Empirical Industrial Organization (ECO 310)
Fall 2019. Victor Aguirregabiria

SOLUTION TO PROBLEM SET #1

Due on Friday, October 11, 2019

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, and the output table with the estimation 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 Friday, October 11, 2019.**
6. **Problem sets should be written individually.**

The total number of marks is 150.

For this problem set, you need to use the Stata datafile

eco310_ps1_production_function_2019.dta. This datafile contains a panel data set of 575 manufacturing firms in Spain during 9 years: $1982 - 1990$**. It includes the following variables:**

Name	Description	Notation
year	**Year**	**t**
id	**Firm identification number**	**i**
y	**Logarithm of real annual output**	$\log(Y)$
lt	**Logarithm of annual total employment**	$\log\left(L^P + L^T\right)$
lp	**Logarithm of annual "permanent" employment**	$\log\left(L^P\right)$
k	**Logarithm of real capital stock at the beginning of year**	$\log(K)$
ik	**Investment rate (annual investment / capital beginning year)**	**I/K**
rt	**Ratio of temporary employment over permanent employment**	$\mathbf{L}^T/\mathbf{L}^P$
wage	**Ratio of annual wage bill over total employment**	**W**

We consider an extension of the standard Cobb-Douglas production function. This extended version takes into account that: (1) temporary and permanent workers can have different productivity; and (2) that new investments during the year can have different productivity that the capital stock at the beginning of the year. That is:

$$Y_{it} = A_{it}\ \left(L_{it}^P + \lambda_L\ L_{it}^T\right)^{\alpha_L}\ \ (K_{it} + \lambda_K\ I_{it})^{\alpha_K} \tag{1}$$

where L^P and L^T represent permanent and temporary workers, respectively; λ_L is a parameter that measures the productivity of temporary workers relative to permanent workers; K is capital at the beginning of the year; I represents investment during the year; λ_K is a parameter that measures the productivity of new investments relative to old capital; and A, α_L, and α_K have the usual interpretation. We can take the logarithm transformation of this production function and – using standard approximations – obtain the following linear regression model:

$$\log(Y_{it}) \ = \ \alpha_L\ \log(L_{it}^P) + \alpha_L\lambda_L\ \left(\frac{L_{it}^T}{L_{it}^P}\right) + \alpha_K\ \log(K_{it}) + \alpha_K\lambda_K\ \left(\frac{I_{it}}{K_{it}}\right) + \omega_{it} \tag{2}$$

QUESTION 1. (25 points) For the following questions, provide the code in Stata and the table of estimation results.

(a) (10 points) Estimate the parameters of this production function using OLS with time dummies. Present also the estimates of the parameters λ_L and λ_K.

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

$$y \ = \ \beta_0 + \beta_{lp}\ lp + \beta_{rt}\ rt + \beta_k\ k + \beta_{ik}\ ik + time\_dummies + u \tag{3}$$

We account for the time effects γ_t by including time (year) dummies: one for each year, except one.

The parameters of this regression model have the following relationship with the parameters of our model: $\beta_{lp} = \alpha_L$; $\beta_{rt} = \alpha_L \lambda_L$; $\beta_k = \alpha_K$; and $\beta_{ik} = \alpha_K \lambda_K$. Therefore, given the estimated β parameters, we can estimate the parameters of the model as follows:

$$\widehat{\alpha}_L = \widehat{\beta}_{lp}; \quad \widehat{\lambda}_L = \widehat{\beta}_{rt}/\widehat{\beta}_{lp}; \quad \widehat{\alpha}_K = \widehat{\beta}_k; \quad \widehat{\lambda}_K = \widehat{\beta}_{ik}/\widehat{\beta}_k$$

CODE

```
// Question 1(a):  OLS Estimation
reg y lp rt k ik i.year

// Estimate of lambda_L
nlcom (lambda_L: _b[rt]/_b[lp])

// Estimate lambda_{K}
nlcom (lambda_K: _b[ik]/_b[k])
```

OUTPUT

```
. // Question 1(a): OLS Estimation
. reg y lp rt k ik i.year
```

Source	SS	df	MS			
				Number of obs	=	4,600
				F(11, 4588)	=	2323.93
Model	7122.81559	11	647.52869	Prob > F	=	0.0000
Residual	1278.37814	4,588	.278635166	R-squared	=	0.8478
				Adj R-squared	=	0.8475
Total	8401.19374	4,599	1.82674358	Root MSE	=	.52786

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lp	.6228523	.0115313	54.01	0.000	.6002453	.6454593
rt	.5838056	.0343929	16.97	0.000	.516379	.6512322
k	.328945	.0084162	39.08	0.000	.3124453	.3454448
ik	1.251736	.0776801	16.11	0.000	1.099446	1.404026
year						
1984	.0133164	.031132	0.43	0.669	-.0477174	.0743501
1985	.0432564	.0311461	1.39	0.165	-.0178049	.1043177
1986	.0716722	.0311597	2.30	0.021	.0105842	.1327602
1987	.1123845	.0312117	3.60	0.000	.0511945	.1735745
1988	.1488675	.0312636	4.76	0.000	.0875757	.2101592
1989	.2107629	.0313051	6.73	0.000	.1493898	.2721361
1990	.1980076	.031303	6.33	0.000	.1366386	.2593766
_cons	-1.160194	.0480973	-24.12	0.000	-1.254488	-1.0659

```
. // Estimate of λ_{L}
. nlcom (lambda_L: _b[rt]/_b[lp])

    lambda_L:  _b[rt]/_b[lp]
```

y	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lambda_L	.9373099	.0545644	17.18	0.000	.8303656	1.044254

```
.
. // Estimate of λ_{K}
. nlcom (lambda_K: _b[ik]/_b[k])

    lambda_K:  _b[ik]/_b[k]
```

y	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lambda_K	3.805305	.2278115	16.70	0.000	3.358802	4.251807

COMMENTS ON RESULTS

• According to the estimates, the technology is more intensive in labor than in capital: $\widehat{\alpha}_L = 0.622 > 0.328 = \widehat{\alpha}_K$.

• The estimate of λ_L is 0.937. This indicates that temporary labor is slightly less productive than permanent labor.

• The estimate of λ_K is 3.805. This indicates that new capital is almost four times more productive than old capital. This seems economically very implausible.

• 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).

(b) (5 points) Test the null hypothesis $\lambda_L = 1$. Is temporary labor significantly less productive than permanent labor? Explain.

ANSWER. Testing the null hypothesis of $\lambda_L = 1$ is equivalent to testing the null hypothesis of $\beta_{lp} - \beta_{rt} = 0$.

CODE

```
// Question 1(b):  Test null hypothesis lambda_{L}=1
test lp - rt = 0
```

OUTPUT

```
. // Question 1(b): Test null hypothesis λ_{L}=1
. test lp - rt = 0

 ( 1)  lp - rt = 0

       F(  1,  4588) =    1.31
            Prob > F =    0.2520
```

COMMENTS ON RESULTS

- The p-value of the test is 0.25. This is larger than standard significance levels (i.e., 1%, 5%, or 10%). Therefore, we can conclude that the estimated parameter λ_L (if consistent) provides significant evidence that temporary labor is as productive as permanent labor.
- However, we expect the OLS estimator to be biased.

(c) (5 points) Test the null hypothesis $\lambda_K = 1$. Is new capital significantly less productive than old capital? Explain.

ANSWER. Testing the null hypothesis of $\lambda_K = 1$ is equivalent to testing the null hypothesis of $\beta_k - \beta_{ik} = 0$.

CODE

```
// Question 1(c):  Test null hypothesis lambda_{K}=1
test k - ik = 0
```

OUTPUT

```
. // Question 1(c): Test null hypothesis λ_{K}=1
. test k - ik = 0

 ( 1)  k - ik = 0

       F(  1,  4588) =  148.73
            Prob > F =    0.0000
```

COMMENTS ON RESULTS

- The p-value of the test is 0.0000. This is clearly smaller than standard significance levels (i.e., 1%, 5%, or 10%). Therefore, we can conclude that the estimated parameter λ_K (if consistent) provides significant evidence that new capital investments are more productive than old capital.
- However, we expect the OLS estimator to be biased.

(d) (5 points) Test the null hypothesis $\alpha_L + \alpha_K = 1$. Is there significant evidence of decreasing returns to scale? Explain.

ANSWER. Testing the null hypothesis of $\alpha_L + \alpha_K = 1$ is equivalent to testing the null hypothesis of $\beta_{lp} + \beta_k - 1 = 0$.

CODE

```
// Question 1(d):
test lp + k - 1 = 0
```

OUTPUT

```
. // Question 1(d): Test null hypothesis α_{L} + α_{K} = 1
. test lp + k - 1 = 0

 ( 1)  lp + k = 1

       F(  1,  4588) =   54.03
            Prob > F =    0.0000
```

COMMENTS ON RESULTS

• The p-value of the test is 0.0000. This is clearly smaller than standard significance levels (i.e., 1%, 5%, or 10%). Therefore, we can conclude that the estimated parameters α_L and α_K (if consistent) provide significant evidence of decreasing returns to scale: $\alpha_L + \alpha_K < 1$.

• However, we expect the OLS estimator to be biased.

QUESTION 2. (25 points) For the following questions, provide the code in Stata and the table of estimation results.

(a) (10 points) Estimate the parameters of this production function using Fixed Effects estimator with time dummies. Present also the estimates of the parameters λ_L and λ_K.

ANSWER. The Fixed Effects estimator is the OLS estimator in the following transformed model:

$$\widetilde{y}_{it} \quad = \quad \beta_{lp}\, \widetilde{lp}_{it} + \beta_{rt}\, \widetilde{rt}_{it} + \beta_k\, \widetilde{k}_{it} + \beta_{ik}\, \widetilde{ik}_{it} + time\_dummies + \widetilde{u}_{it}$$

where: $\widetilde{y}_{it} = y_{it} - \overline{y}_i$; $\widetilde{lp}_{it} = lp_{it} - \overline{lp}_i$; $\widetilde{rt}_{it} = rt_{it} - \overline{rt}_i$; $\widetilde{k}_{it} = k_{it} - \overline{k}_i$; $\widetilde{ik}_{it} = ik_{it} - \overline{ik}_i$; and $\widetilde{u}_{it} = u_{it} - \overline{u}_i$; and the variables $\overline{y}_i$, $\overline{lp}_i$, $\overline{rt}_i$, $\overline{k}_i$, and $\overline{ik}_i$ are the sample means of the original variables for firm i. 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.

```
// Question 2(a):  FE Estimation
xtreg y lp rt k ik i.year, fe

// Estimate of lambda_{L}
nlcom (lambda_L: _b[rt]/_b[lp])

// Estimate of lambda_{K}
nlcom (lambda_K: _b[ik]/_b[k])
```

OUTPUT

```
. // Question 2(a): FE Estimation
. xtreg y lp rt k ik i.year, fe

Fixed-effects (within) regression               Number of obs     =      4,600
Group variable: id                              Number of groups  =        575

R-sq:                                           Obs per group:
     within  = 0.4545                                         min =          8
     between = 0.8532                                         avg =        8.0
     overall = 0.8407                                         max =          8

                                                F(11,4014)        =     304.08
corr(u_i, Xb)  = 0.1152                         Prob > F          =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          lp |   .6869962    .026036    26.39   0.000     .6359512    .7380412
          rt |   .4668872   .0278912    16.74   0.000     .4122049    .5215694
           k |   .2356506   .0245897     9.58   0.000     .1874411    .2838602
          ik |   .2984293   .0335397     8.90   0.000      .232673    .3641857
             |
        year |
       1984  |   .0146511   .0111352     1.32   0.188      -.00718    .0364823
       1985  |   .0568854   .0111652     5.09   0.000     .0349953    .0787754
       1986  |   .0913719   .0113339     8.06   0.000     .0691511    .1135926
       1987  |   .1462661   .0115336    12.68   0.000     .1236537    .1688784
       1988  |   .1911345   .0117005    16.34   0.000     .1681951    .2140739
       1989  |    .257451   .0118572    21.71   0.000     .2342044    .2806977
       1990  |   .2419957   .0122546    19.75   0.000     .2179698    .2660215
             |
       _cons |  -1.193656   .1216535    -9.81   0.000    -1.432164   -.9551476
-------------+----------------------------------------------------------------
     sigma_u |   .5136066
     sigma_e |  .18874104
         rho |  .88102403   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(574, 4014) = 55.53                  Prob > F = 0.0000
```

```
. // Estimate of λ_{L}
. nlcom (lambda_L: _b[rt]/_b[lp])

    lambda_L:  _b[rt]/_b[lp]
```

y	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lambda_L	.6796066	.0444021	15.31	0.000	.5925801	.7666331

```
.
. // Estimate of λ_{K}
. nlcom (lambda_K: _b[ik]/_b[k])

    lambda_K:  _b[ik]/_b[k]
```

y	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lambda_K	1.266406	.1681618	7.53	0.000	.9368147	1.595997

COMMENTS ON RESULTS

• According to the estimates, the technology is slightly more intensive in labor than in capital: $\widehat{\alpha}_L = 0.686 > 0.235 = \widehat{\alpha}_K$.

• The estimate of λ_L is 0.679. This indicates that temporary labor is only 68% as productive than permanent labor.

• The estimate of λ_K is 1.26. This indicates that new capital is 26% more productive than old capital.

• However, we expect the Fixed Effects estimator can be biased because endogeneity, i.e., correlation between the regressors (observed inputs) and the transitory shock u_{it}.

(b) (10 points) Implement the same tests as in Questions 1(b), 1(c), and 1(d), and answer the same questions. Explain the results.

ANSWER. Testing the null hypothesis of $\lambda_L = 1$ is equivalent to testing the null hypothesis of $\beta_{lp} - \beta_{rt} = 0$. Testing the null hypothesis of $\lambda_K = 1$ is equivalent to testing the null hypothesis of $\beta_k - \beta_{ik} = 0$. Testing the null hypothesis of $\alpha_L + \alpha_K = 1$ is equivalent to testing the null hypothesis of $\beta_{lp} + \beta_k - 1 = 0$.

CODE

```
// Question 2(b):  Test null hypothesis lambda_{L}=1
test lp - rt = 0

// Question 2(b):  Test null hypothesis lambda_{K}=1
test k - ik = 0

// Question 2(b):  Test null hypothesis
test lp + k = 1
```

OUTPUT

```
. // Question 2(b): Test null hypothesis λ_{L}=1
. test lp - rt = 0

 ( 1)  lp - rt = 0

       F(  1,  4014) =   39.72
            Prob > F =    0.0000

.
. // Question 2(b): Test null hypothesis λ_{K}=1
. test k - ik = 0

 ( 1)  k - ik = 0

       F(  1,  4014) =    3.00
            Prob > F =    0.0836

.
. // Question 2(b): Test null hypothesis α_{L} + α_{K} = 1
. test lp + k = 1

 ( 1)  lp + k = 1

       F(  1,  4014) =    7.49
            Prob > F =    0.0062
```

COMMENTS ON RESULTS

Test null hypothesis of $\lambda_L = 1$.

- The p-value of the test is 0.000. This is clearly smaller than standard significance levels (i.e., 1%, 5%, or 10%). Therefore, we can conclude that the estimated parameter λ_L (if consistent) provides significant evidence that temporary labor is less productive than permanent labor.
- However, we expect the FE estimator may be biased.

Test null hypothesis of $\lambda_K = 1$.

- The p-value of the test is 0.0836. This p-value is greater than 5% but smaller than 10%. Therefore, at 5% significance level we cannot reject the null hypothesis that new capital investments are as productive as old capital.
- However, we expect the FE estimator may be biased.

Test null hypothesis of $\alpha_L + \alpha_K = 1$.

- The p-value of the test is 0.0062. This is smaller than standard significance levels (i.e., 1%, 5%, or 10%). Therefore, we can conclude that the estimated parameters α_L and α_K (if consistent) provide significant evidence of decreasing returns to scale: $\alpha_L + \alpha_K < 1$.
- However, we expect the FE estimator may be biased.

(c) (5 points) Test the null hypothesis that all the individual fixed effects η_i are the same. Interpret the results.

ANSWER.

- The test of the null hypothesis of no time-invariant unobserved heterogeneity (η_i are the same 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.
- 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 = 0.8810" that represents the estimate for $Var(\eta_i)/Var(\eta_i + u_{it})$.

CODE

```
// Question 2(c):
// This is the test provided at the bottom of the table
// F test that all u_i=0:  F(574, 4014) = 55.53 Prob > F = 0.0000
```

QUESTION 3. (30 points) For the following questions, provide the code in Stata and the table of estimation results.

(a) (5 points) Using the quasi-first difference transformation, and the Within-groups transformation, obtain the equation of the linear regression model that we estimate to implement the Fixed Effects + Cochrane-Orcutt estimator. The model has 9 parameters, but it imposes 4 restrictions on these parameters. Write the equations for these 4 restrictions on the parameters.

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:

$$\begin{aligned} y_{it} &= \beta_{y\_1}\ y_{it-1} + \beta_{lp}\ lp_{it} + \beta_{lp\_1}\ lp_{it-1} + \beta_{rt}\ rt_{it} + \beta_{rt\_1}\ rt_{it-1} \\ &+ \beta_{k}\ k_{it} + \beta_{k\_1}\ k_{it-1} + \beta_{ik}\ ik_{it} + \beta_{ik\_1}\ ik_{it-1} + \eta_i^* + \gamma_t^* + a_{it} \end{aligned}$$

where: $\beta_{y\_1} = \rho$; $\beta_{lp} = \alpha_L$; $\beta_{lp\_1} = -\rho\alpha_L$; $\beta_{rt} = \alpha_L\lambda_L$; $\beta_{rt\_1} = -\rho\alpha_L\lambda_L$; $\beta_k = \alpha_K$; $\beta_{k\_1} = -\rho\alpha_K$; $\beta_{ik} = \alpha_K\lambda_K$; and $\beta_{ik\_1} = -\rho\alpha_K\lambda_K$. The FE Cochrane-Orcutt estimator is the FE estimator in this equation.

The model implies four restrictions on the parameter estimates β:

$$\begin{aligned} \beta_{y\_1} &= -\beta_{lp\_1}/\beta_{lp} \\ \beta_{y\_1} &= -\beta_{rt\_1}/\beta_{rt} \\ \beta_{y\_1} &= -\beta_{k\_1}/\beta_{k} \\ \beta_{y\_1} &= -\beta_{ik\_1}/\beta_{ik} \end{aligned}$$

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

(b) (10 points) Estimate the parameters of this production function using Fixed Effects - Cochrane Orcutt estimator with time dummies. Present also the estimates of the parameters λ_L and λ_K.

ANSWER. Now, there are two possible ways to estimate the parameter λ_L. We have that $\lambda_L = \beta_{rt}/\beta_{lp}$, and we also have that $\lambda_L = \beta_{rt\_1}/\beta_{lp\_1}$. We report both estimates. Similarly, there are two possible ways to estimate the parameter λ_K. We have that $\lambda_K = \beta_{ik}/\beta_k$, and we also have that $\lambda_K = \beta_{ik\_1}/\beta_{k\_1}$.

CODE

```
// Question 3(b):  FE + CO Estimation
xtreg y l.y lp l.lp rt l.rt k l.k ik l.ik i.year, fe
```

```
// There are two possible estimates of lambda_{L}
// Estimate 1:
nlcom (lambda_L1:  _b[rt]/_b[lp])

// Estimate 2:
nlcom (lambda_L2:  _b[l.rt]/_b[l.lp])

// There two possible estimates of lambda_{K}
// Estimate 1:
nlcom (lambda_K1:  _b[ik]/_b[k])

// Estimate 2:
nlcom (lambda_K2:  _b[l.ik]/_b[l.k])
```

OUTPUT

```
. // Question 3(b): FE + CO Estimation
. xtreg y l.y lp l.lp rt l.rt k l.k ik l.ik i.year, fe

Fixed-effects (within) regression               Number of obs     =      4,025
Group variable: id                              Number of groups  =        575

R-sq:                                           Obs per group:
     within  = 0.6072                                         min =          7
     between = 0.9804                                         avg =        7.0
     overall = 0.9694                                         max =          7

                                                F(15,3435)        =     353.94
corr(u_i, Xb)  = 0.7391                         Prob > F          =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           y |
         L1. |   .5649997   .0149714    37.74   0.000      .535646    .5943534
             |
          lp |
         --. |   .4870191   .0347136    14.03   0.000     .4189576    .5550805
         L1. |  -.2128538   .0362182    -5.88   0.000    -.2838651   -.1418425
             |
          rt |
         --. |   .3307219   .0293135    11.28   0.000     .2732483    .3881955
         L1. |  -.0834903    .029104    -2.87   0.004    -.1405532   -.0264274
             |
           k |
         --. |   .2136776   .1150788     1.86   0.063    -.0119521    .4393074
         L1. |   -.177349   .1166075    -1.52   0.128     -.405976     .051278
             |
          ik |
         --. |   .1761647   .0296279     5.95   0.000     .1180747    .2342547
         L1. |  -.0792125   .1028361    -0.77   0.441    -.2808386    .1224137
             |
        year |
        1985 |   .0349842   .0090857     3.85   0.000     .0171703    .0527981
        1986 |   .0487159   .0101665     4.79   0.000     .0287829    .0686489
        1987 |   .0888771   .0096815     9.18   0.000      .069895    .1078591
        1988 |   .1046603   .0098431    10.63   0.000     .0853613    .1239593
        1989 |   .1445524   .0102765    14.07   0.000     .1244037    .1647011
        1990 |   .0882059     .01093     8.07   0.000      .066776    .1096359
             |
       _cons |  -.2256015   .1260913    -1.79   0.074    -.4728229      .02162
-------------+----------------------------------------------------------------
     sigma_u |  .28136976
     sigma_e |  .15310189
         rho |  .77155845   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(574, 3435) = 2.30                   Prob > F = 0.0000
```

```
. // There are two possible estimates of λ_{L}
. //    Estimate 1:
. nlcom (lambda_L1: _b[rt]/_b[lp])

   lambda_L1:  _b[rt]/_b[lp]

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   lambda_L1 |   .6790738   .0707176     9.60   0.000     .5404699    .8176777
------------------------------------------------------------------------------

. //    Estimate 2:
. nlcom (lambda_L2: _b[l.rt]/_b[l.lp])

   lambda_L2:  _b[l.rt]/_b[l.lp]

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   lambda_L2 |   .3922425   .1430725     2.74   0.006     .1118255    .6726595
------------------------------------------------------------------------------
```

```
. // There are two possible estimates of λ_{K}
. //    Estimate 1:
. nlcom (lambda_K1: _b[ik]/_b[k])

   lambda_K1:  _b[ik]/_b[k]

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   lambda_K1 |   .8244414   .4611221     1.79   0.074    -.0793412    1.728224
------------------------------------------------------------------------------

. //    Estimate 2:
. nlcom (lambda_K2: _b[l.ik]/_b[l.k])

   lambda_K2:  _b[l.ik]/_b[l.k]

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   lambda_K2 |   .4466474   .3153468     1.42   0.157     -.171421    1.064716
------------------------------------------------------------------------------
```

COMMENTS ON RESULTS

• The estimated model shows significance evidence of serial correlation in the transitory shock u_{it}. The estimate of the autoregressive parameter ρ is 0.5649 (t-ratio = 37.74), which is relatively large and significantly different to zero.

• The technology is more intensive in labor than in capital: $\widehat{\alpha}_L = 0.487 > 0.213 = \widehat{\alpha}_K$.

• The two estimates of λ_L are 0.679 (s.e. = 0.070) and 0.392 (s.e. = 0.143). This indicates that temporary labor is less productive than permanent labor.

• The two estimates of λ_K are 0.824 (s.e. = 0.461) and 0.446 (s.e. = 0.315). This indicates that – during its first year in a firm – new capital investments are less productive than old capital. This corresponds to the hypothesis of *time to build*. New capital investments require some time for their installation and for being fully productive in a firm. This implies that during their first year in the firm, new capital is less productive than old – already installed – capital.

(c) (10 points) Implement the same tests as in Questions 1(b), 1(c), and 1(d), and answer the same questions. Explain the results.

ANSWER. Since we have two estimates of the parameter λ_L, we have two possible tests for the null hypothesis that $\lambda_L = 1$: test of $\beta_{lp} - \beta_{rt} = 0$ and test of $\beta_{lp\_1} - \beta_{rt\_1} = 0$. We test these restrictions both separately and jointly.

Similarly, we have two possible tests for the null hypothesis that $\lambda_K = 1$: test of $\beta_k - \beta_{ik} = 0$ and test of $\beta_{k\_1} - \beta_{ik\_1} = 0$. We test these restrictions both separately and jointly.

The same comment applies to the test of the null hypothesis of $\alpha_L + \alpha_K = 1$. We have that $\alpha_L = \beta_{lp}$ and $\alpha_K = \beta_k$ such that we can implement the test $\beta_{lp} + \beta_k - 1 = 0$. But we also have that $\alpha_L = -\beta_{lp\_1}/\beta_{y\_1}$ and $\alpha_K = -\beta_{k\_1}/\beta_{y\_1}$ such that we can implement the test $-\beta_{lp\_1}/\beta_{y\_1} - \beta_{k\_7}/\beta_{y\_1} - 1 = 0$. Note that this is equivalent to $\beta_{y\_1} + \beta_{lp\_1} + \beta_{k\_1} = 0$. We test these restrictions separately and jointly.

CODE

```
// Question 3(c):  Test null hypothesis lambda_{L}=1
test lp - rt = 0
test l.lp - l.rt = 0
test (lp - rt = 0) (l.lp - l.rt = 0)

// Question 3(c):  Test null hypothesis lambda_{K}=1
test k - ik = 0
test l.k - l.ik = 0
test (k - ik = 0) (l.k - l.ik = 0)
```

```
// Question 3(c):  Test null hypothesis
test lp + k = 1
test l.lp + l.k + l.y = 0
test (lp + k = 1) (l.lp + l.k + l.y = 0)
```

OUTPUT

```
. // Question 3(c): Test null hypothesis lambda_{L}=1
. test lp - rt = 0

 ( 1)  lp - rt = 0

       F(  1,  3435) =   14.14
            Prob > F =    0.0002

. test l.lp - l.rt = 0

 ( 1)  L.lp - L.rt = 0

       F(  1,  3435) =    9.05
            Prob > F =    0.0026

. test (lp - rt = 0) (l.lp - l.rt = 0)

 ( 1)  lp - rt = 0
 ( 2)  L.lp - L.rt = 0

       F(  2,  3435) =    7.46
            Prob > F =    0.0006
```

```
. // Question 3(c): Test null hypothesis lambda_{K}=1
. test k - ik = 0

 ( 1)  k - ik = 0

       F(  1,  3435) =    0.10
            Prob > F =    0.7505

. test l.k - l.ik = 0

 ( 1)  L.k - L.ik = 0

       F(  1,  3435) =    6.76
            Prob > F =    0.0094

. test (k - ik = 0) (l.k - l.ik = 0)

 ( 1)  k - ik = 0
 ( 2)  L.k - L.ik = 0

       F(  2,  3435) =    3.77
            Prob > F =    0.0231
```

```
. // Question 3(c): Test null hypothesis alpha_{L} + alpha_{K} = 1
. test lp + k = 1

 ( 1)  lp + k = 1

       F(  1,  3435) =    6.19
            Prob > F =    0.0129

. test l.lp + l.k + l.y = 0

 ( 1)  L.y + L.lp + L.k = 0

       F(  1,  3435) =    2.07
            Prob > F =    0.1501

. test (lp + k = 1) (l.lp + l.k + l.y = 0)

 ( 1)  lp + k = 1
 ( 2)  L.y + L.lp + L.k = 0

       F(  2,  3435) =   11.67
            Prob > F =    0.0000
```

COMMENTS ON RESULTS

Tests null hypothesis of $\lambda_L = 1$.

- The two individual tests and the joint test have p-values smaller than 1%. Therefore, we can conclude that the estimated parameter λ_L (if consistent) provides significant evidence that temporary labor is less productive than permanent labor.

Tests null hypothesis of $\lambda_K = 1$.

- These tests present significant evidence in favor of the null hypothesis of $\lambda_K = 1$. The first individual test has a p-value of 75%, which represents strong evidence in favor of the null hypothesis of $\lambda_K = 1$. In contrast, the second test has a p-value of only 1%, which is still a standard significance level at which we can accept – or not reject – a null hypothesis. When we test the two restrictions together using the the joint F test, we get a p-value 2.3%. Again this p-value is greater that 1%, such that we cannot reject the null hypothesis that new investments – during their first year of operation – are as productive as old capital.

Test null hypothesis of $\alpha_L + \alpha_K = 1$.

- The evidence from these tests is not conclusive. The individual tests have significance levels of 1.3% and 15%, respectively. Based on each of these tests, we cannot reject the null hypothesis of constant returns to scale. However, when we consider the joint F-test of the two restrictions, the p-value becomes practically equal to zero such that we reject the null hypothesis. How is this possible? Mathematically – or geometrically – the 99% confidence region for the joint test provides narrower intervals than the 99% confidence regions of the two individual tests. Therefore, there is some evidence of decreasing returns to scale – that is, $\alpha_L + \alpha_K < 1$.

(d) (5 points) Test the 4 restrictions on the parameters implied by the Cochrane-Orcutt model. Test each restriction separately, and also the 4 joint restrictions (5 different tests).

ANSWER. As explained in Question 3(a), the Cochrane-Orcutt model implies four restrictions on the β parameters:

$$\begin{aligned} \beta_{y\_1} &= -\beta_{lp\_1}/\beta_{lp} \\ \beta_{y\_1} &= -\beta_{rt\_1}/\beta_{rt} \\ \beta_{y\_1} &= -\beta_{k\_1}/\beta_{k} \\ \beta_{y\_1} &= -\beta_{ik\_1}/\beta_{ik} \end{aligned}$$

We can test these restrictions separately and jointly.

CODE

```
// Question 3(d):  CO Restrictions
```

```
testnl _b[l.y] + _b[l.lp]/_b[lp] = 0
testnl _b[l.y] + _b[l.rt]/_b[rt] = 0
testnl _b[l.y] + _b[l.k]/_b[k] = 0
testnl _b[l.y] + _b[l.ik]/_b[ik] = 0
testnl (_b[l.y] + _b[l.lp]/_b[lp] = 0) (_b[l.y] + _b[l.rt]/_b[rt] = 0) (_b[l.y]
+ _b[l.k]/_b[k] = 0) (_b[l.y] + _b[l.ik]/_b[ik] = 0)
```

OUTPUT

```
. // Question 3(d): CO Restrictions
. testnl _b[l.y] + _b[l.lp]/_b[lp] = 0

  (1)  _b[l.y] + _b[l.lp]/_b[lp] = 0

               chi2(1) =        5.19
           Prob > chi2 =        0.0227


. testnl _b[l.y] + _b[l.rt]/_b[rt] = 0

  (1)  _b[l.y] + _b[l.rt]/_b[rt] = 0

               chi2(1) =       15.36
           Prob > chi2 =        0.0001

. testnl _b[l.y] + _b[l.k]/_b[k]  = 0

  (1)  _b[l.y] + _b[l.k]/_b[k] = 0

               chi2(1) =        3.25
           Prob > chi2 =        0.0715

. testnl _b[l.y] + _b[l.ik]/_b[ik] = 0

  (1)  _b[l.y] + _b[l.ik]/_b[ik] = 0

               chi2(1) =        0.04
           Prob > chi2 =        0.8443

. testnl (_b[l.y] + _b[l.lp]/_b[lp] = 0) (_b[l.y] + _b[l.rt]/_b[rt] = 0) (_b[l.y] + _b[l.k]/_b[k] = 0) (_b[l.y] + _

  (1)  _b[l.y] + _b[l.lp]/_b[lp] = 0
  (2)  _b[l.y] + _b[l.rt]/_b[rt] = 0
  (3)  _b[l.y] + _b[l.k]/_b[k] = 0
  (4)  _b[l.y] + _b[l.ik]/_b[ik] = 0

               chi2(4) =       20.81
           Prob > chi2 =        0.0003
```

COMMENTS ON RESULTS

• The test for the null hypothesis $\beta_{y\_1} = -\beta_{lp\_1}/\beta_{lp}$ – that is, $\rho = -(-\rho\alpha_L)/\alpha_L$ – has a p-value of 2.2%. Therefore, it cannot be rejected with a 1% significance level.

• The test for the null hypothesis $\beta_{y\_1} = -\beta_{rt\_1}/\beta_{rt}$ – that is, $\rho = -(-\rho\alpha_L\lambda_L)/\alpha_L\lambda_L$ – has a p-value of 0.0%. Therefore, this restriction is clearly rejected at any standard significance level..

• The test for the null hypothesis $\beta_{y\_1} = -\beta_{k\_1}/\beta_k$ – that is, $\rho = -(-\rho\alpha_K)/\alpha_K$ – has a p-value of 7.1%. Therefore, it cannot be rejected with a 5% significance level, though it is rejected with a 10% significance level.

• The test for the null hypothesis $\beta_{y\_1} = -\beta_{ik\_1}/\beta_{ik}$ – that is, $\rho = -(-\rho\alpha_K\lambda_K)/\alpha_K\lambda_K$ – has a p-value of 84.4%. Therefore, we cannot reject this restriction. However, the reason why we cannot reject this null hypothesis is because the estimate of the parameter β_{ik_1} (the parameter associated to variable `l.ik`) is very imprecise and this implies that the paper has very low power, i.e., small probability of rejecting the null hypothesis when the hypothesis is false.

• The joint F-test of the four restrictions has a p-value of 0.0%. Therefore, we can reject the restrictions on the parameters associated to the Cochrane-Orcutt model.

In summary, the estimation of the Fixed-Effects + Cochrane-Orcutt provides: (1) parameter estimates that are sensible from an economic point of view; (2) strong evidence of serial correlation in the transitory shock u_{it}; and (3) strong evidence of time-invariant unobserved heterogeneity across firms. However, the restrictions of the Cochrane-Orcutt model are rejected. A possible explanation for this rejection is that the shocks a_{it} could be correlated with the observed inputs, such that there is still an endogeneity problem.

QUESTION 4. (30 points) For the following questions, provide the code in Stata and the table of estimation results.

(a) (10 points) Estimate the parameters of this production function using Arellano-Bond estimator with time dummies and non-serially correlated transitory shock. Present also the estimates of the parameters λ_L and λ_K.

ANSWER. For the Arellano-Bond estimator, we first consider the model transformed in first differences:

$$\Delta y_{it} = \beta_{lp\_1}\,\Delta lp_{it} + \beta_{rt}\,\Delta rt_{it} + \beta_k\,\Delta k_{it} + \beta_{ik}\,\Delta ik_{it} + \Delta\gamma_t + \Delta u_{it} \tag{4}$$

The parameters of this regression model have the following relationship with the parameters of our model: $\beta_{lp} = \alpha_L$; $\beta_{rt} = \alpha_L\lambda_L$; $\beta_k = \alpha_K$; and $\beta_{ik} = \alpha_K\lambda_K$. Then, we estimate this equation using an Instrumental Variables estimator (actually, a GMM estimator) where the instruments are the dependent variable and the regressors at period $t-2$ and before $t-2$. Under the assumption that u_{it} is not serially correlated, these instruments are not correlated with the error term Δu_{it}. Therefore, it is important to test for serial correlation in u_{it}. We cannot obtain consistent estimates (residuals) for u_{it}, but we can obtain consistent residuals for Δu_{it}. Therefore, we test for first order serial correlation if u_{it} by testing second order serial correlation in the residuals $\widehat{\Delta u}_{it}$.

CODE

```
// Question 4(a):  Arellano-Bond no AR
xtabond2 y lp rt k ik i.year, gmm(y lp rt k ik, lag(2 .))  iv(i.year) robust noleveleq

// Estimate of lambda_{L}
nlcom (lambda_L: _b[rt]/_b[lp])

// Estimate of lambda_{K}
nlcom (lambda_K: _b[ik]/_b[k])
```

OUTPUT

```
. // Question 4(a): Arellano-Bond no AR
. xtabond2 y lp rt k ik i.year, gmm(y lp rt k ik, lag(2 .)) iv(i.year) robust noleveleq
Favoring space over speed. To switch, type or click on mata: mata set matafavor speed, 
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate robust weighting matrix for Hansen test.
  Difference-in-Sargan/Hansen statistics may be negative.

Dynamic panel-data estimation, one-step difference GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      4025
Time variable : year                            Number of groups   =       575
Number of instruments = 133                     Obs per group: min =         7
Wald chi2(13) =    527.16                                      avg =      7.00
Prob > chi2   =     0.000                                      max =         7
------------------------------------------------------------------------------
             |               Robust
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          lp |   .2490154    .156911     1.59   0.113    -.0585244    .5565552
          rt |   .2212789   .1420364     1.56   0.119    -.0571073     .499665
           k |   .1712355   .1280831     1.34   0.181    -.0798028    .4222738
          ik |   .3467822   .1571265     2.21   0.027     .0388199    .6547445
             |
        year |
       1982  |          0  (empty)
       1983  |  -.2766518   .0315032    -8.78   0.000     -.338397   -.2149066
       1984  |   -.263344   .0289686    -9.09   0.000    -.3201215   -.2065665
       1985  |   -.223624   .0247147    -9.05   0.000    -.2720639    -.175184
       1986  |  -.1835018   .0163203   -11.24   0.000    -.2154891   -.1515145
       1987  |  -.1225478    .011294   -10.85   0.000    -.1446836    -.100412
       1988  |  -.0705741   .0076689    -9.20   0.000    -.0856048   -.0555433
       1989  |          0  (omitted)
       1990  |  -.0154142   .0103484    -1.49   0.136    -.0356967    .0048684
------------------------------------------------------------------------------
Instruments for first differences equation
  Standard
    D.(1982b.year 1983.year 1984.year 1985.year 1986.year 1987.year 1988.year
    1989.year 1990.year)
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/8).(y lp rt k ik)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -3.14  Pr > z =  0.002
Arellano-Bond test for AR(2) in first differences: z =  -0.23  Pr > z =  0.815
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(120)   = 891.21  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(120)   = 174.94  Prob > chi2 =  0.001
  (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 1
    Hansen test excluding group:     chi2(113)   = 170.85  Prob > chi2 =  0.000
    Difference (null H = exogenous): chi2(7)     =   4.09  Prob > chi2 =  0.770
```

```
. nlcom (lambda_L: _b[rt]/_b[lp])

    lambda_L:  _b[rt]/_b[lp]
```

y	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lambda_L	.8886151	.8678463	1.02	0.306	-.8123323	2.589563

```
.
. // Estimate of lambda_{K}
. nlcom (lambda_K: _b[ik]/_b[k])

    lambda_K:  _b[ik]/_b[k]
```

y	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lambda_K	2.025177	1.203062	1.68	0.092	-.3327819	4.383136

COMMENTS ON RESULTS

• The estimated parameters $\widehat{\alpha}_L = 0.249$ and $\widehat{\alpha}_K = 0.171$ are very small and imply very strong decreasing returns to scale: $\widehat{\alpha}_L + \widehat{\alpha}_K = 0.420 <<< 1$. This seems quite implausible from an economic point of view.

• The estimate of λ_L is 0.888. This indicates that temporary labor is less productive than permanent labor.

• The estimate of λ_K is 2.025. This indicates that new capital is more than two times more productive than old capital. This seems economically very implausible.

• Therefore, from the point of the economic interpretation of the parameter estimates, the Arellano-Bond estimator provides implausible estimates, especially compared to the FE + CO estimates.

• Furthermore, the parameter estimates are very imprecise. This indicates that the instruments are very weak.

(b) (10 points) Implement the same tests as in Questions 1(b), 1(c), and 1(d), and answer the same questions. Explain the results.

ANSWER. Testing the null hypothesis of $\lambda_L = 1$ is equivalent to testing the null hypothesis of $\beta_{lp} - \beta_{rt} = 0$. Testing the null hypothesis of $\lambda_K = 1$ is equivalent to testing the null hypothesis of $\beta_k - \beta_{ik} = 0$. Testing the null hypothesis of $\alpha_L + \alpha_K = 1$ is equivalent to testing the null hypothesis of $\beta_{lp} + \beta_k - 1 = 0$.

CODE

```
// Question 4(b):  Test null hypothesis lambda_{L}=1
test lp - rt = 0

// Question 4(b):  Test null hypothesis lambda_{K}=1
test k - ik = 0

// Question 4(b):  Test null hypothesis
test lp + k = 1
```

OUTPUT

```
. // Question 4(b): Test null hypothesis lambda_{L}=1
. test lp - rt = 0

 ( 1)  lp - rt = 0

           chi2(  1) =    0.01
         Prob > chi2 =    0.9039

.
. // Question 4(b): Test null hypothesis lambda_{K}=1
. test k - ik = 0

 ( 1)  k - ik = 0

           chi2(  1) =    1.85
         Prob > chi2 =    0.1735

.
. // Question 4(b): Test null hypothesis alpha_{L} + alpha_{K} = 1
. test lp + k = 1

 ( 1)  lp + k = 1

           chi2(  1) =   19.28
         Prob > chi2 =    0.0000
```

COMMENTS ON RESULTS

Test null hypothesis of $\lambda_L = 1$.

• The p-value of the test is 90%. Therefore, we cannot reject the null hypothesis.

• However, this is mostly because the parameter estimate is very imprecise (large standard error) such that the test has very low power. We cannot reject the null hypothesis of $\lambda_L = 1$, but note that we cannot reject either the null hypothesis of $\lambda_L = 0$ (`testnl _b[rt]/_b[lp] = 0`): the p-value is 30%.

• That is, these estimates do not contain any economically meaningful information about the parameter λ_L.

Test null hypothesis of $\lambda_K = 1$.

• The p-value of the test is 17%. Therefore, we cannot reject the null hypothesis.

• However, this is mostly because the parameter estimate is very imprecise (large standard error) such that the test has very low power. We cannot reject the null hypothesis of $\lambda_K = 1$, but similarly we cannot reject either the null hypothesis of $\lambda_K = 0$ (`testnl _b[ik]/_b[k] = 0`): the p-value is 9.2%.

• That is, these estimates do not contain any economically meaningful information about the parameter λ_K.

Test null hypothesis of $\alpha_L + \alpha_K = 1$.

• The p-value of the test is 0.0000. Therefore, we can reject the null hypothesis of constant returns to scale.

• However, the point estimates are quite implausible and very imprecise.

(c) (5 points) Test for the null-hypothesis of no serial correlation in the transitory shocks. Explain the results.

ANSWER. We cannot obtain consistent estimates (residuals) for u_{it}, but we can obtain consistent residuals for Δu_{it}. Therefore, we test for first order serial correlation if u_{it} by testing second order serial correlation in the residuals $\widehat{\Delta u_{it}}$. This is the Arellano-Bond test for AR(2) in the output of the `xtabond2` command.

CODE

```
// Question 4(c):  Test for the null-hypothesis of no serial correlation
// In the table estimates, the line:
// Arellano-Bond test for AR(2) in first differences:  z = -0.23 Pr > z = 0.815
```

COMMENT RESULTS

• The null hypothesis of the test is No Serial Correlation. The p-value of the test is 81.5%. Therefore, we cannot reject the null hypothesis that u_{it} is not serially correlated.

(d) (5 points) Test for the over-identification restrictions of this IV estimator. Explain the results.

ANSWER. In this Arellano-Bond IV estimator, we have more instruments than parameters to estimate. This means that we have more moment restrictions than parameters to estimate. The null hypothesis of the test of over-identifying restrictions is that all the moment restrictions implied by the set of instruments are valid. Therefore, the null hypothesis is that all the instruments are valid.

The test is a Chi-square test where the number of degrees of freedom is equal to the number of instruments (moment restrictions) and the number of parameters. The value of the Chi-square statistic can be obtained as the number of observations times the R-square coefficient from a regression where the dependent variable is the residuals of the IV regression and the explanatory variables are all the instruments used in the IV regression.

The `xtabond2` command reports two tests of over-identifying: Sargan test and Hansen test. Sargan's statistic is a special case of Hansen's statistic under the restriction that the residuals are homoskedastic. Therefore, Hansen's test is more robust that Sargan test. Here we use Hansen test.

CODE

```
// Question 4(d):  Test of validity of over-identifying restrictions
// In the table estimates, the line:
// Hansen test of overid.  restrictions:  chi2(120) = 174.94 Prob > chi2 = 0.001
```

COMMENT RESULTS

• The null hypothesis of the Hansen test is that all the moment conditions are correct such that all the instruments are valid. The p-value of the test is 0.1%. Therefore, we can reject the null hypothesis under a significance level of 1%. We reject the validity of the over-identifying restrictions.

QUESTION 5. (40 points) For the following questions, provide the code in Stata and the table of estimation results.
(a) (10 points) Estimate the parameters of this production function using Arellano-Bond estimator with time dummies and with an AR(1) transitory shock. Present also the estimates of the parameters λ_L and λ_K.

ANSWER.

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. 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:

$$\begin{aligned} y_{it} &= \beta_{y\_1}\, y_{it-1} + \beta_{lp}\, lp_{it} + \beta_{lp\_1}\, lp_{it-1} + \beta_{rt}\, rt_{it} + \beta_{rt\_1}\, rt_{it-1} \\ &+ \beta_{k}\, k_{it} + \beta_{k\_1}\, k_{it-1} + \beta_{ik}\, ik_{it} + \beta_{ik\_1}\, ik_{it-1} + \eta_i^* + \gamma_t^* + a_{it} \end{aligned}$$

where: $\beta_{y\_1} = \rho$; $\beta_{lp} = \alpha_L$; $\beta_{lp\_1} = -\rho\alpha_L$; $\beta_{rt} = \alpha_L\lambda_L$; $\beta_{rt\_1} = -\rho\alpha_L\lambda_L$; $\beta_k = \alpha_K$; $\beta_{k\_1} = -\rho\alpha_K$; $\beta_{ik} = \alpha_K\lambda_K$; and $\beta_{ik\_1} = -\rho\alpha_K\lambda_K$. We transform the model in first differences to eliminate the individual effects η_i^*:

$$\begin{aligned} \Delta y_{it} &= \beta_{y\_1}\, \Delta y_{it-1} + \beta_{lp}\, \Delta lp_{it} + \beta_{lp\_1}\, \Delta lp_{it-1} + \beta_{rt}\, \Delta rt_{it} + \beta_{rt\_1}\, \Delta rt_{it-1} \\ &+ \beta_{k}\, \Delta k_{it} + \beta_{k\_1}\, \Delta k_{it-1} + \beta_{ik}\, \Delta ik_{it} + \beta_{ik\_1}\, \Delta ik_{it-1} + \Delta\gamma_t^* + \Delta a_{it} \end{aligned}$$

Then, we estimate this equation using an Instrumental Variables estimator (actually, a GMM estimator) where the instruments are the dependent variable and all the regressors at period $t-2$ and before $t-2$.

Under the assumption that a_{it} is not serially correlated, these instruments are not correlated with the error term Δa_{it}. Therefore, it is important to test for serial correlation in a_{it}. We cannot obtain consistent estimates (residuals) for a_{it}, but we can obtain consistent residuals for Δa_{it}. Therefore, we test for first order serial correlation if a_{it} by testing second order serial correlation in the residuals $\widehat{\Delta a}_{it}$.

The assumption that u_{it} is AR(1) implies four restrictions on the parameter estimates β:

$$\begin{aligned} \beta_{y\_1} &= -\beta_{lp\_1}/\beta_{lp} \\ \beta_{y\_1} &= -\beta_{rt\_1}/\beta_{rt} \\ \beta_{y\_1} &= -\beta_{k\_1}/\beta_{k} \\ \beta_{y\_1} &= -\beta_{ik\_1}/\beta_{ik} \end{aligned}$$

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

CODE

```
// Question 5(a):  Arellano-Bond with AR(1)
```

```
xtabond2 y l.y lp l.lp rt l.rt k l.k ik l.ik i.year, gmm(y lp rt k ik, lag(2 .))
iv(i.year) robust noleveleq

// There are two possible estimates of lambda_{L}
// Estimate 1:
nlcom (lambda_L1:  _b[rt]/_b[lp])

// Estimate 2:
nlcom (lambda_L2:  _b[l.rt]/_b[l.lp])

// There are two possible estimates of lambda_{K}
// Estimate 1:
nlcom (lambda_K1:  _b[ik]/_b[k])

// Estimate 2:
nlcom (lambda_K2:  _b[l.ik]/_b[l.k])
```

OUTPUT

```
. // Question 5(a): Arellano-Bond with AR(1)
. xtabond2 y l.y lp l.lp rt l.rt k l.k ik l.ik i.year, gmm(y lp rt k ik, lag(2 .)) iv(i.year) robust nolevel
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 test.
  Difference-in-Sargan/Hansen statistics may be negative.

Dynamic panel-data estimation, one-step difference GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      3450
Time variable : year                            Number of groups   =       575
Number of instruments = 129                     Obs per group: min =         6
Wald chi2(18) =   2094.47                                      avg =      6.00
Prob > chi2   =     0.000                                      max =         6
------------------------------------------------------------------------------
             |               Robust
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           y |
         L1. |   .6375806   .0495834    12.86   0.000     .5403989    .7347622
             |
          lp |
         --. |   .5506398   .1745856     3.15   0.002     .2084583    .8928212
         L1. |  -.4357561   .1707874    -2.55   0.011    -.7704932   -.1010189
             |
          rt |
         --. |   .2261455   .1156302     1.96   0.050    -.0004856    .4527766
         L1. |  -.0379607   .0734848    -0.52   0.605    -.1819882    .1060667
             |
           k |
         --. |   .5554058   .2770229     2.00   0.045      .012451    1.098361
         L1. |  -.4518527   .2736703    -1.65   0.099    -.9882367    .0845313
             |
          ik |
         --. |   .1534665   .1346816     1.14   0.255    -.1105046    .4174375
         L1. |  -.3090006   .2422174    -1.28   0.202     -.783738    .1657369
             |
        year |
       1982  |          0  (empty)
       1983  |          0  (omitted)
       1984  |  -.1232722   .0227753    -5.41   0.000    -.1679109   -.0786334
       1985  |   -.093154   .0201115    -4.63   0.000    -.1325719   -.0537361
       1986  |   -.094705   .0214358    -4.42   0.000    -.1367184   -.0526915
       1987  |  -.0549582   .0154987    -3.55   0.000    -.0853351   -.0245813
       1988  |   -.038297   .0103658    -3.69   0.000    -.0586135   -.0179805
       1989  |          0  (omitted)
       1990  |  -.0728178   .0123588    -5.89   0.000    -.0970405    -.048595
------------------------------------------------------------------------------
Instruments for first differences equation
  Standard
    D.(1982b.year 1983.year 1984.year 1985.year 1986.year 1987.year 1988.year
    1989.year 1990.year)
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/8).(y lp rt k ik)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -7.00  Pr > z =  0.000
Arellano-Bond test for AR(2) in first differences: z =   1.05  Pr > z =  0.292
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(111)  = 224.42  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(111)  = 155.94  Prob > chi2 =  0.003
  (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 1990.year)
    Hansen test excluding group:     chi2(105)  = 148.47  Prob > chi2 =  0.003
    Difference (null H = exogenous): chi2(6)    =   7.48  Prob > chi2 =  0.279
```

```
. // There two possible estimates of lambda_{L}
. //    Estimate 1:
. nlcom (lambda_L1: _b[rt]/_b[lp])

   lambda_L1:  _b[rt]/_b[lp]
```

y	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lambda_L1	.410696	.2571408	1.60	0.110	-.0932907	.9146827

```
. //    Estimate 2:
. nlcom (lambda_L2: _b[l.rt]/_b[l.lp])

   lambda_L2:  _b[l.rt]/_b[l.lp]
```

y	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lambda_L2	.0871147	.1633543	0.53	0.594	-.2330538	.4072831

```
. // There two possible estimates of lambda_{K}
. //    Estimate 1:
. nlcom (lambda_K1: _b[ik]/_b[k])

   lambda_K1:  _b[ik]/_b[k]
```

y	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lambda_K1	.2763141	.2626409	1.05	0.293	-.2384525	.7910808

```
. //    Estimate 2:
. nlcom (lambda_K2: _b[l.ik]/_b[l.k])

   lambda_K2:  _b[l.ik]/_b[l.k]
```

y	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lambda_K2	.6838525	.1583133	4.32	0.000	.3735642	.9941408

COMMENTS ON RESULTS

• The estimated parameters $\widehat{\alpha}_L = 0.550$ and $\widehat{\alpha}_K = 0.555$ are reasonable and close to constant returns to scale.

• There are two estimates of λ_L: $\widehat{\beta}_{rt}/\widehat{\beta}_{lp} = 0.410$ ($s.e. = 0.257$); and $\widehat{\beta}_{rt\_1}/\widehat{\beta}_{lp\_1} = 0.087$ ($s.e. = 0.163$). This indicates that temporary labor is less productive than permanent labor.

• There are two estimates of λ_K: $\widehat{\beta}_{ik}/\widehat{\beta}_k = 0.276$ ($s.e. = 0.262$); and $\widehat{\beta}_{ik\_1}/\widehat{\beta}_{k\_1} = 0.683$ ($s.e. = 0.158$). This indicates that new investments during their first year of operation are less productive than old capital (time to build hypothesis).

• From the point of their economic interpretation, these estimates seem sensible.

• However, the estimates are imprecise: standard errors are large. This indicates that the instruments are weak.

(b) (10 points) Implement the same tests as in Questions 1(b), 1(c), and 1(d), and answer the same questions. Explain the results.

ANSWER. Since we have two estimates of the parameter λ_L, we have two possible tests for the null hypothesis that $\lambda_L = 1$: test of $\beta_{lp} - \beta_{rt} = 0$ and test of $\beta_{lp\_1} - \beta_{rt\_1} = 0$. We test these restrictions both separately and jointly.

Similarly, we have two possible tests for the null hypothesis that $\lambda_K = 1$: test of $\beta_k - \beta_{ik} = 0$ and test of $\beta_{k\_1} - \beta_{ik\_1} = 0$. We test these restrictions both separately and jointly.

The same comment applies to the test of the null hypothesis of $\alpha_L + \alpha_K = 1$. We have that $\alpha_L = \beta_{lp}$ and $\alpha_K = \beta_k$ such that we can implement the test $\beta_{lp} + \beta_k - 1 = 0$. But we also have that $\alpha_L = -\beta_{lp\_1}/\beta_{y\_1}$ and $\alpha_K = -\beta_{k\_1}/\beta_{y\_1}$ such that we can implement the test $-\beta_{lp\_1}/\beta_{y\_1} - \beta_{k\_7}/\beta_{y\_1} - 1 = 0$. Note that this is equivalent to $\beta_{y\_1} + \beta_{lp\_1} + \beta_{k\_1} = 0$. We test these restrictions separately and jointly.

CODE

```
// Question 5(b):  Test null hypothesis lambda_{L}=1
test lp - rt = 0
test l.lp - l.rt = 0
test (lp - rt = 0) (l.lp - l.rt = 0)

// Question 5(b):  Test null hypothesis lambda_{K}=1
test k - ik = 0
test l.k - l.ik = 0
test (k - ik = 0) (l.k - l.ik = 0)

// Question 5(b):  Test null hypothesis
```

```
test lp + k = 1
test l.lp + l.k + l.y = 0
test (lp + k = 1) (l.lp + l.k + l.y = 0)
```

OUTPUT

```
. // Question 5(b): Test null hypothesis lambda_{L}=1
. test lp - rt = 0

 ( 1)  lp - rt = 0

           chi2(  1) =    2.21
         Prob > chi2 =    0.1369

. test l.lp - l.rt = 0

 ( 1)  L.lp - L.rt = 0

           chi2(  1) =    5.61
         Prob > chi2 =    0.0179

. test (lp - rt = 0) (l.lp - l.rt = 0)

 ( 1)  lp - rt = 0
 ( 2)  L.lp - L.rt = 0

           chi2(  2) =    5.62
         Prob > chi2 =    0.0602
```

```
. // Question 5(b): Test null hypothesis lambda_{K}=1
. test k - ik = 0

 ( 1)  k - ik = 0

           chi2(  1) =    1.90
         Prob > chi2 =    0.1681

. test l.k - l.ik = 0

 ( 1)  L.k - L.ik = 0

           chi2(  1) =    5.05
         Prob > chi2 =    0.0246

. test (k - ik = 0) (l.k - l.ik = 0)

 ( 1)  k - ik = 0
 ( 2)  L.k - L.ik = 0

           chi2(  2) =    5.05
         Prob > chi2 =    0.0799

. // Question 5(b): Test null hypothesis alpha_{L} + alpha_{K} = 1
. test lp + k = 1

 ( 1)  lp + k = 1

           chi2(  1) =    0.10
         Prob > chi2 =    0.7509

. test l.lp + l.k + l.y = 0

 ( 1)  L.y + L.lp + L.k = 0

           chi2(  1) =    0.54
         Prob > chi2 =    0.4633

. test (lp + k = 1) (l.lp + l.k + l.y = 0)

 ( 1)  lp + k = 1
 ( 2)  L.y + L.lp + L.k = 0

           chi2(  2) =    3.48
         Prob > chi2 =    0.1756
```

COMMENTS ON RESULTS

Tests null hypothesis of $\lambda_L = 1$.

• The two individual tests have p-values of 13.7% and 1.8%, respectively. The joint test has a p-value of 6.0%. Therefore, using significance levels of 1% or 5%, we cannot reject the null hypothesis that $\lambda_L = 1$. However, part of the reason is that the estimate of the parameters β_{rt} and β_{rt_1} are quite imprecise such that these tests have low power.

Tests null hypothesis of $\lambda_K = 1$.

• The two individual tests have p-values of 16.8% and 2.5%, respectively. The joint test has a p-value of 8.0%. Therefore, using significance levels of 1% or 5%, we cannot reject the null hypothesis that $\lambda_K = 1$. However, part of the reason is that the estimate of the parameters β_{ik} and β_{ik_1} are quite imprecise such that these tests have low power.

Test null hypothesis of $\alpha_L + \alpha_K = 1$.

• The two individual tests have p-values of 75% and 46%, respectively. The joint test has a p-value of 17.5%. Therefore, there is strong evidence in favor of the null hypothesis of constant returns to scale.

(c) (10 points) Test the 4 restrictions on the parameters implied by the Cochrane-Orcutt model. Test each restriction separately, and also the 4 joint restrictions (5 different tests).

ANSWER. The Cochrane-Orcutt model implies four restrictions on the β parameters:

$$\begin{aligned}
\beta_{y_1} &= -\beta_{lp_1}/\beta_{lp} \\
\beta_{y_1} &= -\beta_{rt_1}/\beta_{rt} \\
\beta_{y_1} &= -\beta_{k_1}/\beta_{k} \\
\beta_{y_1} &= -\beta_{ik_1}/\beta_{ik}
\end{aligned}$$

We can test these restrictions separately and jointly.

CODE

```
// Question 5(c):  CO Restrictions
testnl _b[l.y] + _b[l.lp]/_b[lp] = 0
testnl _b[l.y] + _b[l.rt]/_b[rt] = 0
testnl _b[l.y] + _b[l.k]/_b[k] = 0
testnl _b[l.y] + _b[l.ik]/_b[ik] = 0
testnl (_b[l.y] + _b[l.lp]/_b[lp] = 0) (_b[l.y] + _b[l.rt]/_b[rt] = 0) (_b[l.y]
+ _b[l.k]/_b[k] = 0) (_b[l.y] + _b[l.ik]/_b[ik] = 0)
```

OUTPUT

```
. // Question 5(c): CO Restrictions
. testnl _b[l.y] + _b[l.lp]/_b[lp] = 0

  (1)  _b[l.y] + _b[l.lp]/_b[lp] = 0

               chi2(1) =       1.12
           Prob > chi2 =       0.2893

. testnl _b[l.y] + _b[l.rt]/_b[rt] = 0

  (1)  _b[l.y] + _b[l.rt]/_b[rt] = 0

               chi2(1) =       1.70
           Prob > chi2 =       0.1917

. testnl _b[l.y] + _b[l.k]/_b[k]  = 0

  (1)  _b[l.y] + _b[l.k]/_b[k] = 0

               chi2(1) =       1.29
           Prob > chi2 =       0.2568

. testnl _b[l.y] + _b[l.ik]/_b[ik] = 0

  (1)  _b[l.y] + _b[l.ik]/_b[ik] = 0

               chi2(1) =       0.37
           Prob > chi2 =       0.5404

. testnl (_b[l.y] + _b[l.lp]/_b[lp] = 0) (_b[l.y] + _b[l.rt]/_b[rt] = 0) (_b[l.y] + _b[l.k]/_b[k] = 0) (_b[l.y] +

  (1)  _b[l.y] + _b[l.lp]/_b[lp] = 0
  (2)  _b[l.y] + _b[l.rt]/_b[rt] = 0
  (3)  _b[l.y] + _b[l.k]/_b[k] = 0
  (4)  _b[l.y] + _b[l.ik]/_b[ik] = 0

               chi2(4) =       5.45
           Prob > chi2 =       0.2445
```

COMMENTS ON RESULTS

- All these test – of the individual restrictions and the joint test – have p-values greater than 19%. Therefore, we cannot reject the restrictions imposed by the assumption that the transitory shock u_{it} follows an AR(1) process.

(d) (5 points) Test for the null-hypothesis of no serial correlation in the transitory shocks. Explain the results.

ANSWER. We cannot obtain consistent estimates (residuals) for a_{it}, but we can obtain consistent residuals for Δa_{it}. Therefore, we test for first order serial correlation if a_{it} by testing second order serial correlation in the residuals $\widehat{\Delta a}_{it}$. This is the Arellano-Bond test for AR(2) in the output of the `xtabond2` command.

CODE

```
// Question 5(d):  Test for the null-hypothesis of no serial correlation
// In the table estimates, the line:
// Arellano-Bond test for AR(2) in first differences:  z = 1.05 Pr > z = 0.292
// We cannot reject null hypothesis of no correlation
```

COMMENT RESULTS

• The null hypothesis of the test is No Serial Correlation. The p-value of the test is 29.2%. Therefore, we cannot reject the null hypothesis that a_{it} is not serially correlated.

• Taking into account: (1) the result of this test; (2) the very significant estimate of the parameter $\widehat{\beta}_{y\_1} = \widehat{\rho} = 0.637$ ($s.e. = 0.049$); and (3) the No Rejection of the four restrictions from the Cochrane-Orcutt model; provide significant evidence in favor of the assumption that the transitory shock u_{it} follows an AR(1) process.

(e) (5 points) Test for the over-identification restrictions of this IV estimator. Explain the results.

ANSWER. In this Arellano-Bond IV estimator, we have more instruments than parameters to estimate. This means that we have more moment restrictions than parameters to estimate. The null hypothesis of the test of over-identifying restrictions is that all the moment restrictions implied by the set of instruments are valid. Therefore, the null hypothesis is that all the instruments are valid.

The test is a Chi-square test where the number of degrees of freedom is equal to the number of instruments (moment restrictions) and the number of parameters. The value of the Chi-square statistic can be obtained as the number of observations times the R-square coefficient from a regression where the dependent variable is the residuals of the IV regression and the explanatory variables are all the instruments used in the IV regression.

The `xtabond2` command reports two tests of over-identifying: Sargan test and Hansen test. Sargan's statistic is a special case of Hansen's statistic under the restriction that the residuals are homoskedastic. Therefore, Hansen's test is more robust that Sargan test. Here we use Hansen test.

CODE

```
// Question 5(e):  Test of validity of over-identifying restrictions
// In the table estimates, the line:
// Hansen test of overid.  restrictions:  chi2(111) = 155.94 Prob > chi2 = 0.003
```

COMMENT RESULTS

• The null hypothesis of the Hansen test is that all the moment conditions are correct such that all the instruments are valid. The p-value of the test is 0.3%. Therefore, we can reject the null hypothesis under a significance level of 1%. However, this p-value has improved with respect to the 0.1% in the model where u_{it} is assumed not serially correlated.

COMMENTS ON OVERALL RESULTS

• The estimation results using the Arellano-Bond estimator with AR(1) error have some good / interesting properties.

• 1. The test of serial correlation (from the new residuals) cannot reject the null.

• 2. The test of OIR improves with respect to the model without AR(1).

• 3. The CO restrictions cannot be rejected.

• 4. The null of CRS cannot be rejected.

• 5. The estimates α_L and α_K are reasonable.

• 6. The estimates of λ_K implies that new investment is less productive that installed capital – time-to-build – which is a common restriction typically imposed in PF estimation.

• 7. The estimate of λ_L is consistent with temporary workers being less productive than permanent workers.