm has a fixed cost (the cost of fixed land) that is exogenous and equal to 2.**

\*\*\* NOTE. In this question, there is a third input that is land. We can consider that the production function is:  $Y = A L^{\alpha_L} K^{\alpha_K} (Land)^{\alpha_{Land}}$ . Since the Land input is the same for all the firms, we can consider, without loss of generality, that firms employ one unit of land. That is  $Land = 1$  such that  $Y = A L^{\alpha_L} K^{\alpha_K} (1)^{\alpha_{Land}} = A L^{\alpha_L} K^{\alpha_K}$ . Since Land is fixed, it is not chosen optimally to minimize costs. Therefore,  $W_{Land} * Land = W_{Land} = 2$  is fixed and can be simply added to the cost function from labor and capital we have obtained above. That is,

$$C(Y) = \alpha \left( \frac{Y}{A} \right)^{\frac{1}{\alpha}} \left( \frac{W_L}{\alpha_L} \right)^{\frac{\alpha_L}{\alpha}} \left( \frac{W_K}{\alpha_K} \right)^{\frac{\alpha_K}{\alpha}} + 2$$

**Q2.3. (5 points) Using these primitives, write the expression for the profit function of a firm (revenue minus cost) as a function of the market price,  $P$ , the firm's output,  $Y_i$ , and its TFP,  $A_i$ .**

ANSWER. A firm's revenue is  $P Y_i$ . To obtain a firm's cost function, note that:  $\alpha = \alpha_L + \alpha_K = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$ ;  $\frac{1}{\alpha} = 2$ ;  $\frac{W_L}{\alpha_L} = \frac{1/4}{1/4} = 1$ ; and  $\frac{W_K}{\alpha_K} = \frac{1/4}{1/4} = 1$ . Therefore, we have:

$$C(Y_i) = \frac{1}{2} \left( \frac{Y_i}{A_i} \right)^2 (1)^{\frac{\alpha_L}{\alpha}} (1)^{\frac{\alpha_K}{\alpha}} + 2 = \frac{1}{2} \left( \frac{Y_i}{A_i} \right)^2 + 2$$

Putting together the revenue and the cost function, we have the profit function:

$$\Pi_i = P Y_i - \frac{1}{2} \left( \frac{Y_i}{A_i} \right)^2 - 2$$

**Q2.4. (5 points)** Using the condition "price equal to marginal cost", obtain the optimal amount of output of a firm as a function of the market price,  $P$ , and the firm's TFP,  $A_i$ . Explain your derivation.

ANSWER. The marginal cost of a firm in this industry is  $MC_i = C'(Y_i) = \frac{Y_i}{A_i^2}$ . Therefore, the marginal condition of profit maximization is  $P = \frac{Y_i}{A_i^2}$ . Solving for output, we get:

$$Y_i = P A_i^2$$

A firm's optimal amount of output increases with market price and with its own TFP.

**Q2.5. (5 points)** A firm is active in the market (i.e., it finds optimal to produce a positive amount of output) only if its profit is greater or equal than zero. Using this condition show that a firm is active in this industry only if its TFP satisfies the condition  $A_i \geq 2/P$ . Explain your derivation.

ANSWER. Solving the condition  $Y_i = P A_i^2$  into the profit function, we get:

$$\begin{aligned} \Pi_i &= P * P * A_i^2 - \frac{1}{2} \left( \frac{P A_i^2}{A_i} \right)^2 - 2 \\ &= \frac{1}{2} (P A_i)^2 - 2 \end{aligned}$$

Then, the condition  $\Pi_i \geq 0$  is equivalent to  $\frac{1}{2} (P A_i)^2 - 2 \geq 0$ . Operating in this inequality, we obtain that this condition is equivalent to  $(P A_i)^2 \geq 4$ , and to  $P A_i \geq 2$ , and finally (given that market price is strictly positive) to  $A_i \geq 2/P$ .

Let  $(P^*, Q^*, Y_1^*, Y_2^*, \dots, Y_N^*)$  the equilibrium price, total output, and individual firms' outputs. Based on the previous results, the market equilibrium can be characterized by the following conditions:

- (i) The demand equation holds:  $P^* = 100 - Q^*$ .
- (ii) Total output is equal to the sum of firms' individual outputs:  $Q^* = Y_1^* + \dots + Y_N^*$ .
- (iii) Firm  $i$  is active ( $Y_i^* > 0$ ) if and only if its total profit is greater than zero:  $Y_i^* > 0$  if and only if  $A_i \geq 2/P^*$ .
- (iv) For firms with  $Y_i^* > 0$ , the optimal amount of output is given by the condition  $P^* = MC_i(Y_i^*)$ , where  $MC_i(\cdot)$  represents the marginal cost function for firm  $i$ .

**Q2.6. (5 points)** Combine conditions (i) to (iv) to show that the equilibrium price can be written as the solution to this equation:

$$P^* = 100 - P^* \left[ \sum_{i=1}^N A_i^2 1\{A_i \geq 2/P^*\} \right]$$

where  $1\{x\}$  is the indicator function that is defined as  $1\{x\} = 1$  if condition  $x$  is true, and  $1\{x\} = 0$  if condition  $x$  is false. Explain your derivation.

ANSWER. Plugging condition (ii) into (i), we get:  $P^* = 100 - (Y_1^* + Y_2^* + \dots + Y_N^*)$ . Combining conditions (iii) and (iv), we have that the equilibrium amount of output for firm  $i$  is:  $Y_i^* = P^* A_i^2 1\{A_i \geq 2/P^*\}$ . Plugging this expression into the previous condition for  $P^*$ , we get:

$$P^* = 100 - P^* \left[ \sum_{i=1}^N A_i^2 1\{A_i \geq 2/P^*\} \right]$$

Suppose that the subindex  $i$  sorts firms by their TFP such that firm 1 is the most efficient, then firm 2, etc. That is,  $A_1 > A_2 > A_3 > \dots$ .

**Q2.7. (5 points)** Suppose that  $A_1 = 7$ ,  $A_2 = 5$ , and  $A_3 = 1$ . Obtain the equilibrium price, total output, and output of each individual firm in this industry. [Hint: Start with the conjecture that only firms 1 and 2 produce in equilibrium. Then, confirm this conjecture. Note that we do not need to know the values of  $A_4, A_5$ , etc].

ANSWER. Suppose for the moment that only firms 1 and 2 produce in equilibrium. Then, the equilibrium price satisfies the condition  $P^* = 100 - P^*[A_1^2 + A_2^2] = 100 - P^*[49 + 25]$ . Solving for  $P^*$ , we get  $P^* = 100/75 = 4/3$ . For firms 1 and 2 to be the only firms active in the market we need that these conditions hold:

$$A_1 \geq 2/P^*, \quad A_2 \geq 2/P^*, \quad \text{and} \quad A_3 < 2/P^*$$

Since  $A_i < A_3$  for any  $i > 3$ , it is clear that third condition implies that  $A_i < 2/P^*$  for any  $i > 3$ , so we do not need to check these other conditions. Now, given that  $P^* = 4/3$  we have that the threshold value for the productivity of an active firm is  $2/P^* = 3/2$ . Therefore, the three conditions are:

$$A_1 \geq 3/2, \quad A_2 \geq 3/2, \quad \text{and} \quad A_3 < 3/2$$

Since  $A_1 = 7$ ,  $A_2 = 5$ , and  $A_3 = 1$ , it is clear that the three conditions hold and the conjecture is confirmed.

The total industry output is  $Q^* = 100 - P^* = 98.66$ .

Firm 1's output is  $Y_1^* = P^* A_1 = 49 * (4/3) = 65.33$ .

Firm 2's output is  $Y_2^* = P^* A_2 = 25 * (4/3) = 33.33$ .  
And  $Y_i^* = 0$  for any firm  $i \geq 3$ .

**Q2.8. (5 points) Explain why the most efficient firm, with the largest TFP, does not produce all the output of the industry.**

ANSWER. Since the cost function  $C(Y_i)$  is strictly convex with respect to output, we have that the profit function is strictly concave. That is, even if this firm is a monopolist, there is a maximum amount of output that is willing to offer in the market, and is not willing to offer more than this amount. In our model, the maximum amount that the most efficient firm is willing to offer given a price  $P$  is  $Y_1 = P A_1 = P 7$ . If this firm were the only one in the market, the equilibrium price would be the solution to  $P = 100 - P * 7$ , which implies  $P = 12.5$ . Given this price, the threshold value for TFP for the decision of market entry is  $2/P = 2/12.5 = 0.16$ . Given this threshold value, we have that  $A_2 = 5 > 0.16$ , such that firm 2 is willing to be active in the market and produce a positive amount of output.

---

**QUESTION 3.** [100 points]. The datafile `blundell_bond_2000_production_function.dta` contains annual information on sales, labor, and capital for 509 firms for the period 1982-1989 (8 years). Consider a Cobb-Douglas production function in terms of labor and capital. Use this dataset to implement the following estimators and hypothesis tests. Provide the code in STATA and the table of estimation results.

**Q3.1.** (10 points) (a) OLS with time dummies. (b) Test the null hypothesis  $\alpha_L + \alpha_K = 1$ . Comment the results.

ANSWER. **MODEL:** The regression model is:

$$y_{it} = \alpha_0 + \alpha_L \ell_{it} + \alpha_K k_{it} + \gamma_t + \omega_{it}$$

where the variables are in logarithms. We account for the time effects  $\gamma_t$  by including time (year) dummies: one for each year, except one.

#### CODE

```
gen logy = ln(sales)
gen logn = ln(labor)
gen logk = ln(capital)
xtset id year
reg logy logn logk i.year
test logn + logk = 1
```

#### COMMENTS ON RESULTS

- (a) According to the estimates, the technology is more intensive in labor than in capital.
- (b) The test of Constant Returns to Scale (CRS) rejects the null hypothesis under the standard significance levels (p-value is smaller than 1%). There is evidence of decreasing returns to scale.
- (c) However, we expect the OLS estimator to be biased because endogeneity, i.e., correlation between the regressors (observed inputs) and the error term (unobserved TFP).

## TABLE OF RESULTS

```
. // -----
. // 3. Question 2.1: OLS estimator
. // -----
. reg logy logn logk i.year
```

Source	SS	df	MS	Number of obs	=	4,072
Model	15946.3907	9	1771.82119	F(9, 4062)	=	14254.66
Residual	504.897075	4,062	.124297655	Prob > F	=	0.0000
				R-squared	=	0.9693
				Adj R-squared	=	0.9692
Total	16451.2878	4,071	4.04109255	Root MSE	=	.35256

logy	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
logn	.5578836	.0098286	56.76	0.000	.5386142	.577153
logk	.4322828	.0081396	53.11	0.000	.4163247	.4482409
year						
1983	-.0568626	.022107	-2.57	0.010	-.1002045	-.0135206
1984	-.050041	.0221342	-2.26	0.024	-.0934362	-.0066458
1985	-.0875714	.0221985	-3.94	0.000	-.1310926	-.0440503
1986	-.092866	.0222691	-4.17	0.000	-.1365256	-.0492063
1987	-.0580931	.0223043	-2.60	0.009	-.1018218	-.0143644
1988	-.0211632	.0223277	-0.95	0.343	-.0649378	.0226114
1989	-.0382923	.0224365	-1.71	0.088	-.0822802	.0056957
_cons	3.046843	.0315266	96.64	0.000	2.985033	3.108652

```
. test logn + logk = 1
```

```
( 1)  logn + logk = 1
```

```
      F( 1, 4062) =    9.29
      Prob > F =    0.0023
```

**Q3.2. (10 points)** (a) **Fixed Effects estimator with time dummies.** (b) **Test the null hypothesis of no time-invariant unobserved heterogeneity:  $\eta_i = \eta$  for every firm  $i$ .** Comment the results.

ANSWER. **MODEL:** The regression model is:

$$y_{it} = \alpha_0 + \alpha_L \ell_{it} + \alpha_K k_{it} + \gamma_t + \eta_i + u_{it}$$

where the variables are in logarithms. We eliminate the individual effect by transforming the model in deviations with respect to firm-means:

$$\tilde{y}_{it} = \alpha_L \tilde{\ell}_{it} + \alpha_K \tilde{k}_{it} + \eta_i + \tilde{u}_{it}$$

where  $\tilde{y}_{it} = y_{it} - \bar{y}_i$ ,  $\tilde{\ell}_{it} = \ell_{it} - \bar{\ell}_i$ ,  $\tilde{k}_{it} = k_{it} - \bar{k}_i$ , and  $\tilde{u}_{it} = u_{it} - \bar{u}_i$ , and the variables  $\bar{y}_i$ ,  $\bar{\ell}_i$ , and  $\bar{k}_i$  are the sample means of log-output, log-labor, and log-capital for firm  $i$ , respectively. We apply OLS to this model. We account for the time effects  $\gamma_t$  by including time (year) dummies: one for each year, except one.

**CODE.** The command `xtreg ..., fe` implements this estimator. We don't need to transform the variables, the command makes this transformation for us.

```
xtreg logy logn logk i.year, fe
test logn + logk = 1
```

## COMMENTS ON RESULTS

(a) According to the estimates, the technology is more intensive in labor than in capital. The test of Constant Returns to Scale (CRS) clearly rejects the null hypothesis under the standard significance levels (p-value is smaller than 1%). There is evidence of decreasing returns to scale.

(b) The test of the null hypothesis of no time-invariant unobserved heterogeneity ( $\eta_i = 0$  for every firm  $i$ ) is the F-test at the bottom of the table of estimates. The p-value of this test is practically zero. Therefore, we clearly reject the null hypothesis of no time-invariant unobserved heterogeneity.

(c) In fact, most of the variance of the error term  $\eta_i + u_{it}$  is accounted by the time-invariant component  $\eta_i$ . This is shown by the parameter "rho = 0.8948" that represents the estimate for  $Var(\eta_i)/Var(\eta_i + u_{it})$ .

(d) However, we expect the FE estimator to be biased. This estimator may control for the endogeneity due to the fixed effect  $\eta_i$  but not for the endogeneity problem due to the correlation between the regressors (observed inputs) and the transitory shock  $u_{it}$ .

## TABLE OF RESULTS

```
. // -----
. // 4. Question 2.2: Fixed Effects estimation
. // -----
. xtreg logy logn logk i.year, fe
```

Fixed-effects (within) regression

Group variable: id

Number of obs = 4,072

Number of groups = 509

R-sq:

within = 0.7379

between = 0.9706

overall = 0.9661

Obs per group:

min = 8

avg = 8.0

max = 8

corr(u\_i, Xb) = 0.5988

F(9,3554) = 1111.47

Prob > F = 0.0000

logy	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
logn	.6544609	.0144048	45.43	0.000	.6262184	.6827034
logk	.2329072	.013637	17.08	0.000	.2061702	.2596443
year						
1983	-.0376406	.0093042	-4.05	0.000	-.0558828	-.0193985
1984	-.0076445	.0096071	-0.80	0.426	-.0264805	.0111914
1985	-.0234513	.0100955	-2.32	0.020	-.0432449	-.0036578
1986	-.0136103	.0105543	-1.29	0.197	-.0343034	.0070829
1987	.0314121	.0108748	2.89	0.004	.0100907	.0527335
1988	.0753576	.0111072	6.78	0.000	.0535805	.0971347
1989	.0764164	.0118166	6.47	0.000	.0532485	.0995844
_cons	3.863804	.0529288	73.00	0.000	3.76003	3.967578
sigma_u	.42922318					
sigma_e	.14715329					
rho	.89482518	(fraction of variance due to u_i)				

F test that all u\_i=0: F(508, 3554) = 38.90 Prob > F = 0.0000

```
. test logn + logk = 1
```

```
( 1) logn + logk = 1
```

F( 1, 3554) = 121.32

Prob > F = 0.0000

**Q3.3. (10 points)** (a) **Fixed Effects - Cochrane Orcutt estimator with time dummies.** (b) **Test the two over-identifying restrictions of the model. Comment the results.**

**ANSWER. MODEL:** The Cochrane-Orcutt estimator is applied to eliminate the serial correlation in the transitory shock  $u_{it}$ . Suppose that  $u_{it}$  follows an AR(1) process such that  $u_{it} = \rho u_{it-1} + a_{it}$ , where  $a_{it}$  is not serially correlated. Then, we can obtain the a quasi-first difference transformation of the model (equation at period  $t$  minus  $\rho$  times equation at period  $t - 1$ ). This implies the following equation:

$$y_{it} = \beta_1 y_{it-1} + \beta_2 \ell_{it} + \beta_3 \ell_{it-1} + \beta_4 k_{it} + \beta_5 k_{it-1} + \eta_i^* + \gamma_t^* + a_{it}$$

with  $\beta_1 = \rho$ ,  $\beta_2 = \alpha_L$ ,  $\beta_3 = -\rho\alpha_L$ ,  $\beta_4 = \alpha_K$ , and  $\beta_5 = -\rho\alpha_K$ . The FE Cochrane-Orcutt estimator is the FE estimator in this equation.

The model implies two restrictions on the parameter estimates  $\beta$ :

$$-\beta_3/\beta_2 = \beta_1 \quad \text{and} \quad -\beta_5/\beta_4 = \beta_1$$

We can test these nonlinear restrictions separately or jointly using the command "testnl" in Stata.

**CODE.** To implement this estimator, we can use the command `xtreg . . . ., fe` but now we need to include also as regressors the first lags of log-output, log-labor, and log-capital.

```
xtreg logy l.logy logn l.logn logk l.logk i.year, fe
```

The code for the test the restrictions. The following is the code for a test of CRS, and for testing the over-identifying restrictions (OIR) between the parameters. We have implemented both single tests of each OIR and a joint test.

```
test logn + logk = 1
testnl (_b[l.logy] = -_b[l.logn]/_b[logn])
testnl (_b[l.logy] = -_b[l.logk]/_b[logk])
testnl (_b[l.logy] = -_b[l.logn]/_b[logn]) (_b[l.logy] = -_b[l.logk]/_b[logk])
```

## TABLE OF RESULTS

```
. // -----
. // 5. Question 2.3: Fixed Effects - Cochrane-Orcutt estimation
. // -----
. xtreg logy l.logy logn l.logn logk l.logk i.year, fe
```

Fixed-effects (within) regression

Group variable: id

Number of obs = 3,563

Number of groups = 509

R-sq:

within = 0.7825

between = 0.9879

overall = 0.9847

Obs per group:

min = 7

avg = 7.0

max = 7

F(11,3043) = 995.10

Prob > F = 0.0000

corr(u\_i, Xb) = 0.7191

logy	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
logy						
L1.	.4039344	.015273	26.45	0.000	.3739879	.4338808
logn						
--.	.4880013	.0166747	29.27	0.000	.4553065	.5206961
L1.	-.0231194	.0192464	-1.20	0.230	-.0608566	.0146179
logk						
--.	.1765454	.0178288	9.90	0.000	.1415877	.2115032
L1.	-.1305487	.0164086	-7.96	0.000	-.1627218	-.0983757
year						
1984	.0564054	.0077479	7.28	0.000	.0412138	.071597
1985	.0271379	.0081264	3.34	0.001	.0112041	.0430717
1986	.0494812	.0086725	5.71	0.000	.0324767	.0664858
1987	.1033078	.0091382	11.31	0.000	.0853902	.1212255
1988	.1310847	.0094591	13.86	0.000	.1125379	.1496315
1989	.1174383	.0100291	11.71	0.000	.0977739	.1371027
_cons	2.625541	.0799623	32.83	0.000	2.468756	2.782327
sigma_u	.31731619					
sigma_e	.12076713					
rho	.87347826	(fraction of variance due to u_i)				

F test that all u\_i=0: F(508, 3043) = 3.76 Prob > F = 0.0000

```

. test logn + logk = 1

(1)  logn + logk = 1

      F( 1, 3043) = 447.06
      Prob > F = 0.0000

. testnl (_b[l.logy] = -_b[l.logn]/_b[logn])

(1)  _b[l.logy] = -_b[l.logn]/_b[logn]

      chi2(1) = 111.37
      Prob > chi2 = 0.0000

. testnl (_b[l.logy] = -_b[l.logk]/_b[logk])

(1)  _b[l.logy] = -_b[l.logk]/_b[logk]

      chi2(1) = 21.13
      Prob > chi2 = 0.0000

. testnl (_b[l.logy] = -_b[l.logn]/_b[logn]) (_b[l.logy] = -_b[l.logk]/_b[logk]
> )

(1)  _b[l.logy] = -_b[l.logn]/_b[logn]
(2)  _b[l.logy] = -_b[l.logk]/_b[logk]

      chi2(2) = 112.43
      Prob > chi2 = 0.0000

```

## COMMENTS ON RESULTS

(a) According to the estimates, the technology is more intensive in labor than in capital. The magnitude for the estimate of  $\alpha_K$  is too small, and as a result the value of  $\alpha_L + \alpha_K$  is too far away from CRS.

(b) The estimate of  $\rho$  (or  $\beta_1$ ) is significantly different to zero and it is not small. Therefore, there is evidence of serial correlation in the transitory shock.

(c) The test of the null hypothesis of no time-invariant unobserved heterogeneity ( $\eta_i = 0$  for every firm  $i$ ) is the F-test at the bottom of the table of estimates. The p-value of this test is practically zero. Therefore, we clearly reject the null hypothesis of no time-invariant unobserved heterogeneity.

(d) Tests of OIRs. The three tests (for each restriction separately, and for the two restrictions together) are clearly rejected with a p-value practically equal to zero. Therefore, there is clear evidence against the restrictions of this model.

(e) Despite there is evidence of serial correlation in the transitory shock and of time-invariant unobserved heterogeneity, the restrictions of this model/estimator are rejected. A possible explanation is that this estimator is biased/inconsistent because it does not control for the endogeneity due to the correlation between the regressors (observed inputs) and the innovation in the transitory shock,  $a_{it}$ . That is, this method controls for endogeneity due to the fixed effect  $\eta_i$  and to the component  $\rho u_{it-1}$  in the error term (we are controlling for this by including the lagged values  $y_{it-1}$ ,  $\ell_{it-1}$ , and  $k_{it-1}$  as regressors). But it does NOT controls for the endogeneity problem due to component  $a_{it}$  of the error term.

**Q3.4. (10 points) Arellano-Bond estimator with time dummies and non-serially correlated transitory shock. Comment the results.**

ANSWER. **MODEL:** We estimate the model in first differences:

$$\Delta y_{it} = \alpha_L \Delta \ell_{it} + \alpha_K \Delta k_{it} + \Delta \gamma_t + \Delta u_{it}$$

using as instruments  $\ell_{it-2}$  and  $k_{it-2}$  and also lags before  $t - 2$ . These instruments are valid only in  $u_{it}$  is not serially correlated, or what is equivalent, if  $\Delta u_{it}$  is serially correlated of first-order but not of second order or higher.

**CODE.** To implement this estimator, we can use the command `xtabond`.

- Though the model is estimated in first differences, the syntax of the model is such that the dependent and explanatory variables are input in levels, i.e., `logy logn logk i.year`
- The part `gmm(., lag())` of the command determines the set instruments. We use lags  $t - 2$  of log-output, log-labor, and log-capital: `gmm(logy logn logk, lag(2.))`
- The part `iv()` of the command determines the regressors which are exogenous, in our case only the time dummies: `iv(i.year)`
- We use the option `noleveleq` to choose the Arellano-Bond estimator (the default of this command is the System GMM estimator).
- We use the option `robust` to obtain standard errors that are robust of serial correlation and heteroscedasticity.

```
xtabond2 logy logn logk i.year, gmm(logy logn logk, lag(2.)) iv(i.year) robust
noleveleq
test logn + logk = 1
```

#### COMMENTS ON RESULTS

(a) According to the estimates, the coefficient of capital is very small and not significantly different to zero. This is completely implausible.

(b) The test of serial correlation "Arellano-Bond test for AR(2) in first differences" tests for the null hypothesis  $E(\Delta u_{it} \Delta u_{it-2}) = 0$ . The p-value of this test is 0.002. The null hypothesis is clearly rejected. This implies that  $u_{it}$  is serially correlated, and therefore that the instruments  $\ell_{it-2}$  and  $k_{it-2}$  are not valid.

In summary, the estimates do not have economic sense and the restrictions of the model are rejected.

## TABLE OF RESULTS

```
. // -----
. // 6. Question 2.4: Arellano-Bond: No AR(1)
. // -----
. xtabond2 logy logn logk i.year, gmm(logy logn logk, lag(2 .)) iv(i.year) robu
> st nolevelq
Favoring space over speed. To switch, type or click on mata: mata set matafavor
> speed, perm.
Warning: Two-step estimated covariance matrix of moments is singular.
Using a generalized inverse to calculate robust weighting matrix for Hansen t
> est.
Difference-in-Sargan/Hansen statistics may be negative.
```

Dynamic panel-data estimation, one-step difference GMM

Group variable: id	Number of obs	=	3563
Time variable: year	Number of groups	=	509
Number of instruments = 70	Obs per group: min	=	7
Wald chi2(10) = 845.00	avg	=	7.00
Prob > chi2 = 0.000	max	=	7

logy	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
logn	.6944881	.1021823	6.80	0.000	.4942144	.8947618
logk	.0529433	.0802371	0.66	0.509	-.1043186	.2102051
year						
1982	0	(empty)				
1983	-.0182737	.0102597	-1.78	0.075	-.0383824	.0018349
1984	.0352441	.0170495	2.07	0.039	.0018278	.0686604
1985	.0391971	.0242658	1.62	0.106	-.008363	.0867572
1986	.0623552	.0318072	1.96	0.050	.0000143	.1246961
1987	.117962	.03499	3.37	0.001	.0493828	.1865412
1988	.16933	.037151	4.56	0.000	.0965153	.2421447
1989	.1872415	.0432925	4.33	0.000	.1023897	.2720933

Instruments for first differences equation

Standard

D.(1982b.year 1983.year 1984.year 1985.year 1986.year 1987.year 1988.year 1989.year)

GMM-type (missing=0, separate instruments for each period unless collapsed)  
L(2/7).(logy logn logk)

Arellano-Bond test for AR(1) in first differences: z = -5.17 Pr > z = 0.000  
Arellano-Bond test for AR(2) in first differences: z = -3.03 Pr > z = 0.002

Sargan test of overid. restrictions: chi2(60) = 331.50 Prob > chi2 = 0.000  
(Not robust, but not weakened by many instruments.)

Hansen test of overid. restrictions: chi2(60) = 92.16 Prob > chi2 = 0.005  
(Robust, but weakened by many instruments.)

Difference-in-Hansen tests of exogeneity of instrument subsets:

iv(1982b.year 1983.year 1984.year 1985.year 1986.year 1987.year 1988.year 1989.year)

Hansen test excluding group: chi2(53) = 74.33 Prob > chi2 = 0.028  
Difference (null H = exogenous): chi2(7) = 17.84 Prob > chi2 = 0.013

. test logn + logk = 1

( 1) logn + logk = 1

chi2( 1) = 15.11  
Prob > chi2 = 0.0001

**Q3.5. (10 points) Arellano-Bond estimator with time dummies and AR(1) transitory shock. Comment the results.**

ANSWER. **MODEL:** Suppose that  $u_{it}$  follows an AR(1) process such that  $u_{it} = \rho u_{it-1} + a_{it}$ , where  $a_{it}$  is not serially correlated. Since the transitory shock is serially correlated, the Arellano-Bond instruments are not valid in the equation in first differences. However, we can transform the model taking a quasi-first difference (as we did for the Cochrane-Orcutt) estimator, such that we can obtain a model where the "new" transitory shock is  $a_{it}$ , which is not serially correlated. The model after the quasi-first difference is:

$$y_{it} = \beta_1 y_{it-1} + \beta_2 \ell_{it} + \beta_3 \ell_{it-1} + \beta_4 k_{it} + \beta_5 k_{it-1} + \eta_i^* + \gamma_t^* + a_{it}$$

Then, in this model we can apply first differences. The equation that we estimate is:

$$\Delta y_{it} = \beta_1 \Delta y_{it-1} + \beta_2 \Delta \ell_{it} + \beta_3 \Delta \ell_{it-1} + \beta_4 \Delta k_{it} + \beta_5 \Delta k_{it-1} + \Delta \gamma_t^* + \Delta a_{it}$$

with  $\beta_1 = \rho$ ,  $\beta_2 = \alpha_L$ ,  $\beta_3 = -\rho\alpha_L$ ,  $\beta_4 = \alpha_K$ , and  $\beta_5 = -\rho\alpha_K$ . We estimate this model using a instrumental variables (GMM) estimator using as instruments  $\ell_{it-2}$  and  $k_{it-2}$  and also lags before  $t - 2$ . Note that these instruments are valid in this transformed model.

**CODE.** To implement this estimator, we can use the command `xtabond`. The only difference with respect to the code in Question Q2.5 is in the set of regressors. Now, we need to include the first lags of log-output, log-labor, and log-capital as regressors. The set of instruments is the same, and the other options of the command are also the same. Remember that in the syntax of this command the variables are input in levels though the estimation is in first differences (the command makes the transformation for us). We also test the restrictions on the parameters implied by the AR(1) model.

```
xtabond2 logy l.logy logn l.logn logk l.logk i.year, gmm(logy logn logk, lag(2
.)) iv(i.year) robust noleveleq
test logn + logk = 1
testnl (_b[l.logy] = -_b[l.logn]/_b[logn])
testnl (_b[l.logy] = -_b[l.logk]/_b[logk])
testnl (_b[l.logy] = -_b[l.logn]/_b[logn]) (_b[l.logy] = -_b[l.logk]/_b[logk])
```

## COMMENTS ON RESULTS

(a) According to the estimates, the technology is more intensive in labor than in capital. Though the coefficient of capital is now significantly greater than zero, it still seems too small. As a result the value of  $\alpha_L + \alpha_K$  is too far away from CRS.

(b) The estimate of  $\rho$  (or  $\beta_1$ ) is significantly different to zero and it is not small. Therefore, there is evidence of serial correlation in the transitory shock.

(c) The test of serial correlation "Arellano-Bond test for AR(2) in first differences" tests for the null hypothesis  $E(\Delta a_{it} \Delta a_{it-2}) = 0$ . The p-value of this test is 0.173. The null hypothesis cannot be rejected using the standard significance levels (i.e., p-value is greater than 10%). This implies that we cannot reject that  $a_{it}$  is not serially correlated, and therefore that the instruments  $\ell_{it-2}$  and  $k_{it-2}$  are valid. That is, it seems that by including the lagged values  $y_{it-1}$ ,  $\ell_{it-1}$ , and  $k_{it-1}$  as regressors we have been able to control for serial correlation in the transitory shock.

(d) The model implies two restrictions on the parameter estimates  $\beta$ :

$$-\beta_3/\beta_2 = \beta_1 \quad \text{and} \quad -\beta_5/\beta_4 = \beta_1$$

We can test these nonlinear restrictions separately or jointly using the command "testnl" in Stata. We have implemented both the single and the joint tests. The test of  $H_0 : -\beta_3/\beta_2 = \beta_1$  has a p-value of 0.0143, such that we cannot reject the null at 1% but we reject it at 5% significance level. The test of  $H_0 : -\beta_5/\beta_4 = \beta_1$  has a p-value of 0.178, such that we cannot reject the null at the standard significance levels. The joint test has a p-value of 0.0227, which again implies that we cannot reject the null at 1% but we reject it at 5% significance level.

In summary, accounting for AR(1) transitory shocks has improved significantly the estimates and the specification tests. However, the restrictions of the model are only marginally "accepted" and, most importantly, the estimate for the coefficient of capital seems implausibly small.

```

. // -----
. // 7. Question 2.5: Arellano-Bond: AR(1)
. // -----
. xtabond2 logy l.logy logn l.logn logk l.logk i.year, gmm(logy logn logk, lag(
> 2 .)) iv(i.year) robust nolevelq
Favoring space over speed. To switch, type or click on mata: mata set matafavor
> speed, perm.
Warning: Two-step estimated covariance matrix of moments is singular.
Using a generalized inverse to calculate robust weighting matrix for Hansen t
> est.
Difference-in-Sargan/Hansen statistics may be negative.

Dynamic panel-data estimation, one-step difference GMM

```

Group variable: id	Number of obs	=	3054
Time variable : year	Number of groups	=	509
Number of instruments = 69	Obs per group: min	=	6
Wald chi2(13) = 1340.23	avg	=	6.00
Prob > chi2 = 0.000	max	=	6

logy	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
logy					
L1.	.3264209	.0521606	6.26	0.000	.2241881 .4286538
logn					
--.	.5128697	.0892432	5.75	0.000	.3379563 .6877831
L1.	.0726061	.0927269	0.78	0.434	-.1091353 .2543475
logk					
--.	.1318291	.1179029	1.12	0.264	-.0992562 .3629145
L1.	-.2066298	.0949589	-2.18	0.030	-.3927458 -.0205137
year					
1982	0	(empty)			
1983	-.091299	.0200443	-4.55	0.000	-.1305851 -.0520129
1984	-.0236894	.0151667	-1.56	0.118	-.0534155 .0060367
1985	-.0347948	.0098768	-3.52	0.000	-.0541531 -.0154366
1986	0	(omitted)			
1987	.0639394	.00849	7.53	0.000	.0472994 .0805794
1988	.1015617	.0105429	9.63	0.000	.080898 .1222254
1989	.0986976	.0134461	7.34	0.000	.0723437 .1250514

```

Instruments for first differences equation
Standard
D.(1982b.year 1983.year 1984.year 1985.year 1986.year 1987.year 1988.year
1989.year)
GMM-type (missing=0, separate instruments for each period unless collapsed)
L(2/7).(logy logn logk)

Arellano-Bond test for AR(1) in first differences: z = -6.21 Pr > z = 0.000
Arellano-Bond test for AR(2) in first differences: z = -1.36 Pr > z = 0.173

Sargan test of overid. restrictions: chi2(56) = 213.56 Prob > chi2 = 0.000
(Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(56) = 98.39 Prob > chi2 = 0.000
(Robust, but weakened by many instruments.)

Difference-in-Hansen tests of exogeneity of instrument subsets:
iv(1982b.year 1983.year 1984.year 1985.year 1986.year 1987.year 1988.year 198
> 9.year)
Hansen test excluding group: chi2(50) = 78.58 Prob > chi2 = 0.006
Difference (null H = exogenous): chi2(6) = 19.81 Prob > chi2 = 0.003

. test logn + logk = 1
( 1) logn + logk = 1
      chi2( 1) = 12.94
      Prob > chi2 = 0.0003

. testnl (_b[l.logy] = -_b[l.logn]/_b[logn])
(1) _b[l.logy] = -_b[l.logn]/_b[logn]
      chi2(1) = 6.00
      Prob > chi2 = 0.0143

. testnl (_b[l.logy] = -_b[l.logk]/_b[logk])
(1) _b[l.logy] = -_b[l.logk]/_b[logk]
      chi2(1) = 1.81
      Prob > chi2 = 0.1780

. testnl (_b[l.logy] = -_b[l.logn]/_b[logn]) (_b[l.logy] = -_b[l.logk]/_b[logk]
> )
(1) _b[l.logy] = -_b[l.logn]/_b[logn]
(2) _b[l.logy] = -_b[l.logk]/_b[logk]
      chi2(2) = 7.57
      Prob > chi2 = 0.0227

```

**Q3.6. (10 points) Blundell-Bond system estimator with time dummies and non-serially correlated transitory shock. Comment the results.**

ANSWER. **MODEL:** Now, we estimate the parameters of the model by combining two sets of moment restrictions: (i) the ones from the Arellano-Bond IV (GMM) estimator; and (ii) the ones from the Blundell-Bond IV (GMM) estimation of the equation in levels:

$$y_{it} = \alpha_L \ell_{it} + \alpha_K k_{it} + \gamma_t + (\eta_i + u_{it})$$

using as instruments  $\Delta \ell_{it-1}$  and  $\Delta k_{it-1}$ . Remember that this system estimator uses both the Arellano-Bond instruments in the equation in first differences, and the Blundell-Bond instruments in the equation in levels. For the validity of these instruments (either AB or BB) we need the transitory shock  $u_{it}$  not being serially correlated.

**CODE.** To implement this estimator, we can use the command `xtabond`. The only difference with respect to the code in Question Q2.4 is that we should not include the option `noleveleq`. Note that the syntax of this command is such that, though we input the list of instruments in levels, the command understands that this instruments should be in levels for the equation in first-differences, and the instruments should be in first differences for the equation in levels.

```
xtabond2 logy logn logk i.year, gmm(logy logn logk, lag(2 .)) iv(i.year) robust
test logn + logk = 1
```

### COMMENTS ON RESULTS

(a) According to the estimates, the technology is more intensive in labor than in capital. But now both  $\alpha_L$  and  $\alpha_K$  are larger and the value of  $\alpha_L + \alpha_K$  is very close CRS. In fact, the test of CRS has a p-value of 0.7794.

(b) The test of serial correlation "Arellano-Bond test for AR(2) in first differences" tests for the null hypothesis  $E(\Delta u_{it} \Delta u_{it-2}) = 0$ . Similarly as for the Arellano-Bond estimator without AR(1), the p-value of this test is practically zero. The null hypothesis is clearly rejected. This implies that  $u_{it}$  is serially correlated, and therefore neither AB nor BB instruments are valid.

```

. // -----
. // 8. Question 2.6: System GMM: No AR(1)
. // -----
. xtabond2 logy logn logk i.year, gmm(logy logn logk, lag(2.)) iv(i.year) robu
> st
Favoring space over speed. To switch, type or click on mata: mata set matafavor
> speed, perm.
Warning: Two-step estimated covariance matrix of moments is singular.
Using a generalized inverse to calculate robust weighting matrix for Hansen t
> est.
Difference-in-Sargan/Hansen statistics may be negative.

Dynamic panel-data estimation, one-step system GMM

```

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
logy						
logn	.7359973	.0860694	8.55	0.000	.5673045	.9046902
logk	.2724414	.0702628	3.88	0.000	.1347289	.410154
year						
1982	0	(empty)				
1983	-.0457584	.0094766	-4.83	0.000	-.0643322	-.0271846
1984	-.0259176	.0147461	-1.76	0.079	-.0548195	.0029842
1985	-.046355	.0214732	-2.16	0.031	-.0884417	-.0042683
1986	-.0387785	.0281867	-1.38	0.169	-.0940235	.0164665
1987	.0013652	.0299436	0.05	0.964	-.0573233	.0600537
1988	.0415917	.031848	1.31	0.192	-.0208291	.1040126
1989	.038114	.0373453	1.02	0.307	-.0350814	.1113095
_cons	3.586964	.2323373	15.44	0.000	3.131591	4.042337

```

Instruments for first differences equation
Standard
D.(1982b.year 1983.year 1984.year 1985.year 1986.year 1987.year 1988.year
1989.year)
GMM-type (missing=0, separate instruments for each period unless collapsed)
L(2/7).(logy logn logk)
Instruments for levels equation
Standard
1982b.year 1983.year 1984.year 1985.year 1986.year 1987.year 1988.year
1989.year
_cons
GMM-type (missing=0, separate instruments for each period unless collapsed)
DL.(logy logn logk)

Arellano-Bond test for AR(1) in first differences: z = -6.52 Pr > z = 0.000
Arellano-Bond test for AR(2) in first differences: z = -3.52 Pr > z = 0.000

Sargan test of overid. restrictions: chi2(78) =2480.33 Prob > chi2 = 0.000
(Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(78) = 190.38 Prob > chi2 = 0.000
(Robust, but weakened by many instruments.)

Difference-in-Hansen tests of exogeneity of instrument subsets:
GMM instruments for levels
Hansen test excluding group: chi2(60) = 89.01 Prob > chi2 = 0.009
Difference (null H = exogenous): chi2(18) = 101.36 Prob > chi2 = 0.000
iv(1982b.year 1983.year 1984.year 1985.year 1986.year 1987.year 1988.year 198
> 9.year)
Hansen test excluding group: chi2(71) = 135.63 Prob > chi2 = 0.000
Difference (null H = exogenous): chi2(7) = 54.75 Prob > chi2 = 0.000

. test logn + logk = 1

( 1) logn + logk = 1

chi2( 1) = 0.08
Prob > chi2 = 0.7794

```

**Q3.7. (10 points) Blundell-Bond system estimator with time dummies and AR(1) transitory shock. Comment the results.**

ANSWER. **MODEL:** Suppose that  $u_{it}$  follows an AR(1) process such that  $u_{it} = \rho u_{it-1} + a_{it}$ , where  $a_{it}$  is not serially correlated. Since the transitory shock is serially correlated, the Arellano-Bond and the Blundell-Bond instruments are not valid in the equation in first differences. However, we can transform the model taking a quasi-first difference (as we did for the Cochrane-Orcutt) estimator, such that we can obtain a model where the "new" transitory shock is  $a_{it}$ , which is not serially correlated. The model after the quasi-first difference is:

$$y_{it} = \beta_1 y_{it-1} + \beta_2 \ell_{it} + \beta_3 \ell_{it-1} + \beta_4 k_{it} + \beta_5 k_{it-1} + \eta_i^* + \gamma_t^* + a_{it}$$

and in first differences,

$$\Delta y_{it} = \beta_1 \Delta y_{it-1} + \beta_2 \Delta \ell_{it} + \beta_3 \Delta \ell_{it-1} + \beta_4 \Delta k_{it} + \beta_5 \Delta k_{it-1} + \Delta \gamma_t^* + \Delta a_{it}$$

with  $\beta_1 = \rho$ ,  $\beta_2 = \alpha_L$ ,  $\beta_3 = -\rho\alpha_L$ ,  $\beta_4 = \alpha_K$ , and  $\beta_5 = -\rho\alpha_K$ . The system GMM estimator of this model consists of using Arellano-Bond instruments in the equation in first differences, and the Blundell-Bond instruments in the equation in levels. For the validity of these instruments (either AB or BB) we need the shock  $a_{it}$  not being serially correlated.

**CODE.** To implement this estimator, we can use the command `xtabond`. The only difference with respect to the code in Question Q2.5 is that we should not include the option `noleveleq`. Note that the syntax of this command is such that, though we input the list of instruments in levels, the command understands that this instruments should be in levels for the equation in first-differences, and the instruments should be in first differences for the equation in levels. We also include a test of CRS, and tests of the restrictions implied by the AR(1) process.

```
xtabond2 logy l.logy logn l.logn logk l.logk i.year, gmm(logy logn logk, lag(2
.)) iv(i.year) robust
test logn + logk = 1
testnl (_b[l.logy] = -_b[l.logn]/_b[logn])
testnl (_b[l.logy] = -_b[l.logk]/_b[logk])
testnl (_b[l.logy] = -_b[l.logn]/_b[logn]) (_b[l.logy] = -_b[l.logk]/_b[logk])
```

**COMMENTS ON RESULTS**

(a) Now, according to the estimates, the technology is more intensive in capital than labor. Both parameters are large. The value of  $\alpha_L + \alpha_K$  is very close to 1 and the null hypothesis of CRS has a p-value of 0.8456.

(b) The estimate of  $\rho$  (or  $\beta_1$ ) is quite large (0.7035) significantly different to zero. There is evidence of strong serial correlation in the transitory shock.

(c) The test of serial correlation "Arellano-Bond test for AR(2) in first differences" tests for the null hypothesis  $E(\Delta a_{it} \Delta a_{it-2}) = 0$ . The p-value of this test is 0.461. The null hypothesis cannot be rejected using the standard significance levels (i.e., p-value is way greater than 10%). This implies that we cannot reject that  $a_{it}$  is not serially correlated, and therefore that AB and BB instruments are valid. That is, it seems that by including

the lagged values  $y_{it-1}$ ,  $\ell_{it-1}$ , and  $k_{it-1}$  as regressors we have been able to control for serial correlation in the transitory shock.

(d) The model implies two restrictions on the parameter estimates  $\beta$ :

$$-\beta_3/\beta_2 = \beta_1 \quad \text{and} \quad -\beta_5/\beta_4 = \beta_1$$

We can test these nonlinear restrictions separately or jointly using the command "testnl" in Stata. We have implemented both the single and the joint tests. The test of  $H_0 : -\beta_3/\beta_2 = \beta_1$  has a p-value of 0.1051, such that we cannot reject the null at 10%. The test of  $H_0 : -\beta_5/\beta_4 = \beta_1$  has a p-value of 0.0047, such that we marginally reject it at 1% significance level. The joint test has a p-value of 0.0169, which implies that we cannot reject the null at 1% but we reject it at 5% significance level.

In summary, accounting for AR(1) transitory shocks has improved significantly the estimates and the specification tests. The restrictions of the model are marginally "accepted" at 1% significance level. Most importantly, in contrast to the AB -with-AR(1), the estimate for the coefficient of capital is now plausibly, and the null hypothesis of CRS cannot be rejected.

```

. // -----
. // 9. Question 2.7: System GMM: AR(1)
. // -----
. xtabond2 logy l.logy logn l.logn logk l.logk i.year, gmm(logy logn logk, lag(
> 2 .)) iv(i.year) robust
Favoring space over speed. To switch, type or click on mata: mata set matafavor
> speed, perm.
Warning: Two-step estimated covariance matrix of moments is singular.
Using a generalized inverse to calculate robust weighting matrix for Hansen t
> est.
Difference-in-Sargan/Hansen statistics may be negative.

```

Dynamic panel-data estimation, one-step system GMM

Group variable: id	Number of obs	=	3563
Time variable : year	Number of groups	=	509
Number of instruments = 88	Obs per group: min	=	7
Wald chi2(13) = 1.73e+06	avg	=	7.00
Prob > chi2 = 0.000	max	=	7

	logy	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
	logy						
	L1.	.703552	.0409546	17.18	0.000	.6232825	.7838215
	logn						
	--.	.4484997	.0929987	4.82	0.000	.2662257	.6307738
	L1.	-.238936	.0932566	-2.56	0.010	-.4217155	-.0561565
	logk						
	--.	.5333291	.1094111	4.87	0.000	.3188874	.7477709
	L1.	-.4439966	.0953875	-4.65	0.000	-.6309527	-.2570404
	year						
	1982	0	(empty)				
	1983	.9693206	.1561941	6.21	0.000	.6631857	1.275455
	1984	1.017554	.1556511	6.54	0.000	.7124836	1.322625
	1985	.9759269	.1577675	6.19	0.000	.6667082	1.285146
	1986	1.003888	.1582892	6.34	0.000	.6936465	1.314129
	1987	1.050145	.1577235	6.66	0.000	.7410126	1.359277
	1988	1.065216	.1602027	6.65	0.000	.7512248	1.379208
	1989	1.016383	.1622965	6.26	0.000	.6982878	1.334478
	_cons	0	(omitted)				

Instruments for first differences equation

Standard

D.(1982b.year 1983.year 1984.year 1985.year 1986.year 1987.year 1988.year 1989.year)

GMM-type (missing=0, separate instruments for each period unless collapsed)

L(2/7).(logy logn logk)

Instruments for levels equation

Standard

1982b.year 1983.year 1984.year 1985.year 1986.year 1987.year 1988.year 1989.year

\_cons

GMM-type (missing=0, separate instruments for each period unless collapsed)

DL.(logy logn logk)

Arellano-Bond test for AR(1) in first differences:	z = -9.69	Pr > z = 0.000
Arellano-Bond test for AR(2) in first differences:	z = -0.74	Pr > z = 0.461

Sargan test of overid. restrictions: chi2(74) = 302.67 Prob > chi2 = 0.000  
(Not robust, but not weakened by many instruments.)

Hansen test of overid. restrictions: chi2(74) = 135.97 Prob > chi2 = 0.000  
(Robust, but weakened by many instruments.)

Difference-in-Hansen tests of exogeneity of instrument subsets:

GMM instruments for levels

Hansen test excluding group: chi2(56) = 92.25 Prob > chi2 = 0.002

Difference (null H = exogenous): chi2(18) = 43.71 Prob > chi2 = 0.001

iv(1982b.year 1983.year 1984.year 1985.year 1986.year 1987.year 1988.year 1989.year)

> 9.year)

Hansen test excluding group: chi2(68) = 124.89 Prob > chi2 = 0.000

Difference (null H = exogenous): chi2(6) = 11.08 Prob > chi2 = 0.086

```

. test logn + logk = 1

(1)  logn + logk = 1

      chi2( 1) =      0.04
    Prob > chi2 =      0.8456

. testnl (_b[l.logy] = -_b[l.logn]/_b[logn])

(1)  _b[l.logy] = -_b[l.logn]/_b[logn]

      chi2(1) =      2.63
    Prob > chi2 =      0.1051

. testnl (_b[l.logy] = -_b[l.logk]/_b[logk])

(1)  _b[l.logy] = -_b[l.logk]/_b[logk]

      chi2(1) =      7.98
    Prob > chi2 =      0.0047

. testnl (_b[l.logy] = -_b[l.logn]/_b[logn]) (_b[l.logy] = -_b[l.logk]/_b[logk]
> )

(1)  _b[l.logy] = -_b[l.logn]/_b[logn]
(2)  _b[l.logy] = -_b[l.logk]/_b[logk]

      chi2(2) =      8.17
    Prob > chi2 =      0.0169

```

**Q3.8. (10 points)** Based on the previous results, select your preferred estimates of the production function. Explain your choice.

ANSWER. Taking into account the following criteria: (i) plausible estimates of the parameters (not too small, and not too far from CRS); (ii) validity of the instruments, i.e., accepting (not rejecting) the null hypothesis that the shock in the regression is not serially correlated; and (iii) accepting (not rejecting) the restrictions on the parameters  $\beta$  imposed by the model; the best estimator, as argued above, is the System-GMM with AR(1) transitory shock.

\*\*\* NOTE. Here it is not so important which estimator you choose, but the arguments you use to select that estimator.

**Q3.9. (20 points)** For this question, your favorite estimates according to your answer to Q3.8, and log-TFPs for year 1989. (a) Obtain the median, the percentile 5, and the percentile 95 in the distribution of log-TFP. Suppose that all the firms operate in the same input markets and  $W_L = W_K = 1$ . (b) Present a figure with three marginal cost functions (i.e., output  $Y$  in the horizontal axis and marginal cost  $MC$  in the vertical axis) for the firms with median, percentile 5, and percentile 95 TFPs, respectively. (c) Comment the results.

ANSWER. As argued in my answer to Question 3.8, I use the estimates from the estimator System-GMM with AR(1) transitory shock. Of course, the particular estimator you choose is not important for a correct answer to this question.

Remember, from Question 2.1, that the cost function from a Cobb-Douglas with labor and capital is:

$$C(Y) = \alpha \left( \frac{Y}{A} \right)^{\frac{1}{\alpha}} \left( \frac{W_L}{\alpha_L} \right)^{\frac{\alpha_L}{\alpha}} \left( \frac{W_K}{\alpha_K} \right)^{\frac{\alpha_K}{\alpha}}$$

Therefore, the marginal cost function is:

$$MC(Y) = Y^{\left( \frac{1}{\alpha} - 1 \right)} \left( \frac{1}{A} \right)^{\frac{1}{\alpha}} \left( \frac{W_L}{\alpha_L} \right)^{\frac{\alpha_L}{\alpha}} \left( \frac{W_K}{\alpha_K} \right)^{\frac{\alpha_K}{\alpha}}$$

Also, be careful with using TFP and not logTFP in the expression for the Marginal Cost function. That is, once you get the percentiles for logTFP, you need to use the exponential of these TFPs,  $\exp(\log TFP)$ , in the MC function.

**CODE.** In my case, the System-GMM with AR(1) transitory shock is the last estimator in my code. Therefore, after that estimation, the values `_b[logn]` and `_b[logk]` correspond to the parameter estimates from that estimator. If your favorite estimator is not the last one implemented in your code, you will need to repeat the implementation of your favorite estimator in your code, such that you use the values `_b[logn]` and `_b[logk]` that you want to use.

Note also that, after using the command `summarize`, the outcomes `r(p5)`, `r(p50)`, and `r(p95)` contain percentiles 5, 10, 95 of the variable, respectively.

The following code calculates logTFPs, the percentiles 5, 50, and 95 of this variable in year 1989, and then the marginal cost for the firms with these TFPs and for every possible value of output (sales) in the sample.

```
gen logTFP = logy - _b[logn] * logn - _b[logk] * logk
sum logTFP if year==1989, detail
gen TFPpc05 = exp(r(p5))
gen TFPpc50 = exp(r(p50))
gen TFPpc95 = exp(r(p95))

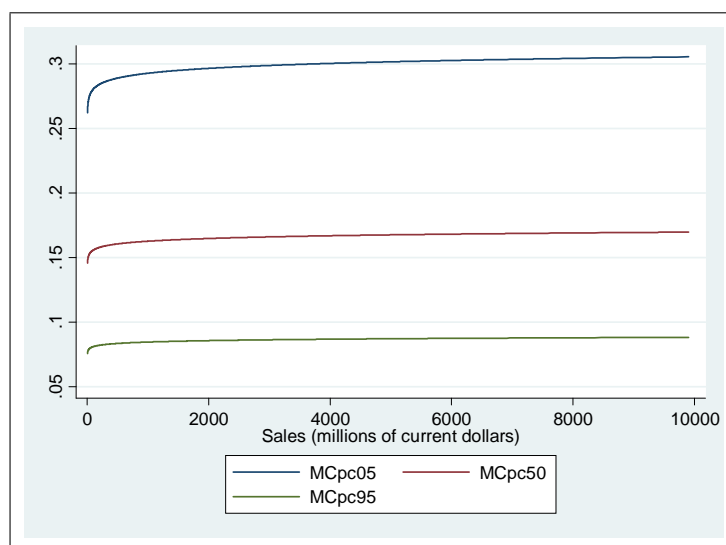
gen common = (sales^(1/(_b[logn]+_b[logk])-1))
* ((1/_b[logn])^(_b[logn]/(_b[logn]+_b[logk])))
* ((1/_b[logk])^(_b[logk]/(_b[logn]+_b[logk])))

gen MCpc05 = (1/TFPpc05)^(1/(_b[logn]+_b[logk])) * common
gen MCpc50 = (1/TFPpc50)^(1/(_b[logn]+_b[logk])) * common
gen MCpc95 = (1/TFPpc95)^(1/(_b[logn]+_b[logk])) * common
```

The following code presents in the same figure the marginal cost functions of the firms with percentiles 5, 50, 95 of TFP.

```
sort sales
line MCpc05 MCpc50 MCpc95 sales if sales<=10000
```

**RESULTS.** The following figure presents the marginal cost functions of the firms with percentiles 5, 50, 95 of TFP. We can see that these marginal costs are quite flat for most levels of output. Only for firm with TFP-5-percentile we can see that the MC increases substantially for levels of output between 0 and 500. The most striking evidence is the large differences in the level of the marginal cost of the three firms.



## DO FILE

```
clear
// -----
// eco310_problem_set_1_question_3_2021.do
// Victor Aguirregabiria
// February 5, 2021
// -----

// -----
// Opening log file to copy results
// -----
capture log close
log using C:\PROBLEM_SETS\problem_set_01_2021\eco310_problem_set_1_question_3_2021.log
replace

// -----
// Reading dataset
// -----
use C:\PROBLEM_SETS\problem_set_01_2021\blundell_bond_2000_production_function.dta

// -----
// Construction of variables
// -----
gen logy = ln(sales)
gen logn = ln(labor)
gen logk = ln(capital)
xtset id year

// -----
// 1. Question 3.1: OLS estimator
// -----
reg logy logn logk i.year
test logn + logk = 1

// -----
// 2. Question 3.2: Fixed Effects estimation
// -----
xtreg logy logn logk i.year, fe
test logn + logk = 1

// -----
// 3. Question 3.3: Fixed Effects - Cochrane-Orcutt estimation
// -----
xtreg logy 1.logy logn 1.logn logk 1.logk i.year, fe
test logn + logk = 1
```

```

testnl (_b[l.logy] = -_b[l.logn]/_b[logn])
testnl (_b[l.logy] = -_b[l.logk]/_b[logk])
testnl (_b[l.logy] = -_b[l.logn]/_b[logn]) (_b[l.logy] = -_b[l.logk]/_b[logk])

// -----
// 4. Question 3.4: Arellano-Bond: No AR(1)
// -----
xtabond2 logy logn logk i.year, gmm(logy logn logk, lag(2 .)) iv(i.year) robust
noleveleq
test logn + logk = 1

// -----
// 5. Question 3.5: Arellano-Bond: AR(1)
// -----
xtabond2 logy l.logy logn l.logn logk l.logk i.year, gmm(logy logn logk, lag(2
.)) iv(i.year) robust noleveleq
test logn + logk = 1
testnl (_b[l.logy] = -_b[l.logn]/_b[logn])
testnl (_b[l.logy] = -_b[l.logk]/_b[logk])
testnl (_b[l.logy] = -_b[l.logn]/_b[logn]) (_b[l.logy] = -_b[l.logk]/_b[logk])

// -----
// 6. Question 3.6: System GMM: No AR(1)
// -----
xtabond2 logy logn logk i.year, gmm(logy logn logk, lag(2 .)) iv(i.year) robust
test logn + logk = 1

// -----
// 7. Question 3.7: System GMM: AR(1)
// -----
xtabond2 logy l.logy logn l.logn logk l.logk i.year, gmm(logy logn logk, lag(2
.)) iv(i.year) robust
test logn + logk = 1
testnl (_b[l.logy] = -_b[l.logn]/_b[logn])
testnl (_b[l.logy] = -_b[l.logk]/_b[logk])
testnl (_b[l.logy] = -_b[l.logn]/_b[logn]) (_b[l.logy] = -_b[l.logk]/_b[logk])

// -----
// 8. Question 3.9: Figure of Marginal costs
// -----
gen logTFP = logy - _b[logn] * logn - _b[logk] * logk
sum logTFP if year==1989, detail

// After using the command summarize, the outcomes r(p5), r(p50), and r(p95)
contain percentiles 5, 10, 95 of the variable, respectively
gen TFPpc05 = exp(r(p5))

```

```

gen TFPpc50 = exp(r(p50))
gen TFPpc95 = exp(r(p95))
gen common = (sales^(1/(_b[logn]+_b[logk])-1))*((1/_b[logn])^(_b[logn]/(_b[logn]+_b[logk])))

gen MCpc05 = (1/TFPpc05)^(1/(_b[logn]+_b[logk])) * common
gen MCpc50 = (1/TFPpc50)^(1/(_b[logn]+_b[logk])) * common
gen MCpc95 = (1/TFPpc95)^(1/(_b[logn]+_b[logk])) * common
sort sales
line MCpc05 MCpc50 MCpc95 sales if sales<=10000

log close

```