МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«Национальный исследовательский ядерный университет «МИФИ»
(НИЯУ МИФИ)
Институт Финансовых Технологий и Экономической Безопасности
Кафедра Финансового мониторинга

Лабораторная работа №2
по дисциплине «Эконометрика»

Выполнил студент группы С20-702: Нуритдинходжаева А.А.
Проверил: Домашова Д.В.

Москва 2023

Оглавление

1. Постановка задачи

Рассмотрим процедуру исследования линейной модели множественной регрессии на наличие или отсутствие гетероскедастичности на основе информации о результатах голосования по поправкам в Конституцию (y), задолженности по кредитам предоставленным кредитным организациями физическим лицам (x_4), средств (вкладов) физических лиц на рублевых счетах в Сбербанке России (x_5), среднедушевые доходы населения (x_7).

$$y\text{^}=86{,}03118 – 0{,}12325x_4-0{,}10977x_5+0{,}00031x_7$$

Исходные данные приведены в приложении А.

Для оценки параметров регрессионной модели воспользуемся методом пошаговой регрессии (методом исключения переменных). Процедура построения уравнения множественной регрессии более подробно рассмотрена в лабораторной работе №1.

Результаты оценивания представлены в таблице 1.

Таблица 1Результат оценивания параметров регрессионной модели

	Regression Summary for Dependent Variable: y (исходные in Данные) R= ,58004301 R?= ,33644989 Adjusted R?= ,31187396 F(3,81)=13,690 p<,00000 Std/Error of estimate: 7,8739					
N=85	b*	Std.Err. of b*	b	Std.Err. of b	t(81)	p-value
Intercept			86,03118	2,413853	35,64061	0,000000
x4	-0,664366	0,155096	-0,12325	0,028772	-4,28359	0,000050
x5	-0,519770	0,146013	-0,10977	0,030836	-3,55974	0,000624
x7	0,663058	0,192380	0,00031	0,000090	3,44661	0,000903

```
OLS_model = s.OLS(Y,X).fit()
OLS_model.summary()
```

OLS Regression Results

Dep. Variable:	Y	R-squared:	0.336
Model:	OLS	Adj. R-squared:	0.312
Method:	Least Squares	F-statistic:	13.69
Date:	Thu, 18 May 2023	Prob (F-statistic):	2.63e-07
Time:	19:27:48	Log-Likelihood:	-293.96
No. Observations:	85	AIC:	595.9
Df Residuals:	81	BIC:	605.7
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	86.0312	2.414	35.641	0.000	81.228	90.834
X4	-0.1232	0.029	-4.284	0.000	-0.180	-0.066
X5	-0.1098	0.031	-3.560	0.001	-0.171	-0.048
X7	0.0003	8.97e-05	3.447	0.001	0.000	0.000

Omnibus:	0.223	Durbin-Watson:	2.292
Prob(Omnibus):	0.895	Jarque-Bera (JB):	0.213
Skew:	0.112	Prob(JB):	0.899
Kurtosis:	2.902	Cond. No.	1.49e+05

По данным Приложения А:

1. построить МНК-оценки коэффициентов линейной модели множественной регрессии;

2. исследовать регрессионные остатки на гетероскедастичность, используя тесты Спирмена, Голдфелда-Квандта, Глейзера;
3. построить ОМНК-оценки параметров регрессионной модели.

2. Визуальный анализ

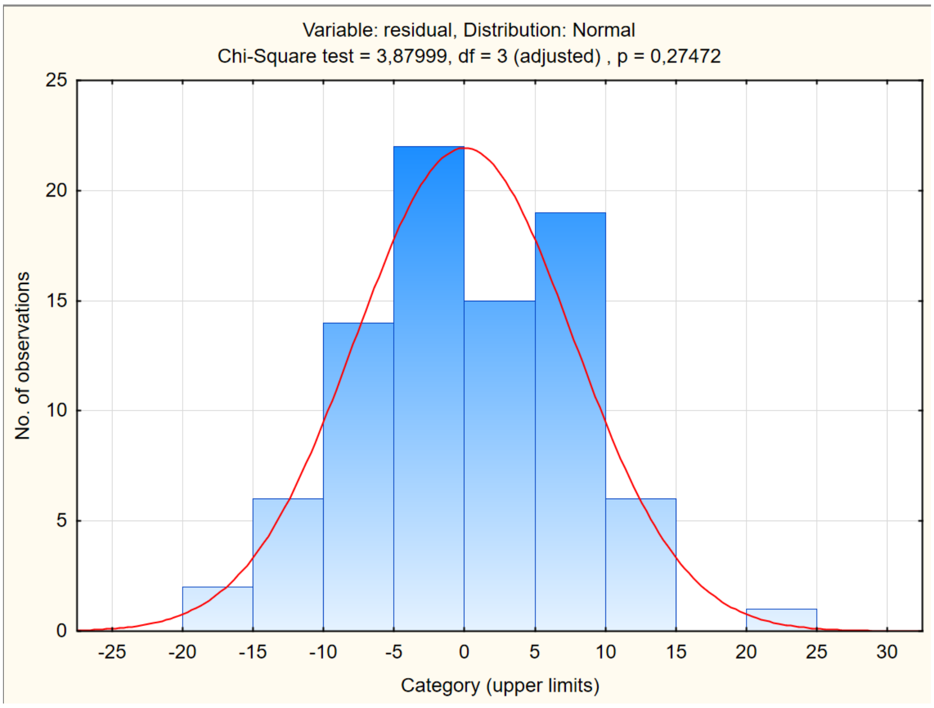


Рисунок 1 Гистограмма распределения регрессионных остатков

```
plt.figure(figsize=(16, 8))
histData = plt.hist(E, bins=18)
range_ = arange(min(E), max(E), 0.005)
coefY = len(E) * (histData[1][1] - histData[1][0])
plt.plot(range_,
        [norm(E.mean(),E.std()).pdf(x) * coefY for x in range_],
        color='r')
plt.xticks(histData[1])
KS_maxD, KS_PValue = kstest(E, cdf='norm', args=(E.mean(), E.std()))
plt.title("Histogram of the distribution of regression residues\n" +
          "Distribution: Normal\n" +
          f"Kolmogorov-Smirnov test = {KS_maxD}, p-value = {KS_PValue}", fontsize=15)
plt.ylabel("No. of observations", fontsize=12)
plt.xlabel("Category (upper limits)", fontsize=12)
plt.grid()
plt.show()
```

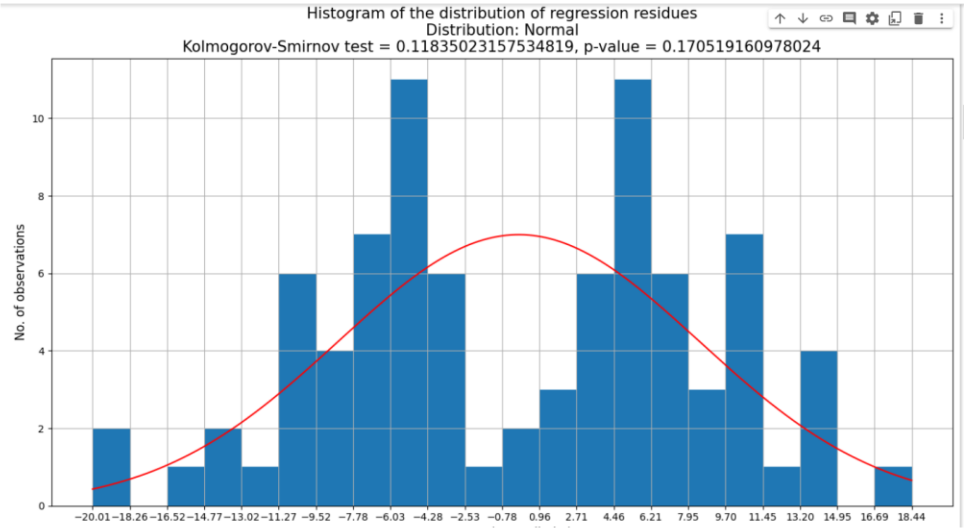


Рисунок 1.2 Гистограмма распределения регрессионных остатков

Результаты формальной проверки гипотезы о нормальном характере распределения регрессионных остатков позволяют ее не отвергнуть, и есть смыл проводить дальнейший анализ построенного уравнения множественной регрессии.

Поскольку построили значимую регрессионную модель, то следующим этапом является исследование регрессионных остатков на наличие/отсутствие гетероскедастичности.

Наличие гетероскедастичности можно предположить по графику зависимости остатков от упорядоченных по возрастанию значений той объясняющей переменной, вариацией которой возможно порождается гетероскедастичность. Для построения графика можно воспользоваться MS Excel.

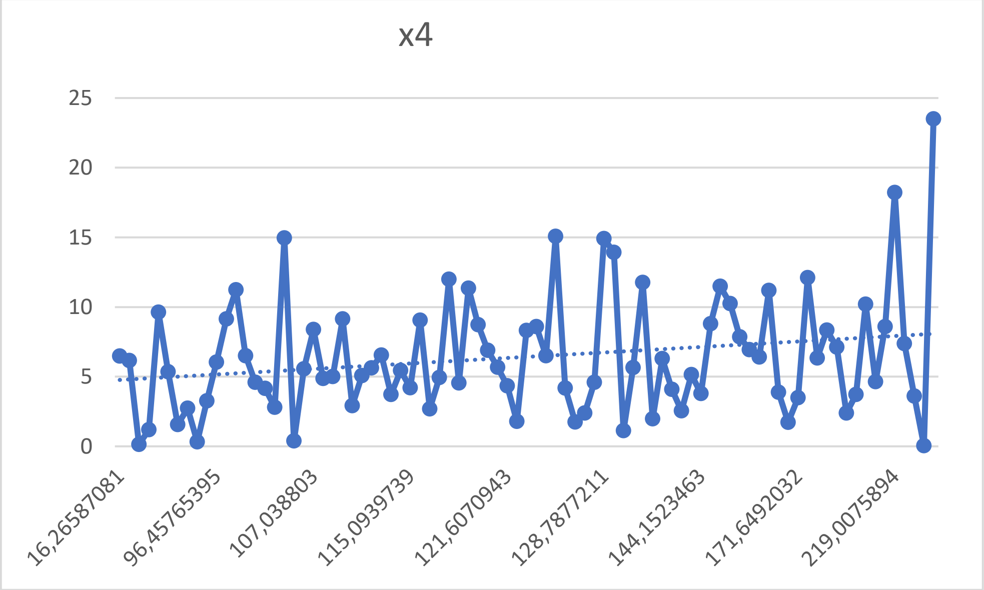


Рисунок 3 График зависимости модуля значений регрессионных остатков и значений объясняющей переменной x4

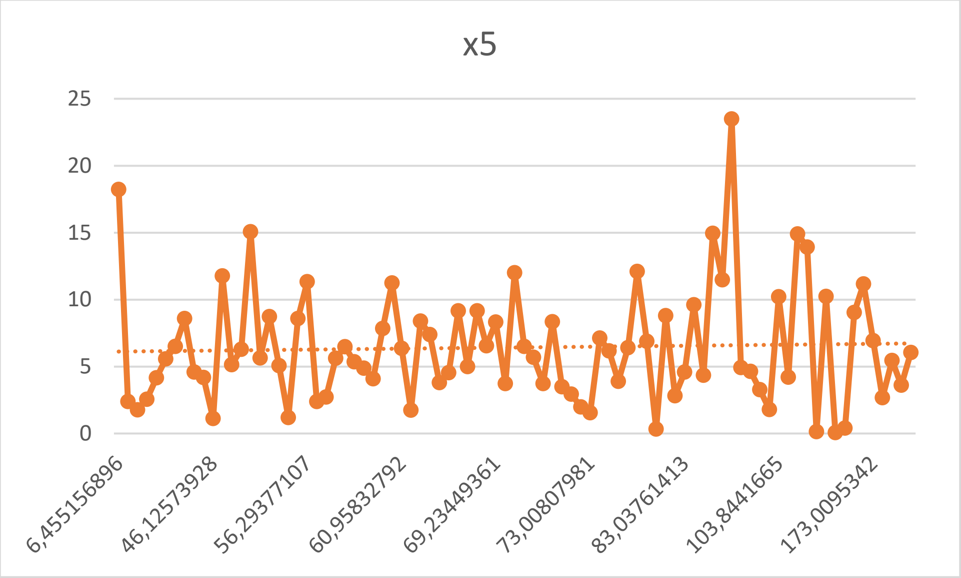


Рисунок 4 График зависимости модуля значений регрессионных остатков и значений объясняющей переменной x5

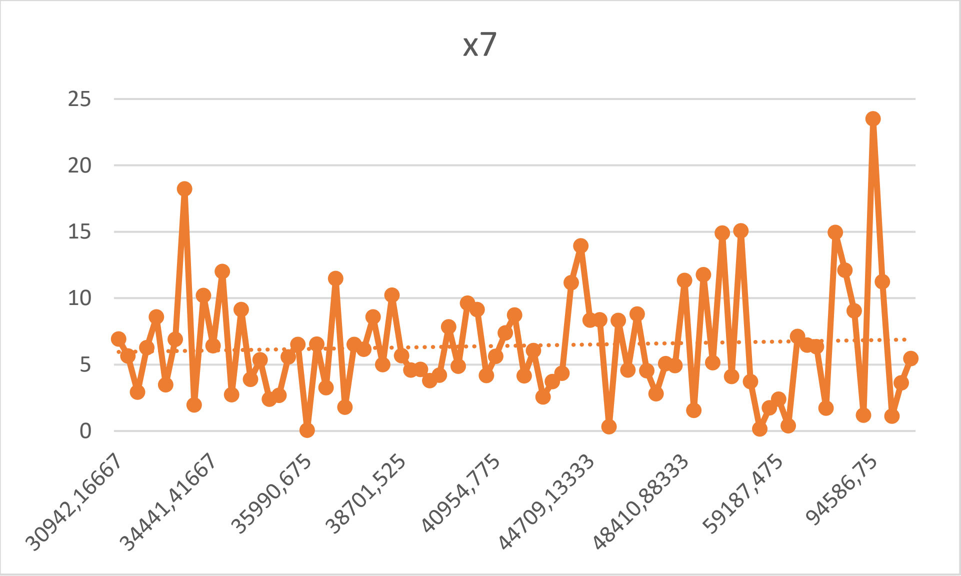
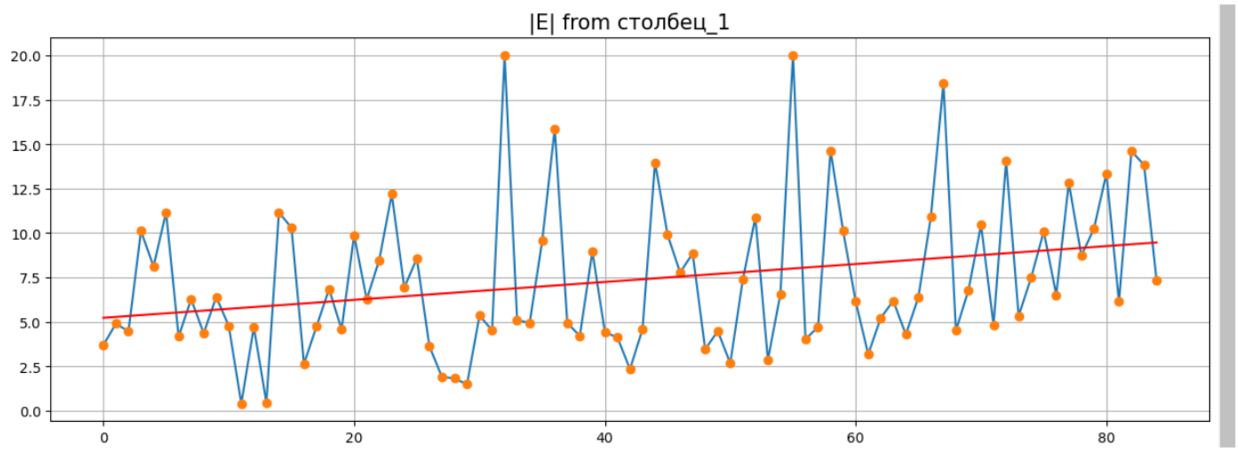
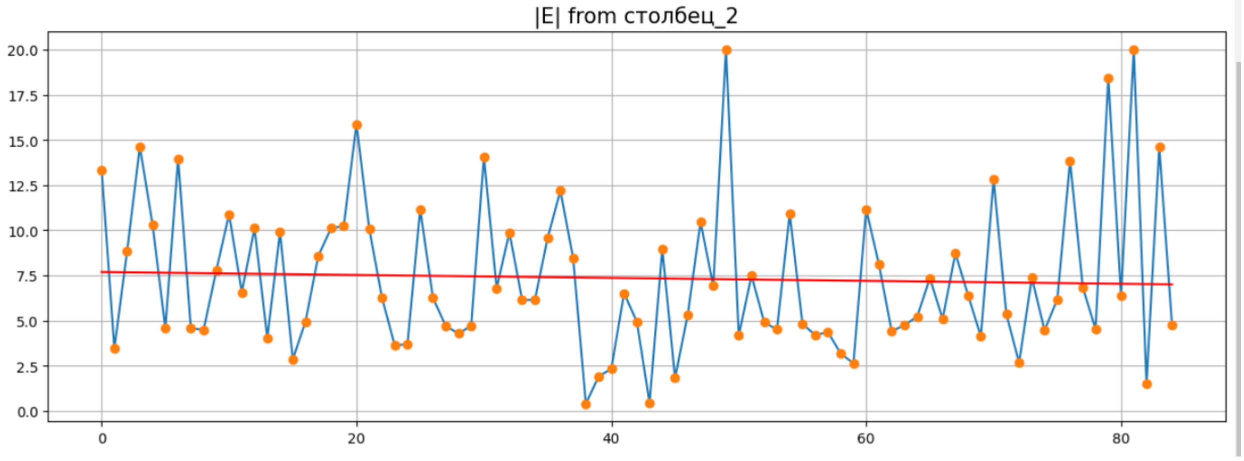


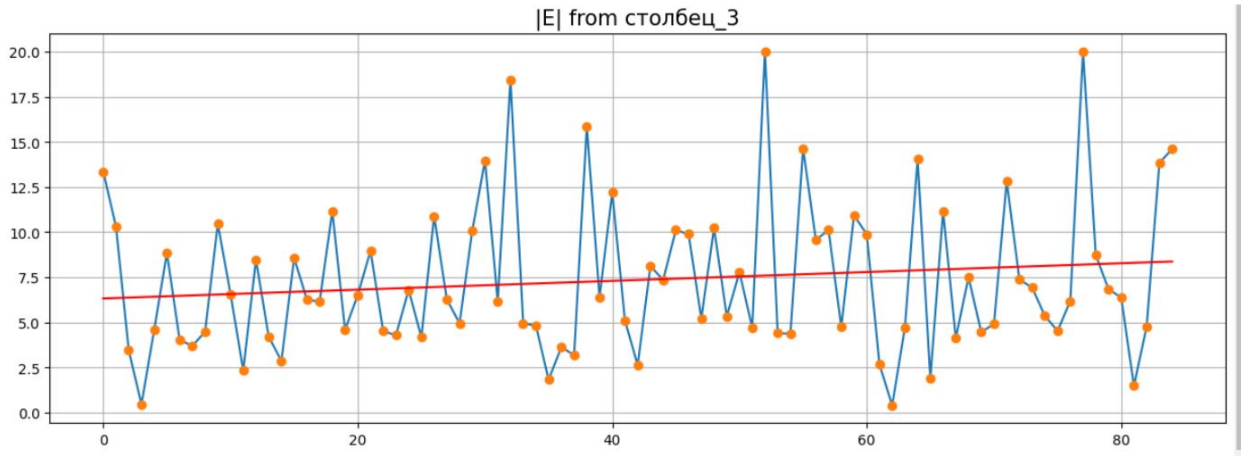
Рисунок 5 График зависимости модуля значений регрессионных остатков и значений объясняющей переменной x7

```
maxz = 0
for var in range(1, k+1):
    Xsort = np.copy(X[:, var])
    Epl = abs(E[Xsort.argsort()])
    plt.figure(figsize=(15, 5))
    plt.grid()
    plt.title(f"|E| from столбец_{var}", fontsize=15)
    plt.plot(sorted(Xsort),Epl)
    Xgraph = [i for i in range(85)]
    pylab.plot(sorted(Xsort), Epl, 'o')
```

```
z = np.polyfit(sorted(Xsort), Epl, 1)
p = np.poly1d(z)
pylab.plot(sorted(Xsort),p(sorted(Xsort)), "r-")
print (f"y = {z[0]:6f}x + ({z[1]:6f})")
if abs(z[0]) > maxz:
    maxz = z[0]; #определение наибольшего наклона по всем уравнениям тренда
    maxk = var;
```







На графиках видно, что модули регрессионных остатков имеют тенденцию к росту при увеличении значений объясняющей переменной. Следовательно, можно заподозрить гетероскедастичность.

Кроме визуального анализа, существуют различные критерии (тесты), с помощью которых выявляется гетероскедастичность.

3. Тест ранговой корреляции Спирмена

При использовании теста предполагается, что дисперсии регрессионных остатков прямо- или обратно-пропорциональны значениям объясняющей переменной x_4. Значения x_i, e_i упорядочиваются по возрастанию переменной. Далее вычисляют значение коэффициента ранговой корреляции Спирмена:

$$\rho_{x,e} = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3 - n}$$

где d_i – разность между рангами значений e_i и x_i.

Выдвигается нулевая гипотеза H_0: $\rho_{x,e}=0$ – отсутствие гетероскедастичности, и альтернативная H_1: $\rho_{x,e}\neq 0$, о наличии гетероскедастичности. Гипотеза проверяется с помощью статистики:

$$|t| = \frac{|\rho_{x,e}|\sqrt{n-2}}{\sqrt{1-\rho_{x,e}^2}} \in St(n\text{-}2)|H_0,$$

т.е. при выполнении нулевой гипотезы статистика $|t|$ будет иметь распределение Стьюдента с n-2 степенями свободы, или, иными словами, $t_{расч}<t_{крит}$, p-value>0,05. И если это так, то гипотеза об отсутствии гетероскедастичности отвергается; в противном случае она принимается.

Результат проведения теста в программе Statistica представлен в таблице 2.

Pair of Variables	Spearman Rank Order Correlations (Лист3 in данные_лаб2) MD pairwise deleted Marked correlations are significant at p <,05000			
	Valid N	Spearman R	t(N-2)	p-value
residual & x4	85	1,000000		

Таблица 2 Результаты оценивания теста ранговой корреляции Спирмена

Во втором столбце данного окна определяется оценка коэффициента ранговой корреляции Спирмена R= 1,000000, который свидетельствует о возможном наличии прямой зависимости, поэтому в данном случае нулевая гипотеза отклоняется и можно сделать вывод, о гетероскедастичности по факторной переменной X_4 существует.

```
maxk = 1

X_investigated = X[:,maxk]

rho, pval = scipy.stats.spearmanr(E, X_investigated)

print(scipy.stats.spearmanr(E, X_investigated))
```

```
SignificanceResult(statistic=-0.16210670314637485,
pvalue=0.13827533057721583)
```

4. Тест Голдфелда-Квандта

При использовании теста предполагается, что дисперсии регрессионных остатков прямо или обратно пропорциональны значению объясняющей переменной, вариации которой порождаются гетероскедастичностью.

Все наши наблюдения упорядочиваются по возрастанию элемента x_4. Упорядоченная выборка разбивается на 3 подвыборки. Оцениваются коэффициенты уравнений регрессии для первых n' и последних n'' наблюдений, причём

$$n'=n''\approx\frac{3}{8}n$$

Выдвигается нулевая гипотеза H_0: $\sigma_1^2=\sigma_2^2=\ldots=\sigma_n^2$ (нет гетероскедастичности), и альтернативная ей гипотеза $\exists i\neq j$ $\sigma_i^2=\sigma_j^2$ (есть гетероскедастичность). Сосчитаем остаточные дисперсии (т.е. суммы квадратов отклонений: $Q'=(e')^Te'$, $Q''=(e'')^T(e'')$) для уравнения регрессии по первым и по последним данным. Если предположение о прямой пропорциональности D и объясняющей переменной, верно, то Q' << Q'', т.к. они обязаны быть расположенными по возрастанию. И если это так, то есть основание говорить о гетероскедастичности.

Effect	Analysis of Variance; DV: y (Лист2 in данные_лаб2) Include condition: v0<=32				
	Sums of Squares	df	Mean Squares	F	p-value
Regress.	724,421	1	724,4209	21,48559	0,000065
Residual	1011,498	30	33,7166		
Total	1735,919				

Таблица 3 Оценка значения Q' для первых 32 наблюдений

Effect	Analysis of Variance; DV: y (Лист2 in данные_лаб2,1) Include condition: V0>=52				
	Sums of Squares	df	Mean Squares	F	p-value
Regress.	1,922	1	1,92241	0,021102	0,885413
Residual	2915,251	32	91,10159		
Total	2917,173				

Таблица 4 Оценка значения Q'' для последних 32 наблюдений

Тест проводится с помощью статистики:

$$F=\frac{\max(Q',Q'')}{\min(Q',Q'')}$$

которая в случае справедливости H_0 имеет распределение Фишера со степенями свободы n'-k-1. Проверка гипотезы: $F_{набл}>F_{крит}$, то гипотеза H_0 об отсутствии гетероскедастичности отклоняется. Узнать расчётное и критическое значение можем с помощью Microsoft Excel:

Q'	1011,498
Q''	2915,251
Fнабл	2,882112
Fкрит	1,840872

Таблица 5 Результаты расчетов для теста Голдфелда-Квандта

```
def het_gq(y, x):
```

```
    nobs, nvar = x.shape
    sizeSubsample = round(3 * nobs / 8)
    fval, _, _ = het_goldfeldquandt(y=y, x=x, idx=1, split=sizeSubsample, drop=nobs - 2 * sizeSubsample)
    if fval < 1.0:
        fval = 1.0 / fval
        GK = 0 #обратная пропорциональность
    else:
        GK = 1 #прямая пропорциальность
    return (fval, f.ppf(q=0.95, dfn=sizeSubsample - nvar, dfd=sizeSubsample - nvar),GK)
name = ['F statistic', 'F crit']
GK_Fstat, GK_Fcrit, GK = het_gq(y=Y, x=X[:,[0,maxk]])
print(lzip(name, [GK_Fstat, GK_Fcrit]) )
if GK_Fstat > GK_Fcrit :
    print("Т.к. Fнабл > Fкрит, то гипотеза об отсутствии гетероскедастичности отклоняется.")
    Test_GK = 1 #есть гетеро\n",
else:
    Test_GK = 0 #нет гетеро
    print("Т.к. Fнабл < Fкрит, то гипотеза об отсутствии гетероскедастичности принимается.")
```

```
[('F statistic', 2.3443539035738654), ('F crit', 1.8608114354760754)]
Т.к. Fнабл > Fкрит, то гипотеза об отсутствии гетероскедастичности отклоняется.
```

Расчётное значение статистики больше критического, поэтому нулевая гипотеза об отсутствии гетероскедастичности отклоняется, можно подозревать наличие гетероскедастичности. По результатам этого теста можно также построить ковариационную матрицу остатков, т.к. он выявил возможность гетероскедастичности. Т.к. Q'' > Q', подтвердилась прямая зависимость дисперсий регрессионных остатков от значений переменной X_4, и оценка матрицы Σ_0:

$$\hat{\sum} = \begin{pmatrix} x_{1j}^2 & 0 & 0\ldots & \\ 0 & x_{2j}^2 & 0\ldots & \\ \ldots & \ldots & \ldots & \\ 0 & 0 & \ldots & x_{nj}^2 \end{pmatrix}$$

12954,63744	0	0	0	0	0	0
0	23393,69	0	0	0	0	0
0	0	29463,45	0	0	0	0
0	0	0	14390,07	0	0	0
0	0	0	0	11949,94	0	0
0	0	0	0	0	9581,255	0
0	0	0	0	0	0	10999,41
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

Рисунок 6Фрагмент ковариационной матрицы остатков

```
Sigm = np.zeros((n,n)) # матрица n x n, заполненная нулями

if Gleyser == 1 :
   for i in range(0, n):
      for j in range(0, n):
         if i==j :
            Sigm[i][j]= (Rsqrd[umax][1] + 
Rsqrd[umax][2]*abs(X[i][maxk])**Rsqrd[umax][0])**2
else: print("Нет гетероскедастичности")

Sigmd = pd.DataFrame(data=Sigm)
Sigmd
```

	0	1	2	3	4	5	6	7	8	9	...	75	76	77	78	79	80	81	82	83	
0	34.202749	0.00000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.0
1	0.000000	41.83338	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.0
2	0.000000	0.00000	47.659269	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.0
3	0.000000	0.00000	0.000000	35.075467	0.000000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.0
4	0.000000	0.00000	0.000000	0.000000	33.625901	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
80	0.000000	0.00000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	29.364153	0.000000	0.000000	0.000000	0.0
81	0.000000	0.00000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000	36.073875	0.000000	0.000000	0.0
82	0.000000	0.00000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	81.602077	0.000000	0.0
83	0.000000	0.00000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	271.160166	0.0
84	0.000000	0.00000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.000000	33.4

5. Тест Глейзера

Позволяет более точно аппроксимировать зависимость между регрессионными остатками и объясняющей переменной, варьированием которой обуславливается гетероскедастичность. В качестве такой аппроксимации берётся линейная регрессия зависимости абсолютных значений остатков от значений объясняющей переменной в степени γ. Предположим, что:

$$|e_i|=\alpha+\beta^*|x_{il}|^{\gamma}+\sigma_i$$

Выдвигается гипотеза H_0: α=0, β =0 (нет гетероскедастичности).

Значения гамма варьируются, как правило, на отрезке [-3; 3] с шагом 0,5. Для каждого уравнения проверяют значимость; если присутствует хотя бы одно значимое уравнение, то есть гетероскедастичность.

Рассмотрим результаты построения регрессии при различных значениях γ:

γ	b_0	S_{b0}	b_1	S_{b1}	R^2	p-value	Fрасч
-3	-0.057809	0.109581	-9155.22	-0.52754	0.00334184	0.599223	0,27830
-2.5	-0.046314	0.109646	-1755.65	4156.438	0.00214498	0.673830	0,17842
-2	-0.032167	0.109707	-288.104	982.5861	0.00103474	0.770093	0,08597
-1.5	-0.014674	0.109752	-30.8413	230.6753	0.00021532	0.893963	0,01788
-1	-0.007350	0.109761	3.716589	55.50485	0.00005402	0.946775	0,00448
-0.5	0.035149	0.109696	5.058145	15.78596	0.00123545	0.749455	0,10267
0.5	-0.100906	0.109204	-0.185426	0.200674	0.01018202	0.358158	0,85380
1	-0.129205	0.108844	-0.010727	0.009037	0.01669403	0.238588	1,4091
1.5	-0.151177	0.108503	-0.000681	0.000489	0.02285454	0.167248	1,9413
2	-0.167532	0.108213	-0.000043	0.000028	0.02806684	0.125387	2,3968
2.5	-0.179265	0.107986	-0.000003	0.000002	0.03213610	0.100672	2,7559
3	-0.187147	0.107825	-0.000000	0.000000	0.03502393	0.086336	3,0125

Поскольку все $F_{расч} < F_{крит} = 3.98$, соответственно все уравнения регрессии являются не значимыми и гетероскедастичность отсутствует.

```
Xgm = np.zeros((n,13))   # это матрица со столбцом из 1 и X в различных степенях
gamma = -3
for j in range(0, 13):
    for i in range(0, n):
        Xgm[i][j] = abs(X[i][maxk])**gamma
    gamma = gamma + 0.5
Xgm = np.hstack([np.ones((Xgm.shape[0], 1)), Xgm])
Xgmd = pd.DataFrame(data=Xgm)
Rsqrd = np.zeros((13,6))
u = 0
gamma = -3
for i in range(1,14):
        Rsqrd[u][0] = gamma
        result = s.OLS(abs(E), Xgm[:, [0, i]]).fit()
        Rsqrd[u][1] = result.params[0]
        Rsqrd[u][2] = result.params[1]
        Rsqrd[u][3] = result.fvalue
        Rsqrd[u][4] = result.f_pvalue
        Rsqrd[u][5] = result.rsquared
        u += 1 # u нужно для того, чтобы идти по строкам при составлении таблицы в отчете
        gamma += 0.5
```

d = {"gamma": Rsqrd[::,0], "b0": Rsqrd[::,1], "b1": Rsqrd[::,2], "F-stat": Rsqrd[::,3], "p-value": Rsqrd[::,4], "R^2": Rsqrd[::,5]}

df = pd.DataFrame(data=d)

df

	gamma	b0	b1	F-stat	p-value	R^2
0	-3.0	6.466272	-8.149623e+03	0.220370	0.639989	2.648033e-03
1	-2.5	6.492191	-2.502401e+03	0.363275	0.548336	4.357731e-03
2	-2.0	6.547887	-7.609907e+02	0.603552	0.439435	7.219212e-03
3	-1.5	6.690690	-2.311805e+02	1.016466	0.316289	1.209842e-02
4	-1.0	7.127461	-7.227453e+01	1.730806	0.191930	2.042712e-02
5	-0.5	8.916359	-2.644453e+01	2.900760	0.092280	3.376874e-02
6	0.0	3.211198	3.211198e+00	NaN	NaN	3.330669e-16
7	0.5	0.958439	4.859161e-01	6.239886	0.014465	6.992261e-02
8	1.0	3.242626	2.413977e-02	7.664731	0.006944	8.453928e-02
9	1.5	4.199999	1.392931e-03	8.767109	0.003998	9.553650e-02
10	2.0	4.770702	8.284071e-05	9.706218	0.002522	1.046987e-01
11	2.5	5.149363	4.951495e-06	10.592507	0.001644	1.131769e-01
12	3.0	5.414040	2.945260e-07	11.454937	0.001091	1.212741e-01

```
Fcr = f.ppf(q = 1 - 0.05, dfn = 1, dfd = 85 - 1 - 1)
print("Fcrit = ", Fcr )
umax = 0
Gleyser = 0;
for u in range(0, 13): # цикл для выявления значимости уравнения\n",
    if Rsqrd[u][3] > Fcr:
        print("Уравнение с гамма = ", Rsqrd[u][0], "значимо")
        Gleyser = 1; #значит, что Глейзер выполнился
    else: print("Уравнение с гамма = ", Rsqrd[u][0], "не значимо")
```

```
Fcrit =  3.9559610071544236
Уравнение с гамма =  -3.0 не значимо
Уравнение с гамма =  -2.5 не значимо
Уравнение с гамма =  -2.0 не значимо
Уравнение с гамма =  -1.5 не значимо
Уравнение с гамма =  -1.0 не значимо
Уравнение с гамма =  -0.5 не значимо
Уравнение с гамма =  0.0 не значимо
Уравнение с гамма =  0.5 значимо
Уравнение с гамма =  1.0 значимо
Уравнение с гамма =  1.5 значимо
Уравнение с гамма =  2.0 значимо
Уравнение с гамма =  2.5 значимо
Уравнение с гамма =  3.0 значимо
```

6. Расчет оценок коэффициентов b ОМНК

Для борьбы с гетероскедастичностью строится модель регрессии с другими коэффициентами – полученными по обобщенному МНК – которые являются эффективными. Оценки вычисляются следующим образом:

$$b_{ОМНК}=(X^T\sum\nolimits^{-1}_0X)^{-1}(X^T\sum\nolimits^{-1}_0Y)$$

bомнк0	89,93432493
bомнк4	-0,116673154
bомнк5	-0,035047958
bомнк7	9,09284E-05

Рисунок 7 Оценки ОМНК

```
Xt = X.T #транспонированная матрица X
Sigm_inv = np.linalg.inv(Sigm) #обратная к сигма-матрице
XtSinv = np.dot(Xt,Sigm_inv) #их произведение
XtSinvY = np.dot(XtSinv,Y) #домноженное на Y
XtSinvXinv = np.linalg.inv(np.dot(XtSinv, X))#обратная к произведению
Bomnk = np.dot(XtSinvXinv, XtSinvY)
S = ((Y-X.dot(Bomnk)).T.dot(Sigm_inv).dot(Y-X.dot(Bomnk)))*(1/(n-k-1)) #матрица стандартных ошибок
Eb = S*np.linalg.inv(np.dot(X.T,Sigm_inv).dot(X)) #ковариационная матрица для bомнк
d = {"coef": Bomnk[:,0], "Std": np.sqrt(np.diagonal(Eb))}
df = pd.DataFrame(data=d)
print('Стандартная ошибка:', *S)
df
```

Стандартная ошибка: [1.38423368]

	coef	Std
0	89.839535	2.577748
1	-0.142876	0.028141
2	-0.073721	0.030934
3	0.000218	0.000094

=МУМНОЖ(МОБР(МУМНОЖ(МУМНОЖ(ТРАНСП(B2:E86); МОБР(Лист5!F2:CL86)); B2:E86)); МУМНОЖ(МУМНОЖ(ТРАНСП(bОМНК!B2:E86); МОБР(Лист5!F2:CL86)); A2:A86))

Рисунок 8Матричная формула для вычисления b ОМНК

Получили новое уравнение регрессии:

$$y\hat{} = 89{,}93 – 0{,}1167x_4 – 0{,}0351x_5 + 9{,}093x_7$$

которое не только не имеет незначимых признаков, но и не подвержено гетероскедастичности остатков.

7. Вывод

1. Оценка уравнения регрессии имеет вид: $y\hat{} = 89{,}93 – 0{,}1167x_4 – 0{,}0351x_5 + 9{,}093x_7$

2. При увеличении задолженностей по кредитам предоставленным кредитным организациями физическим лицам на 1, результаты голосования по поправкам в Конституцию уменьшатся на 0,1167.

3. При увеличении средств (вкладов) физических лиц на рублевых счетах в Сбербанке России на 1 результаты голосования по поправкам в Конституцию уменьшатся на 0,0351.

4. При увеличении среднедушевого дохода населения на 1 результаты голосования по поправкам в Конституцию увеличатся на 9,093.

Приложение А

Исходные данные:

Наименование	y	x4	x5	x7
Алтайский край	71,91	114	59	32729
Амурская область	70,42	153	74	59187
Архангельская область без автономного ок	65,78	172	121	57076
Астраханская область	86,7	120	56	41615
Белгородская область	81,76	109	84	41349
Брянская область	87,55	98	60	35578
Владимирская область	71,18	105	73	37708
Волгоградская область	83,39	102	58	38073
Вологодская область	71,16	129	77	45453
Воронежская область	80,07	109	93	40730
г. Москва	65,29	178	245	111395
Еврейская автономная область	77,30	122	67	50039
Забайкальский край	72,92	132	61	50193
Ивановская область	76,23	94	71	31881
Иркутская область	64,28	146	83	54318
Кабардино-Балкарская Республика	90,48	59	29	31172
Калининградская область	72,5	132	81	41298
Калужская область	70,17	139	86	47854
Камчатский край	61,76	201	163	92667
Карачаево-Черкесская Республика	84,8	77	30	32128
Кемеровская область — Кузбасс	87,29	123	63	48411
Кировская область	70,89	124	60	35390
Костромская область	67,92	107	67	34989
Краснодарский край	88,92	117	66	40955
Красноярский край	70,10	152	72	60026
Курганская область	69,58	121	51	35356
Курская область	73,92	107	59	40206
Ленинградская область	79,12	159	67	52513
Липецкая область	78,12	105	71	40185
Магаданская область	62,03	242	219	108460
Московская область	78,96	190	106	63414
Мурманская область	62,54	196	172	75495
Ненецкий автономный округ	44,42	216	185	44854
Нижегородская область	79,31	116	91	41162
Новгородская область	71,44	118	71	39473
Новосибирская область	67,59	155	66	46444
Омская область	62,09	124	56	40655
Оренбургская область	73,60	136	50	38146
Орловская область	79,74	115	77	35174
Пензенская область	85,25	108	71	36058
Пермский край	70,75	141	85	45178

Приморский край	**78,86**	144	93	55076
Псковская область	**78,74**	108	69	34727
Республика Адыгея (Адыгея)	**84,55**	95	36	35633
Республика Алтай	**66,16**	105	36	38702
Республика Башкортостан	**88,68**	134	50	41647
Республика Бурятия	**71,95**	115	42	44709
Республика Дагестан	**90,34**	26	21	32180
Республика Ингушетия	**85,89**	16	12	31230
Республика Калмыкия	**84,00**	147	40	33167
Республика Карелия	**70,46**	169	107	49303
Республика Коми	**65,08**	195	118	59721
Республика Крым	**90,07**	31	46	37993
Республика Марий Эл	**75,76**	104	50	35202
Республика Мордовия	**85,60**	101	46	34441
Республика Саха (Якутия)	**58,34**	219	73	82555
Республика Северная Осетия — Алания	**82,83**	87	48	32603
Республика Татарстан (Татарстан)	**82,81**	136	60	44854
Республика Тыва	**96,79**	129	23	46092
Республика Хакасия	**69,52**	118	47	47148
Ростовская область	**83,54**	113	65	39045
Рязанская область	**79,1**	122	80	40282
Самарская область	**80,55**	121	86	42729
Санкт-Петербург	**77,66**	195	136	75920
Саратовская область	**82,24**	104	60	37089
Сахалинская область	**74,84**	212	173	94587
Свердловская область	**65,97**	146	75	48347
Севастополь	**84,6**	34	172	39127
Смоленская область	**71,97**	116	75	36097
Ставропольский край	**84,85**	96	62	37090
Тамбовская область	**87,41**	101	71	33689
Тверская область	**71,87**	127	77	39692
Томская область	**64,86**	131	61	50481
Тульская область	**83,06**	124	71	43892
Тюменская область без автономных округ	**85,57**	176	74	55770
Удмуртская Республика	**68,92**	142	58	38564
Ульяновская область	**71,16**	115	69	35991
Хабаровский край	**62,28**	158	104	59247
Ханты-Мансийский автономный округ —	**69**	271	92	85211
Челябинская область	**69,54**	123	56	43477
Чеченская Республика	**97,92**	25	6	30942
Чувашская Республика — Чувашия	**72,57**	126	65	35531
Чукотский автономный округ	**80,3**	231	228	130483
Ямало-Ненецкий автономный округ	**89,16**	335	136	115987
Ярославская область	**70,69**	108	90	41389