yt = α1 + α2ln(xt) + α3kt + α4at +et

yt = weekly expenditure on transport.

xt = weekly income.

kt = number of children.

at = number of adults.

ln(xt) = lnx = log(xt).

1. 𝜹y/𝜹a = α4 = αt = 10
2. 𝜹y/𝜹k = α3 = αk = 2
3. 𝜹y/𝜹 (lnx) = (𝜹y/𝜹lnx) \* (𝜹lnx/𝜹x) = αlnx \* 1/x = 0.03 (from 3 cent).

Since x = 800; αlnx \* 1/800 = 0.03;

🡪 𝜹y/(lnx) = α2 = αln(x) = 24.

Three alternative variance assumptions:

1. Var(et) = 𝜎2 (Homoscedastic) (Run regression without transforming variables).
2. Var(et) = 𝜎2 ln(xt)2(Heteroscedastic) (Transform data by dividing all variables by lnx).
3. Var(et) = 𝜎2 ln(xt)4(Heteroscedastic) (Transform data by dividing all variables by lnx2).

(a).

Null and Alternative Hypothesis under three different conditions:

(A) Ho: αa = 10; Ha: αa ≠ 10.

(B) Ho: αk = 2; Ha: αk ≠ 2.

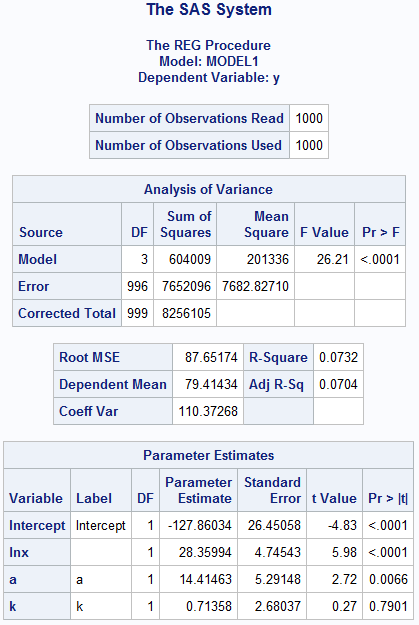
(c) Ho: αln(x) = 24; Ha: αln(x) ≠ 24.

Data Management Process:

|  |
| --- |
| **PROC** **IMPORT** OUT= WORK.bm  DATAFILE= "C:\Users\bmishra\Dropbox\Ph.D. Courseworks\Semest  er II, Spring 2019\Econometric Methods\Homeworks\Homework 4\HW4-DATA.xls"  DBMS=EXCEL REPLACE;  RANGE="hw9$";  GETNAMES=YES;  MIXED=NO;  SCANTEXT=YES;  USEDATE=YES;  SCANTIME=YES;  **RUN**;  **data** bm; set bm;  /\* Condition I \*/  y = trport;  a = a;  k = k;  x = x;  lnx = log(x);  /\* White Test \*/  a2 = a\*\***2**;  k2 = k\*\***2**;  lnx2 = lnx\*\***2**;  /\* Condition II \*/  ylnx = trport/lnx;  intlnx = **1**/lnx;  alnx = a/lnx;  klnx = k/lnx;  /\* Condition III \*/  ylnx2 = trport/lnx2;  intlnx2 = **1**/lnx2;  alnx2 = a/lnx2;  klnx2 = k/lnx2;  **run**;  **proc** **print**;  **run**; |

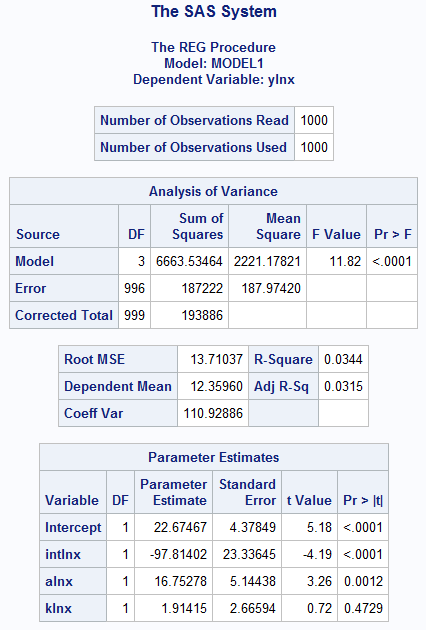
(b). Under Condition (I): Var(et) = 𝜎2

|  |
| --- |
| SAS Code:  **proc** **reg** data = bm;  model y = lnx a k;  **run**; |



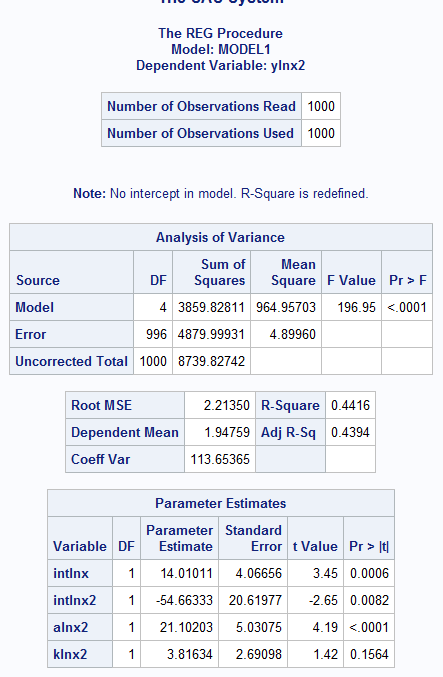
(b): Under Condition (II): Var(et) = 𝜎2 ln(xt)2

|  |
| --- |
| **proc** **reg** data = bm;  model ylnx = intlnx alnx klnx; /\* Condition II \*/  test alnx = **10**;  test klnx = **2**;  test intlnx = **24**;  **run**; |



(b): Under condition (III): Var(et) = 𝜎2 ln(xt)4

|  |
| --- |
| **proc** **reg** data = bm;  model ylnx2 = intlnx intlnx2 alnx2 klnx2 /noint; /\* Condition III \*/  test alnx2 = **10**;  test klnx2 = **2**;  test intlnx = **24**;  **run**; |



**Sensitivity Analysis:**

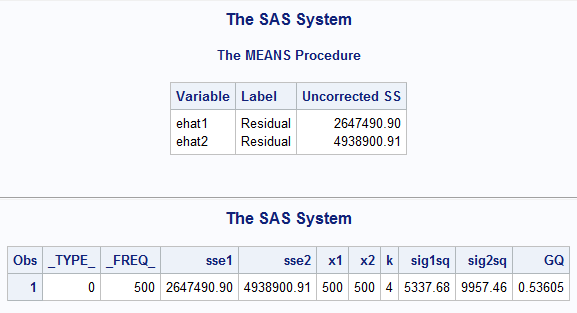
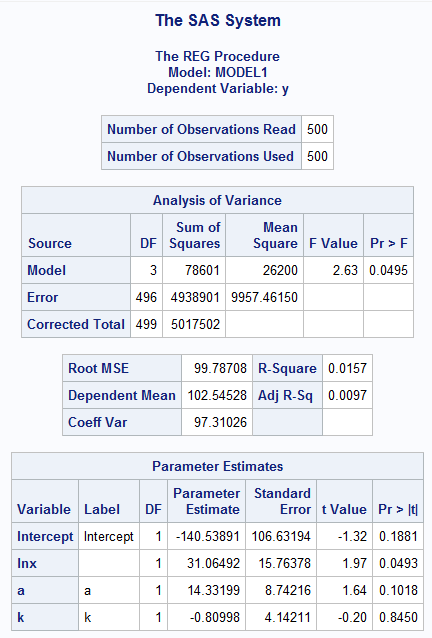
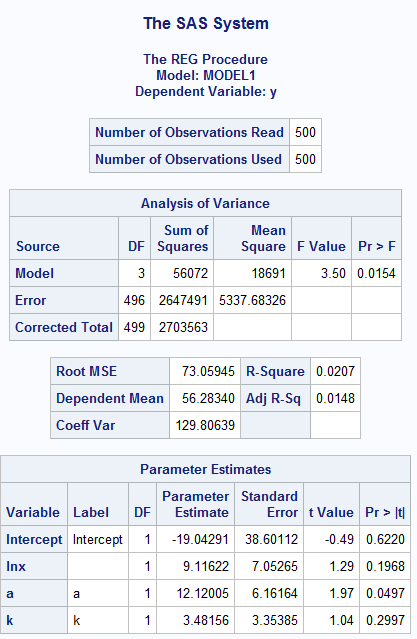
Under first condition (homoscedasticity), the intercept is negative i.e. opposite to other model and variable related to weekly income is positive in first model and negative in other two models (heteroscedasticity). Other variables are consistently positive in all three models. In rest two cases corrected log (x) for heteroscedasticity has negative sign and other variables have positive sign. The significance of variables at 1% is consistent for all variables in all three models. Magnitude wise, the weekly income, number of children and number of adults are consistent in all three models. Income has highest coefficient and number of children has lowest.

(c). Condition I: Var(et) = 𝜎2 (Homoscedastic).

|  |
| --- |
| /\*GQ Test Data Preparation \*/  **data** bm;  set bm;  **proc** **sort**;  by descending x;  **run**;  **data** bm1; set bm;/\* Newly created data = bm1 and bm1 contains first 500 cases \*/  if x le **610**;  **run**;  **data** bm2; set bm; /\* Newly created data = bm1 and bm1 contains last 500 cases \*/  if x ge **611**;  **run**;  **proc** **reg** data = bm1;  model y = lnx a k; /\* Condition I \*/ /\* Change this model for different conditions \*/  output out = out1 r = ehat1;  **run**;  **proc** **reg** data = bm2;  model y = lnx4 k a; /\* Change this model for different conditions \*/  output out = out2 r = ehat2;  **run**;  /\* G-Q Test \*/  **data** bmout;  merge out1 out2;  keep ehat1 ehat2;  **run**;  **proc** **means** uss data = bmout;  var ehat1 ehat2;  output out = out3 uss = sse1 sse2;  **run**;  **data** bmout1; set out3;  x1 = **500**; x2 = **500**; k = **4**;  sig1sq = sse1/(x1-k); sig2sq = sse2/(x2-k);  GQ = sig1sq/sig2sq;  **run**;  **proc** **print**;  **run**; |

Null Hypothesis: Heteroscedasticity does not exist i.e. homoscedastic variance.

Alternative hypothesis: Heteroscedasticity does exist.



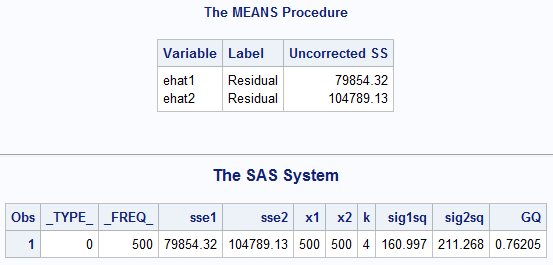
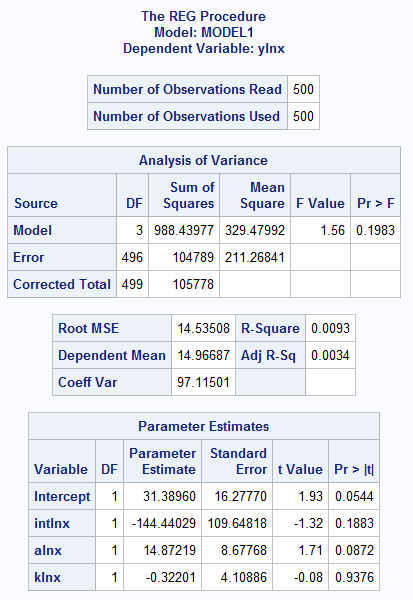
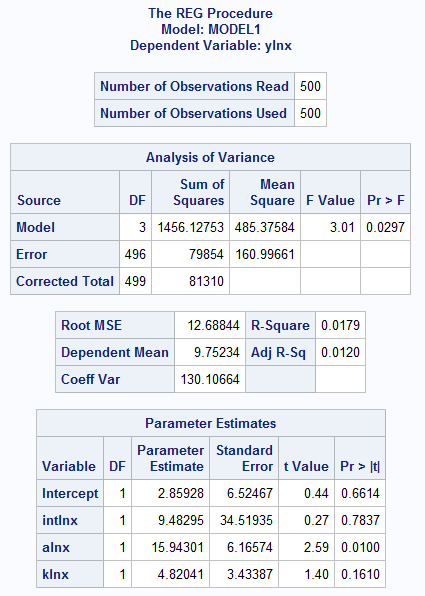
GQ = 1/0.53605 = 1.8655.

Fcric (496, 496) α= 0.05 = 1.233 (Excel: =F.INV.RT(0.01,496, 496)]

Since GQ > Fcric, We reject null hypothesis that there is no heteroscedasticity. There might be problem of heteroscedasticity. This is due to weekly income.

(c). Condition II: Var(et) = 𝜎2 ln(xt)2

|  |
| --- |
| /\*GQ Test Data Preparation \*/  **data** bm;  set bm;  **proc** **sort**;  by descending x;  **run**;  **data** bm1; set bm;/\* Newly created data = bm1 and bm1 contains first 500 cases \*/  if x le **610**;  **run**;  **data** bm2; set bm; /\* Newly created data = bm1 and bm1 contains last 500 cases \*/  if x ge **611**;  **run**;  **proc** **reg** data = bm1;  model ylnx = intlnx alnx klnx; /\* Condition II \*/ /\* Change this model for different conditions \*/  output out = out1 r = ehat1;  **run**;  **proc** **reg** data = bm2;  model ylnx = intlnx alnx klnx; /\* Condition II \*/ /\* Change this model for different conditions \*/  output out = out2 r = ehat2;  **run**;  /\* G-Q Test \*/  **data** bmout;  merge out1 out2;  keep ehat1 ehat2;  **run**;  **proc** **means** uss data = bmout;  var ehat1 ehat2;  output out = out3 uss = sse1 sse2;  **run**;  **data** bmout1; set out3;  x1 = **500**; x2 = **500**; k = **4**;  sig1sq = sse1/(x1-k); sig2sq = sse2/(x2-k);  GQ = sig1sq/sig2sq;  **run**;  **proc** **print**;  **run**; |



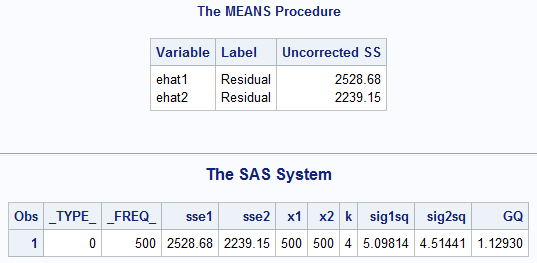
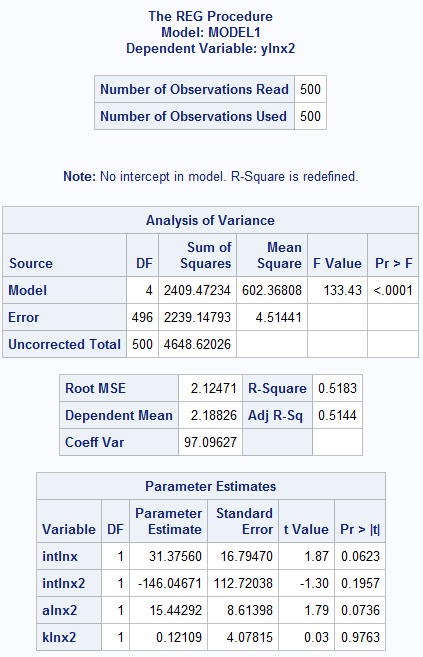
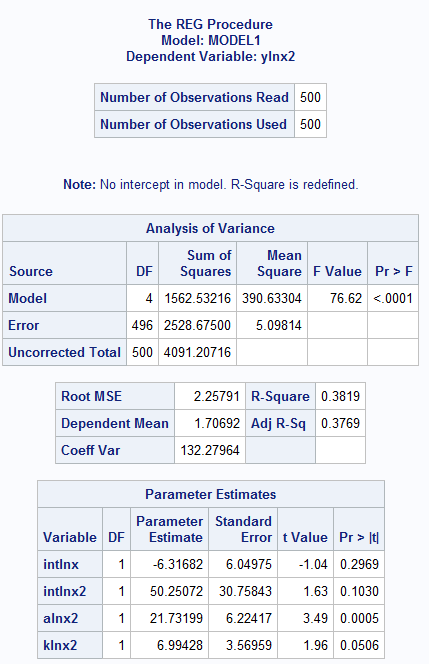
GQ = 1/0.76205 = 1.312

Fcric (496, 496) α= 0.05 = 1.233 (Excel: =F.INV.RT(0.01,496, 496)]

Since GQ > Fcric, We reject null hypothesis that there is no heteroscedasticity. There might be problem of heteroscedasticity. This is due to weekly income.

(c). Condition III: Var(et) = 𝜎2 ln(xt)4

|  |
| --- |
| /\*GQ Test Data Preparation \*/  **data** bm;  set bm;  **proc** **sort**;  by descending x;  **run**;  **data** bm1; set bm;/\* Newly created data = bm1 and bm1 contains first 500 cases \*/  if x le **610**;  **run**;  **data** bm2; set bm; /\* Newly created data = bm1 and bm1 contains last 500 cases \*/  if x ge **611**;  **run**;  **proc** **reg** data = bm1;  model ylnx2 = intlnx intlnx2 alnx2 klnx2 /noint; /\* Condition III \*/ /\* Change this model for different conditions \*/  output out = out1 r = ehat1;  **run**;  **proc** **reg** data = bm2;  model ylnx2 = intlnx intlnx2 alnx2 klnx2 /noint; /\* Condition III \*/ /\* Change this model for different conditions \*/  output out = out2 r = ehat2;  **run**;  /\* G-Q Test \*/  **data** bmout;  merge out1 out2;  keep ehat1 ehat2;  **run**;  **proc** **means** uss data = bmout;  var ehat1 ehat2;  output out = out3 uss = sse1 sse2;  **run**;  **data** bmout1; set out3;  x1 = **500**; x2 = **500**; k = **4**;  sig1sq = sse1/(x1-k); sig2sq = sse2/(x2-k);  GQ = sig1sq/sig2sq;  **run**;  **proc** **print**;  **run**; |



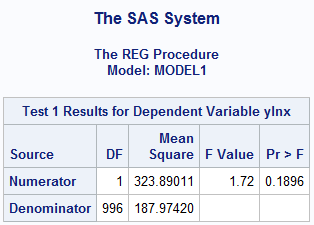
GQ = 1.12930

Fcric (496, 496) α= 0.05 = 1.233 (Excel: =F.INV.RT(0.01,496, 496)]

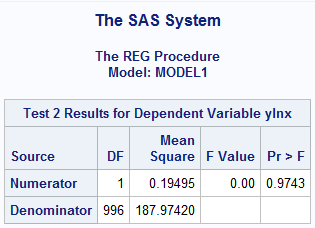
Since GQ < Fcric, We fail to reject null hypothesis. Thus, there is no heteroscedasticity i.e. homoscedasticity. The heteroscedasticity problem still appears in equations under first and second conditions.

(d): Under Condition (II): Var(et) = 𝜎2 ln(xt)2

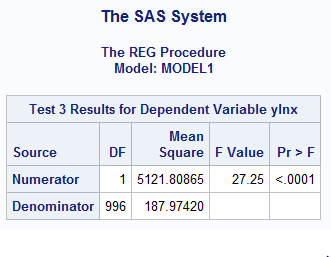
|  |
| --- |
| **proc** **reg** data = bm;  model ylnx = intlnx alnx klnx; /\* Condition II \*/  test alnx = **10**;  test klnx = **2**;  test intlnx = **24**;  **run**; |



Since p value is above 0.05, we fail to reject null hypothesis that adding an adult to a household increases household expenditure on transport by $10 per week.



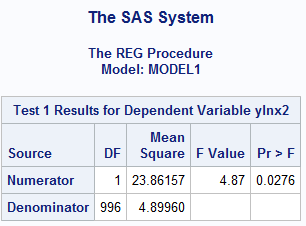
Since p value is above 0.05, we fail to reject null hypothesis that adding a child to a household increases household expenditure on transport by $2 per week.



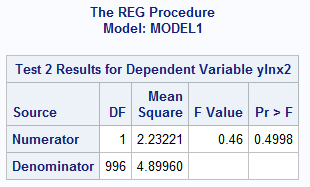
Since p value is below 0.05, we reject null hypothesis that for a household with a weekly income of $800, an incremental increase in income increases household expenditure on transport by 3 cents per one dollar.

Under condition (III): Var(et) = 𝜎2 ln(xt)4

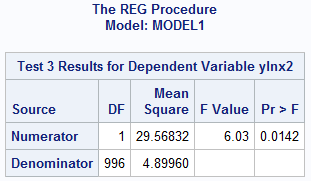
|  |
| --- |
| **proc** **reg** data = bm;  model ylnx2 = intlnx intlnx2 alnx2 klnx2 /noint; /\* Condition III \*/  test alnx2 = **10**;  test klnx2 = **2**;  test intlnx = **24**;  **run**; |



Since p value is below 0.05, we reject null hypothesis that adding an adult to a household increases household expenditure on transport by $10 per week.



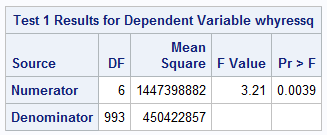
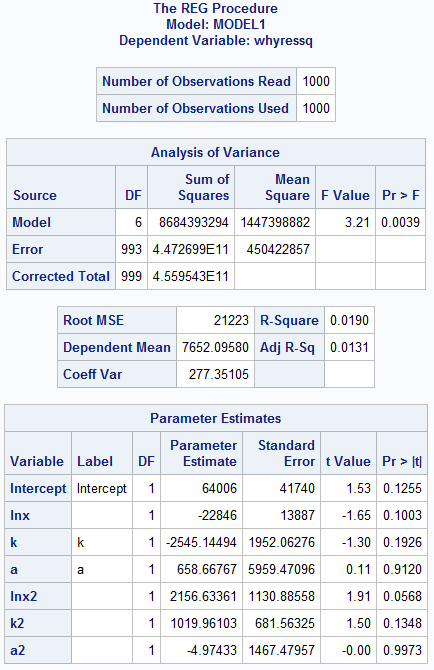
Since p value is above 0.05, we fail to reject null hypothesis that adding a child to a household increases household expenditure on transport by $2 per week.



Since p value is below 0.05, we reject null hypothesis that for a household with a weekly income of $800, an incremental increase in income increases household expenditure on transport by 3 cents per one dollar.

(e).

|  |
| --- |
| /\* White Heteroscedasticity Test \*/  **proc** **reg** data = bm;  model y = lnx a k; /\* Condition I \*/  output out = white  p = whyhatt /\* Predicted Value of dependent Variable y \*/  r = whyresid; /\*Residual values of y \*/  **run**;  **data** white; set white;  whyressq = whyresid\*\***2**;  **run**;  **proc** **reg** data = white;  model whyressq = lnx k a lnx2 k2 a2;  test lnx = k = a = lnx2 = k2 = a2 = **0**;  **run**; |



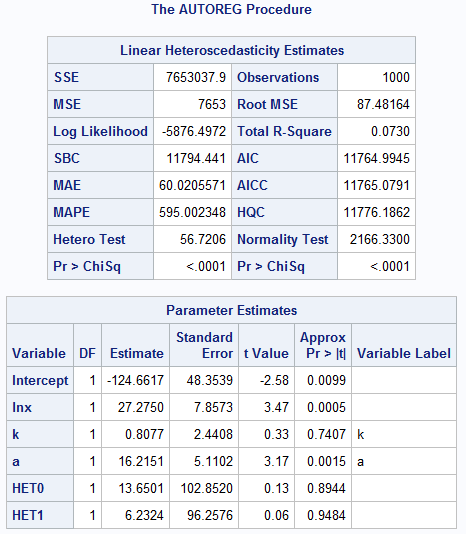
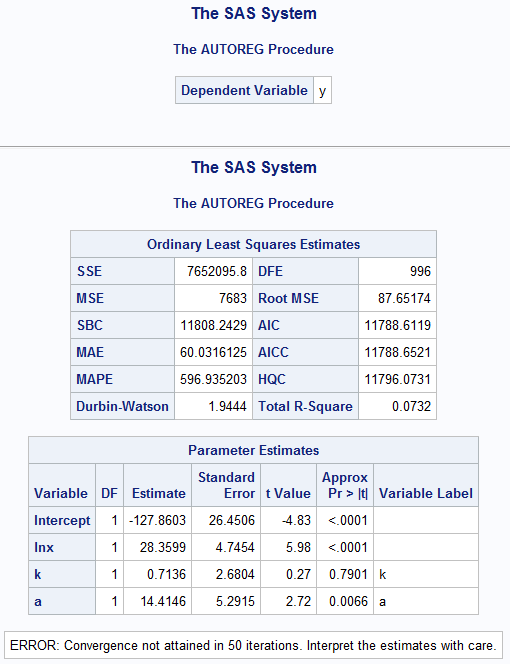
Null Hypothesis: Homoscedasticity or No Heteroscedasticity.

Alternative Hypothesis: Heteroscedasticity.

Conclusion: Since p value is below 0.05, we reject the null hypothesis that there is no homoscedasticity.

(f).

|  |
| --- |
| /\*Variance condition = Linear; Test = Lagrange Multiplier\*/  **proc** **autoreg** data = bm;  model y = lnx k a;  hetero lnx / link = linear test = lm;  **run**; |

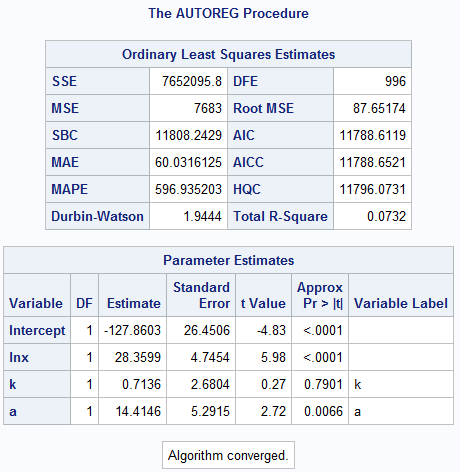
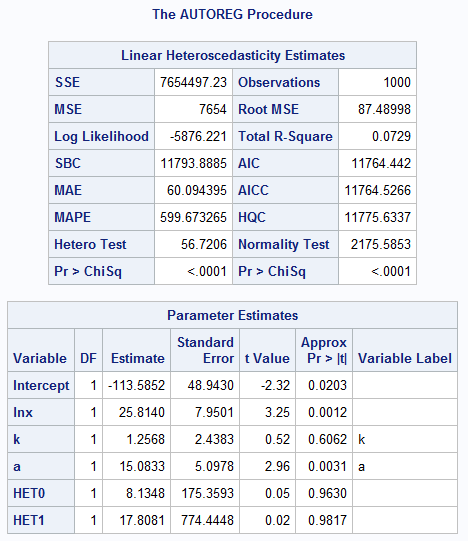


Null Hypothesis: Homoscedasticity.

Alternative Hypothesis: Heteroscedasticity.

Conclusion: Reject null hypothesis. (Looking at the p value)

|  |
| --- |
| /\*Variance condition = Linear; Test = GLS \*/  **proc** **autoreg** data = bm;  model y = lnx k a /method = ml maxiter = **1000**; /\* \*/  hetero lnx / link = linear test = lm; /\*Variance condition = Linear \*/  **run**; |



Null Hypothesis: Homoscedasticity.

Alternative Hypothesis: Heteroscedasticity.

Conclusion: Reject null hypothesis. (Looking at the p value).

|  |
| --- |
| **Complete Code Compilation: Heteroscedasticity:**  **PROC** **IMPORT** OUT= WORK.bm  DATAFILE= "C:\Users\bmishra\Dropbox\Ph.D. Courseworks\Semest  er II, Spring 2019\Econometric Methods\Homeworks\Homework 4\HW4-DATA.xls"  DBMS=EXCEL REPLACE;  RANGE="hw9$";  GETNAMES=YES;  MIXED=NO;  SCANTEXT=YES;  USEDATE=YES;  SCANTIME=YES;  **RUN**;  **data** bm; set bm;  /\* Condition I \*/  y = trport;  a = a;  k = k;  x = x;  lnx = log(x);  /\* White Test \*/  a2 = a\*\***2**;  k2 = k\*\***2**;  lnx2 = lnx\*\***2**;  /\* Condition II \*/  ylnx = trport/lnx;  intlnx = **1**/lnx;  alnx = a/lnx;  klnx = k/lnx;  /\* Condition III \*/  ylnx2 = trport/lnx2;  intlnx2 = **1**/lnx2;  alnx2 = a/lnx2;  klnx2 = k/lnx2;  **run**;  **proc** **print**;  **run**;  **proc** **reg** data = bm;  model y = lnx a k; /\* Condition I \*/  test a = **10**;  test k = **2**;  test lnx = **24**;  **run**;  **proc** **reg** data = bm;  model ylnx = intlnx alnx klnx; /\* Condition II \*/  test alnx = **10**;  test klnx = **2**;  test intlnx = **24**;  **run**;  **proc** **reg** data = bm;  model ylnx2 = intlnx intlnx2 alnx2 klnx2 /noint; /\* Condition III \*/  test alnx2 = **10**;  test klnx2 = **2**;  test intlnx = **24**;  **run**;  /\*GQ Test Data Preparation \*/  **data** bm;  set bm;  **proc** **sort**;  by descending x;  **run**;  **data** bm1; set bm;/\* Newly created data = bm1 and bm1 contains first 500 cases \*/  if x le **610**;  **run**;  **data** bm2; set bm; /\* Newly created data = bm1 and bm1 contains last 500 cases \*/  if x ge **611**;  **run**;  **proc** **reg** data = bm1;  model ylnx2 = intlnx intlnx2 alnx2 klnx2 /noint; /\* Condition III \*/ /\* Change this model for different conditions \*/  output out = out1 r = ehat1;  **run**;  **proc** **reg** data = bm2;  model ylnx2 = intlnx intlnx2 alnx2 klnx2 /noint; /\* Condition III \*/ /\* Change this model for different conditions \*/  output out = out2 r = ehat2;  **run**;  /\* G-Q Test \*/  **data** bmout;  merge out1 out2;  keep ehat1 ehat2;  **run**;  **proc** **means** uss data = bmout;  var ehat1 ehat2;  output out = out3 uss = sse1 sse2;  **run**;  **data** bmout1; set out3;  x1 = **500**; x2 = **500**; k = **4**;  sig1sq = sse1/(x1-k); sig2sq = sse2/(x2-k);  GQ = sig1sq/sig2sq;  **run**;  **proc** **print**;  **run**;  /\* E \*/ /\* White Heteroscedasticity Test \*/  **proc** **reg** data = bm;  model y = lnx a k; /\* Condition I \*/  output out = white  p = whyhatt /\* Predicted Value of dependent Variable y \*/  r = whyresid; /\*Residual values of y \*/  **run**;  **data** white; set white;  whyressq = whyresid\*\***2**;  **run**;  **proc** **reg** data = white;  model whyressq = lnx k a lnx2 k2 a2;  test lnx = k = a = lnx2 = k2 = a2 = **0**;  **run**;  **proc** **print**;  **run**;  /\* F \*/  /\*Variance condition = Linear; Test = Lagrange Multiplier\*/  **proc** **autoreg** data = bm;  model y = lnx k a;  hetero lnx / link = linear test = lm;  **run**;  /\*Variance condition = Linear; Test = GLS \*/  **proc** **autoreg** data = bm;  model y = lnx k a /method = ml maxiter = **1000**; /\* \*/  hetero lnx / link = linear test = lm; /\*Variance condition = Linear \*/  **run**; |