

Введение

В данном домашнем задании мы попробуем узнать, существует ли проблема гетероскедастичности при построении регрессии на имеющихся данных из Росстата. Для этого будет проведено 4 теста: от общих к более специализированным. В случае обнаружении проблемы нами будет проведена корректировка.

Работа построена следующим образом: первая часть посвящена описанию данных, их дескриптивным статистикам и диаграммам рассеяния; вторая часть содержит описание построенной регрессии и предположения касательно потенциальной причины гетероскедастичности; третья часть полностью отведена на проведение необходимых тестов на гетероскедастичность; четвертая – последняя часть – посвящена корректировке и анализу результатов с повторным проведением нескольких тестов.

Часть 1. Данные.

Прежде всего, необходимо познакомиться с набором данных, который у нас есть. В предыдущем задании у нас были следующие данные: безработица (Unemployment); ВРП на душу населения (GDPpercap); удельный вес городского населения (Urbanshare); доля занятого населения с высшим образованием (Higheduc); дамми-переменной, показывающей, является ли регион западным (WEST). Теперь к этому списку нами было добавлено еще несколько переменных, а именно: посевные площади под с/х культуры (Sarea); количество убийств и попыток убийств (Murder); дамми-переменная WEST, помноженная на Higheduc (HWEST); дамми-переменная MWEST, рассчитанная следующим образом:

$$HWEST_i = \begin{cases} 1, & \text{если } murder_i > \overline{murder} (\text{ср.знач по переменной}) \\ 0, & \text{иначе} \end{cases}$$

Посмотрим на дескриптивные статистики основных переменных.

. sum Unemployment Urbanshare Higheduc GDPpercap Sarea Murder

Variable	Obs	Mean	Std. Dev.	Min	Max
Unemployment	85	6.441176	3.583969	1.4	27
Urbanshare	85	70.50941	13.15896	29	100
Higheduc	85	32.20588	4.920302	24.8	49.1
GDPpercap	85	500392.5	620665.2	126738.6	4658124
Sarea	85	941.7482	1195.709	0	5397.9
Murder	85	113.8353	93.76996	5	421

Рис 1. Дескриптивные статистики основных регрессоров

Мы хотели бы отметить, что по двум регионам не было информации касательно посевных площадей, поэтому пропуски были заполнены нулями. Именно поэтому минимальное значение Sarea равно нулю. Также заметим, что переменная Murder имеет довольно-таки крупный разброс.

Теперь обратимся к диаграммам рассеяния зависимой переменной (Unemployment) от всех остальных.

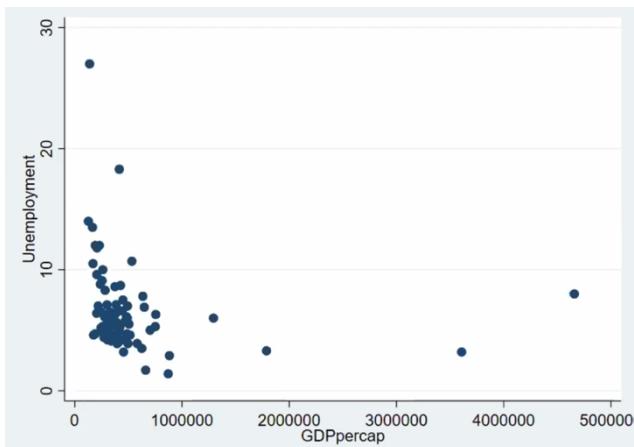


Рис 2. Диаграмма рассеяния для GDPpercap

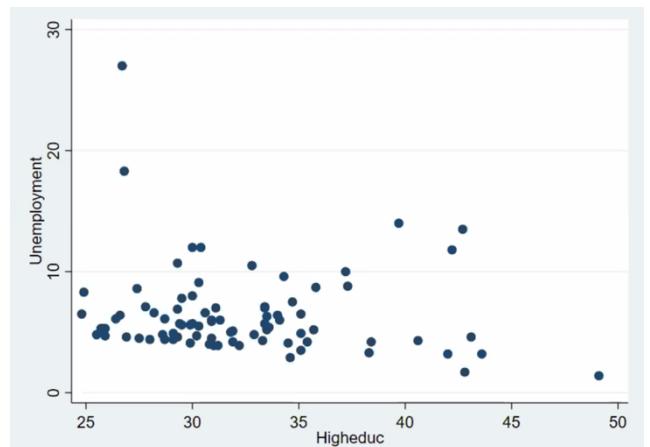


Рис 3. Диаграмма рассеяния для Higheduc

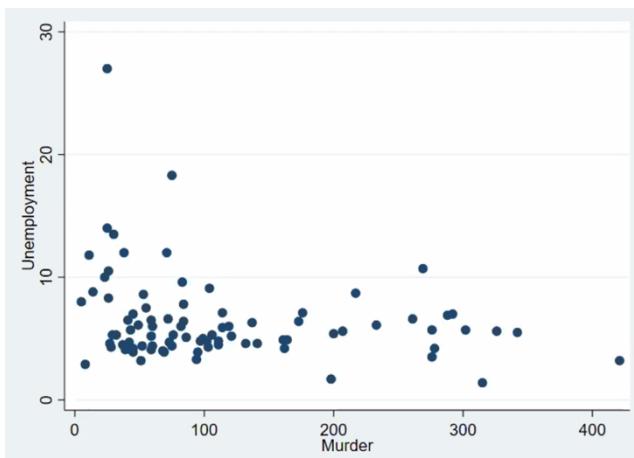


Рис 4. Диаграмма рассеяния для Murder

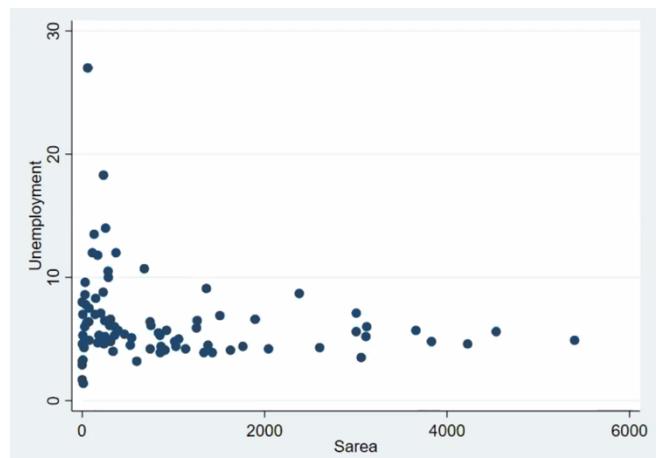


Рис 5. Диаграмма рассеяния для Sarea

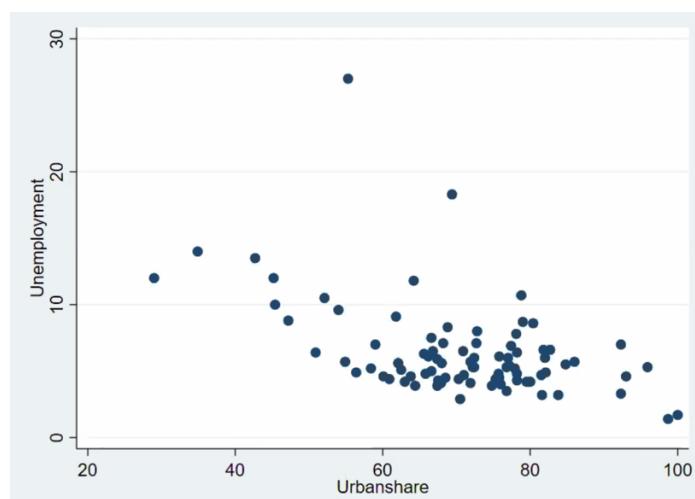


Рис 6. Диаграмма рассеяния для Urbanshare

Часть 2. Линейная регрессия.

Теперь, изучив данные, мы можем построить линейную регрессию. Нами была оценена следующая модель:

$$Unemp = \beta_0 + \beta_1 * Urbshare + \beta_2 * Heduc + \beta_3 * GDP + \beta_4 * Sarea + \beta_5 * Murder \\ + \beta_6 * WEST + \beta_7 * MWEST + \beta_8 * HWEST$$

Результаты можно увидеть ниже:

. reg Unemployment Urbanshare Higheduc GDPpercap Sarea Murder WEST MWEST HWEST						
Source	SS	df	MS	Number of obs	=	85
Model	418.691341	8	52.3364176	F(8, 76)	=	6.02
Residual	660.274542	76	8.68782292	Prob > F	=	0.0000
Total	1078.96588	84	12.8448319	R-squared	=	0.3880
				Adj R-squared	=	0.3236
				Root MSE	=	2.9475

Unemployment	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
Urbanshare	-.1305034	.0278502	-4.69	0.000	-.1859719 -.0750349
Higheduc	-.389514	.1619551	-2.41	0.019	-.7120754 -.0669525
GDPpercap	-1.72e-08	5.80e-07	-0.03	0.976	-1.17e-06 1.14e-06
Sarea	-.0007552	.0003201	-2.36	0.021	-.0013927 -.0001177
Murder	.0006023	.0054091	0.11	0.912	-.010171 .0113756
WEST	-13.05676	5.822562	-2.24	0.028	-24.6534 -1.46012
MWEST	-.829916	1.008295	-0.82	0.413	-2.838109 1.178277
HWEST	.3709216	.179691	2.06	0.042	.013036 .7288073
_cons	30.01622	5.351235	5.61	0.000	19.35832 40.67413

Рис 7. Базовая регрессия

Видно, что R^2 в целом неплохой. Однако, как будет отмечено позже, его можно улучшить. Видим, что коэффициенты перед двумя из добавленных нами переменных оказались значимы (Sarea и Hwest) на 5% уровне значимости. Тем не менее, ровно перед двумя другими переменными (Murder и Mwest) коэффициенты регрессии не значимы на любом разумном уровне.

Данную проблему мы можем попробовать объяснить наличием гетероскедастичности в переменной Murder. Взглянув на диаграмму рассеяния видно, что при малых значениях имеется довольно большой разброс среди данных, в то время как с её увеличением этот разброс падает. Такая ситуация немного необычна, потому что отличается от большинства случаев, где гетероскедастичность проявляется ровно наоборот. Запомним этот момент, сохраним остатки регрессии и перейдем к следующей главе.

Глава 3. Тесты на гетероскедастичность

В данной секции нами было проведено 4 теста на гетероскедастичность: тест Уайта, Броиша-Пагана, Голдфелда-Квандта, Глейзера. Начнем от простого к сложного и постепенно опишем каждый из тестов.

Тест Уайта

Начнем именно с него, поскольку данный тест является, наверно, самым общим из всех четырёх. Гипотезы выглядят следующим образом (в самом простом виде):

$$H_0: \text{гомоскедастичность}$$

$$H_1: \text{гетероскедастичность}$$

Как уже было сказано, это самый общий тест. В нём никак не конкретизируется вид гетероскедастичности. Суть самого теста проста: сохранив остатки самой первой регрессии, строится новая регрессия, где независимой переменной является квадрат остатков, а объясняющими – исходные переменные, их квадраты и попарное перемножение. Построив и оценив новую регрессию, сохраняем R^2 . После этого уже можно рассчитать $\chi^2_{\text{стат}} = n * R^2$, n – число наблюдений. Сама статистика $\sim \chi^2(m - 1)$, где m – число коэффициентов вспомогательной регрессии.

Приведем результаты на рисунке ниже.

White's test for Ho: homoskedasticity against Ha: unrestricted heteroskedasticity			
chi2(39) = 32.07			
Prob > chi2 = 0.7762			
Cameron & Trivedi's decomposition of LM-test			
Source	chi2	df	p
Heteroskedasticity	32.07	39	0.7762
Skewness	6.17	8	0.6277
Kurtosis	1.14	1	0.2848
Total	39.39	48	0.8074

Рис 8. Результаты теста Уайта

Согласно полученному *pvalue*, при любом разумном уровне значимости гипотеза о гомоскедастичности отвергнута не будет.

Тест Брайша-Пагана

Проведем следующий тест на гетероскедастичность. Он чуть более специфицирован, однако также не совсем четко указывает на источник проблемы. Рассмотрим гипотезы:

$$H_0: \sigma_i^2 = \sigma_e^2 \forall i = 1 \dots n$$

$$H_1: \sigma_i^2 \sim f(\alpha_0 + \alpha_1 * Z_1 + \dots + \alpha_r * Z_r) \text{ для некоторых } Z_1 \dots Z_r$$

Его реализация также довольно проста:

1. Оценивается самая первая регрессия
2. Сохраняем значение $a = \widehat{\sigma}^2 = \frac{RSS}{n}$. Сохраняем остатки регрессии e
3. Строим вспомогательную регрессию: $e^2 = \beta_0 + \beta_1 * Z_1 + \dots + \beta_r * Z_r$
4. Сохраняем ESS данной регрессии
5. Считаем $\chi^2_{\text{стат}} = \frac{ESS}{2 * a^2}$

Статистика выше также имеет распределение $\chi^2(r)$. Сравниваем её с критическим значением и делаем соответствующие выводы.

Данный тест был проведен с помощью STATA. Посмотрим на его результаты:

. estat hettest, rhs mtest			
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity			
Ho: Constant variance			
Variable	chi2	df	p
Urbanshare	16.83	1	0.0000 #
Higheduc	14.90	1	0.0001 #
GDPpercap	4.74	1	0.0294 #
Sarea	7.03	1	0.0080 #
Murder	14.26	1	0.0002 #
WEST	0.90	1	0.3438 #
MWEST	17.87	1	0.0000 #
HWEST	0.01	1	0.9244 #
simultaneous	48.79	8	0.0000

unadjusted p-values

Рис 9. Результаты теста Брайша-Пагана

Видим, что p_{value} всего теста равно 0. Это означает, что нами сразу отвергается гипотеза о гомоскедастичности и подтверждается предположение гетероскедастичности. На самом деле, если взглянуть более подробно на его результаты, то мы заметим, что у всех 5 основных переменных выявлена гетероскедастичность. И это соответствует действительности: снова взглянув на диаграммы рассеяния видим, что в самом начале у **каждой** переменной большой разброс, который сильно уменьшается по мере увеличения переменной.

Тест Глейзера

Данный тест является, пожалуй, самым подробным, поскольку учитывает сразу 3 спецификации гетероскедастичности. Опишем гипотезы:

$$H_0: \sigma_i^2 = \sigma_e^2 \forall i = 1 \dots n$$

$$H_1: \sigma_i \sim X^\gamma \text{ для некоторого } X_j; \gamma \in \{-1, 1, 0.5\}$$

Теперь перейдем к его реализации.

1. Оцениваем базовую регрессию и сохраняем остатки e

2. Оцениваем три вспомогательные регрессии:

- $|e_i| = \alpha_0 + \beta * X_i, i = 1..n$
- $|e_i| = \alpha_0 + \beta * \sqrt{X_i}, i = 1..n$
- $|e_i| = \alpha_0 + \beta * \frac{1}{X_i}, i = 1..n$

3. Если окажется, что хоть одна из β значима, то гетероскедастичность есть.

Приведём результаты. Они расположены следующим образом: каждая строка – один из регрессоров. Всего было проведено 15 регрессий (5 регрессоров, 3 регрессии для каждого). Чтобы не засорять общий ход повествования, мы расположили их в разделе **Приложение**. После проведения 15 регрессий можно сделать следующий состоятельный вывод: на 5% уровне значимости гетероскедастичность наблюдается по двум переменным: GDPpercap и

Murder. В случае с ВРП коэффициент вспомогательной регрессии был значим при как раз обратной зависимости, в то время как Murder показал гетероскедастичность на двух оставшихся спецификациях.

Тест Голдфелда-Квандта

Данный тест предполагает проверку на единственный вид гетероскедастичности, что является его серьезным минусом. Запишем гипотезы и обсудим почему:

$$H_0: \sigma_i^2 = \sigma_e^2 \quad \forall i = 1 \dots n$$

$$H_1: \sigma_i \sim X_j \text{ для некоторого } j$$

В нашей ситуации, когда мы предполагаем совершенно другой тип зависимости дисперсии от переменных, данный тест, по сути, будет бесполезен, потому что он не сможет отразить требуемую спецификацию. Тем не менее, проведем его и опишем суть:

1. Выборка сортируется по «подозрительной переменной» (от малых значений к большим)
2. Берется 3/8 наблюдений от начала и 3/8 наблюдений от конца выборки. Для каждой из них строится отдельная регрессия по своим взятым наблюдениям. Сохраним RSS_1 и RSS_2 ($RSS_2 > RSS_1$).
3. Посчитаем следующую статистику: $F_{\text{стат}} = \frac{RSS_2/(n_2-k)}{RSS_1/(n_1-k)}$, где n_1 и n_2 – число наблюдений для соответствующей выборки; k – количество коэффициентов в регрессии.
4. Данная статистика имеет F-распределение с параметрами $\sim F(n_2 - k; n_1 - k)$.

Осталось сравнить полученное значение с критическим и сделать вывод. Приведём наши результаты. Мы решили провести тест по двум переменным, которые более всего подозрительны: GDPpercap и Murder. Сначала рассмотрим результаты по GDPpercap:

. reg Unemployment Urbanshare Higheduc GDPpercap Sarea Murder WEST MWEST HWEST in 1/28						
Source	SS	df	MS	Number of obs	=	28
Model	360.003493	8	45.0004367	F(8, 19)	=	3.57
Residual	239.363292	19	12.598068	Prob > F	=	0.0108
Total	599.366786	27	22.1987698	R-squared	=	0.6006
				Adj R-squared	=	0.4325
				Root MSE	=	3.5494

Unemployment	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
Urbanshare	-.1093058	.0561602	-1.95	0.067	-.2268505 .0082388
Higheduc	-.091135	.5312374	-0.17	0.866	-1.203028 1.020758
GDPpercap	-.0000492	.0000171	-2.87	0.010	-.0000133
Sarea	-.0007611	.0007343	-1.04	0.313	-.002298 .0007758
Murder	-.019834	.0210596	-0.94	0.358	-.0639121 .0242442
WEST	3.338164	16.77095	0.20	0.844	-.31.76383 38.44016
MWEST	.5842062	2.373405	0.25	0.808	-4.383387 5.551799
HWEST	-.1397041	.5572835	-0.25	0.805	-.1.306112 1.026704
_cons	32.12996	16.86665	1.90	0.072	-3.172335 67.43225

Рис 10. Регрессия для первой выборки

. reg Unemployment Urbanshare Higheduc GDPpercap Sarea Murder WEST MWEST HWEST in 58/85						
Source	SS	df	MS	Number of obs	=	28
Model	64.2782929	8	8.03478662	F(8, 19)	=	3.09
Residual	49.4713499	19	2.60375526	Prob > F	=	0.0289
Total	113.749643	27	4.21294974	R-squared	=	0.5651
				Adj R-squared	=	0.3820
				Root MSE	=	1.6136

Unemployment	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
Urbanshare	-.0061626	.0384901	-0.16	0.874	-.0867233 .0743981
Higheduc	-.510272	.1589367	-3.21	0.005	-.8429384 -.1776136
GDPpercap	5.20e-07	3.90e-07	1.33	0.198	-2.95e-07 1.34e-06
Sarea	-.0004902	.0003739	-1.31	0.206	-.0012727 .0002924
Murder	.0000135	.005122	0.00	0.998	-.010707 .010734
WEST	-10.67175	6.585976	-1.62	0.122	-24.45635 3.112861
MWEST	.1812746	.9845004	0.18	0.856	-1.879388 2.241858
HWEST	.3125479	.1969491	1.59	0.129	-.0996713 .7247671
_cons	22.57142	5.921415	3.81	0.001	10.17775 34.96508

Рис 11. Регрессия для второй выборки

```
. display invFtail(18,18,0.05)
2.2171971
```

Рис 12. Критическое значение

Как видим, тестовая статистика $\approx \frac{239}{49} \approx 4.87$, что заметно превышает критическое значение

значение, а значит, гетероскедастичность есть. Повторим эксперимент для Murder:

```
. reg Unemployment Higheduc Urbanshare GDPpercap Murder Sarea WEST MWEST HWEST in 1/30
note: MWEST omitted because of collinearity
```

Source	SS	df	MS	Number of obs	=	30
				F(7, 22)	=	3.14
Model	329.230741	7	47.032963	Prob > F	=	0.0185
Residual	329.243925	22	14.965633	R-squared	=	0.5000
				Adj R-squared	=	0.3409
Total	658.474667	29	22.706023	Root MSE	=	3.8685

```
. scalar rss1=e(rss)
disp rss1
329.24393
```

Рис 13. Регрессия для первой выборки

```
. reg Unemployment Higheduc Urbanshare GDPpercap Murder Sarea WEST MWEST HWEST in 55/85
note: MWEST omitted because of collinearity
```

Source	SS	df	MS	Number of obs	=	31
				F(7, 23)	=	4.79
Model	58.1120302	7	8.3017186	Prob > F	=	0.0019
Residual	39.8699053	23	1.73347414	R-squared	=	0.5931
				Adj R-squared	=	0.4692
Total	97.9819355	30	3.26606452	Root MSE	=	1.3166

```
. scalar rss2=e(rss)
disp rss2
39.869905
```

Рис 14. Регрессия для второй выборки

Тестовая статистика приблизительно равна $\frac{329}{39} \approx 8.43$, критическое значение $F(20, 21, 0.05) = 2.096$. Снова получается, что статистика больше, и гипотеза о гомоскедастичности отвергается.

Глава 4. Корректировка.

Так как несколько тестов обнаружили гетероскедастичность, необходимо провести корректировку. Нами была выбрана переменная Murder, поскольку в предыдущих заданиях нами было показано, что проблему с GDPpercap можно частично решить путем поиска более оптимальной функциональной формы.

Ввиду того, что мы предполагаем обратную зависимость дисперсии, а именно $\sigma_i^2 \sim \frac{1}{Murder}$, то нам потребуется домножить все переменные на \sqrt{Murder} . Введем новые переменные и проведем 2 теста: Уайта и Брайша-Пагана.

```
. gen sMurder = sqrt(Murder)
. gen new_Unemployment = Unemployment*sMurder
.
. gen new_Higheduc = Higheduc*sMurder
.
. gen new_Urbanshare = Urbanshare*sMurder
.
. gen new_Murder = Murder*sMurder
.
. gen new_Sarea = Sarea*sMurder
.
. gen new_cons = 1*sMurder
```

Рис. 15. Создание новых переменных

. reg new_Unemployment new_Higheduc new_Urbanshare new_Murder new_Sarea new_cons					
Source	SS	df	MS	Number of obs	= 85
Model	42709.1067	5	8541.82134	F(5, 79)	= 15.57
Residual	43345.9886	79	548.683399	Prob > F	= 0.0000
Total	86055.0953	84	1024.46542	R-squared	= 0.4963
				Adj R-squared	= 0.4644
				Root MSE	= 23.424

new_Unemployment	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
new_Higheduc	-.1722108	.052112	-3.30	0.001	-.2759371 -.0684846
new_Urbanshare	-.0796529	.0246888	-3.23	0.002	-.1287946 -.0305111
new_Murder	.0052082	.0043243	1.20	0.232	-.0033991 .0138154
new_Sarea	-.0004833	.000205	-2.36	0.021	-.0008913 -.0000754
new_cons	15.4252	2.735556	5.64	0.000	9.980216 20.87019
_cons	16.36025	12.91647	1.27	0.209	-9.349341 42.06984

Рис. 16. Регрессия с новыми переменными

```
. estat imtest, white
```

White's test for Ho: homoskedasticity
against Ha: unrestricted heteroskedasticity

chi2(20) = **16.14**
Prob > chi2 = **0.7081**

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	16.14	20	0.7081
Skewness	7.08	5	0.2150
Kurtosis	3.92	1	0.0477
Total	27.14	26	0.4023

```
. estat hettest, rhs mtest
```

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance

Variable	chi2	df	p
new_Higheduc	0.17	1	0.6817 #
new_Urbans~e	0.02	1	0.8765 #
new_Murder	0.20	1	0.6551 #
new_Sarea	0.21	1	0.6459 #
new_cons	0.17	1	0.6827 #
simultaneous	7.45	5	0.1896

unadjusted p-values

Рис. 17. Тест Уайта после проведения корректировки

Рис. 18. Тест Брайша-Пагана после проведения корректировки

Как видно из полученных результатов, во-первых, регрессия стала лучше: коэффициент детерминации стал равен почти 0.5; во-вторых, заметим, что *p_value* для Murder и MWEST уменьшился довольно сильно – это тоже хороший знак; в-третьих, тест Брайша-Пагана теперь показывает, что у нас нет гетероскедастичности.

Жизнь – это то, что проходит мимо нас, пока мы боремся с гетероскедастичностью (с) Джон Леннон

Приложение

. reg modres iGDPpercap

Source	SS	df	MS
Model	34.7939405	1	34.7939405
Residual	372.028244	83	4.482268
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
iGDPpercap	453811.9	162882	2.79 0.007
_cons	.3893616	.5321273	0.73 0.466

Рис 19. Вспомогательная регрессия на $\frac{1}{GDP}$

. reg modres sGDPpercap

Source	SS	df	MS
Model	7.13562238	1	7.13562238
Residual	399.686562	83	4.81550075
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
sGDPpercap	-.0010967	.0009009	-1.22 0.227
_cons	2.446433	.6373019	3.84 0.000

Рис 20. Вспомогательная регрессия на $\sqrt{GDPpercap}$

. reg modres GDPpercap

Source	SS	df	MS
Model	3.59310764	1	3.59310764
Residual	403.229077	83	4.85818165
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
GDPpercap	-3.33e-07	3.87e-07	-0.86 0.392
_cons	1.89353	.307811	6.15 0.000

Рис 21. Вспомогательная регрессия на $GDPpercap$

. reg modres iHigheduc

Source	SS	df	MS
Model	10.3920769	1	10.3920769
Residual	396.430108	83	4.77626636
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
iHigheduc	79.39972	53.82847	1.48 0.144
_cons	-.790747	1.723127	-0.46 0.648

Рис 22. Вспомогательная регрессия на $\frac{1}{Higheduc}$

. reg modres sHigheduc

Source	SS	df	MS
Model	8.69672316	1	8.69672316
Residual	398.125461	83	4.79669231
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
sHigheduc	-.7633007	.5668769	-1.35 0.182
_cons	6.046709	3.217039	1.88 0.064

Рис 23. Вспомогательная регрессия на $\sqrt{Higheduc}$

. reg modres Higheduc

Source	SS	df	MS
Model	8.07862521	1	8.07862521
Residual	398.743559	83	4.80413927
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
Higheduc	-.0630286	.0486045	-1.30 0.198
_cons	3.756677	1.583301	2.37 0.020

Рис 24. Вспомогательная регрессия на $Higheduc$

. reg modres iMurder

Source	SS	df	MS
Model	13.0018395	1	13.0018395
Residual	393.820345	83	4.74482343
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
iMurder	14.41585	8.708592	1.66 0.102
_cons	1.442588	.2920562	4.94 0.000

Рис 25. Вспомогательная регрессия на $\frac{1}{Murder}$

. reg modres sMurder

Source	SS	df	MS
Model	25.4823613	1	25.4823613
Residual	381.339823	83	4.5944557
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
sMurder	-.1323219	.0561861	-2.36 0.021
_cons	3.028078	.5994698	5.05 0.000

Рис 26. Вспомогательная регрессия на \sqrt{Murder}

. reg modres Murder

Source	SS	df	MS
Model	20.0731081	1	20.0731081
Residual	386.749076	83	4.65962743
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
Murder	-.0052132	.0025117	-2.08 0.041
_cons	2.320232	.3695553	6.28 0.000

Рис 27. Вспомогательная регрессия на $Murder$

. reg modres iSarea

Source	SS	df	MS
Model	1.6674768	1	1.6674768
Residual	396.751679	81	4.89816888
Total	398.419156	82	4.8587702
modres	Coef.	Std. Err.	t P> t
iSarea	-.0924824	.1585061	-0.58 0.561
_cons	1.730785	.2464059	7.02 0.000

Рис 28. Вспомогательная регрессия на $\frac{1}{Sarea}$

. reg modres sSarea

Source	SS	df	MS
Model	5.83586457	1	5.83586457
Residual	400.98632	83	4.83116048
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
sSarea	-.0145471	.0132357	-1.10 0.275
_cons	2.088218	.4061773	5.14 0.000

Рис 29. Вспомогательная регрессия на \sqrt{Sarea}

. reg modres Sarea

Source	SS	df	MS
Model	5.1348563	1	5.1348563
Residual	401.687328	83	4.83960636
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
Sarea	-.0002068	.0002007	-1.03 0.306
_cons	1.921517	.3044275	6.31 0.000

Рис 30. Вспомогательная регрессия на $Sarea$

. reg modres iUrbanshare

Source	SS	df	MS
Model	9.16583394	1	9.16583394
Residual	397.656351	83	4.79104037
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
iUrbanshare	88.24795	63.80189	1.38 0.170
_cons	.4185094	.9752046	0.43 0.669

Рис 31. Вспомогательная регрессия на $\frac{1}{Urbanshare}$

. reg modres sUrbanshare

Source	SS	df	MS
Model	13.3626168	1	13.3626168
Residual	393.459568	83	4.74047672
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
sUrbanshare	-.4850719	.2889159	-1.68 0.097
_cons	5.780586	2.426023	2.38 0.019

Рис 32. Вспомогательная регрессия на $\sqrt{Urbanshare}$

. reg modres Urbanshare

Source	SS	df	MS
Model	14.2885133	1	14.2885133
Residual	392.533671	83	4.72932134
Total	406.822185	84	4.84312124
modres	Coef.	Std. Err.	t P> t
Urbanshare	-.0313424	.0180318	-1.74 0.086
_cons	3.936721	1.293105	3.04 0.003

Рис 33. Вспомогательная регрессия на $Urbanshare$