Quantitative methods for economics, finance and management

2019/2020

LIUC

Università Cattaneo

LAB 10

02.12.2019

LIUC

Università Cattaneo

Lab-10: Heteroskedasticity and Autocorrelation

$$y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \dots + \beta_k X_{ik} + u_i$$

u_i IS THE "CLASSICAL" ERROR TERM IF:

1) $E[u_i] = 0 \quad \forall i = 1, 2, \dots N$

2) $Var(u_i) = \sigma_u^2 = \sigma^2 \quad \forall i = 1, 2, \dots N$

3) $Cov(u_i, u_j) = 0 \quad \forall i, j$ WITH $i \neq j$

IF $Var(u_i) = \sigma_i^2 \neq \sigma_u^2 = \sigma^2$ WE HAVE HETEROSKEDASTICITY

IF $Cov(u_i\ u_j)$ OR $Cov(u_t\ u_\tau) \neq 0$ WE HAVE AUTOCORRELATION

HOW TO DEAL WITH THEM?

IN BOTH CASES WE NEED TO ADJUST THE MODEL IN ORDER TO OBTAIN THE CLASSICAL ERROR TERMS.

HETEROSK. $\rightarrow$ WE USE WLS/GLS ESTIMATOR

AUTOCORR. $\rightarrow$ WE USE QUASI-DIFFERENCES METHOD

Lab-10: Linear regression model

Dependent Variable: CONSUMPTION
Method: Least Squares
Date: 12/02/18 Time: 20:23
Sample: 1 100
Included observations: 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.668497	1.331578	4.256977	0.0000
INCOME	0.700288	0.021752	32.19466	0.0000

R-squared	0.913618	Mean dependent var	42.77938
Adjusted R-squared	0.912737	S.D. dependent var	22.56622
S.E. of regression	6.666147	Akaike info criterion	6.651759
Sum squared resid	4354.877	Schwarz criterion	6.703862
Log likelihood	-330.5879	Hannan-Quinn criter.	6.672846
F-statistic	1036.496	Durbin-Watson stat	2.566340
Prob(F-statistic)	0.000000		

- $C_i = \beta_0 + \beta_1 I_i + u_i$
- $\hat{\beta}_0 = 5,66$ AND $\hat{\beta}_1 = 0,70$
- SPECIFICATION IS IN "LEVEL-LEVEL"
- $\hat{\beta}_1 = 0,70$
- IF INCOME INCREASES BY 1 \$ (1 UNIT) THEN CONSUMPTION RISES BY 0,70 \$ (KEEPING THE OTHER VARIABLES CONSTANT)

NOW, SUPPOSE A "LOG-LOG" SPECIFICATION; THIS IMPLIES THAT $\hat{\beta}_0$ AND $\hat{\beta}_1$ HAVE ELASTICITY INTERPRETATION. EXAMPLE:
IF INCOME INCREASES BY 10% THEN CONSUMPTION INCREASES BY 7% (KEEPING THE OTHER VARIABLES CONSTANT)

Lab-10: Linear regression model

- $R^2 = 0,91 \cong 91\%$; MEANING THAT 91,36% OF THE TOTAL VARIATION OF THE DEPENDENT VARIABLE (CONSUMPTION) IS EXPLAINED BY THE MODEL (VARIATION OF THE FITTED VALUES)

- $H_0: \beta_1 = 0$ VS $H_1: \beta_1 \neq 0$

$$t_{CALC}(\beta_1) = \left| \frac{\hat{\beta}_1 - \beta_1 / H_0}{Se(\hat{\beta}_1)} \right| = \frac{\hat{\beta}_1 - 0}{SE(\hat{\beta}_1)} = \frac{0,700288}{0,021752} = 32,19$$

$$|t(\beta_1)| > t_{CRITICAL}^{\alpha = 5\%} \Rightarrow 32,19 > 1,96 \cong 2$$; H_0 IS REJECTED AT 5% (WE)

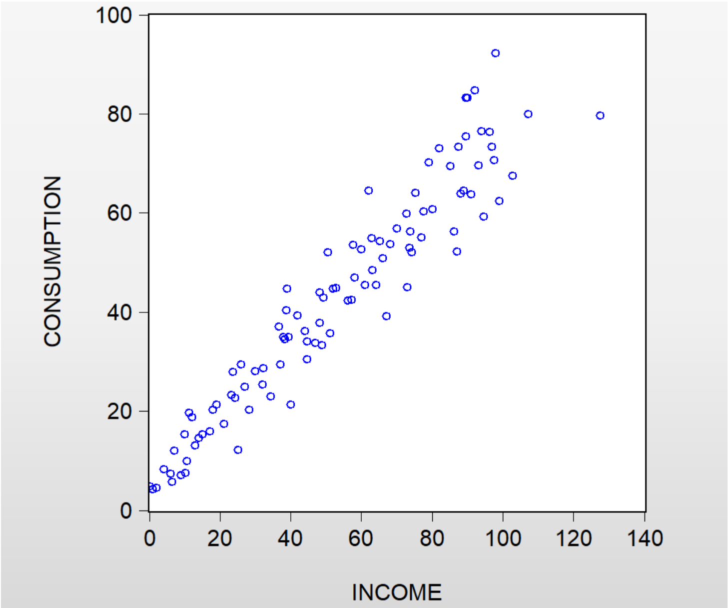
- $r_{CI} = \rho_{CI} = 0,95 \Rightarrow$ STRONG LINEAR RELATIONSHIP BETWEEN C_i AND I_i

Lab-10: Heteroskedasticity

- Considering the consumption-income model. Thus, the reference equation is

$$C_i = \beta_1 + \beta_2 I_i + u_i$$

- The scatterplot between consumption and income shows the empirical relationship between the actual values.



- There is empirical evidence that the consumption variability increases with the level of income. This fact suggests the existence of heteroskedaticity.
- The condition in which some observations have a larger (or smaller) variance than others describes what is usually called heteroscedasticity phenomena.

Lab-10: The White Test

- The White test is based on a different auxiliary regression where the squared residuals are regressed on: **the model explanatory variables all their squares.**

1. The asymptotic chi-squared White test-statistic is obtained by the product of the number of observations times the R-squared of the auxiliary regression.
2. The F-version for small samples is obtained by setting to zero all the explanatory variables of the auxiliary regression (i.e. by looking at the F-test for the overall significance of the auxiliary regression).

→ $y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \cdots + \beta_k X_{ik} + u_i$; $Var(u_i) = \sigma_i^2$ → THE VARIANCE CHANGES ACROSS INDIVIDUALS

TEST THE PRESENCE OF HETER. WE NEED A TEST EQUATION, THAT IS:

$\sigma_i^2 = \gamma_1 \cdot 1 + \gamma_2 X_{i2} + \gamma_3 X_{i3} + \cdots + \gamma_k X_{ik} + \beta_2 X_{i2}^2 + \beta_3 X_{i3}^2 + \cdots + \beta_k X_{ik}^2$

→ IT IS NOT KNOWN; FOR THIS REASON WE NEED A PROXY FOR σ_i^2; THAT IS $\hat{u}_i^2$

WHERE $\hat{u}_i$ IS THE RESIDUAL OF THE MAIN EQUATION

- IN PRACTICE WE NEED TO FOLLOW THESE 4 STEPS:

1) USE OLS TO ESTIMATE $\beta_1, \beta_2 \ldots$; THAT IS TO GET $\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_K$

Lab-10: The White Test

2) $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_k X_{ik}$

3) $\hat{u}_i = y_i - \hat{y}_i$ (NOTICE THAT $y_i = \hat{y} + \hat{u}_i$)

y_i (ACTUAL VALUES) $\hat{y}_i$ (FITTED VALUES)

4) $\hat{u}_i^2 \quad \forall i = 1, 2, \dots N$; THUS FROM THE FOLLOWING TEST EQUATION:

$$\sigma_i^2 = \gamma_1 + \gamma_2 X_{i2} + \dots + \gamma_k X_{ik} + \rho_2 X_{i2}^2 + \dots + \rho_k X_{ik}^2$$

WE MOVE TO THE AUXILLARY REGRESSION, THAT IS:

$$\hat{u}_i^2 = \gamma_1 + \gamma_2 X_{i2} + \dots + \gamma_k X_{ik} + \rho_2 X_{i2}^2 + \dots + \rho_k X_{ik}^2 + \varepsilon_i$$

↳ SINCE $\hat{u}_i^2$ IS A PROXY FOR σ_i^2 WE ARE ADDING AN ERROR TERM

- WE HAVE 2 WAYS TO CONSTRUCT THE WHITE TEST:

1) USE $N \cdot R_A^2 \overset{H_0}{\sim} \chi^2(k-1)$ AND WE CHECK THE P-VALUE

2) USE THE F-TEST

Lab-10: The White Test

THE TEST TAKES THE FOLLOWING FORM

$$\hat{u}_i^2 = \gamma_1 + \gamma_2 X_{i2} + \dots + \gamma_k X_{ik} + \rho_2 X_{i2}^2 + \dots + \rho_k X_{ik}^2 + \varepsilon_i$$

$H_0: \gamma_1 = \gamma_2 = \dots = \gamma_k = \rho_2 = \rho_3 = \dots = \rho_k = 0$

H_1: AT LEAST ONE COEFFICIENT DIFFERENT FROM ZERO

- THE WHITE TEST IS THE F-TEST $\sim F(G, N-(2(k-1)+1))$
  WHERE $G = 2(k-1) \Rightarrow$ # OF RESTRICTIONS
- $C_i = \beta_1 + \beta_2 I_i + u_i$; $Var(u_i) = \sigma_i^2 \neq \sigma^2 \Rightarrow$ OLS ESTIMATES IS NOT BLUE.

Equation: EQ2WAUXREGR Workfile: QMEFM_LAB10_021219::heter...

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: UHAT2
Method: Least Squares
Date: 12/02/19 Time: 15:07
Sample: 1 100
Included observations: 100

AUXILIARY REGRESSION

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	18.40369	17.58896	1.046321	0.2980
INCOME	-0.402642	0.729608	-0.551861	0.5823
INCOME^2	0.012403	0.006527	1.900252	0.0604

R-squared	0.210069	Mean dependent var	43.54877
Adjusted R-squared	0.193781	S.D. dependent var	67.52709
S.E. of regression	60.63236	Akaike info criterion	11.07708
Sum squared resid	356599.4	Schwarz criterion	11.15523
Log likelihood	-550.8538	Hannan-Quinn criter.	11.10871
F-statistic	12.89773	Durbin-Watson stat	1.995534
Prob(F-statistic)	0.000011		

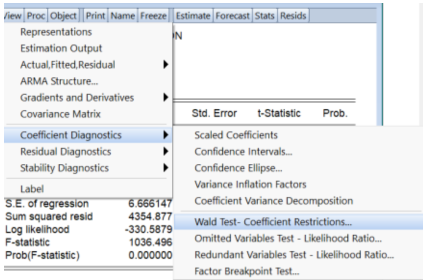
THE TEST EQ. BECOMES

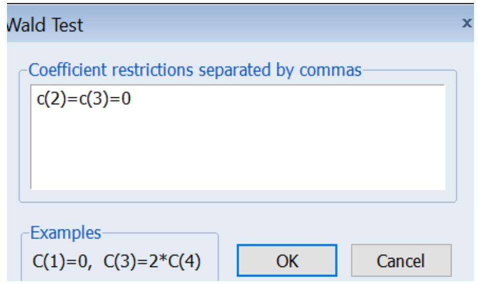
$$\hat{u}_i^2 = \gamma_1 + \gamma_2 I_i + \rho_2 I_i^2 + \varepsilon_i$$

$H_0: \gamma_2 = \rho_2 = 0$

$H_1: \exists$ AT LEAST 1 COEFF. DIFFERENT FROM ZERO

Lab-10: The White Test





Wald Test:
Equation: EQ2WAUXREGR

Test Statistic	Value	df	Probability
F-statistic	12.89773	(2, 97)	0.0000
Chi-square	25.79547	2	0.0000

Null Hypothesis: C(2)=C(3)=0
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(2)	-0.402642	0.729608
C(3)	0.012403	0.006527

Restrictions are linear in coefficients.

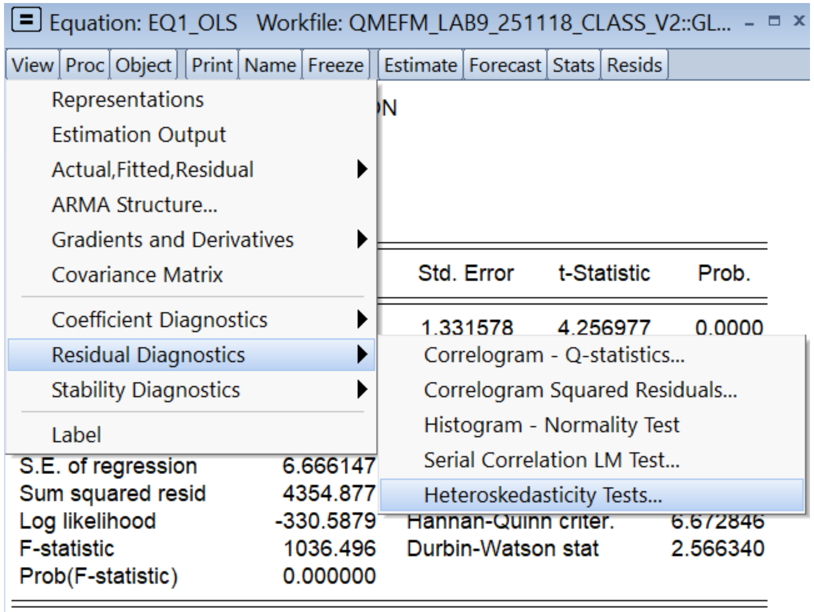
F-TEST = 12,89 = WHITE TEST

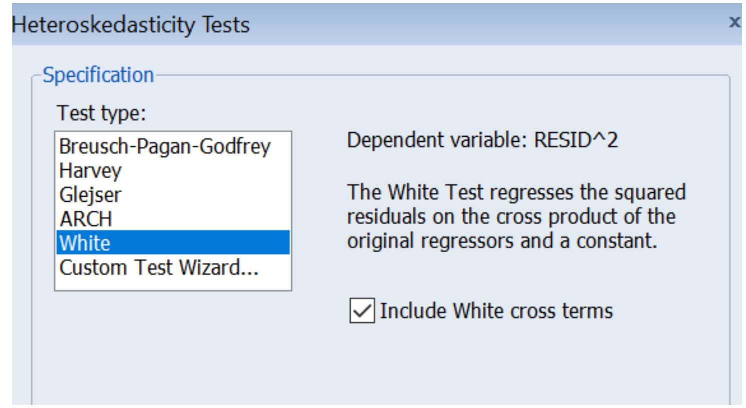
P-VALUE WHITE TEST = 0.00000... < { 1% ; 5% ; 10% } LEVELS ; THEREFORE :

Ho IS REJECTED AT 1%, 5% AND 10% LEVELS.

WE HAVE HETEROSK. IN OUR MODEL.

Lab-10: The White Test





Heteroskedasticity Test: White

F-statistic	12.89773	Prob. F(2,97)	0.0000
Obs*R-squared	21.00686	Prob. Chi-Square(2)	0.0000
Scaled explained SS	24.01169	Prob. Chi-Square(2)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 12/02/18 Time: 21:49
Sample: 1 100
Included observations: 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	18.40369	17.58896	1.046321	0.2980
INCOME^2	0.012403	0.006527	1.900252	0.0604
INCOME	-0.402642	0.729608	-0.551861	0.5823

R-squared	0.210069	Mean dependent var	43.54877
Adjusted R-squared	0.193781	S.D. dependent var	67.52709
S.E. of regression	60.63236	Akaike info criterion	11.07708
Sum squared resid	356599.4	Schwarz criterion	11.15523
Log likelihood	-550.8538	Hannan-Quinn criter.	11.10871
F-statistic	12.89773	Durbin-Watson stat	1.995534
Prob(F-statistic)	0.000011		