

Empirical Industrial Organization (ECO 310)
Fall 2018. Victor Aguirregabiria

Problem Set #1

Due on Monday, October 15th, 2018

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 midnight of Monday, October 15th, 2018.
6. Problem sets should be written individually.

The total number of marks is 130.

QUESTION 1. [50 points]. Consider an industry for an homogeneous product. Firms use capital and labor to produce output according to a Cobb-Douglas technology with parameters α_L and α_K and Total Factor Productivity (TFP) A .

Q1.1. (5 points) Write the expression for this Cobb-Douglas production function (PF).

ANSWER. Let Y , L , and K be the physical amounts of output, labor, and capital, respectively. Then, the expression for the production function is: $Y = A L^{\alpha_L} K^{\alpha_K}$

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. Capital is a fixed input such that the fixed cost for a firm, say i , is $FC_i = W_K K_i$. The variable cost function, $VC(Y)$, is defined as the minimum cost of labor to produce an amount of output Y .

Q1.2. (5 points) Derive the expression for the variable cost function of a firm in this industry. Explain your derivation. [Hint: Given that capital is fixed and there is only one variable input, the minimization problem is trivial. The PF implies that there is only one possible amount of labor that give us a certain amount of output].

ANSWER. The variable cost function $VC(Y)$ is defined as:

$$\begin{aligned} VC(Y) &= \min_L W_L L \\ \text{subject to} & : Y = AL^{\alpha_L} K^{\alpha_K} \end{aligned}$$

However, this constrained minimization problem is trivial because the constrain (the production function) already give us the optimal amount of labor as a function of output, TFP, capital, and parameters. That is, taking capital (and TFP) fixed, there is only one amount of labor that can generate a given amount of output. Solving for L in the equation $Y = AL^{\alpha_L} K^{\alpha_K}$, we have that:

$$L = \left(\frac{Y}{AK^{\alpha_K}} \right)^{1/\alpha_L}$$

Plugging this expression in the variable cost $W_L L$, we get the cost function:

$$VC(Y) = W_L \left(\frac{Y}{AK^{\alpha_K}} \right)^{1/\alpha_L}$$

NOTES ON GRADING: Some of you may have considered the minimization of the Lagrangian:

$$\min_{L, \lambda} W_L L - \lambda [Y - AL^{\alpha_L} K^{\alpha_K}]$$

with the first order conditions (f.o.c.) of optimality:

$$\text{f.o.c. with respect to } L: W_L - \lambda \alpha_L Y / L = 0$$

$$\text{f.o.c. with respect to } \lambda: Y - AL^{\alpha_L} K^{\alpha_K} = 0$$

This is perfectly correct, and you should obtain the same result as above. Note that the f.o.c. with respect to λ is the production function, and this equation give us the optimal amount of labor as a function of output, capital, TFP, and the parameters. In this case, with only one input, the other f.o.c. (with respect to labor) is only needed to obtain the Lagrange multiplier λ .

If you have obtained the variable cost function by minimizing the cost with respect to capital and labor, then your answer is incorrect because we are asking for the variable cost function when capital is fixed, not for the total cost function when both labor and capital are chosen optimally.

Q1.3. (5 points) Using the expressions for the fixed cost and for the variable cost function in Q1.2:

(a) Explain how an increase in the amount of capital affects the fixed cost and the variable cost of a firm.

(b) Explain how an increase in TFP affects the fixed cost and the variable cost.

ANSWER. Fixed cost: $FC = W_K K$. Variable cost: $VC(Y) = W_L \left(\frac{Y}{AK^{\alpha_K}} \right)^{1/\alpha_L}$.

(a) An increase in K implies an increase in the fixed cost and a reduction in the variable cost. There is a trade-off in the choice of capital. This trade-off implies that there is an amount of capital that minimizes the total cost. But in this exercise we are assuming that, in the short run, capital is fixed and it can be different to the amount of capital that minimizes total cost.

(b) An increase in A does not have any effect on the fixed cost, but it reduces the variable cost.

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 $\alpha_L = \alpha_K = 1/2$, and the value of input prices are $W_L = 1/2$ and $W_K = 2$. Remember that firms' capital stocks are fixed (exogenous), and for simplicity suppose that all the firms have the same capital stock $K = 1$.

Q1.4. (5 points) Using these primitives, write the expression for the profit function of a firm (revenue, minus variable cost, minus fixed cost) as a function of the market price, P , the firm's output, Y_i , and its TFP, A_i .

ANSWER. Revenue is equal to $P Y_i$. The fixed cost is $FC = W_K K_i = 2$. The variable cost is: $VC_i(Y_i) = W_L \left(\frac{Y_i}{A_i K_i^{\alpha_K}} \right)^{1/\alpha_L} = \frac{1}{2} \left(\frac{Y_i}{A_i} \right)^2$. Therefore, the profit function is:

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

Q1.5. (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 $MC_i = \frac{Y_i}{A_i^2}$. Therefore, the marginal condition of profit maximization is $P = \frac{Y_i}{A_i^2}$. Solving for output, we get:

$$Y_i = P A_i^2$$

Q1.6. (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:

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

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 $P A_i \geq 2$, and 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; (ii) total output is equal to the sum of firms' individual outputs; (iii) firm i is active ($Y_i^* > 0$) if and only if its total profit is greater than zero; and (iv) for firms with $Y_i^* > 0$, the optimal amount of output is given by the condition price is equal to marginal cost.

Q1.7. (5 points) Write conditions (i) to (iv) for this particular industry.

ANSWER. The mathematical representation of conditions (i) to (iv) is:

- (i) $P^* = 100 - Q^*$.
- (ii) $Q^* = Y_1^* + Y_2^* + \dots + Y_N^*$.
- (iii) $Y_i^* > 0$ if and only if $A_i \geq 2/P^*$.
- (iv) If $A_i \geq 2/P^*$, then $Y_i^* = P^* A_i^2$

Q1.8. (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, $A_1 > A_2 > A_3 > \dots$

Q1.9. (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 $i \geq 3$.

Q1.10. (5 points) Explain why the most efficient firm, with the largest TFP, does not produce all the output of the industry.

ANSWER. Since the variable cost function is convex, we have that the profit function is strictly concave. This implies that the most efficient firm has a finite amount of output that is willing to produce: $Y_1 = P A_1 = P 7$. If this firm were a monopolist 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 2. [80 points].

The Stata 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.

Q2.1. (10 points) OLS with time dummies. Test the null hypothesis $\alpha_L + \alpha_K = 1$. Provide the code in Stata and the table of estimation results. 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 γ_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
```


Q2.2. (10 points) Fixed Effects estimator with time dummies. Test the null hypothesis of no time-invariant unobserved heterogeneity: $\eta_i = 0$ for every firm i . Provide the code in Stata and the table of estimation results. 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 γ_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 η_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 η_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      Number of obs   =      4,072
Group variable: id                    Number of groups =       509

R-sq:                                  Obs per group:
    within = 0.7379                    min =           8
    between = 0.9706                   avg =          8.0
    overall = 0.9661                   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

```

Q2.3. (10 points) Fixed Effects - Cochrane Orcutt estimator with time dummies. Test the two over-identifying restrictions of the model. Provide the code in Stata and the table of estimation results. 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 ρ 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 β :

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

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])
```



```

. 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 α_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 ρ (or β_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 η_i and to the component ρu_{it-1} in the error term (we are controlling for this by including the lagged values y_{it-1} , ℓ_{it-1} , and k_{it-1} as regressors). But it does NOT control for the endogeneity problem due to component a_{it} of the error term.

Q2.4. (10 points) Arellano-Bond estimator with time dummies and non-serially correlated transitory shock. Provide the code in Stata and the table of estimation results. 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 ℓ_{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 Δu_{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 ℓ_{it-2} and k_{it-2} are not valid.

In summary, the estimates do nor 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 198
> 9.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

```


Q2.5. (10 points) Arellano-Bond estimator with time dummies and AR(1) transitory shock. Provide the code in Stata and the table of estimation results. 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 an instrumental variables (GMM) estimator using as instruments ℓ_{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 ρ (or β_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 ℓ_{it-2} and k_{it-2} are valid. That is, it seems that by including the lagged values y_{it-1} , ℓ_{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 β :

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

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(
v 2 .)) iv(i.year) robust nolevelq
Favoring space over speed. To switch, type or click on mata: mata_set matafavor
v speed, perm.
Warning: Two-step estimated covariance matrix of moments is singular.
Using a generalized inverse to calculate robust weighting matrix for Hansen t
v 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
v 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]
v )
(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

```

Q2.6. (10 points) Blundell-Bond system estimator with time dummies and non-serially correlated transitory shock. Provide the code in Stata and the table of estimation results. 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 Δk_{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 α_L and α_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

```

Group variable: id                Number of obs   =    4072
Time variable : year              Number of groups =    509
Number of instruments = 89        Obs per group: min =     8
Wald chi2(10) = 2754.89          avg =    8.00
Prob > chi2 = 0.000              max =     8

```

	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

```

Q2.7. (10 points) Blundell-Bond system estimator with time dummies and AR(1) transitory shock. Provide the code in Stata and the table of estimation results. 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 1.logy logn 1.logn logk 1.logk i.year, gmm(logy logn logk, lag(2
.)) iv(i.year) robust
test logn + logk = 1
testnl (_b[1.logy] = -_b[1.logn]/_b[logn])
testnl (_b[1.logy] = -_b[1.logk]/_b[logk])
testnl (_b[1.logy] = -_b[1.logn]/_b[logn]) (_b[1.logy] = -_b[1.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 ρ (or β_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} , ℓ_{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 β :

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

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 198
> 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

```

Q2.8. (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 β imposed by the model; the best estimator, as argued above, is the System-GMM with AR(1) transitory shock.

DO FILE

```
clear
// -----
// eco310_problem_set_1_question_2_2018.do
// Victor Aguirregabiria
// October 1st, 2018
// -----
// -----
// 1. Reading dataset
// -----
use c:\PROBLEM_SETS\blundell_bond_2000_production_function.dta
// -----
// 2. Construction of variables
// -----
gen logy = ln(sales)
gen logn = ln(labor)
gen logk = ln(capital)
xtset id year
// -----
// 3. Question 2.1: OLS estimator
// -----
reg logy logn logk i.year
test logn + logk = 1
// -----
// 4. Question 2.2: Fixed Effects estimation
// -----
xtreg logy logn logk i.year, fe
test logn + logk = 1
// -----
// 5. Question 2.3: Fixed Effects - Cochrane-Orcutt estimation
// -----
xtreg logy l.logy logn l.logn logk l.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])
// -----
// 6. Question 2.4: 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
// -----
// 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 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])
// -----
// 8. Question 2.6: 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
// -----
// 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
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])

```