

УДК 330.4(075.8)

**Ковальчук А. В.**

**Научный руководитель: Золотухина Л. С.**

## ПРОВЕРКА НА ГЕТЕРОСКЕДАСТИЧНОСТЬ ВРЕМЕННЫХ РЯДОВ

На сегодняшний день прогнозирования социально-экономических явлений часто проводятся с применением методов прогнозирования, основанных на анализе временных рядов. Однако качество прогноза существенно снижается в случае нарушения третьего условия Гаусса-Маркова, а именно гомоскедастичность остатков модели. Для диагностики гетероскедастичности разработано довольно много тестов и критериев для них. К наиболее популярным тест Глейзера, тест Голдфелда-Квандта и тест Уайта.

**Таблица 1 – Исходные данные**

Дата	Расходы на конечное потребление государственных организаций в текущих ценах, млрд. руб.	Чистые налоги, млрд. руб.
1 квартал 2009	521,06	392,08
2 квартал 2009	579,71	409,80
3 квартал 2009	532,15	528,03
4 квартал 2009	654,76	600,55
1 квартал 2010	584,21	377,56
2 квартал 2010	682,60	472,97
3 квартал 2010	658,35	581,93
4 квартал 2010	805,94	654,43
1 квартал 2011	831,46	456,57
2 квартал 2011	940,84	795,20
3 квартал 2011	943,34	989,86
4 квартал 2011	1 416,30	1 377,97
1 квартал 2012	1 616,23	1 216,60
2 квартал 2012	1 951,71	1 711,03
3 квартал 2012	1 619,10	1 786,47
4 квартал 2012	2 242,31	2 018,22
1 квартал 2013	1 982,31	1 617,47
2 квартал 2013	2 232,71	2 016,55
3 квартал 2013	2 047,44	2 181,45
4 квартал 2013	2 744,76	2 142,53
1 квартал 2014	2 389,41	2 042,05
2 квартал 2014	2 737,21	2 191,63
3 квартал 2014	2 565,29	2 450,26
4 квартал 2014	3 391,18	2 716,18
1 квартал 2015	2 933,77	2 961,22
2 квартал 2015	3 339,08	3 000,05
3 квартал 2015	3 097,49	2 831,77
4 квартал 2015	4 045,79	3 149,49
1 квартал 2016	3 369,00	3 106,00
2 квартал 2016	3 900,20	3 354,60
3 квартал 2016	3 357,10	3 302,80
4 квартал 2016	4 236,20	3 171,80
1 квартал 2017	3 677,20	3 375,10
2 квартал 2017	4 229,60	3 450,60
3 квартал 2017	3 577,80	3 378,50
4 квартал 2017	5 068,00	3 803,20
1 квартал 2018	4 424,10	4 108,03
2 квартал 2018	4 865,80	4 367,33
3 квартал 2018	4 223,50	4 487,10
4 квартал 2018	5 739,10	4 570,64

Разберем эти тесты на примере. Все расчеты будут проводиться в программе EViews.

Рассмотрим методику проведения *теста Уайта*. Смысл теста состоит в следующем: если в модели дисперсия остатков каким-то образом зависит от регрессора, то это должно каким-то образом отразиться в остатках обычной регрессии исходной модели. Выдвигается две гипотезы:  $H_0$  – гомоскедастичность,  $H_1$  – гетероскедастичность. Далее последовательность действий следующая:

- 1) Оцениваются коэффициенты основной регрессии

$$Y = \beta_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon.$$

- 2) Сохраняются остатки регрессии  $\hat{\varepsilon}_i$ ;

- 3) Оценивается регрессия квадратов остатков на все регрессоры, их квадраты, попарные произведения и константу:

$$\hat{\varepsilon}_i = \alpha_1 + \sum_{i=2}^k \alpha_i X_i + \sum_{i=2}^k \beta_{i2} X_i^2 + \sum_{i,j=2, i < j}^k \gamma_{ij} X_i X_j + u.$$

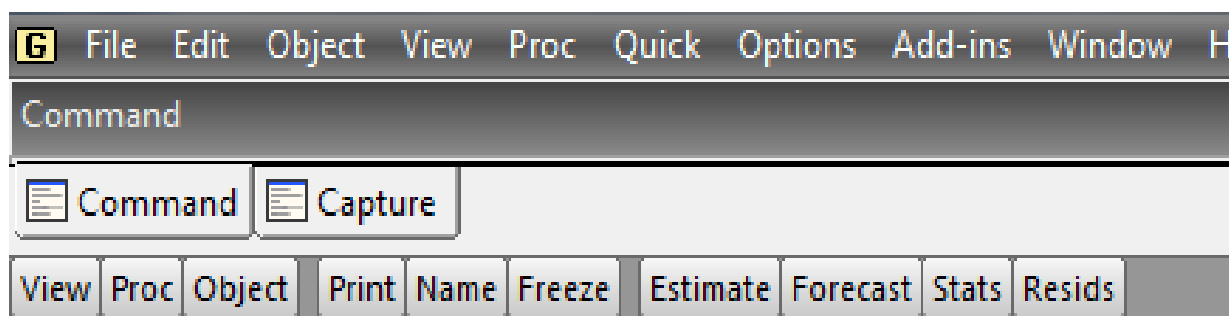
- 4) В последней оцененной регрессии находится коэффициент множественной детерминации  $R^2$ ;

- 5) Вычисляется тестовая статистика по формуле  $nR^2$ . При выполнении нулевой гипотезы тестовая статистика имеет распределение «хи-квадрат» с  $m-1$  степенями свободы;

- 6) Сравнивается полученное значение тестовой статистики с критическим  $\chi^2_{\alpha}(m-1)$ . Если значение тестовой статистики выше критического, отвергается нулевая гипотеза.

Преимуществом теста Уайта является его универсальность, однако он не является конструктивным. Если гетероскедастичность выявлена, то тест Уайта не указывает на функциональную форму гетероскедастичности.

Проведем тест Уайта для нашей модели регрессии при помощи программы EViews. Результаты вычислений представлены на рисунке 1.



#### Heteroskedasticity Test: White

F-statistic	9.642029	Prob. F(2,37)	0.0004
Obs*R-squared	13.70481	Prob. Chi-Square(2)	0.0011
Scaled explained SS	16.75742	Prob. Chi-Square(2)	0.0002

**Рисунок 1 – Результаты проведения теста Уайта**

Значение F-статистики 9,642 выше критического уровня, поскольку  $PF=0,004 < 0,05$ , следовательно, остатки модели гетероскедастичны.

Применение *теста Глейзера*, основывается оценивании регрессионной зависимости модулей отклонений  $|\varepsilon_i|$  от функциональных форм каждого регрессора:

$$|\hat{\varepsilon}| = \alpha + \beta X_j + u; \quad |\hat{\varepsilon}| = \alpha + \beta \sqrt{X_j} + u; \quad |\hat{\varepsilon}| = \alpha + \frac{\beta}{X_j} + u; \quad j \in 2, \dots, k.$$

Если согласно t-статистике коэффициент  $\beta$  значим хотя бы в одной из трех регрессий, то имеет место гетероскедастичность.

Возвращаясь к примеру, проверим остатки модели на гомоскедастичность с помощью теста Глейзера. Результат представлен на рисунке 2.

Heteroskedasticity Test: Glejser				
F-statistic	18.71055	Prob. F(1,38)	0.0001	
Obs*R-squared	13.19723	Prob. Chi-Square(1)	0.0003	
Scaled explained SS	15.19092	Prob. Chi-Square(1)	0.0001	
Test Equation:				
Dependent Variable: ARESID				
Method: Least Squares				
Date: 04/12/20 Time: 13:22				
Sample: 1 40				
Included observations: 40				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	28.16124	54.29222	0.518697	0.6070
NALOGI	0.092185	0.021312	4.325569	0.0001
R-squared	0.329931	Mean dependent var	231.3030	
Adjusted R-squared	0.312297	S.D. dependent var	207.7625	
S.E. of regression	172.2929	Akaike info criterion	13.18498	
Sum squared resid	1128025	Schwarz criterion	13.26942	
Log likelihood	-261.6995	Hannan-Quinn criter.	13.21551	
F-statistic	18.71055	Durbin-Watson stat	2.537217	
Prob(F-statistic)	0.000106			

**Рисунок 2 – Результат проведения теста Глейзера**

Сравнивая рассчитанную t-статистику с табличной (табличное значение критерия Стьюдента, соответствующее доверительной вероятности  $\gamma=0,95$  и числу степеней свободы  $v = n-m-1 = 40-1-1=38$ , равно  $t_{\text{крит}} = 2,024$ ), получаем, что коэффициент при переменной  $x$  является статистически значимым. Это свидетельствует о наличии в модели гетероскедастичности.

Одним из наиболее популярных тестов является *тест Голдфелда-Квандта*. Как правило, его применяют, когда есть предположение о прямой зависимости дисперсии ошибки от величины независимой переменной в линейной модели. Для проведения этого теста предполагается, что стандартное отклонение  $\sigma_i = \sigma(\varepsilon_i)$  пропорционально значению переменной  $x$  в этом наблюдении:  $\sigma_i^2 = \sigma^2 x_i^2$ , также предполагается, что остатки имеют нормальное распределение и автокорреляция отсутствует. Все  $n$  наблюдений упорядочивают по величине  $x$ . Упорядоченная выборка делится примерно на три равные части объемом  $k$ ,  $n-2k$ ,  $k$  соответственно. Например, при  $n=30$ ,  $k=11$ , при  $n=60$ ,  $k=22$ . Для каждой из выборок объема  $k$  оценивается свое уравнение регрессии и находятся суммы отклонений  $S_1 = \sum_{i=1}^k e_i^2$  и  $S_3 = \sum_{i=n-k+1}^n e_i^2$  соответственно. Далее по F-таблице находится критическое значение  $F_{\alpha, k-m-1}$ , где

$\alpha = 1 - \gamma$  – уровень значимости,  $m$  – число факторов модели. Статистика  $F = \frac{S_2}{S_1}$ , при  $S_3 > S_1$  и  $F = \frac{S_1}{S_2}$ , при  $S_3 < S_1$ . Если расчетное значение меньше критического, то на уровне значимости  $\alpha$  принимается гипотеза об отсутствии гетероскедастичности, в противном случае гипотеза об отсутствии гетероскедастичности принимается.

В случае множественной регрессии тест проводят для каждого фактора или для того фактора, который наиболее тесно связан с  $\sigma_i$ .

Проверим остатки модели в нашем примере по тесту Голдфелда-Квандта:

Heteroskedasticity Test: Breusch-Pagan-Godfrey				
F-statistic	14.90495	Prob. F(1,38)	0.0004	
Obs*R-squared	11.26923	Prob. Chi-Square(1)	0.0008	
Scaled explained SS	13.77934	Prob. Chi-Square(1)	0.0002	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 04/12/20 Time: 13:33				
Sample: 1 40				
Included observations: 40				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-48389.98	43113.18	-1.122394	0.2687
NALOGI	65.33605	16.92340	3.860693	0.0004
R-squared	0.281731	Mean dependent var	95587.17	
Adjusted R-squared	0.262829	S.D. dependent var	159351.4	
S.E. of regression	136817.0	Akaike info criterion	26.53938	
Sum squared resid	7.11E+11	Schwarz criterion	26.62383	
Log likelihood	-528.7876	Hannan-Quinn criter.	26.56991	
F-statistic	14.90495	Durbin-Watson stat	2.407428	

**Рисунок 10 – Результат проведения теста Голдфелда-Квандта**

Сравним рассчитанную t-статистику с табличной:  $t_{\text{расч}}=3,861 > t_{\text{крит}}=2,024$ , это говорит о значимости коэффициента при независимой переменной  $x$ . Что в свою очередь свидетельствует о гетероскедастичности остатков в модели регрессии.

Таким образом, наличие гетероскедастичности не позволяет получить эффективные оценки, что зачастую приводит к необоснованным выводам по их качеству. Обнаружение гетероскедастичности достаточно трудоемкая проблема, и для ее решения разработано несколько методов. Все они используют в качестве нулевой гипотезы  $H_0$  гипотезу об отсутствии гетероскедастичности. В работе были рассмотрены три теста, позволяющие определить наличие гетероскедастичности в остатках. Все они показали, что в остатках исходных исследуемых данных присутствует гетероскедастичность. Для дальнейшего эффективного прогнозирования по исследуемой регрессионной модели необходима ее корректировка и устранение гетероскедастичности.

#### Список цитированных источников

1. Харин, Ю.С. Экономическое моделирование / Ю.С. Харин, В.И. Малюгин, А.Ю. Харин. – Мн.: БГУ, 2003.
2. Эконометрика и экономико-математические методы и модели: учеб. пособие / Г.О. Читая [и др.] ; под ред. Г.О Читая, С.Ф. Миксюк. – Минск : БГЭУ, 2018. – 511 с.
3. Национальный статистический комитет Республики Беларусь [Электронный ресурс]. – Режим доступа: <http://belstat.gov.by/>. – Дата доступа: 06.04.2020.