

Übung zu Kapitel 8

1

$$V_i \sim \mathcal{N}(0, 36) \quad i = 1, \dots, 1000$$

$$U_i = V_i \cdot \sqrt{x_i}$$

$$\text{Var}(U_i | X) = \text{Var}(V_i \cdot \sqrt{x_i} | X)$$

$$= (\sqrt{x_i})^2 \text{Var}(V_i | X)$$

$$= x_i \cdot \underbrace{36}_{\sigma^2} = \sigma_i^2$$

Modell 1: KQ, benutze die Beobachtungen 1-1000
Abhängige Variable: y

Originalmodell
OLS

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	-0,0157166	2,75678	-0,005701	0,9955
x	2,03601	0,0474703	42,89	1,10e-228 ***

$\Rightarrow \sigma_i^2$ „usg 1“

Mittel abhängige Var.	102,6552	Stdabw. abhängige Var.	72,87111
Summe quad. Residuen	1865784	Stdfehler Regression	43,23798
R-Quadrat	0,648290	Korrigiertes R-Quadrat	0,647937
F(1, 998)	1839,562	P-Wert(F)	1,1e-228

Hilfsregression

Modell 2: KQ, benutze die Beobachtungen 1-1000
Abhängige Variable: usq1

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	-244,923	202,518	-1,209	0,2268
x	41,8562	3,48724	12,00	4,24e-31 ***

Mittel abhängige Var.	1865,784	Stdabw. abhängige Var.	3396,162
Summe quad. Residuen	1,01e+10	Stdfehler Regression	3176,332
R-Quadrat	0,126143	Korrigiertes R-Quadrat	0,125268
F(1, 998)	144,0638	P-Wert(F)	4,24e-31

H_0 verwerfen, Fehlerterme u_i sind heteroskedastisch

$$\sigma_i^2 = x_i \cdot 36$$

$$\Phi(v|x) = \begin{pmatrix} \underline{36} \cdot x_1 & & 0 \\ & \underline{36} \cdot x_2 & \\ & & \ddots \\ 0 & & & \underline{36} \cdot x_n \end{pmatrix} = \sigma^2 \cdot \Omega$$

$$\Rightarrow \Omega = \begin{pmatrix} x_1 & & 0 \\ & x_2 & \\ & & \ddots \\ 0 & & & x_n \end{pmatrix}$$

mit $\sigma^2 = 36$

$$\tilde{y}_i = y_i \cdot \frac{1}{\sqrt{x_i}}$$

$$\tilde{1} = 1 \cdot \frac{1}{\sqrt{x_i}}$$

$$\tilde{x}_i = x_i \cdot \frac{1}{\sqrt{x_i}} = \sqrt{x_i}$$

$$P = \begin{pmatrix} \frac{1}{\sqrt{x_1}} & & 0 \\ & \frac{1}{\sqrt{x_2}} & \\ & & \ddots \\ 0 & & & \frac{1}{\sqrt{x_n}} \end{pmatrix}$$

Modell 3: KQ, benutze die Beobachtungen 1-1000

Abhängige Variable: t_y

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
t_const	-1,01751	0,848630	-1,199	0,2308
t_x	2,05587	0,0313958	65,48	0,0000 ***

Modell mit transformierten Daten
 $\rightarrow$ OLS-Schätzung $\rightarrow \sigma_i^2$: "usq3"

Hilfsregression $usq3 = \alpha_0 + \alpha_1 \tilde{t} + \alpha_2 \tilde{x} + \varepsilon$

Modell 5: KQ, benutze die Beobachtungen 1-1000

Abhängige Variable: usq3

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	28,8106	8,56875	3,362	0,0008 ***
t_const	-2,18066	12,8214	-0,1701	0,8650
t_x	1,04004	0,962156	1,081	0,2800
Mittel abhängige Var.	35,35396	Stdabw. abhängige Var.	49,53799	
Summe quad. Residuen	2444035	Stdfehler Regression	49,51151	
R-Quadrat	0,003069	Korrigiertes R-Quadrat	0,001069	
F(2, 997)	<u>1,534506</u>	P-Wert(F)	0,216071	

$H_0: \alpha_1 = 0$ und $\alpha_2 = 0 \rightarrow F = 1,535 < C_{F_{2,997}} \approx 3 \rightarrow H_0$ nicht
 Verwerfen
 $\tilde{t}$ nicht heterosq.

2

Modell 1: KQ, benutze die Beobachtungen 1-2017

Abhängige Variable: nettfa

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	-10,5710	2,06068	-5,130	3,18e-07 ***
inc	0,820681	0,0609000	13,48	1,08e-39 ***
Mittel abhängige Var.	13,59498	Stdabw. abhängige Var.	47,59058	
Summe quad. Residuen	4188483	Stdfehler Regression	45,59223	
R-Quadrat	0,082673	Korrigiertes R-Quadrat	0,082218	
F(1, 2015)	181,5995	P-Wert(F)	1,08e-39	

Modell 2: KQ, benutze die Beobachtungen 1-2017

Abhängige Variable: usq1

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	-1886,36	1537,68	-1,227	0,2201
inc	134,583	45,4437	2,962	0,0031 ***
Mittel abhängige Var.	2076,590	Stdabw. abhängige Var.	34086,54	
Summe quad. Residuen	2,33e+12	Stdfehler Regression	34021,04	
R-Quadrat	0,004334	Korrigiertes R-Quadrat	0,003840	
F(1, 2015)	8,770638	P-Wert(F)	0,003097	

$$usq1 = \alpha_0 + \alpha_1 inc + \epsilon$$

$H_0: \alpha_1 = 0$ Verwerfen $\leadsto u_i$ aus Modell 1 heteroskedastisch

$$\text{Var}(v_i | X) = \sigma^2 \cdot \text{inc}_i = \sigma_i^2$$

$$\frac{\sigma_i^2}{\sigma^2} = \text{inc}_i \quad (\Omega)$$

$$\frac{1}{\text{inc}_i} \quad (\Omega^{-1})$$

$$\frac{1}{\sqrt{\text{inc}_i}} \quad (P)$$

$$\widetilde{\text{netffa}_i} = \text{netffa}_i / \sqrt{\text{inc}_i}$$

$$\widetilde{L}_i = 1 / \sqrt{\text{inc}_i}$$

$$\widetilde{\text{inc}_i} = \text{inc}_i / \sqrt{\text{inc}_i} = \sqrt{\text{inc}_i}$$

OLS-Regression des transformierten Modells

Modell 3: KQ, benutze die Beobachtungen 1-2017
Abhängige Variable: t_netfta

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
t_const	-9,58070	1,65328	-5,795	7,91e-09 ***
t_inc	0,787052	0,0634814	12,40	4,52e-34 ***
Mittel abhängige Var.	2,180711			
Stdabw. abhängige Var.	7,381265			
Summe quad. Residuen	105019,5			
Stdfehler Regression	7,219338			

$\leadsto \hat{\tilde{u}}_i^2$ "usq3"

Hilfsregression

Modell 4: KQ, benutze die Beobachtungen 1-2017
Abhängige Variable: usq3

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	125,013	474,345	0,2635	0,7922
t_const	-548,898	1238,08	-0,4433	0,6576
t_inc	7,29225	43,6424	0,1671	0,8673

Mittel abhängige Var.	52,06716	Stdabw. abhängige Var.	838,7603
Summe quad. Residuen	1,42e+09	Stdfehler Regression	838,3889
R-Quadrat	0,001876	Korrigiertes R-Quadrat	0,000885
F(2, 2014)	<u>1,893169</u>	P-Wert(F)	<u>0,150862</u>

$$usq3 = \alpha_0 + \alpha_1 \tilde{t} + \alpha_2 \ln \tilde{t}_i + \varepsilon$$

$$H_0: \alpha_1 = 0 \text{ \& } \alpha_2 = 0$$

nicht verwerfen

#3

a)

OLS

klassische SE

Modell 3: KQ, benutze die Beobachtungen 1-2017

Abhängige Variable: nettfa

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	-20,9850	2,47202	-8,489	3,98e-17 ***
inc	0,770583	0,0614520	12,54	8,73e-35 ***
age_25_sq	0,0251267	0,00259339	9,689	9,96e-22 ***
male	2,47793	2,04778	1,210	0,2264
e401k	6,88622	2,12327	3,243	0,0012 ***
Mittel abhängige Var.	13,59498	Stdabw. abhängige Var.	47,59058	
Summe quad. Residuen	3982124	Stdfehler Regression	44,48805	
R-Quadrat	0,127868	Korrigiertes R-Quadrat	0,126134	
F(4, 2012)	73,74763	P-Wert(F)	2,18e-58	

Modell 4: KQ, benutze die Beobachtungen 1-2017

Abhängige Variable: nettfa

Heteroskedastizitäts-robuste Standardfehler, Variante HC1

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	-20,9850	3,49519	-6,004	2,28e-09 ***
inc	0,770583	0,0995719	7,739	1,57e-14 ***
age_25_sq	0,0251267	0,00434415	5,784	8,44e-09 ***
male	2,47793	2,05836	1,204	0,2288
e401k	6,88622	2,28658	3,012	0,0026 ***
Mittel abhängige Var.	13,59498	Stdabw. abhängige Var.	47,59058	
Summe quad. Residuen	3982124	Stdfehler Regression	44,48805	
R-Quadrat	0,127868	Korrigiertes R-Quadrat	0,126134	
F(4, 2012)	28,96073	P-Wert(F)	1,95e-23	

OLS
robuste SE

weighted least squares

$$\Omega = \begin{pmatrix} \text{inc}_1 & & & \\ & \text{inc}_2 & & \\ & & \ddots & \\ & & & \text{inc}_n \\ & 0 & & & \end{pmatrix}$$

$$\Omega^{-1} = \begin{pmatrix} \frac{1}{\text{inc}_1} & & & \\ & \frac{1}{\text{inc}_2} & & \\ & & \ddots & \\ & & & \frac{1}{\text{inc}_n} \\ & 0 & & & \end{pmatrix}$$

$$\Omega^{-\frac{1}{2}} = \begin{pmatrix} \frac{1}{\sqrt{\text{inc}_1}} & & & \\ & \frac{1}{\sqrt{\text{inc}_2}} & & \\ & & \ddots & \\ & & & \frac{1}{\sqrt{\text{inc}_n}} \\ & 0 & & & \end{pmatrix} = P$$

$$\tilde{y} = P \cdot y$$

$$\tilde{X} = P \cdot X$$

a)
WLS
normale se

Modell 7: KQ, benutze die Beobachtungen 1-2017
Abhängige Variable: tnettfa

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
tconst	-16,7025	1,95799	-8,530	2,82e-17	***
tinc	0,740384	0,0643029	11,51	9,38e-30	***
tage_25_sq	0,0175373	0,00193150	9,080	2,53e-19	***
tmale	1,84053	1,56359	1,177	0,2393	
te401k	5,18828	1,70343	3,046	0,0024	***

WLS
robuste se

Modell 8: KQ, benutze die Beobachtungen 1-2017
Abhängige Variable: tnettfa
Heteroskedastizitäts-robuste Standardfehler, Variante HC1

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
tconst	-16,7025	2,24298	-7,447	1,41e-13	***
tinc	0,740384	0,0750557	9,864	1,90e-22	***
tage_25_sq	0,0175373	0,00258488	6,785	1,53e-11	***
tmale	1,84053	1,31081	1,404	0,1604	
te401k	5,18828	1,57186	3,301	0,0010	***

b) Breusch Pagan Test für Modell 3 (OLS normale se)

Modell 9: KQ, benutze die Beobachtungen 1-2017
Abhängige Variable: usq3

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	-4573,55	1848,70	-2,474	0,0134 **
inc	$\hat{\delta}_1$ 112,358	45,9568	2,445	0,0146 **
age_25_sq	$\hat{\delta}_2$ 4,84866	1,93946	2,500	0,0125 **
male	$\hat{\delta}_3$ 2331,25	1531,43	1,522	0,1281
e401k	$\hat{\delta}_4$ 1164,83	1587,89	0,7336	0,4633
Mittel abhängige Var.	1974,280	Stdabw. abhängige Var.	33367,52	
Summe quad. Residuen	2,23e+12	Stdfehler Regression	33270,33	
R-Quadrat	0,007789	Korrigiertes R-Quadrat	0,005817	
F(4, 2012)	3,948695	P-Wert(F)	0,003387	< 1%

$$H_0 : \delta_1 = 0 \ \& \ \delta_2 = 0 \ \& \ \delta_3 = 0 \ \& \ \delta_4 = 0$$

verwerfen mit $\alpha = 1\%$

→ MLR5 verletzt u_i heteroskedastisch

Breusch Pagan Test für Modell 7 (WLS normale se)

Breusch-Pagan-Test auf Heteroskedastizität

KQ, benutze die Beobachtungen 1-2017

Abhängige Variable: skalierte uhat^2

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	2,08411	9,45880	0,2203	0,8256
tconst	-14,7478	24,8743	-0,5929	0,5533
tinc	0,139673	0,867015	0,1611	0,8720
tage_25_sq	0,00949000	0,00453877	2,091	0,0367 **
tmale	3,76456	3,67494	1,024	0,3058
te401k	2,18449	4,02711	0,5424	0,5876

Erklärte Quadratsumme = 2301,27

Teststatistik: LM = 1150,633672,

mit p-Wert = $P(\text{Chi-Quadrat}(5) > 1150,633672) = 0,000000$

Verwerfe H_0 (Homoskedastizität)!

Auch Modelle ^{OLS} 4 & ^{WLS} 8 weisen Heteroskedastizität auf
robuste se

c) $\text{Var}(u_i | x) = \sigma^2 \cdot \text{inc}_i$ ungeeignet.

4
a)

Modell 1: KQ, benutze die Beobachtungen 1-526
Abhängige Variable: lwage

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	0,216854	0,108595	1,997	0,0464 **
educ	0,0979356	0,00762240	12,85	4,96e-33 ***
exper	0,0103469	0,00155514	6,653	7,24e-11 ***
Mittel abhängige Var.	1,623268	Stdabw. abhängige Var.	0,531538	
Summe quad. Residuen	111,3447	Stdfehler Regression	0,461407	
R-Quadrat	0,249343	Korrigiertes R-Quadrat	0,246473	
F(2, 523)	86,86167	P-Wert(F)	2,68e-33	

Breusch Pagan:

Modell 2: KQ, benutze die Beobachtungen 1-526
Abhängige Variable: usql

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	0,0493227	0,0793420	0,6216	0,5344
educ	0,00818592	0,00556910	1,470	0,1422
exper	0,00349776	0,00113622	3,078	0,0022 ***
Mittel abhängige Var.	0,211682	Stdabw. abhängige Var.	0,339611	
Summe quad. Residuen	59,43689	Stdfehler Regression	0,337114	
R-Quadrat	0,018405	Korrigiertes R-Quadrat	0,014651	
F(2, 523)	4,903145	P-Wert(F)	0,007768	

$< 0,01 = 1\%$

H_0 Homoskedastizität verwerfen!

multiplikative Heteroskedastizität

b)

$$\text{Var}(u_i|x) = \sigma_i^2 = e^{\alpha_0 + \alpha_1 \text{educ}_i + \alpha_2 \text{exper}_i}$$

$$\hat{\sigma}_i^2 = e^{\widehat{\ln(\hat{u}_i^2)}}$$

$$\Rightarrow \ln(\sigma_i^2) = \alpha_0 + \alpha_1 \text{educ}_i + \alpha_2 \text{exper}_i$$

Schätze σ_i^2 durch $\hat{\sigma}_i^2 \leftarrow$ aus OLS Regression a)

Hilfsregression $\leadsto$ Modell 3

$$\ln(\hat{u}_i^2) = \alpha_0 + \alpha_1 \text{educ}_i + \alpha_2 \text{exper}_i + \varepsilon_i$$

OLS $\leadsto \hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2$

$$\widehat{\ln(\hat{u}_i^2)} = \hat{\alpha}_0 + \hat{\alpha}_1 \text{educ}_i + \hat{\alpha}_2 \text{exper}_i$$

Modell 3: KQ, benutze die Beobachtungen 1-526
 Abhängige Variable: l_usq1

		Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	α_0	-4,09325	0,479215	-8,542	1,45e-16	***
educ	α_1	0,0732219	0,0336366	2,177	0,0299	**
exper	α_2	0,0257570	0,00686261	3,753	0,0002	***

Mittel abhängige Var.		-2,735073	Stdabw. abhängige Var.		2,061761	
Summe quad. Residuen		2168,254	Stdfehler Regression		2,036124	
R-Quadrat		0,028430	Korrigiertes R-Quadrat		0,024715	
F(2, 523)		7,651982	P-Wert(F)		0,000530	

Fitted values / geschätzte Werte $\hat{\ln(\hat{\sigma}_i^2)}$ $\leadsto$ "yhat3"

c) transformiere Daten

$$\tilde{y}_i = y_i / \sqrt{\hat{\sigma}_i^2}$$

$$\tilde{educ}_i = educ_i / \sqrt{\hat{\sigma}_i^2}$$

$$\tilde{exper}_i = exper_i / \sqrt{\hat{\sigma}_i^2}$$

$$\tilde{const} = 1 / \sqrt{\hat{\sigma}_i^2}$$

Modell 4: KQ, benutze die Beobachtungen 1-526

Abhängige Variable: tlwage

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
tconst	0,252804	0,0996672	2,536	0,0115	**
teduc	0,0916702	0,00727465	12,60	5,55e-32	***
texper	0,0131062	0,00165649	7,912	1,52e-14	***

Breusch - Pagan Test

Breusch-Pagan-Test auf Heteroskedastizität

KQ, benutze die Beobachtungen 1-526

Abhängige Variable: skalierte $\hat{u}^2$

		Koeffizient	Std.-fehler	t-Quotient	p-Wert
const		4,57469	5,01698	0,9118	0,3623
tconst	δ_1	-0,127870	0,461405	-0,2771	0,7818
teduc	δ_2	-0,0454246	0,0455350	-0,9976	0,3189
texper	δ_3	-0,0135413	0,0167212	-0,8098	0,4184

Erklärte Quadratsumme = 13,891

Teststatistik: LM = 6,945489,

mit p-Wert = $P(\text{Chi-Quadrat}(3) > 6,945489) = 0,073656 > 0,05 = 5\%$

$H_0: \delta_1 = 0 \ \& \ \delta_2 = 0 \ \& \ \delta_3 = 0$ (homoskedastische Störterme)

Kann mit $\alpha = 5\%$ nicht zurückgewiesen werden.

White Test

White's Test für Heteroskedastizität

KQ, benutze die Beobachtungen 1-526

Abhängige Variable: $\hat{u}^2$

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	12294,8	13588,0	0,9048	0,3660
tconst	79548,5	76980,3	1,033	0,3019
teduc	-1045,34	1043,94	-1,001	0,3171
texper	-365,862	366,839	-0,9973	0,3191
sq_tconst	-10486,6	10174,1	-1,031	0,3032
X2_X3	-254,501	244,570	-1,041	0,2985
X2_X4	-89,6589	86,0631	-1,042	0,2980
sq_teduc	-2,65905	2,46258	-1,080	0,2807
X3_X4	-1,88859	1,73622	-1,088	0,2772
sq_texper	-0,335569	0,306036	-1,097	0,2734

Warnung: Datenmatrix ist nahe an Singularität!

Unkorrigiertes R-Quadrat = 0,042663

Teststatistik: $TR^2 = 22,440692$,

mit p-Wert = $P(\text{Chi-Quadrat}(9) > 22,440692) = 0,007582 < 0,01 = 1\%$

H_0 verwerfen!