Семінар 10. Регресія з гетероскедастичними збуреннями

Виявлення гетероскедастичності

Критерії виявлення гетероскедастичності поділяються на дві групи: загальні та регресійні.

Загальні критерії виявлення гетероскедастичності

Загальні критерії відрізняються тим, що при їх формуванні не використовуються припущення про характер гетероскедастичності. В цьому полягає їх перевага. Недоліком є те, що такі критерії лише виявляють наявність гетероскедастичності, але не дають інформації для розв'язання проблеми. В цьому початковому курсі ми розглянемо лише один з цієї групи критеріїв, а саме критерій Голдфелда-Квондта.

Критерій Голдфелда-Квондта

Цей критерій використовують тоді, коли всі наявні спостереження можна розбити за деякою ознакою на дві групи. У випадку однієї незалежної змінної спостереження з найменшими значеннями можуть складати одну групу, а другу – спостереження з найбільшими значеннями незалежної змінної. Розбиття можна робити також за значеннями залежної змінної.

Нехай сукупність n спостережень розбита на дві групи об'ємами n_1 і n_2. Частину спостережень з середніми значеннями можна виключити. В цьому випадку $n_1 + n_2 < n$. Для того щоб застосувати критерій Голдфельда-Квондта, необхідно оцінити модель за методом найменших квадратів окремо на кожній підвиборці і знайти

$\hat{\sigma}_1^2$ – оцінку дисперсії збурень за першою групою спостережень та

$\hat{\sigma}_2^2$ – оцінку дисперсії збурень за другою групою спостережень.

У припущенні, що гетероскедастичність відсутня, статистика

$$F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} \qquad (1)$$

має розподіл Фішера з $n_1 - k$, $n_2 - k$ ступенями свободи.

Перевірка гіпотези виконується таким чином.

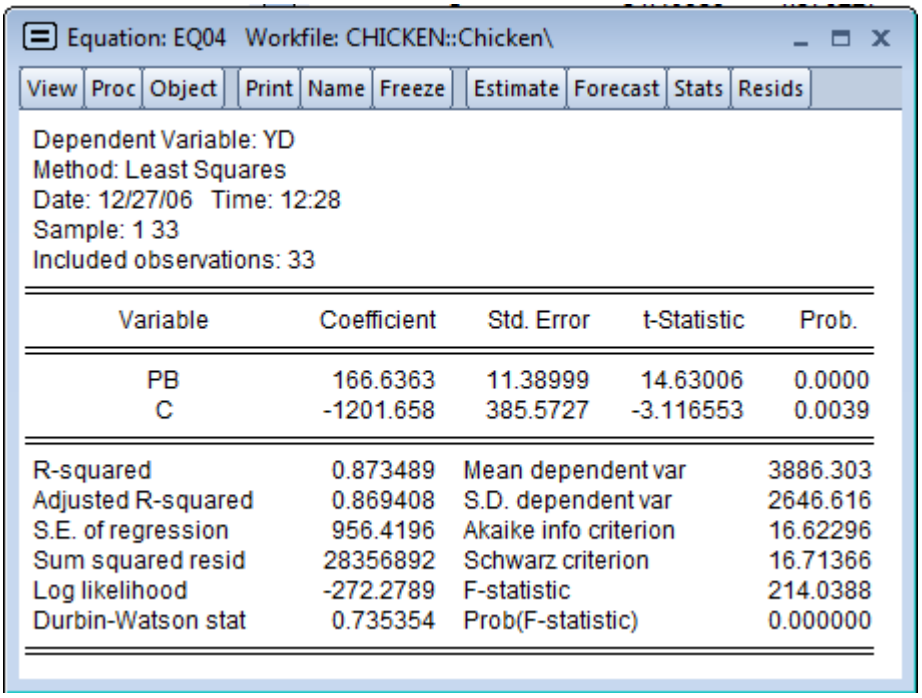
Якщо $\hat{\sigma}_1^2 > \hat{\sigma}_2^2$, то обчислюють статистику (1) і порівнюють її з критичним значенням $F_{кр}(\alpha, n_1 - k, n_2 - k)$, знайденим за вибраним рівнем значущості α в таблиці розподілу Фішера з $n_1 - k$, $n_2 - k$ ступенями свободи. Якщо $\hat{\sigma}_1^2 < \hat{\sigma}_2^2$, то обчислюють статистику

$$F = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2} \tag{2}$$

і порівнюють її з критичним значенням $F_{кр}(\alpha, n_2 - k, n_1 - k)$, знайденим за вибраним рівнем значущості α в таблиці розподілу Фішера з $n_2 - k$, $n_1 - k$ ступенями свободи.

Далі значення F-статистики (1) або (2) порівнюють з табличним. Якщо $F < F_{кр}$, то вважають, що гетероскедастичність відсутня. Якщо $F \geq F_{кр}$, то вважають, що гетероскедастичність має місце.

Нехай необхідно перевірити на наявність гетероскедастичності критерієм Голдфелда-Квондта наступну регресію[1]:



Equation: EQ04 Workfile: CHICKEN::Chicken\

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

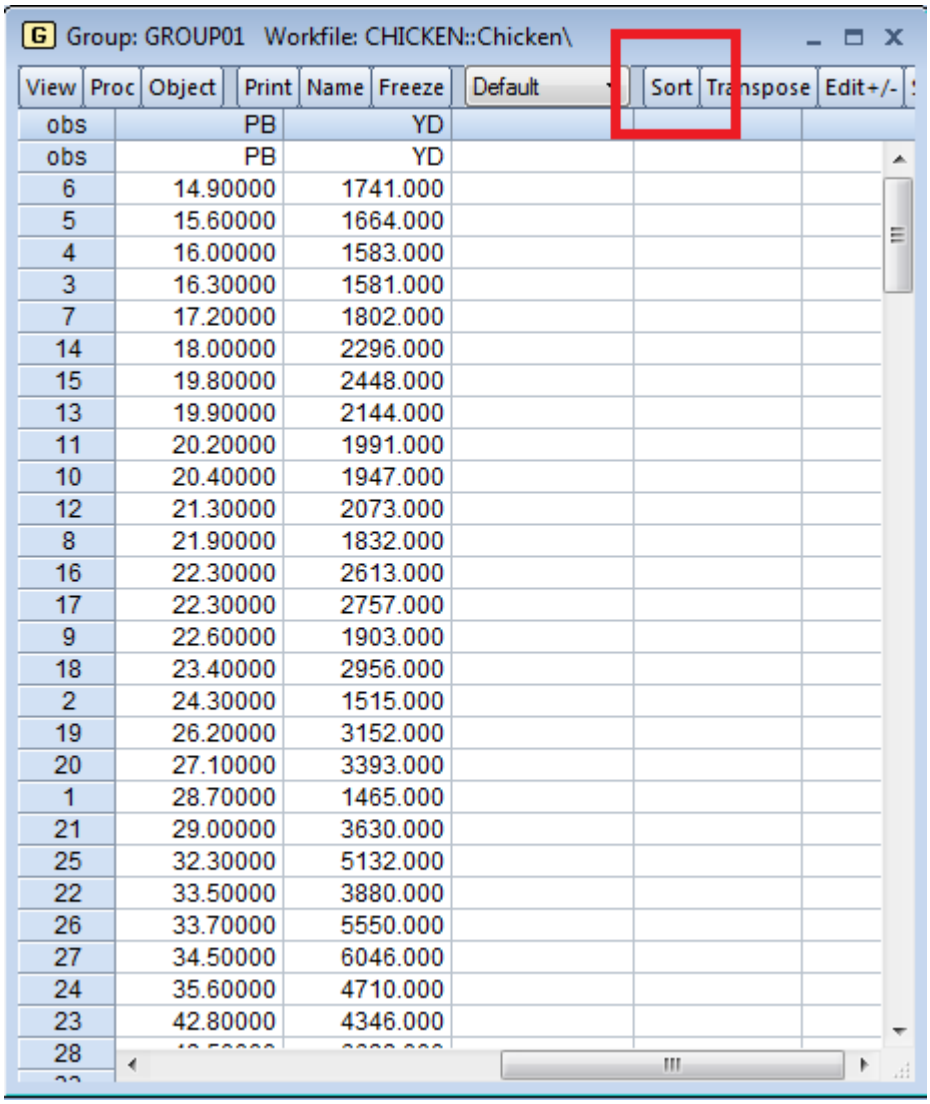
Dependent Variable: YD
Method: Least Squares
Date: 12/27/06 Time: 12:28
Sample: 1 33
Included observations: 33

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PB	166.6363	11.38999	14.63006	0.0000
C	-1201.658	385.5727	-3.116553	0.0039

R-squared	0.873489	Mean dependent var	3886.303
Adjusted R-squared	0.869408	S.D. dependent var	2646.616
S.E. of regression	956.4196	Akaike info criterion	16.62296
Sum squared resid	28356892	Schwarz criterion	16.71366
Log likelihood	-272.2789	F-statistic	214.0388
Durbin-Watson stat	0.735354	Prob(F-statistic)	0.000000

[1] Файл chicken.wf1

Розбиваємо сукупність спостережень на дві групи та проводимо сортування за залежною змінною. Для цього створюємо дві нові змінні YD1 та PB1, заповнюючи їх нульовими значеннями. Відкриваємо змінні YD та PB як групу, сортуємо їх за змінною PB (незалежна змінна),та копіюємо результат до нової групи змінних YD1 та PB1 (для копіювання необхідно включити режим редагування – кнопка Edit):



obs	PB	YD
obs	PB	YD
6	14.90000	1741.000
5	15.60000	1664.000
4	16.00000	1583.000
3	16.30000	1581.000
7	17.20000	1802.000
14	18.00000	2296.000
15	19.80000	2448.000
13	19.90000	2144.000
11	20.20000	1991.000
10	20.40000	1947.000
12	21.30000	2073.000
8	21.90000	1832.000
16	22.30000	2613.000
17	22.30000	2757.000
9	22.60000	1903.000
18	23.40000	2956.000
2	24.30000	1515.000
19	26.20000	3152.000
20	27.10000	3393.000
1	28.70000	1465.000
21	29.00000	3630.000
25	32.30000	5132.000
22	33.50000	3880.000
26	33.70000	5550.000
27	34.50000	6046.000
24	35.60000	4710.000
23	42.80000	4346.000
28		

Group: GROUP02 Workfile: CHICKEN::Chicken\

View | Proc | Object | Print | Name | Freeze | Default | Sort | Transpose | Edit+/- | Smpl

14.9

obs	PB1	YD1
obs	PB1	YD1
1	14.90000	1741.000
2	15.60000	1664.000
3	16.00000	1583.000
4	16.30000	1581.000
5	17.20000	1802.000
6	18.00000	2296.000
7	19.80000	2448.000
8	19.90000	2144.000
9	20.20000	1991.000
10	20.40000	1947.000
11	21.30000	2073.000
12	21.90000	1832.000
13	22.30000	2613.000
14	22.30000	2757.000
15	22.60000	1903.000
16	23.40000	2956.000
17	24.30000	1515.000
18	26.20000	3152.000
19	27.10000	3393.000
20	28.70000	1465.000
21	29.00000	3630.000
22		

Далі оцінюємо регресії, розбивши всі спостереження на дві групи (близько 40% спостережень у кожній):

Equation: UNTITLED Workfile: CHICKEN::Chicken\

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

Dependent Variable: YD1
Method: Least Squares
Date: 08/27/13 Time: 16:26
Sample: 1 13
Included observations: 13

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	361.8733	544.7097	0.664342	0.5202
PB1	86.17985	28.80199	2.992149	0.0123

R-squared	0.448703	Mean dependent var	1978.077
Adjusted R-squared	0.398585	S.D. dependent var	327.0641
S.E. of regression	253.6412	Akaike info criterion	14.05036
Sum squared resid	707672.5	Schwarz criterion	14.13727
Log likelihood	-89.32732	Hannan-Quinn criter.	14.03249
F-statistic	8.952957	Durbin-Watson stat	1.431484
Prob(F-statistic)	0.012251		

Equation: UNTITLED Workfile: CHICKEN::Chicken\

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

Dependent Variable: YD1
Method: Least Squares
Date: 08/27/13 Time: 16:28
Sample: 21 33
Included observations: 13

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-94.74948	1353.960	-0.069980	0.9455
PB1	147.0193	28.77811	5.108717	0.0003

R-squared	0.703496	Mean dependent var	6568.615
Adjusted R-squared	0.676541	S.D. dependent var	2303.086
S.E. of regression	1309.845	Akaike info criterion	17.33384
Sum squared resid	18872623	Schwarz criterion	17.42076
Log likelihood	-110.6700	Hannan-Quinn criter.	17.31598
F-statistic	26.09899	Durbin-Watson stat	1.275689
Prob(F-statistic)	0.000339		

Знаходимо:

$\hat{\sigma}_1^2$=707672,5/(13-2)= 64333,86

$\hat{\sigma}_2^2$=18872623/(13-2)= 1715693

$F = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2}$=1715693/64333,86=26,67 (оскільки $\hat{\sigma}_2^2$ більше за $\hat{\sigma}_1^2$).

Порівнюємо розрахункове значення з критичним (рівень значущості 0,95), яке дорівнює 2,83 (для рівня надійності 0,95 та 11 і 11 степенів свободи). Розрахункове значення F-статистики більше за критичне. Таким чином, можна зробити висновок, що гетероскедастичність присутня.

Регресійні критерії виявлення гетероскедастичності

Регресійні критерії гетероскедастичності будуються на основі припущення, що дисперсія пропорційна функції від деякої відомої змінної:

$$w_i^2 = f(z_i), i = \overline{1,n}.$$

Критерій Глейзера

Застосування цього критерію розглянемо на прикладі моделі

$$y_i = \alpha + \beta x_i + \upsilon_i, i = \overline{1,n}. \quad (3)$$

На 1-му етапі модель (3) оцінюємо за методом найменших квадратів і знаходимо залишки u_i, $i = \overline{1,n}$. На 2-му етапі будуємо регресію модуля залишків відносно однієї з таких функцій:

$$|u_i| = \gamma + \delta x_i + \varepsilon_i, \tag{4}$$

$$|u_i| = \gamma + \delta \sqrt{x_i} + \varepsilon_i, \tag{5}$$

$$|u_i| = \gamma + \delta \frac{1}{x_i} + \varepsilon_i, \tag{6}$$

$$|u_i| = \gamma + \delta \frac{1}{\sqrt{x_i}} + \varepsilon_i. \tag{7}$$

Зауважимо, що замість змінної x можна використати іншу змінну, яка вибирається, як правило, з економічних міркувань. Будується послідовно декілька таких регресій. Далі оцінюємо коефіцієнти регресій (4-7) і вибираємо з них ту, яка має найбільший коефіцієнт детермінації .

На 3-му етапі перевіряємо гіпотезу про значущість обраної моделі. Якщо дана модель з є значущою, то збурення в моделі (3) гетероскедастичні.

Отже, нехай необхідно перевірити наявність гетероскедастичності в наступній моделі:

Equation: EQ04 Workfile: CHICKEN::Chicken\

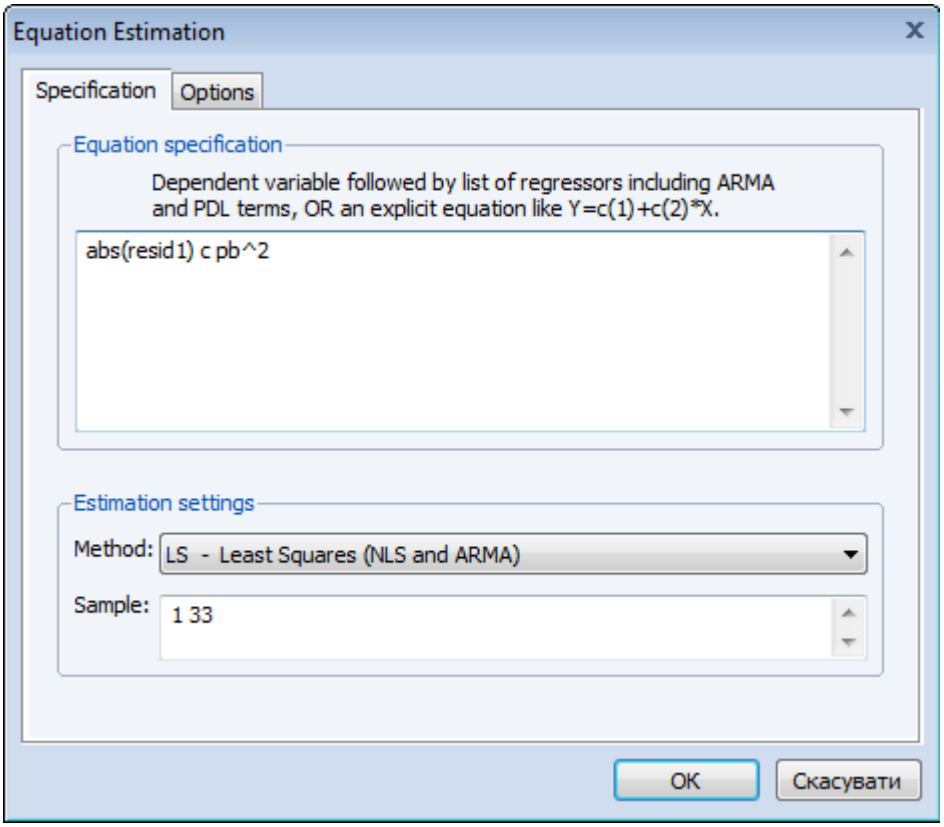
View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: YD
Method: Least Squares
Date: 12/27/06 Time: 12:28
Sample: 1 33
Included observations: 33

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PB	166.6363	11.38999	14.63006	0.0000
C	-1201.658	385.5727	-3.116553	0.0039

R-squared	0.873489	Mean dependent var	3886.303
Adjusted R-squared	0.869408	S.D. dependent var	2646.616
S.E. of regression	956.4196	Akaike info criterion	16.62296
Sum squared resid	28356892	Schwarz criterion	16.71366
Log likelihood	-272.2789	F-statistic	214.0388
Durbin-Watson stat	0.735354	Prob(F-statistic)	0.000000

Для цього побудуємо біля десяти допоміжних регресій (створивши нову змінну *resid1*, в яку скопіювали значення залишків моделі) наступного вигляду (приклад однієї з побудованих моделей):



Серед побудованих моделей обираємо ту, що має найбільший коефіцієнт детермінації. Перевіряємо її на адекватність:

Equation: UNTITLED Workfile: CHICKEN::Chicken\

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

Dependent Variable: ABS(RESID1)
Method: Least Squares
Date: 08/27/13 Time: 16:51
Sample: 1 33
Included observations: 33

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-180.5228	222.9040	-0.809868	0.4242
PB	27.03478	6.584681	4.105708	0.0003

R-squared	0.352235	Mean dependent var	644.9392
Adjusted R-squared	0.331339	S.D. dependent var	676.1716
S.E. of regression	552.9170	Akaike info criterion	15.52698
Sum squared resid	9477232.	Schwarz criterion	15.61768
Log likelihood	-254.1952	Hannan-Quinn criter.	15.55750
F-statistic	16.85684	Durbin-Watson stat	1.530452
Prob(F-statistic)	0.000272		

Дана модель є значущою при рівні значущості 0,05, отже гетероскедастичність в початковій моделі присутня.

Критерій Уайта

Нехай, ми маємо модель

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_{k-1} x_{i,k-1} + \upsilon_i, i = \overline{1,n}. \qquad (8)$$

На 1-му етапі модель (8) оцінюємо за методом найменших квадратів і знаходимо залишки u_i, $i = \overline{1,n}$. На 2-му етапі будуємо регресію квадратів залишків відносно всіх змінних з моделі (8), їх квадратів та попарних добутків. Наприклад, якщо модель (8) має вигляд

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \upsilon_i, i = \overline{1,n}, \qquad (9)$$

то на 2-му етапі будуємо наступну регресію:

$$u_i^2 = \gamma_0 + \gamma_1 x_{i1} + \gamma_2 x_{i2} + \gamma_3 x_{i3} + \gamma_4 x_{i4} + \gamma_5 x_{i1}^2 + \gamma_6 + \gamma_7 x_{i3}^2 + \gamma_8 x_{i4}^2 + \gamma_9 x_{i1} x_{i2} +$$

$$+ \gamma_{10} x_{i1} x_{i3} + \gamma_{11} x_{i1} x_{i4} + \gamma_{12} x_{i2} x_{i3} + \gamma_{13} x_{i2} x_{i4} + \gamma_1 x_{i3} x_{i4} + \varepsilon_i, \ i = \overline{1,n}. \qquad (10)$$

На 3-му етапі перевіряємо гіпотезу про значущість моделі (10). Якщо модель (10) є значущою, то збурення в моделі (9) гетероскедастичні.

Отже, нехай ми маємо наступне регресійне рівняння:

Equation: EQ01 Workfile: CHICKEN::Chicken\

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

Dependent Variable: Y
Method: Least Squares
Date: 08/27/13 Time: 21:20
Sample: 1 33
Included observations: 33

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	31.40959	1.376227	22.82297	0.0000
YD	0.001839	0.000405	4.538711	0.0001
PB	0.247457	0.070428	3.513599	0.0015
PC	-0.819809	0.089305	-9.179902	0.0000

R-squared	0.963632	Mean dependent var	35.87879
Adjusted R-squared	0.959870	S.D. dependent var	9.927763
S.E. of regression	1.988782	Akaike info criterion	4.326134
Sum squared resid	114.7024	Schwarz criterion	4.507529
Log likelihood	-67.38122	Hannan-Quinn criter.	4.387168
F-statistic	256.1347	Durbin-Watson stat	0.753680
Prob(F-statistic)	0.000000		

Для перевірки наявності чи відсутності гетероскедастичності будуємо наступну регресію:

Equation Estimation

Specification | Options

Equation specification

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like Y=c(1)+c(2)*X.

resid^2 c yd pb pc yd^2 pb^2 pc^2 yd*pb yd*pc pb*pc

Estimation settings

Method: LS - Least Squares (NLS and ARMA)

Sample: 1 33

OK | Скасувати

Equation: UNTITLED Workfile: CHICKEN::Chicken\

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

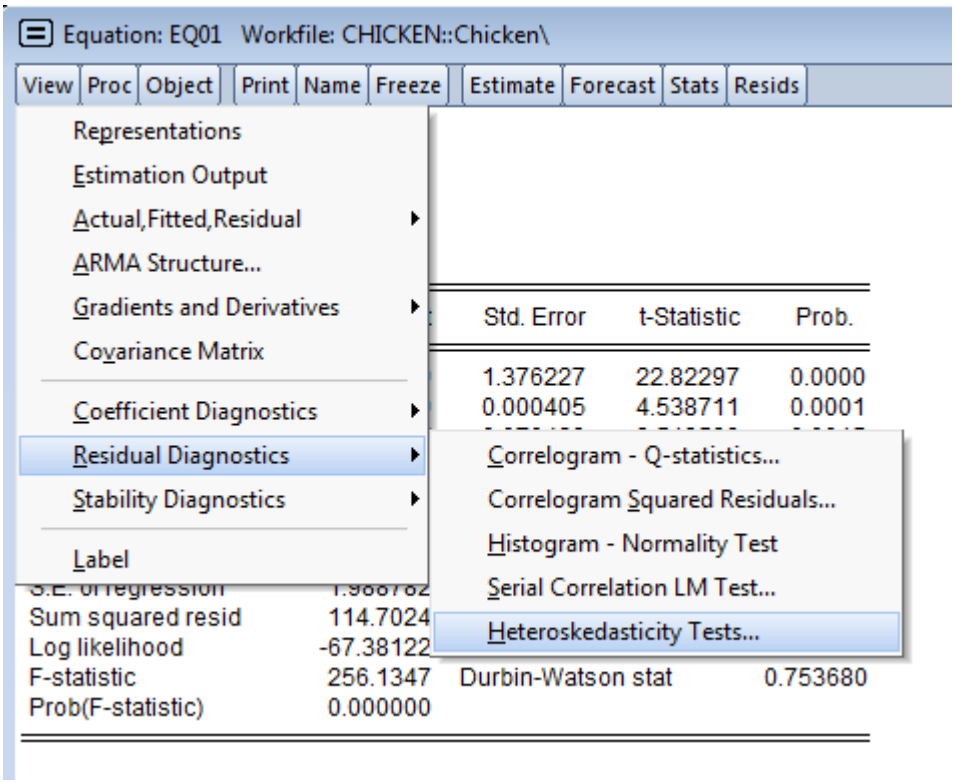
Dependent Variable: RESID^2
Method: Least Squares
Date: 08/27/13 Time: 21:22
Sample: 1 33
Included observations: 33

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-19.94193	16.12156	-1.236972	0.2286
YD	-0.000855	0.006084	-0.140567	0.8894
PB	0.939104	0.905059	1.037617	0.3102
PC	1.608144	1.857038	0.865973	0.3954
YD^2	-1.80E-07	1.14E-06	-0.157509	0.8762
PB^2	-0.007442	0.033422	-0.222681	0.8257
PC^2	-0.032974	0.053848	-0.612360	0.5463
YD*PB	3.19E-05	0.000392	0.081312	0.9359
YD*PC	8.36E-05	0.000338	0.247598	0.8066
PB*PC	-0.033598	0.054039	-0.621736	0.5402

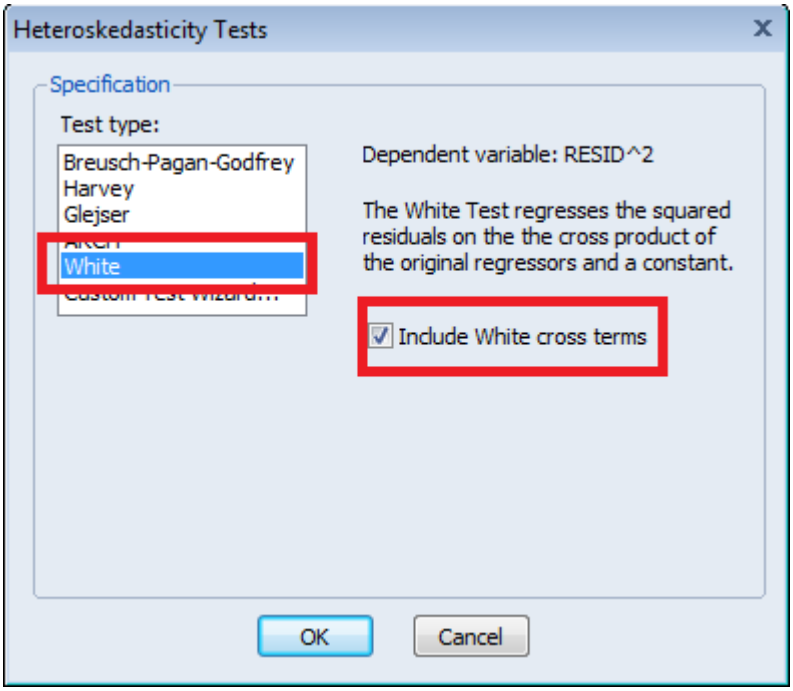
R-squared	0.219018	Mean dependent var	3.475829
Adjusted R-squared	-0.086584	S.D. dependent var	3.367886
S.E. of regression	3.510662	Akaike info criterion	5.594534
Sum squared resid	283.4693	Schwarz criterion	6.048021
Log likelihood	-82.30981	Hannan-Quinn criter.	5.747118
F-statistic	0.716677	Durbin-Watson stat	2.207246
Prob(F-statistic)	0.688712		

Як бачимо з таблиці, дана модель не є значущою, отже гетероскедастичність заданого вигляду в початковій моделі відсутня.

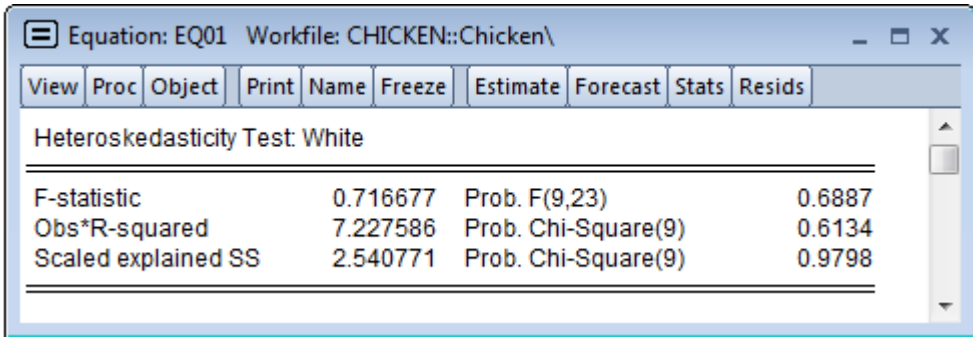
Цю перевірку можна зробити також іншим способом. Для перевірки наявності гетероскедастичності в програмному пакеті Eviews використовується вбудований критерій Уайта з попарними добутками та без них. Для цього в меню необхідно обрати **View→Residual Tests→ Heteroskedasticity tests**.



Обрання тесту Уайта дозволяє також обрати критерій з попарними добутками або без них, що залежить від кількості спостережень в моделі. В моделях з достатньою кількістю спостережень обирають тест Уайта з попарними добутками (**Include White cross terms**), в іншому випадку - тест Уайта без попарних добутків.



Результати перевірки свідчать про відсутність гетероскедастичності збурень в даній моделі (високі значення **Probability** навпроти **F-statistic** та **Obs*R-squared,** більші за 0,05):



Аналогічно здійснюється перевірка наявності гетероскедастичності за іншими критеріями.

Зважений метод найменших квадратів у випадку відомої коваріаційної матриці збурень

Для позбавлення від гетероскедастичності застосовують зважений метод найменших квадратів.

Припустимо, що коваріаційна матриця збурень відома з точністю до коефіцієнта пропорційності, тобто

$$D\upsilon_i = \sigma^2 w_i^2, i = \overline{1,n}, \tag{11}$$

де w_i^2 відомі, а σ^2 – невідомий коефіцієнт пропорційності. В системі (9) почленно розділимо i-те рівняння на $w_i (i = \overline{1,n})$:

$$y_i^* = \beta_0 x_{i0}^* + \beta_1 x_{i1}^* + \ldots + \beta_{k-1} x_{i,k-1}^* + \varepsilon_i, i = \overline{1,n}, \tag{12}$$

де

$$y_i^* = \frac{y_i}{w_i}, i = \overline{1,n}, \tag{13}$$

$$x_{ij}^* = \frac{x_{ij}}{w_j}, j = \overline{0,k-1}, i = \overline{1,n}, \tag{14}$$

$$\varepsilon_i = \frac{\upsilon_i}{w_i}, i = \overline{1,n}. \tag{15}$$

Якщо ми розглядаємо модель з константою (9), то значення змінної x_0^* обчислюються за такою формулою:

$$x_{i0}^* = \frac{1}{w_j}, i = \overline{1,n}. \qquad (16)$$

Зауважимо, що оцінка β_0 - коефіцієнта при змінній x_0^* є оцінкою постійного доданку у вихідній моделі.

Обчислення вагів на основі критерію Глейзера

Нехай, наприклад, виявилось, що допоміжна модель (5) є значущою, тобто в моделі (3) має місце гетероскедастичність. Позначимо через $\hat{\gamma}$ та $\hat{\delta}$ оцінки коефіцієнтів γ і δ в моделі (5). Ваги w_i для підстановки до формул (13) – (15) обчислюються так:

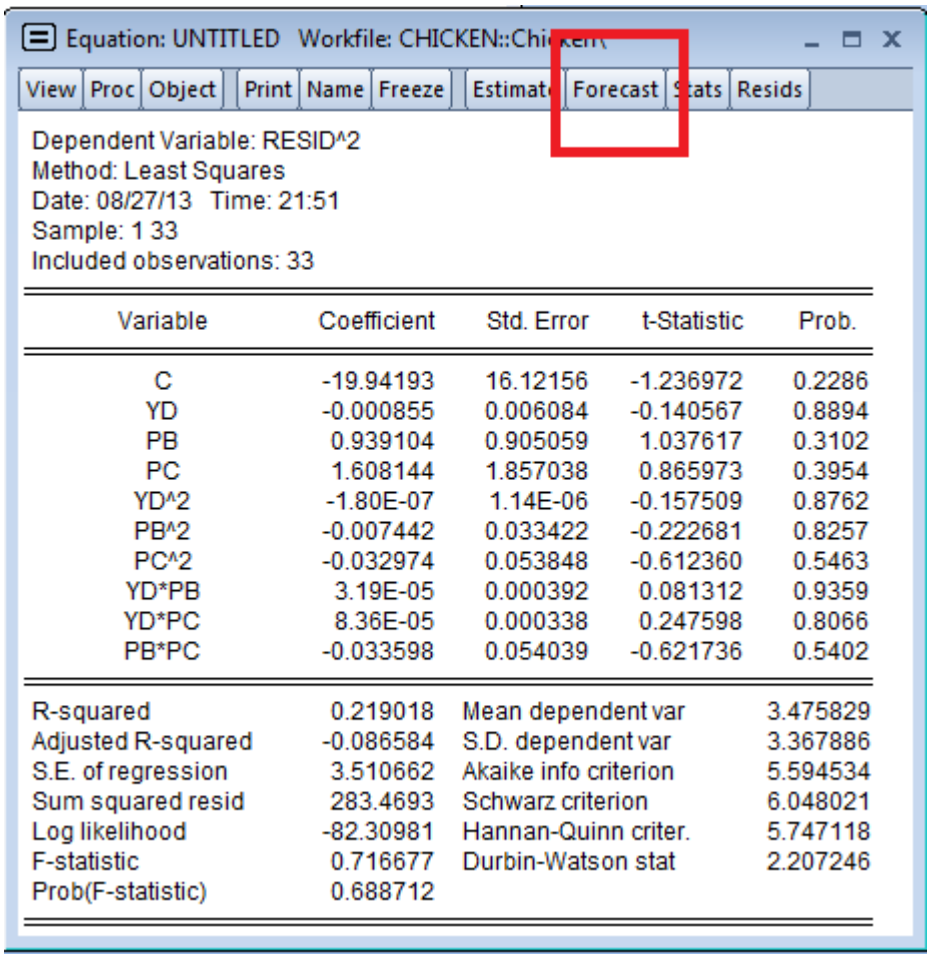
$$w_i = \hat{\gamma} + \hat{\delta}\sqrt{x_i}, \; i = \overline{1,n}.$$

Обчислення вагів на основі критерію Уайта

Припустимо, що допоміжна модель (10) виявилась значущою, тобто в моделі (9) має місце гетероскедастичність. Позначимо через $\hat{u}_i^2$ оцінки u_i^2, знайдені за моделлю (10) так, як в попередньому пункті. Ваги w_i для підстановки до формул (13) – (15) обчислюються так

$$w_i = \sqrt{\hat{u}_i^2}, i = \overline{1,n}.$$

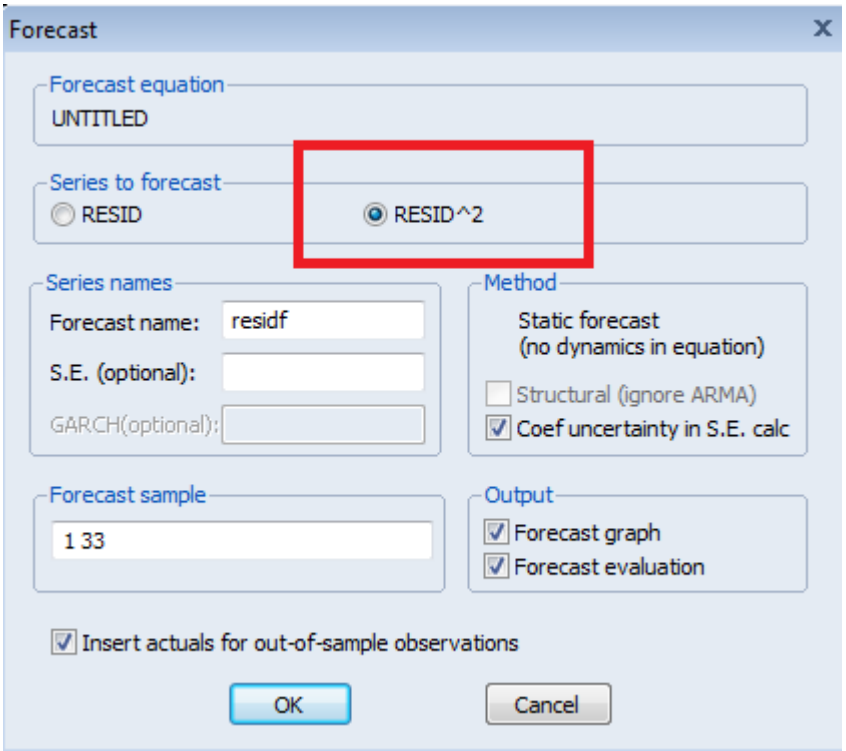
Отже, позбудемося від гетероскедастичності за допомогою зваженого методу найменших квадратів з обчисленням вагів на основі критерію Уайта.



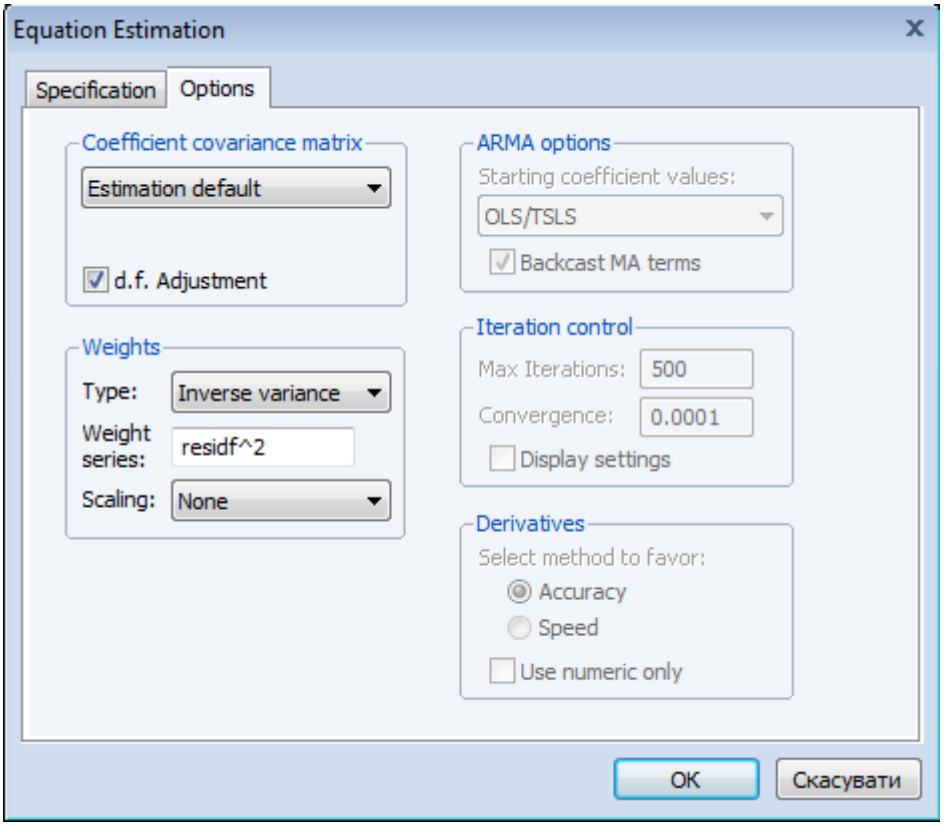
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-19.94193	16.12156	-1.236972	0.2286
YD	-0.000855	0.006084	-0.140567	0.8894
PB	0.939104	0.905059	1.037617	0.3102
PC	1.608144	1.857038	0.865973	0.3954
YD^2	-1.80E-07	1.14E-06	-0.157509	0.8762
PB^2	-0.007442	0.033422	-0.222681	0.8257
PC^2	-0.032974	0.053848	-0.612360	0.5463
YD*PB	3.19E-05	0.000392	0.081312	0.9359
YD*PC	8.36E-05	0.000338	0.247598	0.8066
PB*PC	-0.033598	0.054039	-0.621736	0.5402

R-squared	0.219018	Mean dependent var	3.475829
Adjusted R-squared	-0.086584	S.D. dependent var	3.367886
S.E. of regression	3.510662	Akaike info criterion	5.594534
Sum squared resid	283.4693	Schwarz criterion	6.048021
Log likelihood	-82.30981	Hannan-Quinn criter.	5.747118
F-statistic	0.716677	Durbin-Watson stat	2.207246
Prob(F-statistic)	0.688712		

Якщо були знайдена вибіркова регресійна функція для допоміжного рівняння (ряд residf), то застосувати зважений метод найменших квадратів можна за допомогою вбудованих функцій програмного пакету Eviews.



За ваги в цьому випадку використаємо корінь зі знайденого ряду residf. Для цього за допомогою кнопки меню **Estimate** заходимо в діалогове вікно початкової регресії **Estimation Output** та натискаємо кнопку **Options**. В частині **Weights** обираємо опцію **Inverse variance,** у **Weights series –** записуємо вираз **residf^2**:



Оцінене за допомогою зваженого методу найменших квадратів рівняння матиме наступний вигляд:

Equation: EQ01 Workfile: CHICKEN::Chicken\

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

Dependent Variable: Y
Method: Least Squares
Date: 08/27/13 Time: 22:08
Sample (adjusted): 2 33
Included observations: 32 after adjustments
Weighting series: RESIDF^2
Weight type: Inverse variance (no scaling)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	32.40762	1.801924	17.98501	0.0000
YD	0.001933	0.000460	4.205734	0.0002
PB	0.225282	0.078688	2.862995	0.0079
PC	-0.856664	0.119578	-7.164081	0.0000

Weighted Statistics

R-squared	0.950803	Mean dependent var	69.11401
Adjusted R-squared	0.945531	S.D. dependent var	26.56613
S.E. of regression	4.092057	Akaike info criterion	5.772441
Sum squared resid	468.8580	Schwarz criterion	5.955658
Log likelihood	-88.35906	Hannan-Quinn criter.	5.833172
F-statistic	180.3784	Durbin-Watson stat	0.852907
Prob(F-statistic)	0.000000	Weighted mean dep.	37.80814

Unweighted Statistics

R-squared	0.960630	Mean dependent var	36.32188
Adjusted R-squared	0.956411	S.D. dependent var	9.749466
S.E. of regression	2.035483	Sum squared resid	116.0094
Durbin-Watson stat	0.792976		

Аналогічно застосовується зважений метод НМК для вагів, знайдених за допомогою інших критеріїв.

Самостійна робота

1. Американські вчені Андо та Модільяні зібрали наступні дані про доход і споживання в США:

Інтервал доходу ($)	**Середній доход ($)**	**Середнє споживання ($)**
0-999	556	2760
1000-1999	1622	1930
2000-2999	2664	2740
3000-3999	3587	3515
4000-4999	4535	4350
5000-5999	5538	5320

Інтервал доходу ($)	**Середній доход ($)**	**Середнє споживання ($)**
6000-7499	6585	6250
7500-9999	8582	7460
10000 і вище	14033	11500

Побудуйте регресію середнього споживання по середньому доходу. Використовуючи метод Уайта, перевірте наявність гетероскедастичності у моделі.

2. Вивчимо залежність державних витрат на освіту *EE* від різних факторів, зокрема, валового внутрішнього продукту *GDP*. У файлі educ.xls знаходяться необхідні для цього вибіркові дані по 34 країнам світу, причому в таблицю включені також дані про чисельність населення різних країн.

- Побудуйте регресію витрат на утворення *EE* від валового внутрішнього продукту GDP. Чому ви можете припускати гетероскедастичність? Які її прояви в рівнянні?
- Проведіть тест Голдфелда-Квондта для виявлення гетероскедастичності.
- Використовуйте зважений метод найменших квадратів, узявши як ваги фактор пропорційності – величину населення країни *PP*. Приведіть інтерпретацію отриманого рівняння регресії й оцініте його якість.

3. Для даних з файлу expend.wf1 4 різні регресії та перевірити наявність гетероскедастичності за допомогою критеріїв Глейзера та Уайта.

4. За допомогою комп'ютера розв'язати задачі: **3.4-3.11.**

5. Розв'язати задачі: **3.2.**