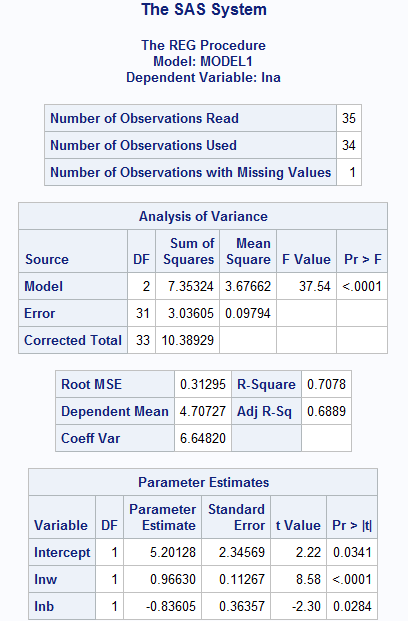
**Part I: Autocorrelation test, GLS and MLE:**

|  |
| --- |
| **PROC** **IMPORT** OUT= WORK.bm  DATAFILE= "C:\Users\bmishra\Dropbox\Ph.D. Courseworks\Semest  er II, Spring 2019\Econometric Methods\Homeworks\Homework 5\HW5-DATA1.xl  s"  DBMS=EXCEL REPLACE;  RANGE="SUGAR$";  GETNAMES=YES;  MIXED=NO;  SCANTEXT=YES;  USEDATE=YES;  SCANTIME=YES;  **RUN**;  /\* Data Management \*/  **data** bm; set bm;  a = a;  lna = log(a);  w = w;  lnw = log(w);  b = b;  lnb = log(b);  p = w/b;  lnp = log(p);  **run**; |

(a).

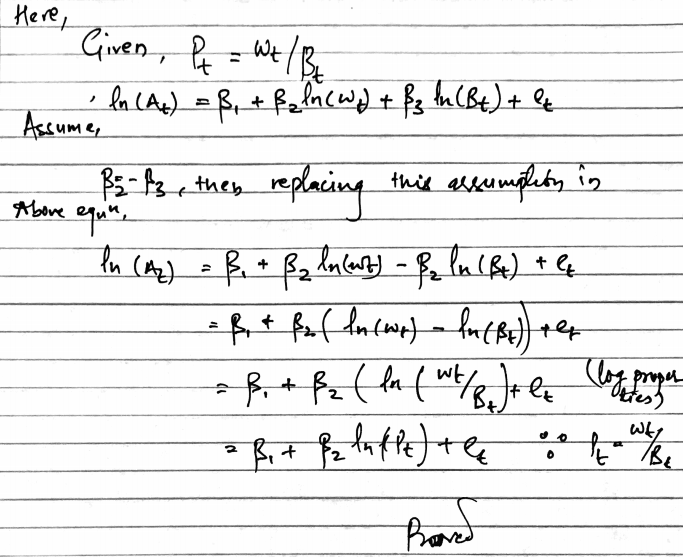
|  |
| --- |
| /\* Part I \*/  **proc** **reg** data = bm; /\* OLS \*/  model lna = lnw lnb / dwprob; /\* Q (a, b, d) \*/  p = lnayhatt;  r = lnaresid;  test lnw + lnb = **0**; /\* Q (d) \*/  **run**;  **proc** **print**;  **run**; |



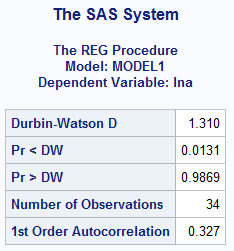
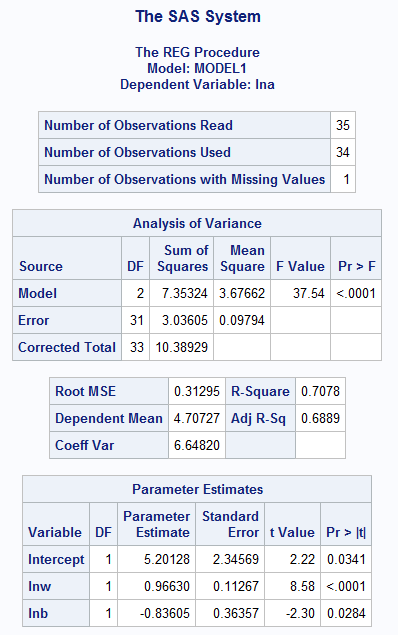
Interpretations:

β2: with the one unit percentage increase in the price of w measured in $/ton and keeping other variable constant the percentage of acreage in wheat in thousands of hectares increase by 0.96630 unit an average.

β3: with the one unit percentage increase in the price of jute (b) measured in $/ton and keeping other variable constant, the percentage of acreage in wheat decrease by 0.836025 unit on an average.



(b).



Note: Pr<DW (DW and d in lecture note are same) is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing negative autocorrelation.

Ist order autocorrelation (⍴) = 0.327

Null Hypothesis: The residuals from an OLS are not auto-correlated (no autocorrelation).

Ho: ⍴ = 0

Alternative Hypothesis: The residuals follows an AR1 process (auto-correlated).

Ha: ⍴ > 0 (because d < 2)

Conclusion: we reject null hypothesis (P = 0.0131) at 5% significance level.. There might be autocorrelation.

d = 1.310. Number of observations (n) = 34 and # of variables (k) (no intercept) = 2. At 5% Significance point dL = 1.333 and dU = 1.580

Since d < dL, we reject null hypothesis at 5% significance level. There might be evidence of autocorrelation.

(c).

|  |
| --- |
| **proc** **autoreg** data = bm;  model lna = lnw lnb / nlag = **1**; /\* Q (c) \*/  test lnw + lnb = **0**; /\* Q (d) \*/  output out = autoreg1 r = auregehatt; /\* Optional Statement \*/  **run**; |

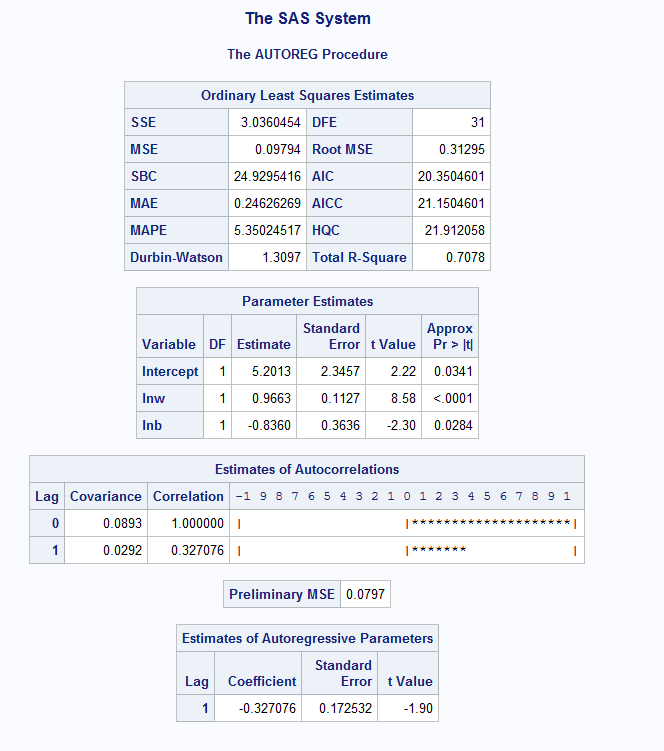


Figure: OLS (Left) & GLS (Right)

* The SSE, MSE SBC, MAE, MAPE, DFE, Root MSE, AIC, AICC, HQC are larger in OLS compared to GLS and higher Total R-squared value in GLS.
* Looking at parameters estimates table, parameter estimates are smaller, standard errors are larger from and t-values are smaller in OLS compared to GLS except for lnb. Both models have parameters with similar significance trend at 95% CI (all significant) but OLS has some non-significant variables at 99% CI.

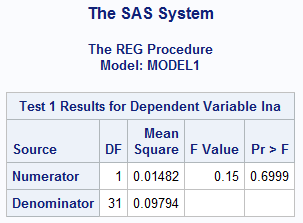


(d).

Estimated result from (b):

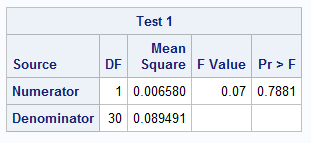
Null hypothesis: Ho: β2 = β3. Alternative Hypothesis Ha: β2 ≠ β3.

Conclusion: We fail to reject null hypothesis at 95% CI. (P = 0.6999)



Null hypothesis: Ho: β2 = β3. Alternative Hypothesis Ha: β2 ≠ β3.

Conclusion: We fail to reject null hypothesis at 95% CI. (P = 0.7881)



Estimated result from (c):

The hypothesis test results do not differ between (b) and (c).

(e). Predicting for first year using GLS result:

Ln(A) = 5.6184 + 1.0127\*lnw – 0.9328 \* lnb + ⍴\*d

⍴ = 0.327 & d = 1.9391

For first year,

Ln(A) =5.6184 + (LN(500)\*1.0127)-(LN(500)\*0.9328)+ (0.327\*1.9391) = 6.7490

For second year,

Ln(A) =5.6184 + (LN(500)\*1.0127)-(LN(500)\*0.9328)+ (0.327\*1.9391)2 = 6.5170

**Part II: Two-stage Least Square Estimation:**

|  |
| --- |
| **PROC** **IMPORT** OUT= WORK.HW5D2  DATAFILE= "C:\Users\bmishra\Dropbox\Ph.D. Courseworks\Semest  er II, Spring 2019\Econometric Methods\Homeworks\Homework 5\HW5-DATA2.tx  t"  DBMS=TAB REPLACE;  GETNAMES=YES;  DATAROW=**2**;  **RUN**;  /\* Part II \*/  **data** hw5d2; set hw5d2;  lny = log(y);  lnk = log(k);  lnw = log(w);  lnl = log(l);  lnp = log(p);  lnr = log(r);  **run**; |

Given:

Y = output; P = output price;

K = capital input; R = capital input price;

L = Labor input; W = labor input price.

Output = *f* (capital input, labor input) ………………….. (1)

Capital input = *f* (output price, output, capital input price) ………………….. (2)

Labor input = *f* (output price, output, labor input price) ………………….. (3)

Endogenous Variables: Y, K, and L. Exogenous variables: P, R, and W.

(a).

In our case, we have M= 3 simultaneous equations to jointly determine the values of M = 3 endogenous variables. At least M-1 = 2 variables must be absent from an equation for estimation of its parameters to be possible. (For an equation to be identified in a system the total number of variables excluded from the equation but included in other equations must be at least equal to one less than the number of equations in the system.)

For equation 1, at least one variable in the system must be absent from equation 1. We have total of three variables (output price, capital input price and labor input price) absent from the system in this equation. This is over identified.

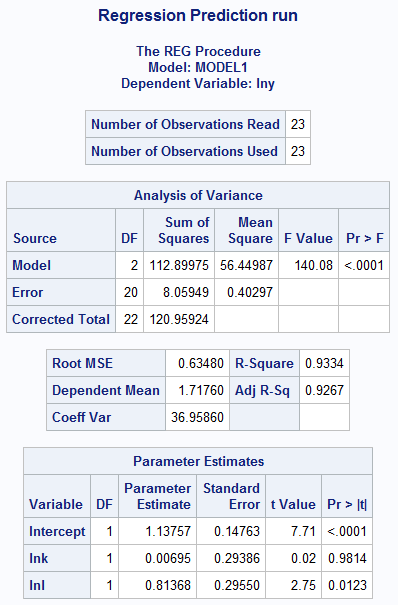
For equation 2, at least two variables should be absent from the equation 2. We have two missing variables (labor input and labor input price) from the system in this equation. This equation is identified.

For Equation 3, at least two variables should be absent from the equation 2 to be identified. We have two missing variables (capital input and capital input price) from the system in this equation. This equation is also identified.

(b).

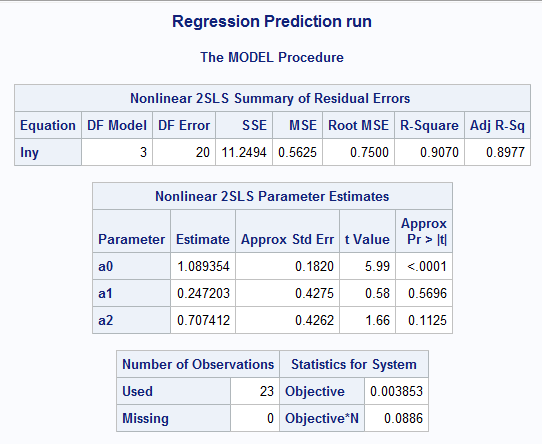
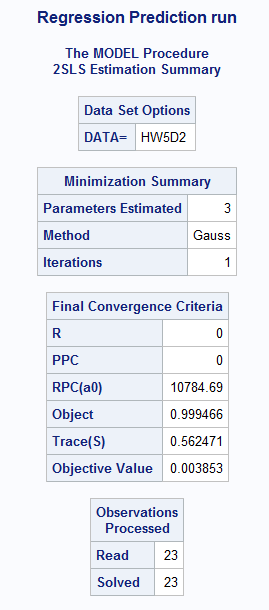
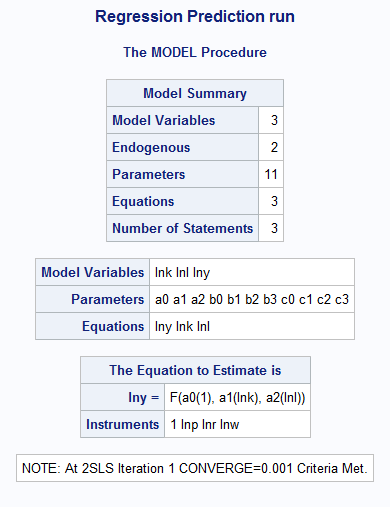
|  |
| --- |
| /\* b \*/  **proc** **reg** data = hw5d2;  model lny = lnk lnl;  **run**; |

Theoretically, output (Y) is dependent upon capital input and labor input. Further capital input is dependent upon output price, output, capital input price and labor input is dependent upon output price, output, and labor input price. However, with the OLS regression above we estimated equation only based on capital input and labor input but their dependence upon other six variables in the system are not accounted. This is not accounted while estimated output equation and thus the prediction is inaccurate/not consistent and also biased.



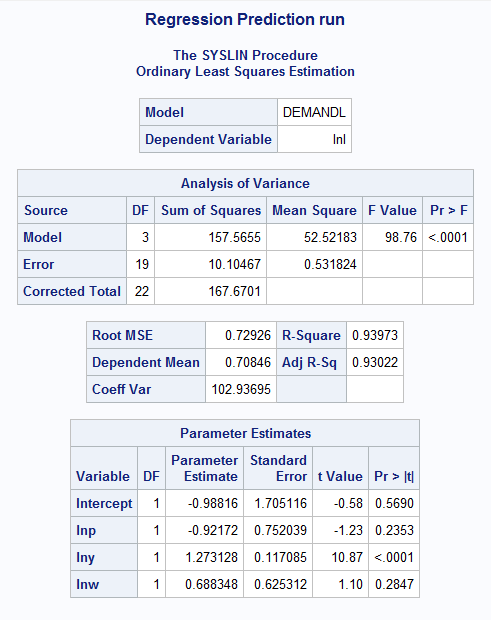
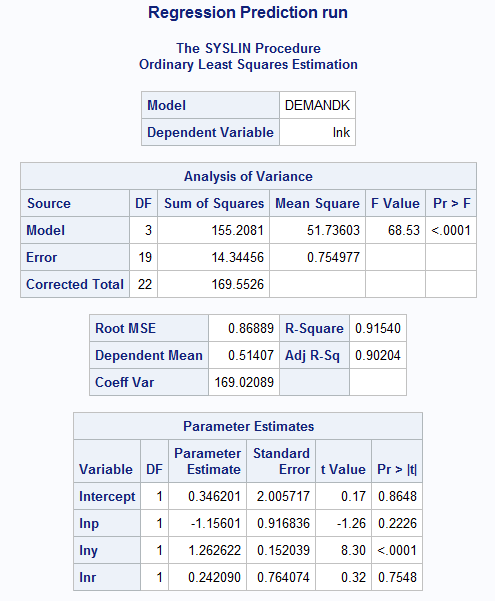
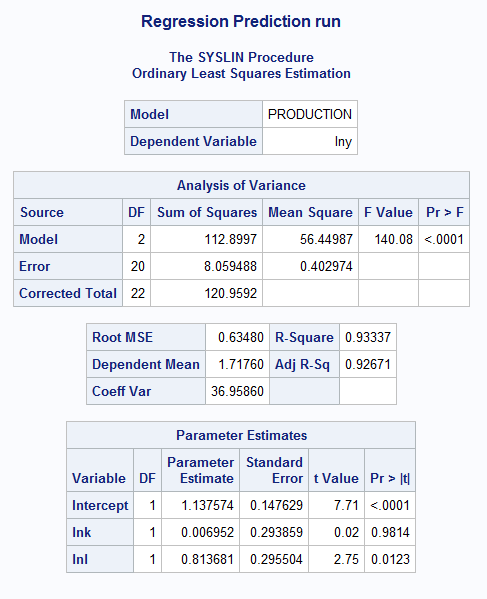
(c).

|  |
| --- |
| /\* Model Procedure: Creates One Production Model \*/  **proc** **model** data = hw5d2;  instruments lnp lnr lnw;  endogenous lnk lnl;  production: lny = a0 + a1\*lnk + a2\*lnl;  demandk: lnk = b0 + b1\*lnp + b2\*lny + b3\*lnr;  demandl: lnl = c0 + c1\*lnp + c2\*lny + c3\*lnw;  fit lny / **2**sls;  **run**; |



Two stage least square estimates are available in nonlinear 2SLS parameter estimates table. Parameters are still biased but they are consistent (not consistent before). Parameters are in same directions in both models but parameter of lnl (labor input) become insignificant (it was significant before). Also, lnk is also insignificant.

(c). Alternative Way using Syslin (Same as Above Extra Work) (Remove before Submission)



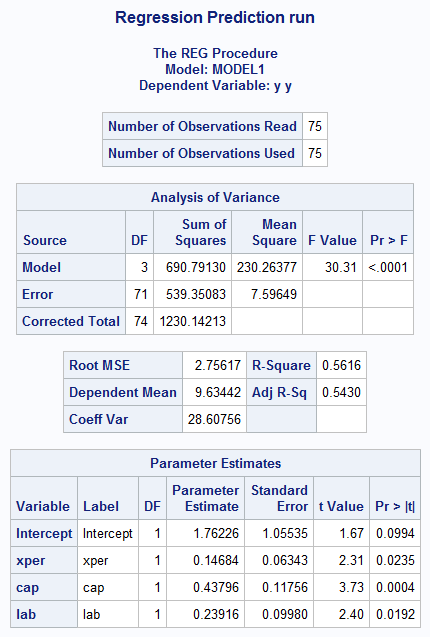
**Question:** Why results are different? Are they supposed to be different or it is because of my error?

|  |
| --- |
| Extra: /\* c: Syslin Procedure: Creates Several Models \*/  **proc** **syslin** data = hw5d2;  instruments lnp lnr lnw;  endogenous lnk lnl;  production: model lny = lnk lnl / overid;  demandk: model lnk = lnp lny lnr / overid;  demandl: model lnl = lnp lny lnw / overid;  **run**; |

**Part III: Instrumental Variable/ Two-stage Least Square Estimation:**

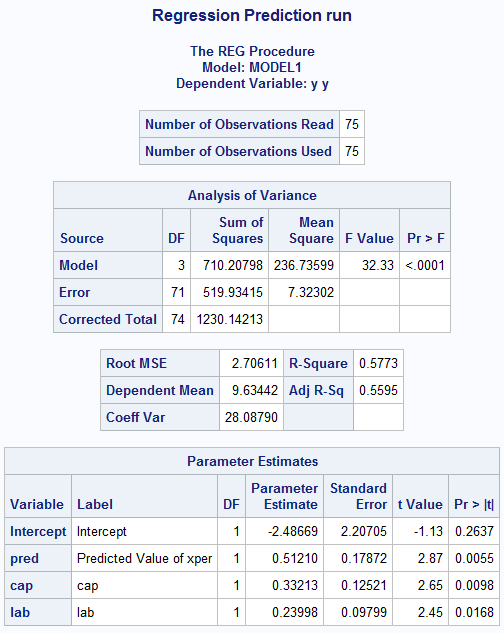
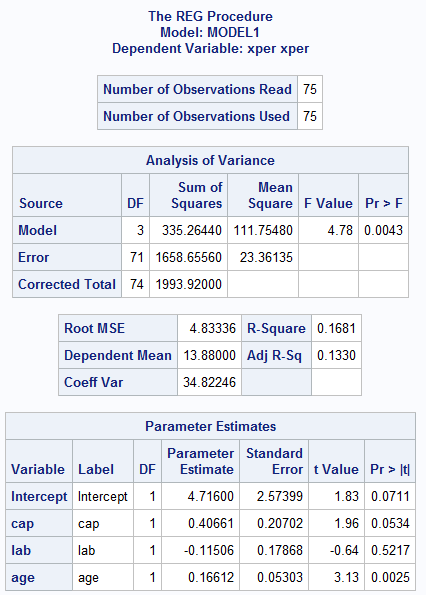
(a).

|  |
| --- |
| **PROC** **IMPORT** OUT= WORK.HW5D3  DATAFILE= "C:\Users\bmishra\Dropbox\Ph.D. Courseworks\Semest  er II, Spring 2019\Econometric Methods\Homeworks\Homework 5\HW5-DATA3.xl  s"  DBMS=EXCEL REPLACE;  RANGE="data";  GETNAMES=YES;  MIXED=NO;  SCANTEXT=YES;  USEDATE=YES;  SCANTIME=YES;  **RUN**;  **proc** **reg** data = hw5d3;  model y = xper cap lab;  **run**; |



(b).

|  |
| --- |
| /\* b \*/  **proc** **reg** data = hw5d3;  model xper = cap lab age;  output out = aut r = resid p = pred;  **run**;  **proc** **reg** data = aut;  model y = pred cap lab;  **run**; |

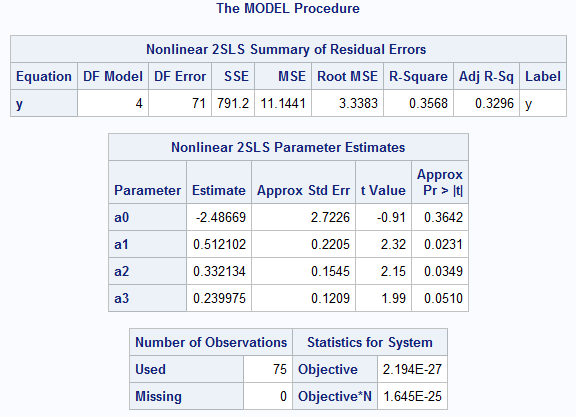
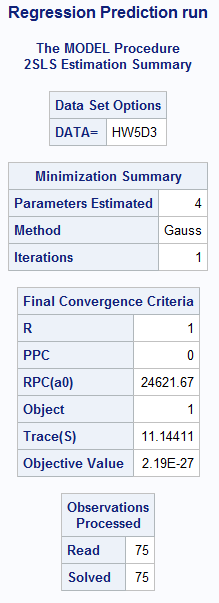
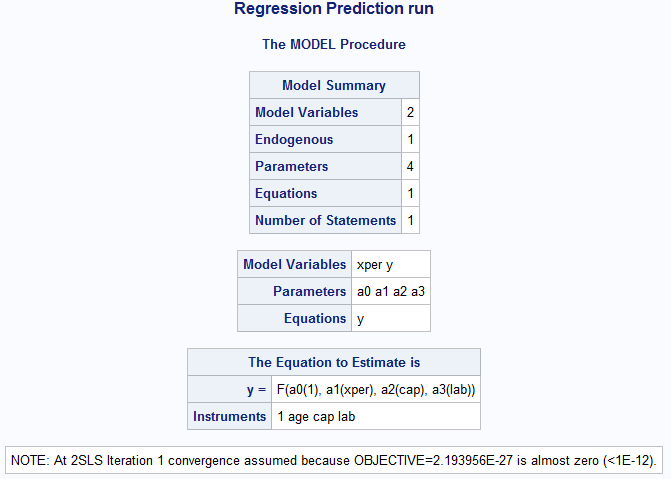
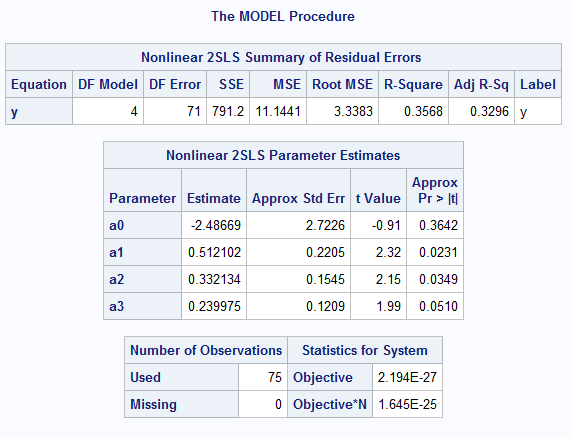
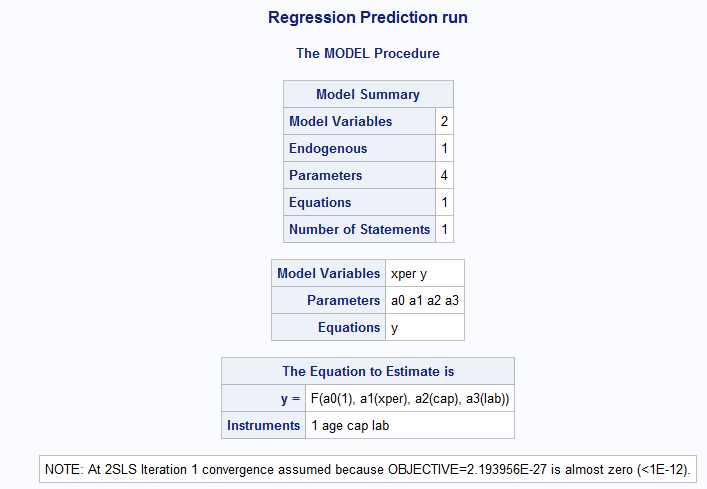
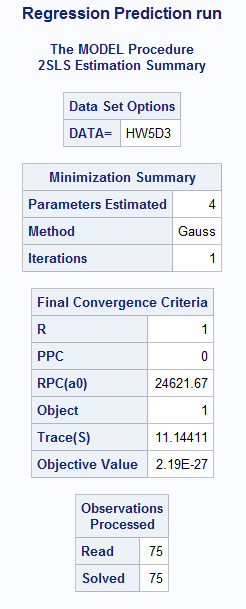


* After checking instrument, lab is not significant (P = 0.5217). This mode is identified.
* After 2SLS, all variables are significant except intercept, the sign of intercept and lab also changed. The estimates also varies.

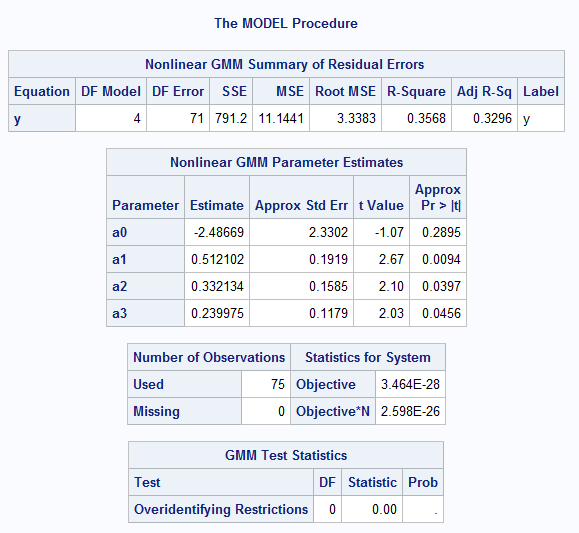
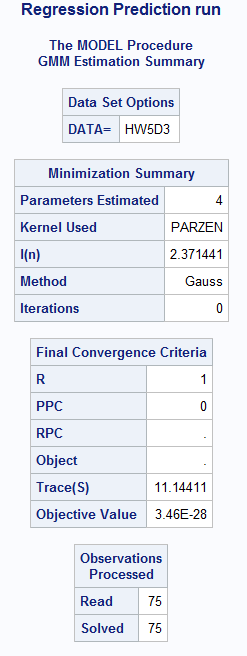
(c).

|  |
| --- |
| /\* c \*/  **proc** **model** data = hw5d3;  instruments age cap lab;  endogenous xper;  BeerProdn: y = a0 + a1\*xper + a2\*cap + a3\*lab;  fit y / **2**sls hausman;  **run**;  **proc** **model** data = hw5d3;  instruments age cap lab;  endogenous xper;  BeerProdn: y = a0 + a1\*xper + a2\*cap + a3\*lab;  fit y / **2**sls gmm;  **run**; |

I ran Hausman test and did not get the test result.



I ran GMM test.



GMM Test Statistics shows DF = 0. Hausman test might not work.

|  |
| --- |
| SAS Code:  Homework 5 Part I:  **PROC** **IMPORT** OUT= WORK.bm  DATAFILE= "C:\Users\bmishra\Dropbox\Ph.D. Courseworks\Semest  er II, Spring 2019\Econometric Methods\Homeworks\Homework 5\HW5-DATA1.xl  s"  DBMS=EXCEL REPLACE;  RANGE="SUGAR$";  GETNAMES=YES;  MIXED=NO;  SCANTEXT=YES;  USEDATE=YES;  SCANTIME=YES;  **RUN**;  **proc** **print**;  **run**;  /\* Data Management \*/  **data** bm; set bm;  lna = log(a);  lnw = log(w);  lnb = log(b);  **run**;  **proc** **print**;  **run**;  /\* Part I \*/  **proc** **reg** data = bm; /\* OLS \*/  model lna = lnb lnw / dwprob; /\* Q (a, b, d) \*/  test lnb + lnw = **0**; /\* Q (d) \*/  **run**;  **proc** **autoreg** data = bm;  model lna = lnw lnb / nlag = **1**; /\* Q (c) \*/  test lnw + lnb = **0**; /\* Q (d) \*/  output out = autoreg1 r = auregehatt;  **run**;  **proc** **autoreg** data = bm;  model lna = lnw lnb / nlag = **1**; /\* Q (c) \*/  test lnw + lnb = **0**; /\* Q (d) \*/  **run**;  Homework 5 Part II:  **PROC** **IMPORT** OUT= WORK.HW5D2  DATAFILE= "C:\Users\bmishra\Dropbox\Ph.D. Courseworks\Semest  er II, Spring 2019\Econometric Methods\Homeworks\Homework 5\HW5-DATA2.tx  t"  DBMS=TAB REPLACE;  GETNAMES=YES;  DATAROW=**2**;  **RUN**;  **proc** **print**;  **run**;  /\* Part II \*/  **data** hw5d2; set hw5d2;  lny = log(y);  lnk = log(k);  lnw = log(w);  lnl = log(l);  lnp = log(p);  lnr = log(r);  **run**;  /\* b \*/  **proc** **reg** data = hw5d2;  model lny = lnk lnl;  **run**;  /\* c: Model Procedure \*/  **proc** **model** data = hw5d2;  instruments lnp lnr lnw;  endogenous lnk lnl;  production: lny = a0 + a1\*lnk + a2\*lnl;  demandk: lnk = b0 + b1\*lnp + b2\*lny + b3\*lnr;  demandl: lnl = c0 + c1\*lnp + c2\*lny + c3\*lnw;  fit lny / **2**sls;  **run**;  /\* c: Syslin Procedure \*/  **proc** **syslin** data = hw5d2;  instruments lnp lnr lnw;  endogenous lnk lnl lny;  production: model lny = lnk lnl / overid;  demandk: model lnk = lnp lny lnr / overid;  demandl: model lnl = lnp lny lnw / overid;  **run**;  Homework 5 Part III:  **PROC** **IMPORT** OUT= WORK.HW5D3  DATAFILE= "C:\Users\bmishra\Dropbox\Ph.D. Courseworks\Semest  er II, Spring 2019\Econometric Methods\Homeworks\Homework 5\HW5-DATA3.xl  s"  DBMS=EXCEL REPLACE;  RANGE="data";  GETNAMES=YES;  MIXED=NO;  SCANTEXT=YES;  USEDATE=YES;  SCANTIME=YES;  **RUN**;  **proc** **print**;  **run**;  /\* Part III \*/  /\* a \*/  **proc** **reg** data = hw5d3;  model y = xper cap lab;  **run**;  /\* b \*/  **proc** **reg** data = hw5d3;  model xper = cap lab age;  output out = aut r = resid p = pred;  **run**;  **proc** **reg** data = aut;  model y = pred cap lab;  **run**;  /\* c \*/  **proc** **model** data = hw5d3;  instruments age cap lab;  endogenous xper;  BeerProdn: y = a0 + a1\*xper + a2\*cap + a3\*lab;  fit y / **2**sls hausman;  **run**;  **proc** **model** data = hw5d3;  instruments age cap lab;  endogenous xper;  BeerProdn: y = a0 + a1\*xper + a2\*cap + a3\*lab;  fit y / **2**sls gmm;  **run**; |