



Methoden der Politikwissenschaft

Regressionsanalyse

Siegfried Schumann



- Backhaus-Formel für den **Regressionskoeffizienten** (S. 16)
 - äquivalent zum Grundkurs!
$$b = \frac{SAP}{SAQ_x}$$
 - aber: Rohwerte!
 - in SPSS: B

- Backhaus-Formel für die **Regressionskonstante** (S. 16)
 - wie im Grundkurs
$$a = \bar{y} - b \cdot \bar{x}$$
 - aber Benennung: b_0 !
 - in SPSS: CONST



- Standardisierte Regressionskoeffizienten:

$$\hat{b} = b \cdot \frac{SD_x}{SD_y}$$

- \hat{b} bei Backhaus: → **Standardisierung**
- entspricht dem Koeffizienten für z-standardisierte Variablen!
- Sinn: Vergleichbarkeit innerhalb einer Gleichung herstellen!
- In SPSS: BETA

Arbeitsblatt 1: Eine unabhängige Variable

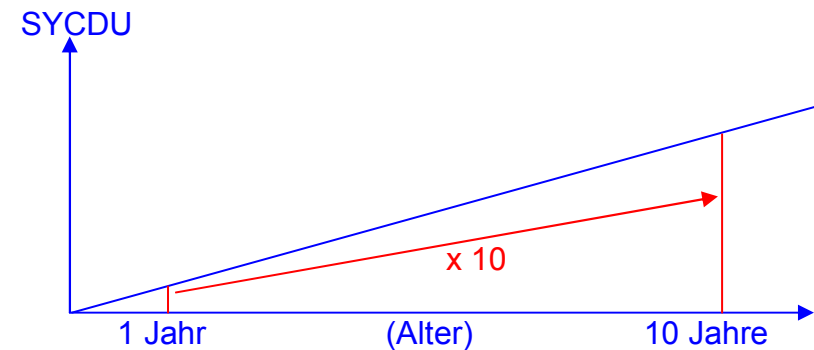
r aus Gleichungen 1 u. 2 errechenbar?

Konstante = ?

Modell Nr.		Regressionsgleichung	B	CONST	BETA	R ²	r
1	Ausgangsgleichung (bivariat)	SYCDU = B x AJHR + CONST	* .04	4.80	.18	.03	.18
2	UV und AV vertauscht	AJHR = B x SYCDU + CONST	***.90	36.21	.18	.03	.18
3	Alter in Jahrzehnten	SYCDU = B x AJHZ + CONST	** .36	4.80	.18	.03	.18
4	Alter + CDU-Symp. z-standard	SYCDUz = B x AJHRz + CONST	.18	.00	.18	.03	.18
5	nur Alter z-standardisiert	SYCDU = B x AJHRz + CONST	.58	6.32	.18	.03	.18
6	nur CDU-Symp. z-standardisiert	SYCDUz = B x AJHR + CONST	.01	-.47	.18	.03	.18

Quelle: Emnid-Umfrage 1989 (WEST)

- * genauer Wert: .036195
- ** genauer Wert: .361951
- *** genauer Wert: .898312





- Regressionsgleichung:

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_jx_j \dots + b_jx_j$$

- Zielfunktion der multiplen Regressionsfunktion:

$$\sum_{k=1}^K e_k^2 = \min \quad e = \text{Residuum}$$

Arbeitsblatt 1: y-Werte für Gleichung $y = 2x_1 + x_2 - 4$



x_2

6	2	4	6	8	10	12	14
5	1	3	5	7	9	11	13
4	0	2	4	6	8	10	12
3	-1	1	3	5	7	9	11
2	-2	0	2	4	6	8	10
1	-3	-1	1	3	5	7	9
0	-4	-2	0	2	4	6	8
	0	1	2	3	4	5	6

Lego-Modell!

x_1



- Regressionsgerade →
- Regressionskoeffizienten:
- Interpretationen:
 - Regressionskoeffizienten
 - Regressionskonstante

analog zum bivariaten Modell

Arbeitsblatt 1: z-Standardisierung / Unabhängigkeit



Hätte R^2 auch kleiner sein können als in vorhergehender Tabelle?

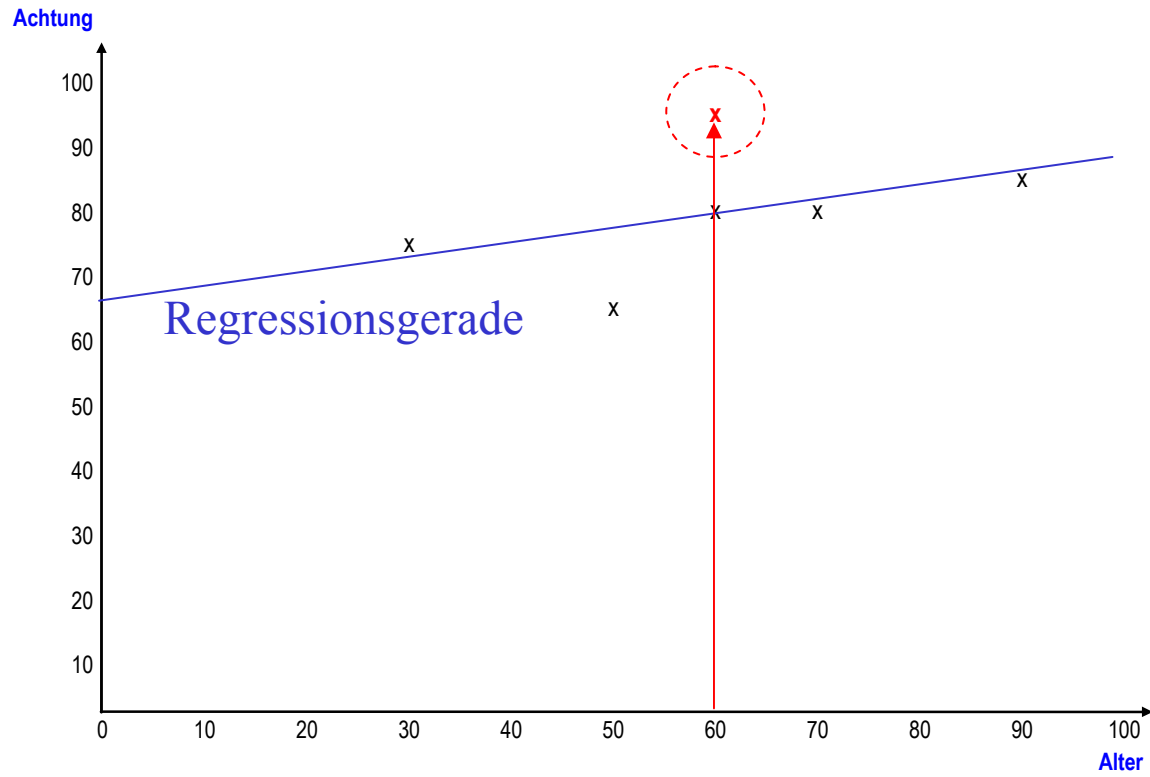
bereits bekannt?

Mo- dell Nr.	AV	Regressionen						Korrelationen		
		B_{AJHR}	B_{LIRE}	CONST	$BETA_{AJHR}$	$BETA_{LIRE}$	R^2	SYCDU x AJHR	SYCDU x LIRE	AJHR x LIRE
1	SYCDU	.022	.92	.46	.11	.61	.40	.18	.63	.11
2	SYCDUZ	.11	.61	.00	.11	.61	.40	SYCDUZ x AJHRz	SYCDUZ x LIREz	AJHRz x LIREz
3	SYCDUZ	.11	.28	-1.52	.11	.61	.40	SYCDUZ x AJHRz	SYCDUZ x LIRE	AJHRz x LIRE
4	SYCDU	.36	.92	1.40	.11	.61	.40	SYCDU x AJHRz	SYCDU x LIRE	AJHRz x LIRE
5	SYCDUZ	.01	.61	-.29	.11	.61	.40	SYCDUZ x AJHR	SYCDUZ x LIREz	AJHR x LIREz
6	SYCDU	.36	1.98	6.33	.11	.61	.40	SYCDU x AJHRz	SYCDU x LIREz	BAJHRz x LIREz
7	SYCDUZ	.01	.28	-1.81	.11	.61	.40	SYCDUZ x AJHR	SYCDUZ x LIRE	AJHR x LIRE

Quelle: Emnid-Umfrage 1989 (WEST)

Nächste Zeile: Welche Werte sind bekannt?

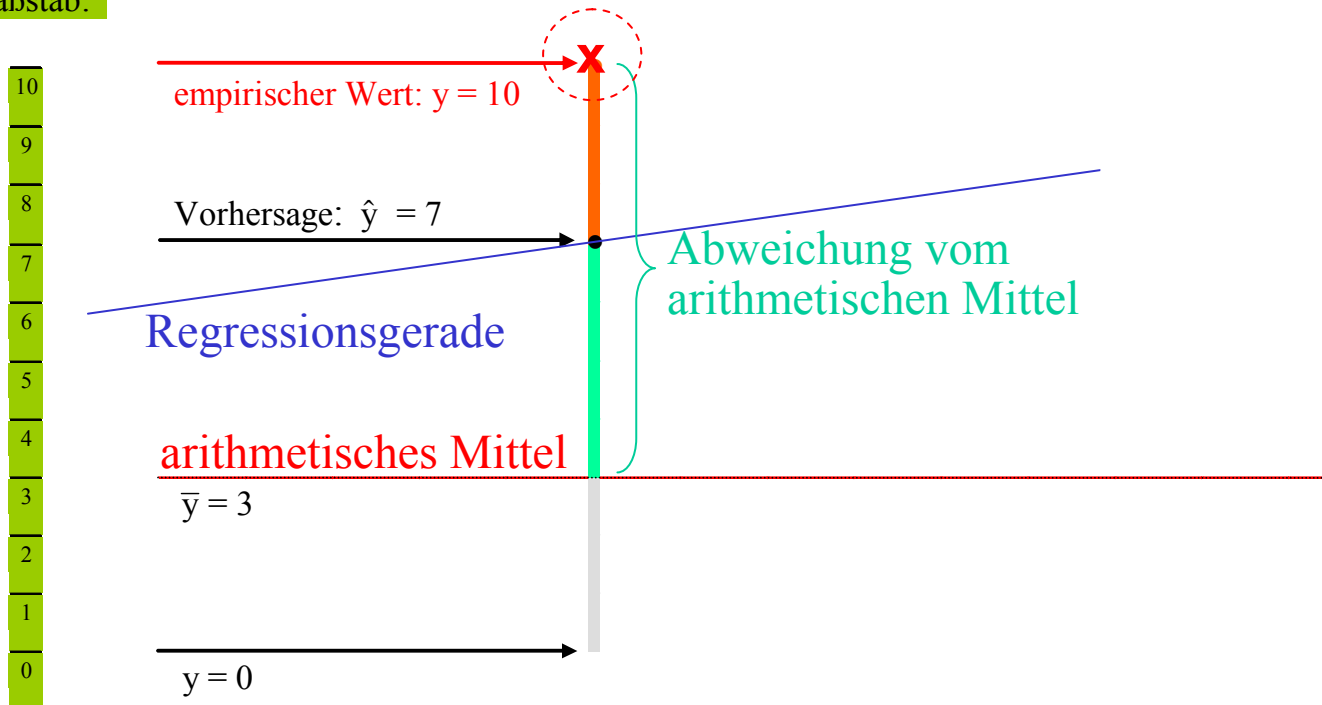
Wdh: Bestimmtheitsmaß R^2 – Zerlegung eines Werts



Wdh: Zerlegung eines y-Wertes



Maßstab:



Werte für Beispiel 2 (Grundkurs)



Wdh: Berechnung von R²

i	Alter (Jahre)			Achtung (Punkte)				\hat{y}_i	$\hat{y}_i - \bar{y}$	$(\hat{y}_i - \bar{y})^2$	$y_i - \hat{y}_i$	$(y_i - \hat{y}_i)^2$
	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x}) \cdot (y_i - \bar{y})$					
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1	50	-10	100	65	-15	225	150	77.75	-2.25	5.0625	-12.75	162.5625
2	30	-30	900	75	-5	25	150	73.25	-6.75	45.5625	1.75	3.0625
3	60	0	0	80	0	0	0	80.00	0	0	0	0
4	70	10	100	80	0	0	0	82.25	2.25	5.0625	-2.25	5.0625
5	90	30	900	85	5	25	150	86.75	6.75	45.5625	-1.75	3.0625
6	60	0	0	95	15	225	0	80.00	0	0	15.00	225.0000
$\bar{x} = 60$			$\bar{y} = 80$									

SAQ_x
= 2000

SAQ_y
= 500

SAP
= 450

$$R^2 = \frac{101.25}{500} = 0.2025$$

Gesamt-Variation:
500

erklärte Variation:
101.25

unerklärte Variation:
398.75



- Prüfung der Regressionsfunktion insgesamt
- Backhaus: $R^2 \rightarrow r^2$
- R^2 wird beeinflusst von:
 - Zahl der Regressoren
 - Größe der Stichprobe

- Korrigiertes Bestimmtheitsmaß:

$$r_{\text{korr}}^2 = r^2 - \frac{J \cdot (1 - r^2)}{K - J - 1}$$

J = Anzahl der Regressoren
K = Zahl der Beobachtungen

- Bei großen Fallzahlen kaum bedeutsam.

Beispiel: $N=K=1000, J=5$:

$$\begin{aligned} r^2 = 0.81 &\rightarrow r_{\text{korr}}^2 = 0.8090 \\ r^2 = 0.49 &\rightarrow r_{\text{korr}}^2 = 0.4874 \\ r^2 = 0.09 &\rightarrow r_{\text{korr}}^2 = 0.0854 \end{aligned}$$

- Anmerkung: $R = r =$ Multipler Korrelationskoeffizient



- F-Test: Hat sich $R^2_{emp} \neq 0$ zufällig ergeben, (Backhaus: $r^2_{emp} \neq 0$)
obwohl in der Grundgesamtheit R^2 gleich „0“ ist?
 - H_0 : **Kein** Zusammenhang zwischen der AV und den UVs (in der Grundgesamtheit)
 - H_1 : Zusammenhang zwischen der AV und den UVs (in der Grundgesamtheit)

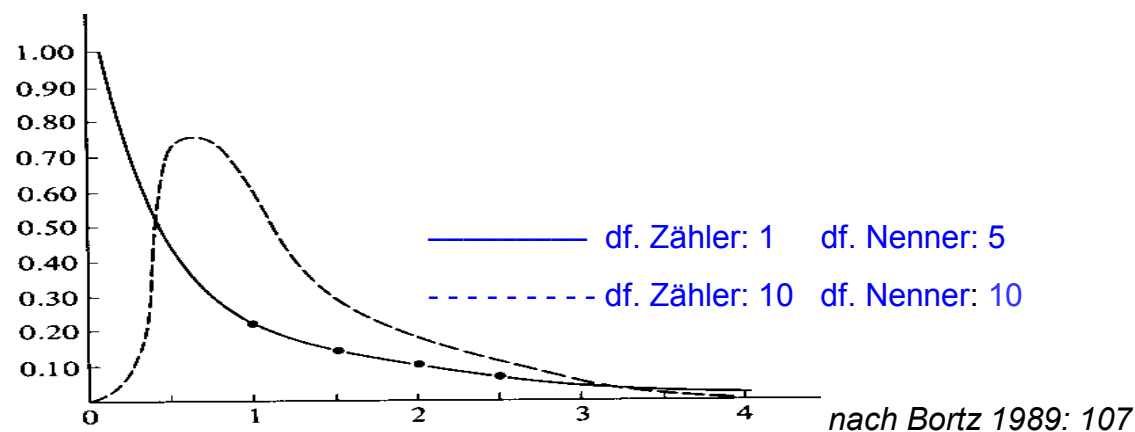
- Testgröße:

$$F_{emp} = \frac{\frac{r^2}{1-r^2}}{\frac{J}{K-J-1}} = \frac{\frac{\text{erklärter Streuungsanteil}}{J}}{\frac{\text{unerklärter Streuungsanteil}}{K-J-1}}$$

Zähler- bzw. Nennerfreiheitsgrade

J = Anzahl der Regressoren
K = Zahl der Beobachtungen

- F-Verteilung:
(Dichte)



F-Tabelle - Vertrauenswahrscheinlichkeit 0.95 / 0.99



Beispiel: J = 1
K = 10

Kritische Werte $F_{\alpha; f_1; f_2}$ der F-Verteilung für $\alpha = 0,05$ und $\alpha = 0,01$ (fett gedruckt)

$F_{emp} = 4.23 \rightarrow$

$f_1 =$ Freiheitsgrade für die Varianz im Zähler von F

$f_1 = J$: Anzahl der UVs

$f_2 = K - J - 1$: Fallzahl - Anzahl der UVs - 1

$f_2 =$ Freiheitsgrade für die Varianz im Nenner von F

	1	2	3	4	5	6	7	8	9	10	11	12	
1	161 4052	200 4999	216 5403	225 5625	230 5764	234 5859	237 5928	239 5981	241 6022	242 6056	243 6082	244 6106	1
2	18,51 98,49	19,00 99,00	19,16 99,17	19,25 99,25	19,30 99,30	19,33 99,33	19,36 99,34	19,37 99,36	19,38 99,38	19,39 99,40	19,40 99,41	19,41 99,42	2
3	10,13 34,12	9,55 30,82	9,28 29,46	9,12 28,71	9,01 28,24	8,94 27,91	8,88 27,67	8,84 27,49	8,81 27,34	8,78 27,23	8,76 27,13	8,74 27,05	3
4	7,71 21,20	6,94 18,00	6,59 16,69	6,39 15,98	6,26 15,52	6,16 15,21	6,09 14,98	6,04 14,80	6,00 14,66	5,96 14,54	5,93 14,45	5,91 14,37	4
5	6,61 16,26	5,79 13,27	5,41 12,06	5,19 11,39	5,05 10,97	4,95 10,67	4,88 10,45	4,82 10,27	4,78 10,15	4,74 10,05	4,70 9,96	4,68 9,89	5
6	5,99 13,74	5,14 10,92	4,76 9,78	4,53 9,15	4,39 8,75	4,28 8,47	4,21 8,26	4,15 8,10	4,10 7,98	4,06 7,87	4,03 7,79	4,00 7,72	6
7	5,59 12,25	4,74 9,55	4,35 8,45	4,12 7,85	3,97 7,46	3,87 7,19	3,79 7,00	3,73 6,84	3,68 6,71	3,63 6,62	3,60 6,54	3,57 6,47	7
8	5,32 11,26	4,46 8,65	4,07 7,59	3,84 7,01	3,69 6,63	3,58 6,37	3,50 6,19	3,44 6,03	3,39 5,91	3,34 5,82	3,31 5,74	3,28 5,67	8
9	5,12 10,56	4,26 8,02	3,86 6,99	3,63 6,42	3,48 6,06	3,37 5,80	3,29 5,62	3,23 5,47	3,18 5,35	3,13 5,26	3,10 5,18	3,07 5,11	9
10	4,96 10,04	4,10 7,56	3,71 6,55	3,48 5,99	3,33 5,64	3,22 5,39	3,14 5,20	3,07 5,06	3,02 4,95	2,97 4,85	2,94 4,78	2,91 4,71	10
11	4,84 9,65	3,98 7,20	3,59 6,22	3,36 5,67	3,20 5,32	3,09 5,07	3,01 4,88	2,95 4,74	2,90 4,63	2,86 4,54	2,82 4,46	2,79 4,40	11
12	4,75 9,33	3,88 6,93	3,49 5,95	3,26 5,41	3,11 5,06	3,00 4,82	2,92 4,65	2,85 4,50	2,80 4,39	2,76 4,30	2,72 4,22	2,69 4,16	12



- Standardfehler der Schätzung $\sqrt{\frac{\sum_k e_k^2}{K - J - 1}}$ Gibt an, welcher mittlere Fehler bei der Verwendung der Regressionsfunktion zur Schätzung der AV gemacht wird. (Backhaus u.a. 2006: 73)

- Im Beispiel: $\sqrt{\frac{1188685}{10 - 1 - 1}} = 385$

- Bezogen auf den Mittelwert von Y (1806.8)
 - beträgt der Standardfehler der Schätzung
 - was nicht als gut bewertet werden kann

$$\longleftrightarrow \boxed{\frac{385}{1806.8} = 0.21}$$

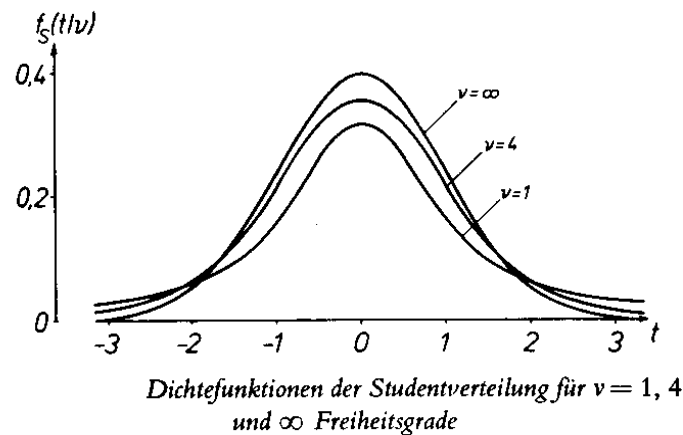
(Backhaus u.a.. 2006: 73)



- Für jede der unabhängigen Variablen (UV) ist H_0 gegen H_1 zu prüfen:
 - H_0 : Der Regressionskoeffizient hat in der Grundgesamtheit den Wert „0“ (kein „Einfluss“ der betreffenden UV auf die AV!)
 - H_1 : Der Regressionskoeffizient hat in der Grundgesamtheit einen Wert \neq „0“ („Einfluss“ der betreffenden UV auf die AV!)

- Testgröße: $t_{\text{emp}} = \frac{b_j}{s_{b_j}}$

- T-Verteilung (Dichte)



(aus: Bley Müller u.a. 1992: 63)

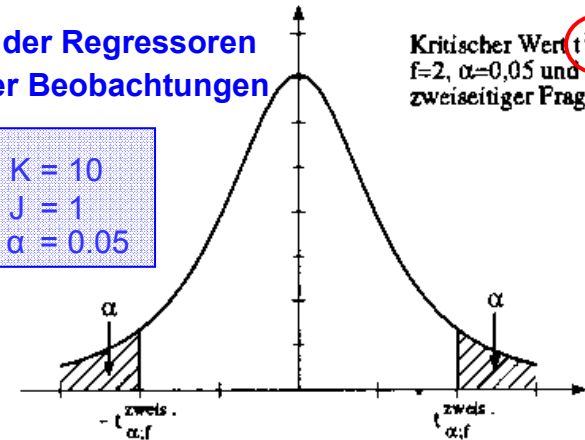
Kritische Werte der t-Verteilung (Wdh.)

f	Irrtumswahrscheinlichkeit α in % für zweiseitige Fragestellungen							
	50	25	10	5	2	1	0,2	0,1
1	1,00	2,41	6,31	12,7	31,82	63,7	318,3	637,0
2	,816	1,60	2,92	4,30	6,97	9,92	22,33	31,6
3	,765	1,42	2,35	3,18	4,54	5,84	10,22	12,9
4	,741	1,34	2,13	2,78	3,75	4,60	7,17	8,61
5	,727	1,30	2,01	2,57	3,37	4,03	5,89	6,86
6	,718	1,27	1,94	2,45	3,14	3,71	5,21	5,96
7	,711	1,25	1,89	2,36	3,00	3,50	4,79	5,40
8	,706	1,24	1,86	2,31	2,90	3,36	4,50	5,04
9	,703	1,23	1,83	2,26	2,82	3,25	4,30	4,78
10	,700	1,22	1,81	2,23	2,76	3,17	4,14	4,59
11	,697	1,21	1,80	2,20	2,72	3,11	4,03	4,44
12	,695	1,21	1,78	2,18	2,68	3,05	3,93	4,32
13	,694	1,20	1,77	2,16	2,65	3,01	3,85	4,22
14	,692	1,20	1,76	2,14	2,62	2,98	3,79	4,14
15	,691	1,20	1,75	2,13	2,60	2,95	3,73	4,07
16	,690	1,19	1,75	2,12	2,58	2,92	3,69	4,01
17	,689	1,19	1,74	2,11	2,57	2,90	3,65	3,96
18	,688	1,19	1,73	2,10	2,55	2,88	3,61	3,92
19	,688	1,19	1,73	2,09	2,54	2,86	3,58	3,88
20	,687	1,18	1,73	2,09	2,53	2,85	3,55	3,85
21	,686	1,18	1,72	2,08	2,52	2,83	3,53	3,82
22	,686	1,18	1,72	2,07	2,51	2,82	3,51	3,79
23	,685	1,18	1,71	2,07	2,50	2,81	3,49	3,77
24	,685	1,18	1,71	2,06	2,49	2,80	3,47	3,74
25	,684	1,18	1,71	2,06	2,49	2,79	3,45	3,72
26	,684	1,18	1,71	2,06	2,48	2,78	3,44	3,71
27	,684	1,18	1,71	2,05	2,47	2,77	3,42	3,69
28	,683	1,17	1,70	2,05	2,47	2,76	3,41	3,67
29	,683	1,17	1,70	2,05	2,46	2,76	3,40	3,66
30	,683	1,17	1,70	2,04	2,46	2,75	3,39	3,65
40	,681	1,17	1,68	2,02	2,42	2,70	3,31	3,55
60	,679	1,16	1,67	2,00	2,39	2,66	3,23	3,46
120	,677	1,16	1,66	1,98	2,36	2,62	3,17	3,37
∞ SNV	,674	1,15	1,64	1,96	2,33	2,58	3,09	3,29
f	25	12,5	5	2,5	1	0,5	0,1	0,05
Irrtumswahrscheinlichkeit α in % für einseitige Fragestellungen								

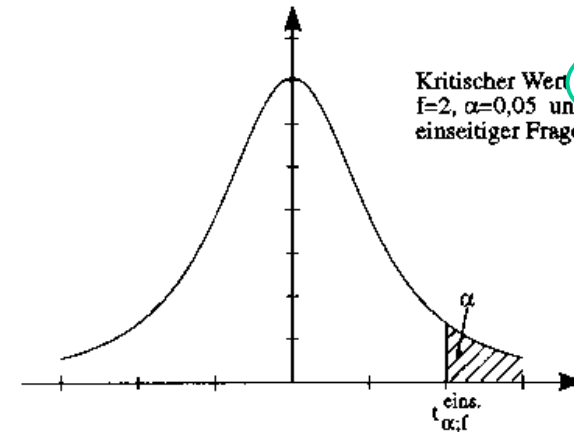
df.: K-J-1

J = Anzahl der Regressoren
K = Zahl der Beobachtungen

Beispiel: K = 10
J = 1
 $\alpha = 0.05$



Kritischer Wert $t_{\alpha/2, f}^{zweis.}$ bei $f=2, \alpha=0,05$ und zweiseitiger Fragestellung.



Kritischer Wert $t_{\alpha, f}^{eins.}$ bei $f=2, \alpha=0,05$ und einseitiger Fragestellung.



- **Berechnung des Konfidenzintervalls (KI):**

$$b_j \pm t \cdot s_{b_j}$$

mit: b_j = Regressionskoeffizient des j-ten Regressors
 s_{b_j} = Standardfehler des j-ten Regressionskoeffizienten
 t = t-Wert für gewählte Vertrauenswahrscheinlichkeit

- **Interpretation?**

- Intervall liegt mit einer Wahrscheinlichkeit von [Vertrauenswahrscheinlichkeit] so, dass es den Regressionskoeffizienten in der Grundgesamtheit umschließt.
- Je größer das KI, desto unsicherer die Schätzung der Steigung des Regressionskoeffizienten.
- Gilt insbesondere bei Vorzeichenwechsel innerhalb des KI!



Prämissen / Prämissenverletzungen (PV) beim linearen Regressionsmodell

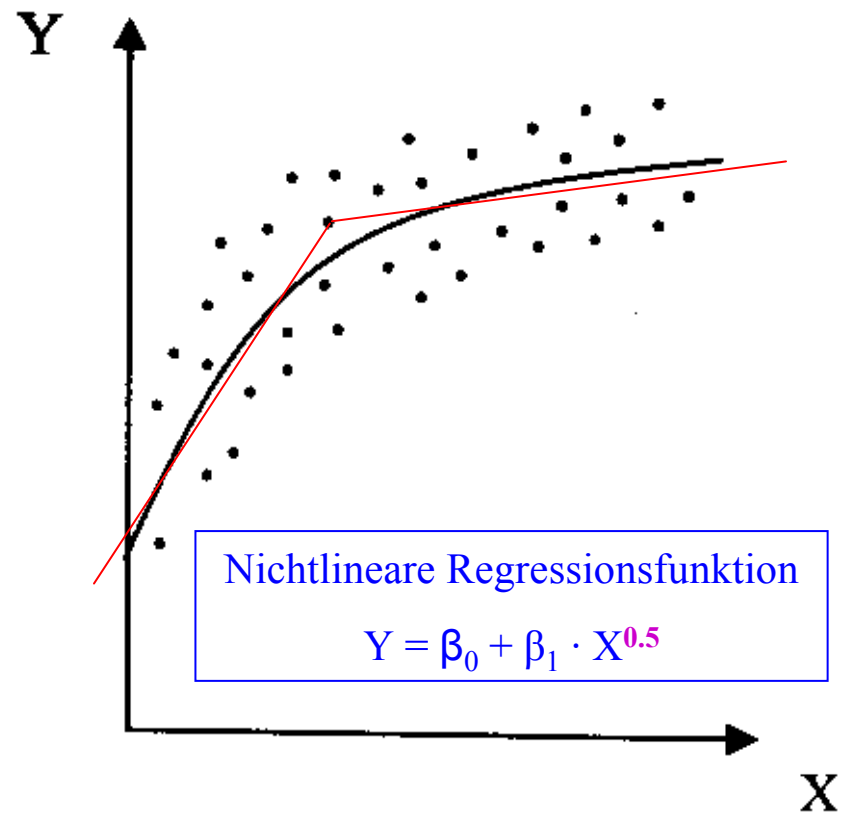
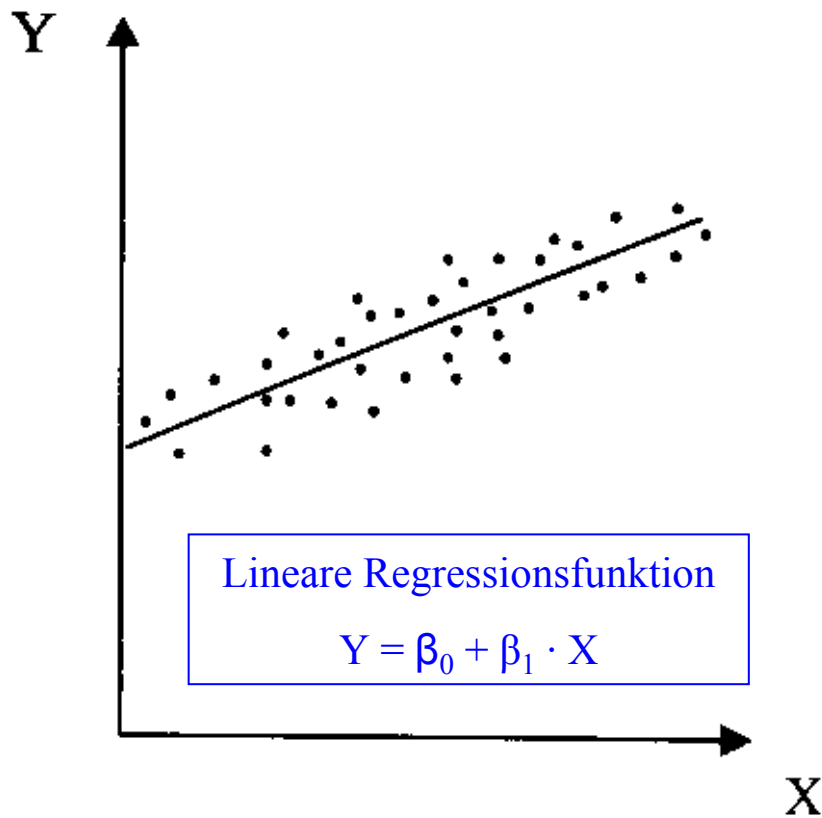


$$y_k = \beta_0 + \sum_{j=1}^J \beta_j \cdot x_{jk} + u_k$$

- Das Modell ist richtig spezifiziert.
 - Es ist linear in den Parametern β_j .
 - Es enthält die relevanten UVs
 - Zahl der zu schätzenden Parameter $(J+1) <$ Zahl der Beobachtungen (K) .
- Die Störgrößen haben den Erwartungswert Null.
- Keine Korrelation zwischen den UVs und der Störgröße
- Störgrößen haben eine konstante Varianz (Homoskedasizität)
- Störgrößen sind unkorreliert (keine Autokorrelation)
- Keine lineare Abhängigkeit zwischen den UVs (keine perfekte Multikollinearität).
- Die Störgrößen sind normalverteilt. (→ relevant bei statistischen Tests!)
- **Praktisches Problem: Ausreichende Varianz der UVs in der Stichprobe!**

PV = Prämissenverletzung!

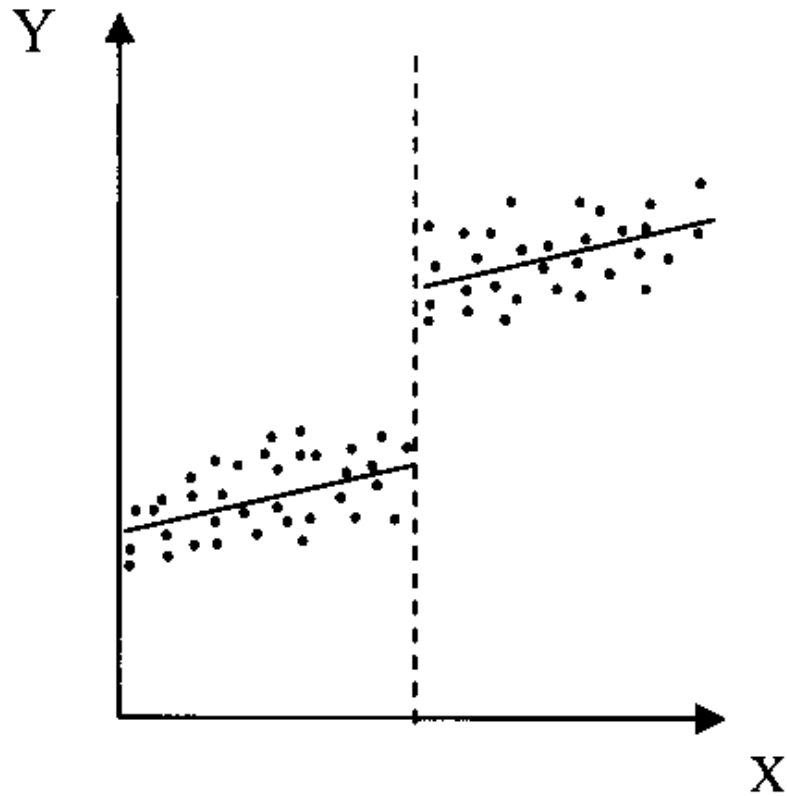
PV: Nicht lineare Regressionsbeziehung



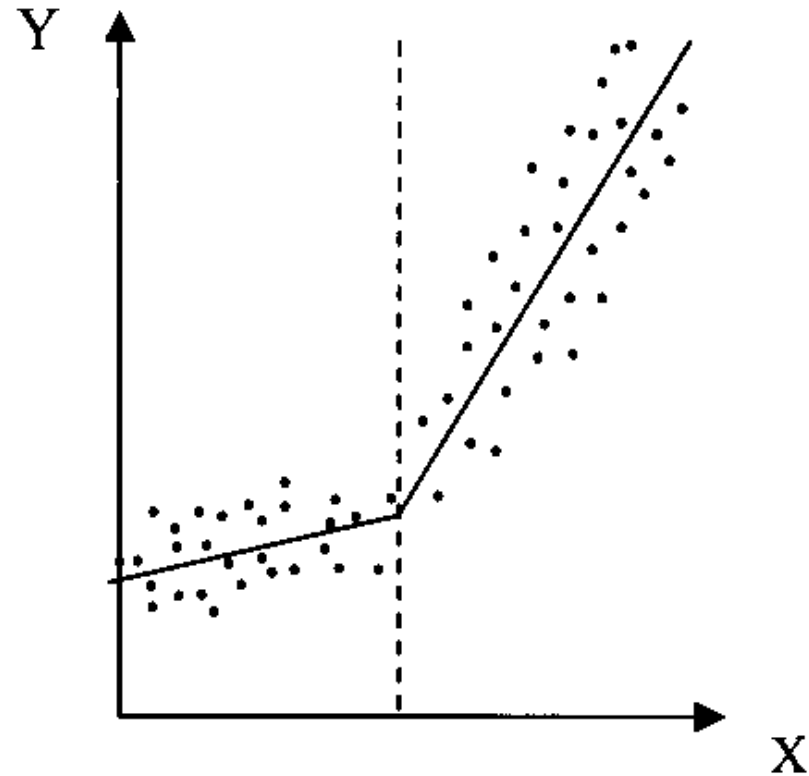
PV: Nicht lineare Regressionsbeziehung II



Strukturbruch:
Niveauänderung



Strukturbruch:
Trendänderung



PV: nicht alle relevanten UVs im Modell enthalten



UVs:		-	-	- -	+	- - +
Sympathie J. Fischer	.24	.35			.23	
Sympathie J. Trittin	.27	.35			.26	
„Irak“ gerechtfertigt	-.07	-.05	-.12	-.13	-.06	-.08
Zuzug Ausländer ↓	-.04	-.11	-.05	-.19	-.03	-.16
AKWs „abschalten“	.04	.07	.07	.16	.04	-.14
Zufried. Grüne in Koalition	.34		.54		.33	
NEO-FFI: OE	.08	.10	.11	.18	.07	.15
NEO-FFI: GW	.02	.00	.01	-.03	.02	-.02
Links-Rechts Selbsteinstufung					-.08	-.24
Erklärte Varianz R²:		.46	.41	.17	.54	.47

Abhängige Variable (AV) jeweils: Sympathie für die Grünen
Zahlenwerte: standardisierte Regressionskoeffizienten
nicht signifikant (unter Signifikanzniveau „.05“)

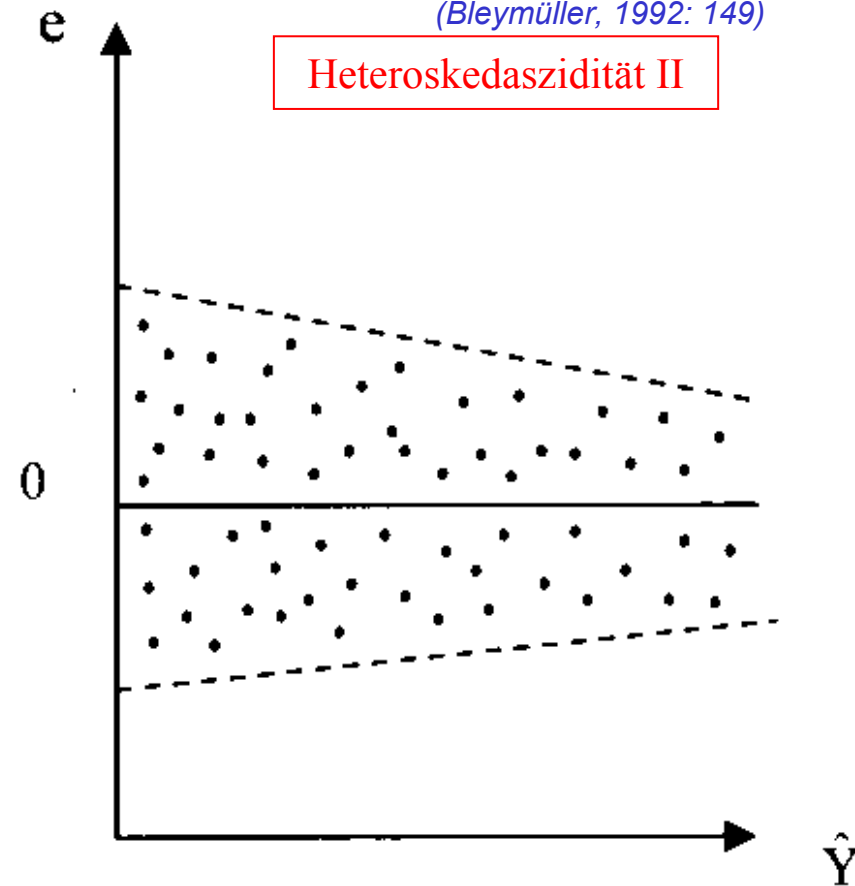
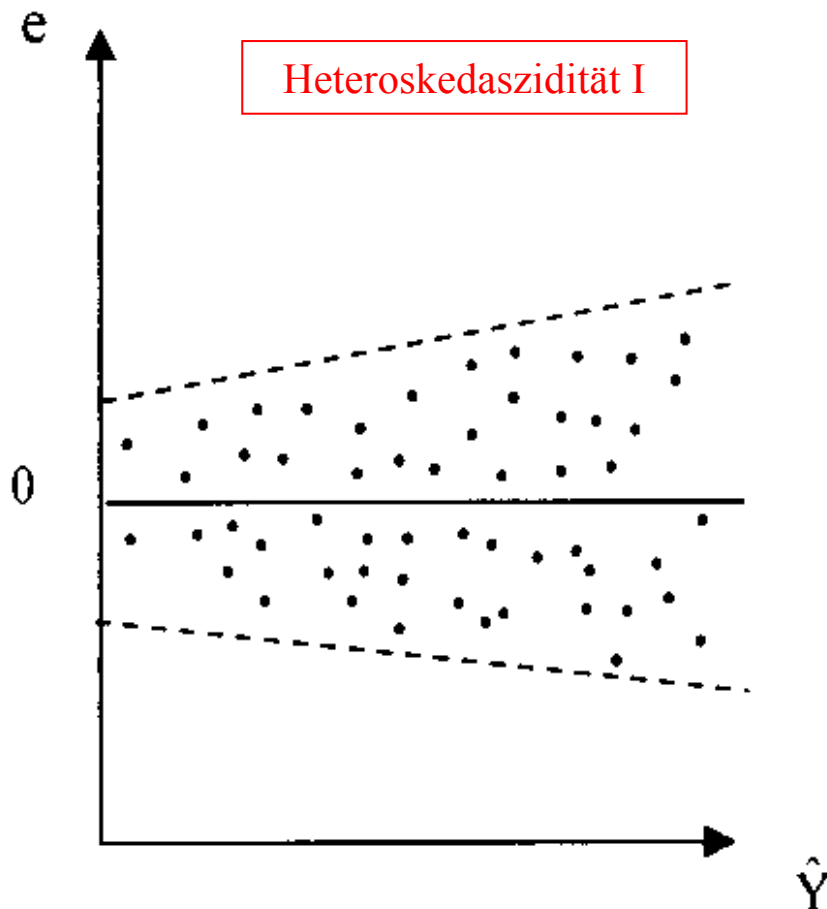
PV: Nicht konstante Varianz der Störgrößen



Homoskedasizität:

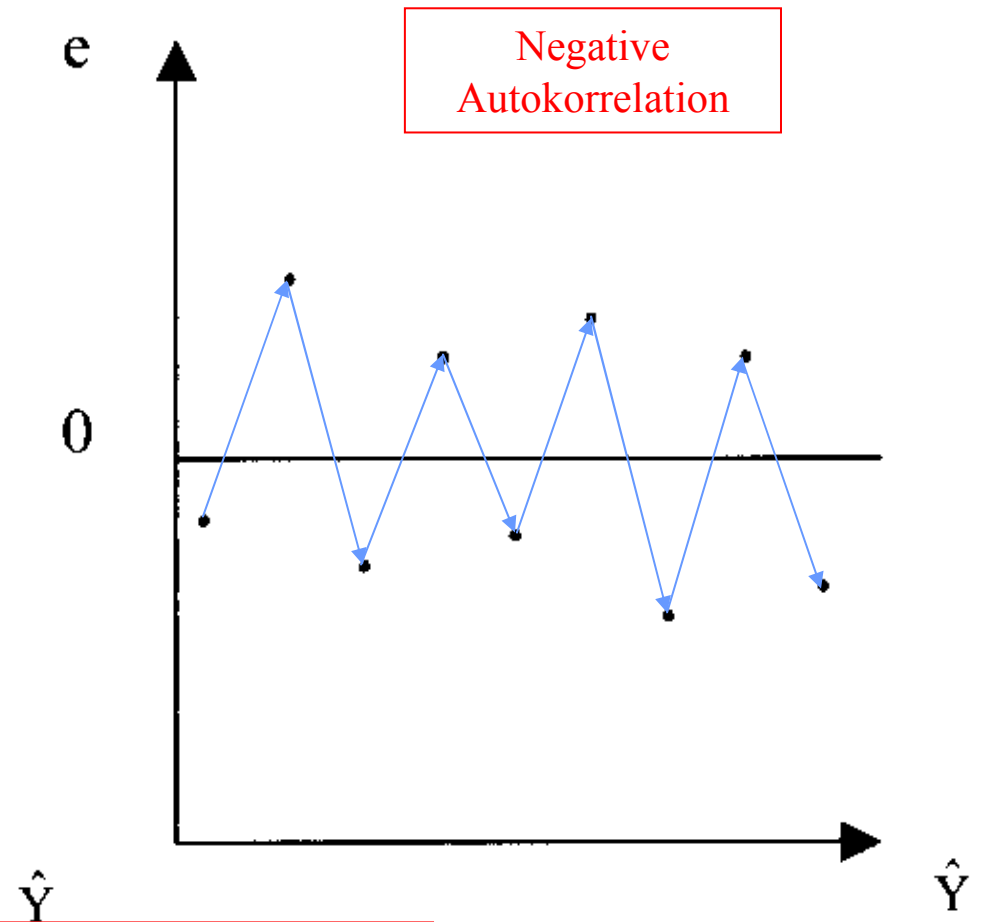
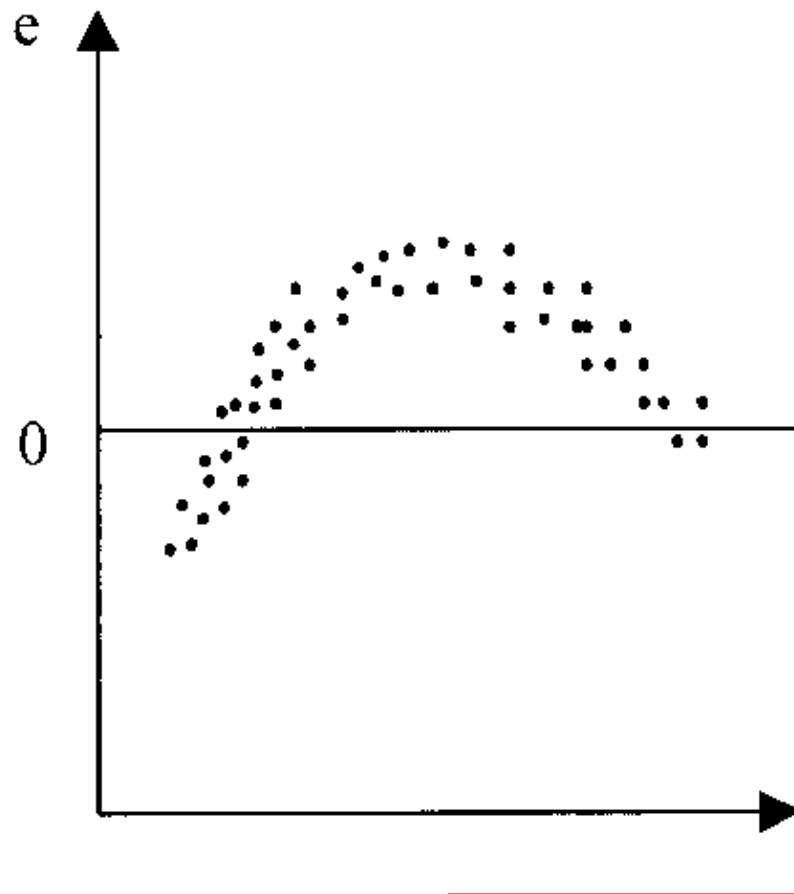
Alle bedingten Verteilungen der u_i haben Erwartungswert 0 und konstante Varianz σ_u^2 .

(Bleymüller, 1992: 149)



Relevant vor allem bei Zeitreihen!

PV: Nicht unkorrelierte Störgrößen (Autokorrelation)



Negative
Autokorrelation

Relevant vor allem bei Zeitreihen!

nach Backhaus u.a. 2006: 87



- Optische Inspektion (Residuen)

Fallweise Diagnose ^a

Fallnummer	Standardisierte Residuen	MENGE	Nichtstandardisierter vorhergesagter Wert	Nichtstandardisierte Residuen
1	1,455	2585,00	2359,0653	225,9347
2	1,066	1819,00	= 1653,4810	+ 165,5190
3	-1,153	1647,00	1826,0974	-179,0974
4	,847	1496,00	1364,3922	131,6078
5	-,558	921,00	1007,6728	-86,6728
6	,962	2278,00	2128,5119	149,4881
7	-,649	1810,00	1910,7293	-100,7293
8	,487	1987,00	1911,2997	75,7003
9	-1,114	1612,00	1785,0727	-173,0727
10	1,486	1913,00	1682,1216	230,8784
11	1,006	2118,00	1961,7850	156,2150
12	-,743	1438,00	1553,4221	-115,4221
13	-1,495	1834,00	2066,2381	-232,2381
14	-,942	1869,00	2015,2797	-146,2797

Beziehung?

Berechnung?

165.5190 : 155.3195
(Standardfehler der Residuen in der Grundgesamtheit!)

(aus Backhaus u.a. 2006: 101)

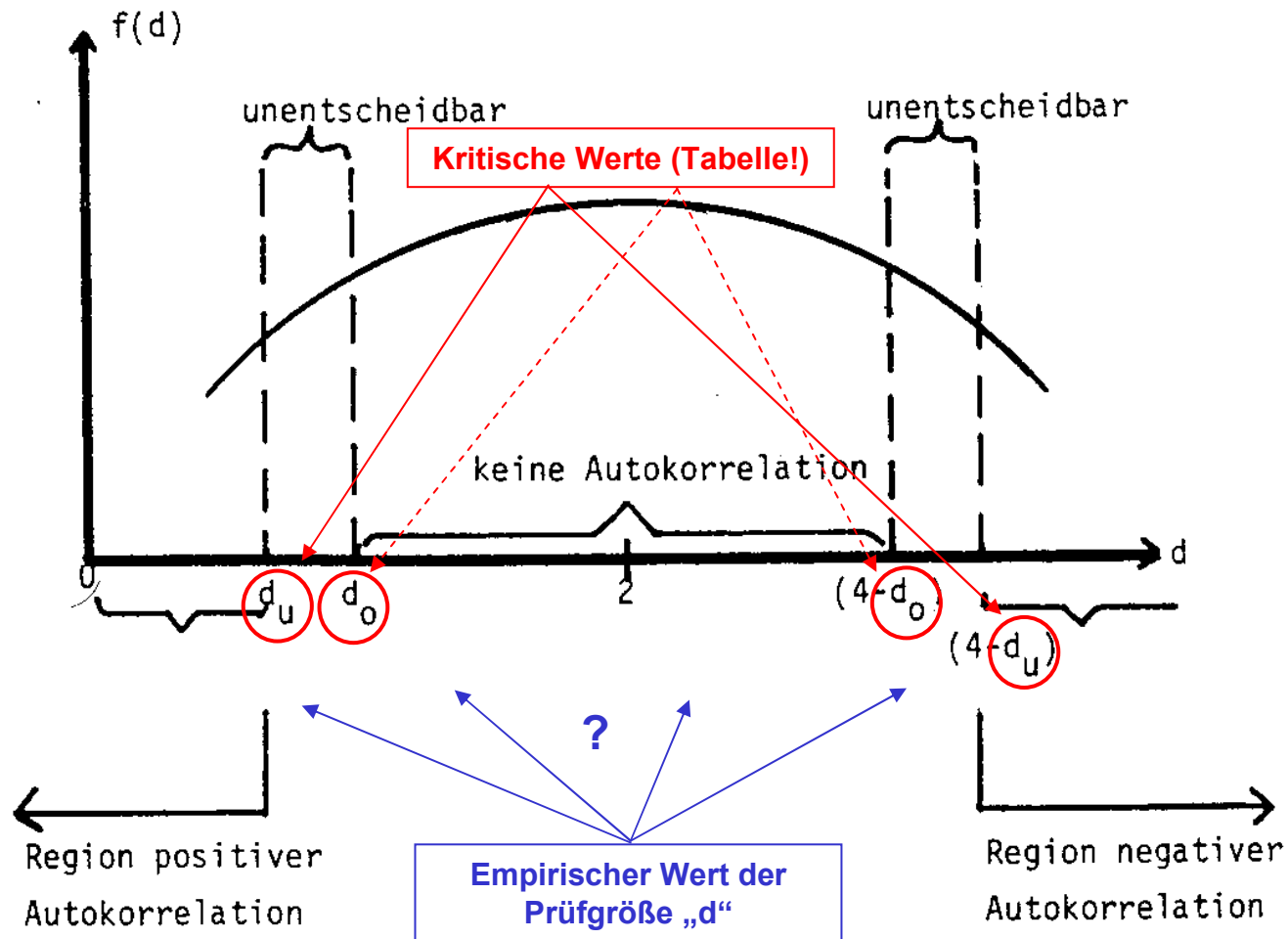
- Durbin-Watson-Test



Test auf Autokorrelation (Durbin-Watson-Formel)



Verteilung des Durbin-Watson-Koeffizienten und seiner kritischen Test-Intervalle



Durbin-Watson Tabelle

K	J=1		J=2		J=3		J=4		J=5	
	d^+_u	d^+_o	d^+_u	d^+_o	d^+_u	d^+_o	d^+_u	d^+_o	d^+_u	d^+_o
15	0,95	1,23	0,83	1,40	0,71	1,61	0,59	1,84	0,48	2,09
16	0,98	1,24	0,86	1,40	0,75	1,59	0,64	1,80	0,53	2,03
17	1,01	1,25	0,90	1,40	0,79	1,58	0,68	1,77	0,57	1,98
18	1,03	1,26	0,93	1,40	0,82	1,56	0,72	1,74	0,62	1,93
19	1,06	1,28	0,96	1,41	0,86	1,55	0,76	1,72	0,66	1,90
20	1,08	1,28	0,99	1,41	0,89	1,55	0,79	1,70	0,70	1,87
21	1,10	1,30	1,01	1,41	0,92	1,54	0,83	1,69	0,73	1,84
22	1,12	1,31	1,04	1,42	0,95	1,54	0,86	1,68	0,77	1,82
23	1,14	1,32	1,06	1,42	0,97	1,54	0,89	1,67	0,80	1,80
24	1,16	1,33	1,08	1,43	1,00	1,54	0,91	1,66	0,83	1,79
25	1,18	1,34	1,10	1,43	1,02	1,54	0,94	1,65	0,86	1,77
26	1,19	1,35	1,12	1,44	1,04	1,54	0,96	1,65	0,88	1,76
27	1,21	1,36	1,13	1,44	1,06	1,54	0,99	1,64	0,91	1,75
28	1,22	1,37	1,15	1,45	1,08	1,54	1,01	1,64	0,93	1,74
29	1,24	1,38	1,17	1,45	1,10	1,54	1,03	1,63	0,96	1,73
30	1,25	1,38	1,18	1,46	1,12	1,54	1,05	1,63	0,98	1,73
31	1,26	1,39	1,20	1,47	1,13	1,55	1,07	1,63	1,00	1,72
32	1,27	1,40	1,21	1,47	1,15	1,55	1,08	1,63	1,02	1,71
33	1,28	1,41	1,22	1,48	1,16	1,55	1,10	1,63	1,04	1,71
34	1,29	1,41	1,24	1,48	1,17	1,55	1,12	1,63	1,06	1,70
35	1,30	1,42	1,25	1,48	1,19	1,55	1,13	1,63	1,07	1,70
36	1,31	1,43	1,26	1,49	1,20	1,56	1,15	1,63	1,09	1,70
37	1,32	1,43	1,27	1,49	1,21	1,56	1,16	1,62	1,10	1,70
38	1,33	1,44	1,28	1,50	1,23	1,56	1,17	1,62	1,12	1,70
39	1,34	1,44	1,29	1,50	1,24	1,56	1,19	1,63	1,13	1,69
40	1,35	1,45	1,30	1,51	1,25	1,57	1,20	1,63	1,15	1,69
45	1,39	1,48	1,34	1,53	1,30	1,58	1,25	1,63	1,21	1,69
50	1,42	1,50	1,38	1,54	1,34	1,59	1,30	1,64	1,26	1,69
55	1,45	1,52	1,41	1,56	1,37	1,60	1,33	1,64	1,30	1,69
60	1,47	1,54	1,44	1,57	1,40	1,61	1,37	1,65	1,33	1,69
65	1,49	1,55	1,46	1,59	1,43	1,62	1,40	1,66	1,36	1,69
70	1,51	1,57	1,48	1,60	1,45	1,63	1,42	1,66	1,39	1,70
75	1,53	1,58	1,50	1,61	1,47	1,64	1,45	1,67	1,42	1,70
80	1,54	1,59	1,52	1,62	1,49	1,65	1,47	1,67	1,44	1,70
85	1,56	1,60	1,53	1,63	1,51	1,65	1,49	1,68	1,46	1,71
90	1,57	1,61	1,55	1,64	1,53	1,66	1,50	1,69	1,48	1,71
95	1,58	1,62	1,56	1,65	1,54	1,67	1,52	1,69	1,50	1,71
100	1,59	1,63	1,57	1,65	1,55	1,67	1,53	1,70	1,51	1,72

**Vertrauenswahrscheinlichkeit:
einseitig 0.975, zweiseitig 0.95**

K = Zahl der Beobachtungen

J = Zahl der Regressoren

d^+_u = unterer Grenzwert
des Unschärfebereichs

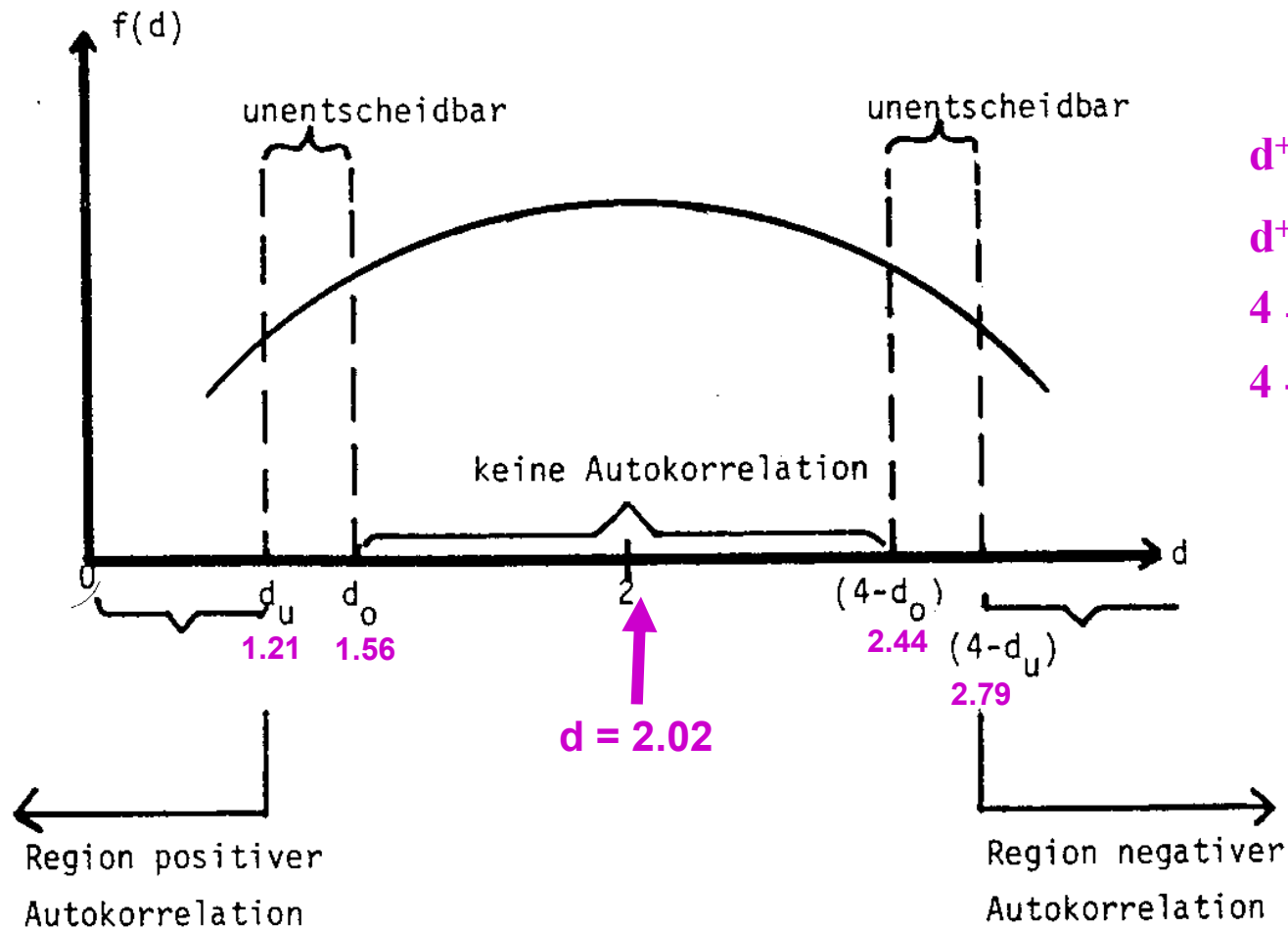
d^+_o = unterer Grenzwert
des Unschärfebereichs

Beispiel:

Test auf Autokorrelation (Durbin-Watson-Formel)



Verteilung des Durbin-Watson-Koeffizienten und seiner kritischen Test-Intervalle



Beispiel:

$$d_u^+ \text{ (hier: } d_u) = 1.21$$

$$d_o^+ \text{ (hier: } d_o) = 1.56$$

$$4 - d_o = 2.44$$

$$4 - d_u = 2.79$$

$$d = 2.02$$

H_0 „keine Autokorrelation“ kann nicht zurückgewiesen werden!

Test auf Autokorrelation: Beispiel 1

positive Autokorrelation

k	e_k	e_{k-1}	$(e_k - e_{k-1})^2$	e_k^2
1	0.3	-	-	0.09
2	0.2	0.3	0.01	0.04
3	0.3	0.2	0.01	0.09
4	0.2	0.3	0.01	0.04
5	0.3	0.2	0.01	0.09
6	0.3	0.3	0.00	0.09
7	0.2	0.3	0.01	0.04
8	0.2	0.2	0.00	0.04
9	-0.3	0.2	0.25	0.09
10	-0.3	-0.3	0.00	0.09
11	-0.2	-0.3	0.01	0.04
12	-0.2	-0.2	0.00	0.04
13	-0.3	-0.2	0.01	0.09
14	-0.2	-0.3	0.01	0.04
15	-0.2	-0.2	0.00	0.04
(K =) 16	-0.3	-0.2	0.01	0.09
		-0.3		

$$d = \frac{\sum_{k=2}^K (e_k - e_{k-1})^2}{\sum_{k=1}^K e_k^2}$$

d = Indexwert für die Prüfung auf Autokorrelation
 e_k = Residualgröße für den Beobachtungswert in Periode k (k = 1, 2, ..., K)

Spaltensumme: Spaltensumme:
 0.34 1.04

d = 0.33

Test auf Autokorrelation: Beispiel 2

negative Autokorrelation

k	e_k	e_{k-1}	$(e_k - e_{k-1})^2$	e_k^2
1	0.3	-	-	0.09
2	-0.3	0.3	0.36	0.09
3	0.2	-0.3	0.25	0.04
4	-0.2	0.2	0.16	0.04
5	0.3	-0.2	0.25	0.09
6	-0.3	0.3	0.36	0.09
7	0.2	-0.3	0.25	0.04
8	-0.2	0.2	0.16	0.04
9	0.3	-0.2	0.25	0.09
10	-0.2	0.3	0.25	0.04
11	0.3	-0.2	0.25	0.09
12	-0.3	0.3	0.36	0.09
13	0.2	-0.3	0.25	0.04
14	-0.3	0.2	0.25	0.09
15	0.2	-0.3	0.25	0.04
(K =) 16	-0.2	0.2	0.16	0.04
		-0.2		

Spaltensumme: 3.81 Spaltensumme: 1.04

$d = 3.66$

$$d = \frac{\sum_{k=2}^K (e_k - e_{k-1})^2}{\sum_{k=1}^K e_k^2}$$

d = Indexwert für die Prüfung auf Autokorrelation
 e_k = Residualgröße für den Beobachtungswert in Periode k ($k = 1, 2, \dots, K$)



- d.h. Forderung: Keine perfekte lineare Abhängigkeit der UVs

- Prüfung auf Multikollinearität:

- 1. Prüfung: Korrelationsmatrix der UVs
- 2. multiple Korrelation R^2

$1 - R^2$ (Toleranz)

Variance Inflation

Factor (VIF) =

Kehrwert der Toleranz

R^2	$1 - R^2$ = (Toleranz)	$\frac{1}{1 - R^2} = \frac{1}{\text{Toleranz}} = \text{VIF}$
1.0	0.0	$\rightarrow \infty$
0.9	0.1	10.000
0.8	0.2	5.000
0.7	0.3	3.333
0.6	0.4	2.500
0.5	0.5	2.000
0.4	0.6	1.667
0.3	0.7	1.429
0.2	0.8	1.250
0.1	0.9	1.111
0.0	1.0	1.000

- Abhilfemöglichkeiten:

- Variablen (UVs) entfernen
- Ersetzung von Variablen durch Faktoren



Beispiele für Regressionsmodelle

Bivariates Modell: Lineare Einfachregression



```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS CI R ANOVA TOL
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT fr4_3
/METHOD=ENTER fr9_5
/CASEWISE DEPENDENT PRED RESID OUTLIRE (0)
```

Modellzusammenfassung^b

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,057 ^a	,003	,003	2,553

- a. Einflußvariablen : (Konstante), fr9_5 BFI-Kurz: offen für neue Erfahrungen ist, vielseitig
- b. Abhängige Variable: fr4_3 Sympathie/Antipathie: Angela Merkel

Koeffizienten^a

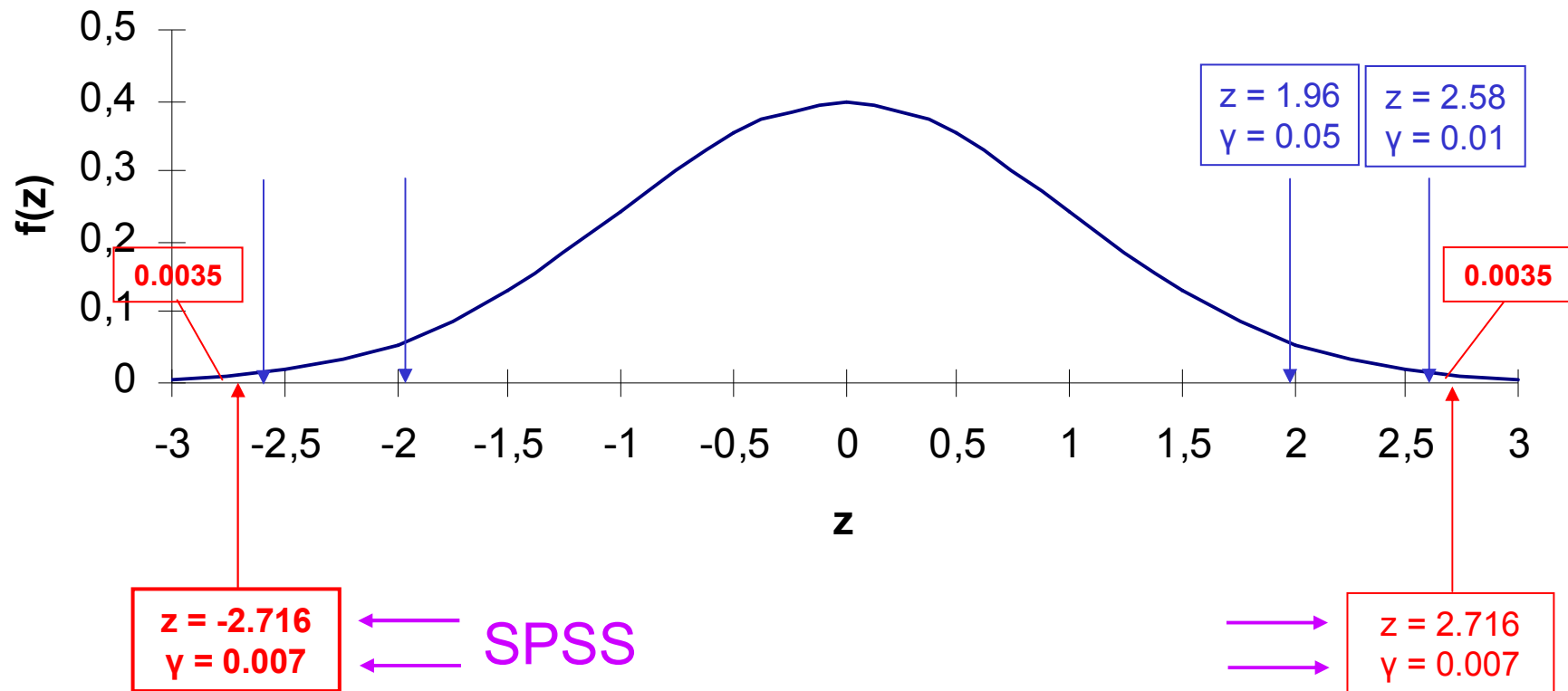
Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	95%-Konfidenzintervall für B		Kollinearitätsstatistik	
		B	Standardfehler	Beta			Untergrenze	Obergrenze	Toleranz	VIF
1	(Konstante)	5,883	,167	0,000	35,170	,000	5,555	6,211	1,000	1,000
1	fr9_5 BFI-Kurz: offen für neue Erfahrungen ist, vielseitig	-,108	,040	-,057	-2,716	,007	-,185	-,030	1,000	1,000

- a. Abhängige Variable: fr4_3 Sympathie/Antipathie: Angela Merkel

Zusammenhang?



t-Verteilung → Standardnormalverteilung



Multiple Regression – Modell 1 (Ausgangsmodell)



Modellzusammenfassung^b

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,337 ^a	,113	,111	2,408

- a. Einflußvariablen : (Konstante), fr40_24 Die meisten Politiker sind korrupt, fr29_10 Issue-Haltung selbst: Kernenergie, fr9_5 BFI-Kurz: offen für neue Erfahrungen ist, vielseitig, fr6_2 Sympathie/Antipathie: katholische Kirche?, fr29_4 Issue-Haltung selbst: Zuzugsmöglichkeit für Ausländer
- b. Abhängige Variable: fr4_3 Sympathie/Antipathie: Angela Merkel

?

REGRESSION

```

/MISSING LISTWISE
/STATISTICS COEFF OUTS CI R ANOVA TOL
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT fr4_3
/METHOD=ENTER fr6_2 fr9_5 fr29_4 fr29_10
fr40_24
/RESIDUALS HIST(ZRESID).
    
```

Y: erklärte Streuung

Y: nicht erklärte Streuung

Y: Gesamt-Streuung SAQ_y

ANOVA^b

Modell	Quadratsumme	K=2262 df ?	Mittel der Quadrate ?	F ?	Signifikanz
1 Regression	1673,113	J 5	=334,623	57,730 Quotient	,000 ^a
Residuen	13076,540	K-J-1 2256	=5,796		
Gesamt	14749,653	K-1 2261			

- a. Einflußvariablen : (Konstante), fr40_24 Die meisten Politiker sind korrupt, fr29_10 Issue-Haltung selbst: Kernenergie, fr9_5 BFI-Kurz: offen für neue Erfahrungen ist, vielseitig, fr6_2 Sympathie/Antipathie: katholische Kirche?, fr29_4 Issue-Haltung selbst: Zuzugsmöglichkeit für Ausländer

- b. Abhängige Variable: fr4_3 Sympathie/Antipathie: Angela Merkel

Fs.: Multiple Regression – Modell 1



$$B_j \cdot \frac{SD_{x_j}}{SD_y}$$

Koeffizienten

Mo dell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signif ikanz	95%- Konfidenzinterv all für B		Kollinearitäts statistik	
		B	Standardfehler	Beta			Untergrenze	Obergrenze	Toleranz	VIF
1	(Konstante)	4,923	,313		15,727	,000	4,309	5,537		
	fr6_2 Sympathie/Antipathie: katholische Kirche?	,243	,019	,259	12,810	,000	,206	,280	,958	1,044
	fr9_5 BFI-Kurz: offen für neue Erfahrungen ist, vielseitig	-,031	,038	-,016	-,822	,411	-,106	,043	,978	1,022
	fr29_4 Issue-Haltung selbst: Zuzugsmöglichkeit für Ausländer	,140	,032	,090	4,410	,000	,078	,203	,934	1,071
	fr29_10 Issue-Haltung selbst: Kernenergie	-,219	,032	-,136	-6,763	,000	-,282	-,155	,966	1,035
	fr40_24 Die meisten Politiker sind korrupt	-,166	,046	-,073	-3,607	,000	-,256	-,076	,953	1,049

a. Abhängige Variable: fr4_3 Sympathie/Antipathie: Angela Merkel

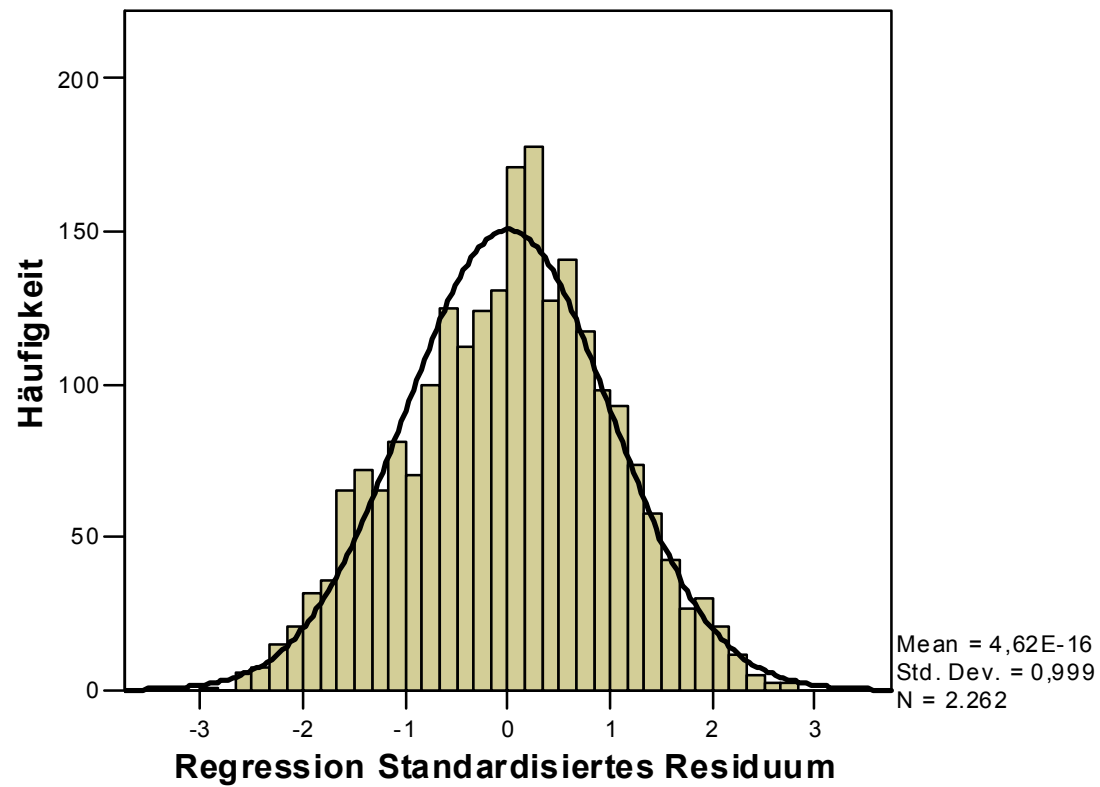
$$\frac{B_j}{SE_{B_j}}$$



Normalverteilung
der Residuen?

Histogramm

Abhängige Variable: Sympathie/Antipathie: Angela Merkel





Residuenstatistik^a

	Minimum	Maximum	Mittelwert	Standardabweichung	N
Nicht standardisierter vorhergesagter Wert	2,90	8,27	5,45	,860	2262
Nicht standardisierte Residuen	-7,116	6,464	,000	2,405	2262
Standardisierter vorhergesagter Wert	-2,961	3,282	,000	1,000	2262
Standardisierte Residuen	-2,956	2,685	,000	,999	2262

a. Abhängige Variable: fr4_3 Sympathie/Antipathie: Angela Merkel

Multiple Regression – Modell 2 (+ LIRE)



Modellzusammenfassung^b

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,416 ^a	.113	,171	2,330

a. Einflußvariablen : (Konstante), fr40_24 Die meisten Politiker sind korrupt, fr29_10 Issue-Haltung selbst: Kernenergie, fr9_5 BFI-Kurz: offen für neue Erfahrungen ist, vielseitig, fr6_2 Sympathie/Antipathie: katholische Kirche?, fr29_4 Issue-Haltung selbst: Zuzugsmöglichkeit für Ausländer, fr28 Links-Rechts Einordnung: selbst

Koeffizienten^a

vorher:
.000

Mo dell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	95%-Konfidenzintervall für B		Kollinearitätsstatistik	
		B	Standardfehler				Untergrenze	Obergrenze	Toleranz	VIF
1	(Konstante)	3,150	,335		9,408	,000	2,493	3,807		
	fr6_2 Sympathie/Antipathie: katholische Kirche?	,196	,019	.259	10,43	,000	,159	,233	,923	1,083
	fr9_5 BFI-Kurz: offen für neue Erfahrungen ist, vielseitig	,006	,037	.090	,154	,877	-,067	,078	,974	1,026
	fr28 Links-Rechts Einordnung: selbst	,355	,028	-.136	12,74	,000	,300	,410	,889	1,125
	fr29_4 Issue-Haltung selbst: Zuzugsmöglichkeit für Ausländer	,064	,032	.041	2,025	.043	,002	,126	,899	1,112
	fr29_10 Issue-Haltung selbst: Kernenergie	-,162	,032	-.101	-5,078	,000	-,225	-,100	,945	1,058
	fr40_24 Die meisten Politiker sind korrupt	-,153	,045	-.068	-3,424	,001	-,241	-,065	,953	1,049

a. Abhängige Variable: fr4_3 Sympathie/Antipathie: Angela Merkel

Multiple Regression – Modell 3 (+ Symp. CDU)



Modellzusammenfassung^a

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,659 ^a	.434	,433	1,928

Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	95%-Konfidenzintervall für B		Kollinearitätsstatistik	
		B	Standardfehler				Untergrenze	Obergrenze	Toleranz	VIF
1	(Konstante)	1,244	,283		4,388	,000	,688	1,800		
	fr5_2 Sympathie/Antipathie: CDU	,585	,018	,599	32,02	,000	,549	,621	,730	1,369
	fr6_2 Sympathie/Antipathie: katholische Kirche?	,077	,016	,083	4,835	,000	,046	,109	,874	1,144
	fr9_5 BFI-Kurz: offen für neue Erfahrungen ist, vielseitig	-,011	,031	-,006	-,364	,716	-,071	,049	,974	1,026
	fr28 Links-Rechts Einordnung: selbst	,062	,025	,046	2,517	,012	,014	,111	,768	1,302
	fr29_4 Issue-Haltung selbst: Zugangsmöglichkeit für Ausländer	,022	,026	,014	,857	,392	-,029	,074	,898	1,114
	fr29_10 Issue-Haltung selbst: Kernenergie	-,029	,027	-,018	-1,066	,287	-,081	,024	,923	1,084
	fr40_24 Die meisten Politiker sind korrupt	-,026	,037	-,012	-,711	,477	-,099	,047	,943	1,061

vorher:
.000
.043
.000
.001

a. Abhängige Variable: fr4_3 Sympathie/Antipathie: Angela Merkel

Multiple Regression – Modell 4 (+ Symp. CSU)



Modellzusammenfassung^b

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,665 ^a	,442	,440	1,917

$$r_{\text{CDU CSU}} = .86$$

Koeffizienten^f

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	95%-Konfidenzintervall für B		Kollinearitätsstatistik	
		B	Standardfehler				Untergrenze	Obergrenze	Toleranz	VIF
		1	(Konstante)	1,389	,284		4,897	,000	,833	1,945
	fr5_2 Sympathie/Antipathie: CDU	,459	,031	,470	14,838	,000	,398	,520	,252	3,964
	fr5_3 Sympathie/Antipathie: CSU	,147	,029	,165	5,037	,000	,090	,204	,236	4,231
	fr6_2 Sympathie/Antipathie: katholische Kirche?	,069	,016	,074	4,327	,000	,038	,101	,864	1,157
	fr9_5 BFI-Kurz: offen für neue Erfahrungen ist, vielseitig	-,011	,031	-,006	-357	,721	-,071	,049	,974	1,026
	fr28 Links-Rechts Einordnung: selbst	,041	,025	,030	1,635	,102	-,008	,090	,744	1,345
	fr29_4 Issue-Haltung selbst: Zuzugsmöglichkeit für Ausländer	,008	,026	,005	,286	,775	-,044	,059	,886	1,129
	fr29_10 Issue-Haltung selbst: Kernenergie	-,016	,027	-,010	-,603	,546	-,069	,036	,915	1,093
	fr40_24 Die meisten Politiker sind korrupt	-,034	,037	-,015	-,908	,364	-,106	,039	,942	1,062

a. Abhängige Variable: fr4_3 Sympathie/Antipathie: Angela Merkel



Residuenstatistik^a

	Minimum	Maximum	Mittelwert	Standardabweichung	N
Nicht standardisierter vorhergesagter Wert	1,73	9,13	5,44	1,687	2223
Nicht standardisierte Residuen	-7,868	8,229	,000	1,925	2223
Standardisierter vorhergesagter Wert	-2,199	2,186	,000	1,000	2223
Standardisierte Residuen	-4,081	4,268	,000	,998	2223

Modell 3

a. Abhängige Variable: fr4_3 Sympathie/Antipathie: Angela Merkel

Residuenstatistik^a

	Minimum	Maximum	Mittelwert	Standardabweichung	N
Nicht standardisierter vorhergesagter Wert	1,88	9,09	5,44	1,702	2220
Nicht standardisierte Residuen	-7,800	7,992	,000	1,914	2220
Standardisierter vorhergesagter Wert	-2,090	2,149	,000	1,000	2220
Standardisierte Residuen	-4,069	4,169	,000	,998	2220

Modell 4

a. Abhängige Variable: fr4_3 Sympathie/Antipathie: Angela Merkel



Standardisierte Regressionskoeffizienten und R² in unterschiedlichen Modellen

			Modell 1		Modell 2		Modell 3		Modell 4	
	Beta	sig.	Beta	sig.	Beta	sig.	Beta	sig.	Beta	sig.
Sympathie für CDU	-	-	-	-	-	-	.599	.000	.470	.000
Sympathie für CSU	-	-	-	-	-	-	-	-	.165	.000
Sympathie katholische Kirche	-	-	.259	.000	.209	.000	.083	.000	.074	.000
Offenheit für neue Erfahrungen	-.057	.007	-.016	.441	.003	.877	-.006	.716	-.006	.721
Links- Rechts-Selbsteinstufung	-	-	-	-	.261	.000	.046	.012	.030	.102
Zuzugsmöglichk. Ausländer ↓	-	-	.090	.000	.041	.043	.014	.392	.005	.775
Kernkraftwerke abschlaten	-	-	-.136	.000	-.101	.000	-.018	.287	-.010	.546
Die meisten Politiker korrupt	-	-	-.073	.000	-.068	.001	-.012	.477	-.015	.364
R²		.003		.113		.173		.434		.442

Quelle: Studie 2003

AV: Sympathie für Angela Merkel



Produktvariablen im linearen Regressionsmodell

Produktvariablen - 1



Koeffizienten^a

Ausgangsmodell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	26,650	,913		29,190	,000		
	PV PartPolitikerVerdross	,710	,105	,123	6,738	,000	,987	1,014
	BI formale Bildung	-,834	,084	-,192	-9,874	,000	,863	1,158
	LR Links-Rechts	,482	,042	,211	11,419	,000	,959	1,043
	oe NEO-FFI: Offenheit f. Erfahrung	-1,975	,171	-,230	-11,565	,000	,830	1,205
	vt NEO-FFI: Vertraeglichkeit	-1,843	,161	-,209	-11,454	,000	,979	1,022

R² = .281

→ a. Abhängige Variable: RE (extrem) rechte Einstellungen

Interaktion: Offenheit x Bildung?

Koeffizienten^a

		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	19,740	2,203		8,960	,000		
	PV PartPolitikerVerdross	,714	,105	,123	6,785	,000	,986	1,014
	BI formale Bildung	,942	,522	,217	1,803	,071	,022	44,532
	LR Links-Rechts	,481	,042	,211	11,431	,000	,959	1,043
	oe NEO-FFI: Offenheit f. Erfahrung	,171	,646	,020	,265	,791	,058	17,321
	vt NEO-FFI: Vertraeglichkeit	-1,793	,161	-,204	-11,130	,000	,971	1,030
	BIOE nur: Produktvariable	-,559	,162	-,555	-3,445	,001	,013	79,698

R² = .285

a. Abhängige Variable: RE (extrem) rechte Einstellungen



Korrelationen

		BI formale Bildung	zBI	oe NEO-FFI: Offenheit f. Erfahrung	zOE	BI x OE	zBI x zOE
BI formale Bildung	Korrelation nach Pearson	1	1,000	,370	,370	,902	,265
	Signifikanz (2-seitig)		,000	,000	,000	,000	,000
	N	2274	2274	2264	2264	2264	2264
zBI	Korrelation nach Pearson	1,000	1	,370	,370	,902	,265
	Signifikanz (2-seitig)	,000		,000	,000	,000	,000
	N	2274	2274	2264	2264	2264	2264
oe NEO-FFI: Offenheit f. Erfahrung	Korrelation nach Pearson	,370	,370	1	1,000	,721	,118
	Signifikanz (2-seitig)	,000	,000		,000	,000	,000
	N	2264	2264	2264	2264	2264	2264
zOE	Korrelation nach Pearson	,370	,370	1,000	1	,721	,118
	Signifikanz (2-seitig)	,000	,000	,000		,000	,000
	N	2264	2264	2264	2264	2264	2264
BI x OE	Korrelation nach Pearson	,902	,902	,721	,721	1	,356
	Signifikanz (2-seitig)	,000	,000	,000	,000		,000
	N	2264	2264	2264	2264	2264	2264
zBI x zOE	Korrelation nach Pearson	,265	,265	,118	,118	,356	1
	Signifikanz (2-seitig)	,000	,000	,000	,000	,000	
	N	2264	2264	2264	2264	2264	2264

BI und OE sind korreliert !

„Explodieren“ des Zusammenhangs durch Produktbildung
 Abhilfe: Mittelwert herausnehmen (zentrieren reicht!)

Produktvariablen - 3



x₂	1.2	-1.44	-1.20	-0.96	-0.72	-0.48	-0.24	0	0.24	0.48	0.72	0.96	1.20	1.44
	1.0	-1.20	-1.00	-0.80	-0.60	-0.40	-0.20	0	0.20	0.40	0.60	0.80	1.00	1.20
	0.8	-0.96	-0.80	-0.64	-0.48	-0.32	-0.16	0	0.16	0.32	0.48	0.64	0.80	0.96
	0.6	-0.72	-0.60	-0.48	-0.36	-0.24	-0.12	0	0.12	0.24	0.36	0.48	0.60	0.72
	0.4	-0.48	-0.40	-0.32	-0.24	-0.16	-0.08	0	0.08	0.16	0.24	0.32	0.40	0.48
	0.2	-0.24	-0.20	-0.16	-0.12	-0.08	-0.04	0	0.04	0.08	0.12	0.16	0.20	0.24
	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	-0.2	0.24	0.20	0.16	0.12	0.08	0.04	0	-0.04	-0.08	-0.12	-0.16	-0.20	-0.24
	-0.4	0.48	0.40	0.32	0.24	0.16	0.08	0	-0.08	-0.16	-0.24	-0.32	-0.40	-0.48
	-0.6	0.72	0.60	0.48	0.36	0.24	0.12	0	-0.12	-0.24	-0.36	-0.48	-0.60	-0.72
	-0.8	0.96	0.80	0.64	0.48	0.32	0.16	0	-0.16	-0.32	-0.48	-0.64	-0.80	-0.96
	-1.0	1.20	1.00	0.80	0.60	0.40	0.20	0	-0.20	-0.40	-0.60	-0.80	-1.00	-1.20
	-1.2	1.44	1.20	0.96	0.72	0.48	0.24	0	-0.24	-0.48	-0.72	-0.96	-1.20	-1.44
		-1.2	-1.0	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1.0	1.2

Werte der z-
standardisierten
Produktvariablen
x₁ · x₂

$$s_{xy} = \frac{SAP}{n}$$

$$r = \frac{s_{xy}}{s_x \cdot s_y}$$

Für z-
standardisierte
Variablen:

$$r = S_{xy}$$

Produktvariablen - 4



Koeffizienten^a

Ausgangsmodell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	26,650	,913		29,190	,000		
	PV PartPolitikerVerdross	,710	,105	,123	6,738	,000	,987	1,014
	BI formale Bildung	-,834	,084	-,192	-9,874	,000	,863	1,158
	LR Links-Rechts	,482	,042	,211	11,419	,000	,959	1,043
	oe NEO-FFI: Offenheit f. Erfahrung	-1,975	,171	-,230	-11,565	,000	,830	1,205
	vt NEO-FFI: Vertraeglichkeit	-1,843	,161	-,209	-11,454	,000	,979	1,022

R² = .281

→ a. Abhängige Variable: RE (extrem) rechte Einstellungen

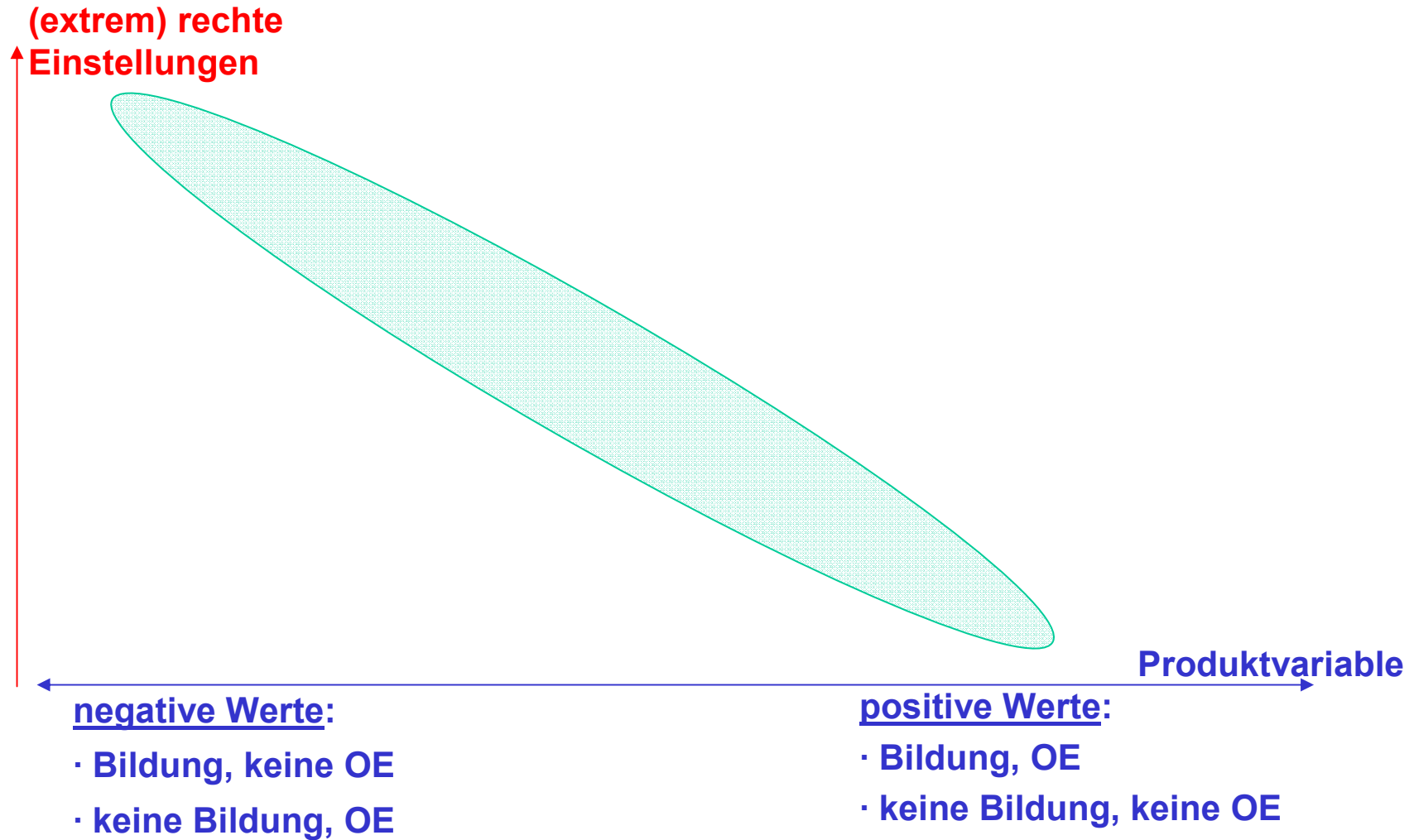
Koeffizienten^a

Ausgangsmodell mit z-standardisierter Produktvariable		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	26,260	,918		23,615	,000		
	PV PartPolitikerVerdross	,714	,105	,123	6,785	,000	,986	1,014
	BI formale Bildung	-,762	,087	-,176	-8,789	,000	,814	1,228
	LR Links-Rechts	,481	,042	,211	11,431	,000	,959	1,043
	OE Offenheit f. Erfahrung	-1,966	,170	-,229	-11,540	,000	,830	1,205
	vt NEO-FFI: Vertraeglichkeit	-1,793	,161	-,204	-11,130	,000	,971	1,030
	zBIOE z - Produktvariable	-,276	,080	-,065	-3,445	,001	,924	1,083

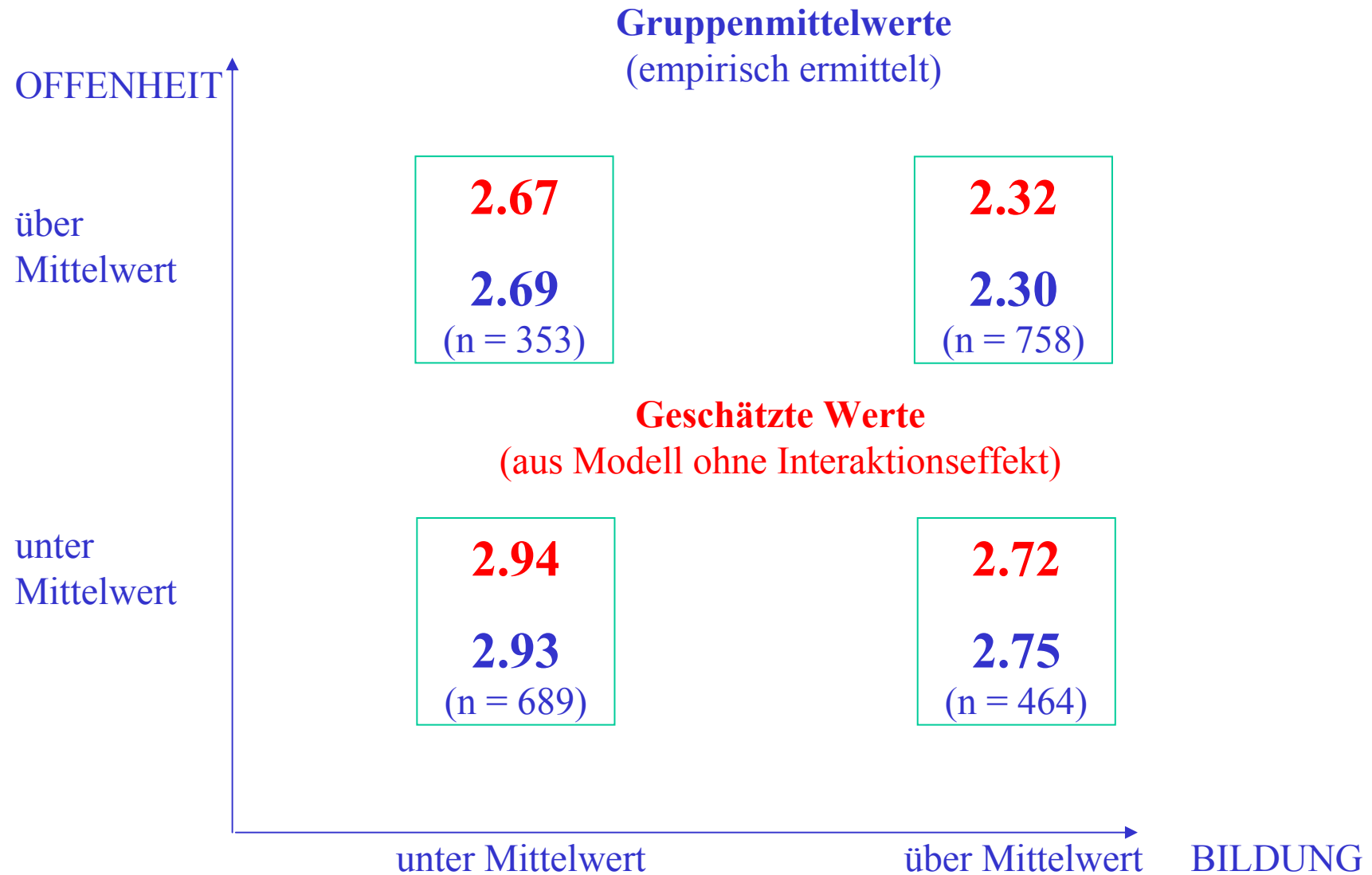
R² = .285

a. Abhängige Variable: RE (extrem) rechte Einstellungen

Interpretation des Interaktionseffekts I



Interpretation des Interaktionseffekts II





Dichotome unabhängige Variablen im linearen Regressionsmodell

Vorstellung der Variablen



NBL = 1 ABL = 0

Deskriptive Statistik

hohe
Werte:
Rechts

	N	Minimum	Maximum	Mittelwert	Standardabweichung
NBL	2261	,00	1,00	,2379	,42592
LIRE	2254	1,00	11,00	5,5475	1,88088
lix extrem linke Einstellungen	2279	1,00	5,00	2,8077	,69296
Gültige Werte (Listenweise)	2213				

dichotome Variable

Umfrage 2003

Deskriptive Statistik

	N	Minimum	Maximum	Mittelwert	Standardabweichung
NBL	1723	,00	,00	,0000	,00000
LIRE	1693	1,00	11,00	5,7106	1,85659
lix extrem linke Einstellungen	1713	1,00	4,88	2,6372	,61792
Gültige Werte (Listenweise)	1684				

ABL

Deskriptive Statistik

	N	Minimum	Maximum	Mittelwert	Standardabweichung
NBL	538	1,00	1,00	1,0000	,00000
LIRE	531	1,00	11,00	5,0339	1,83745
lix extrem linke Einstellungen	536	1,25	5,00	3,3508	,63636
Gültige Werte (Listenweise)	529				

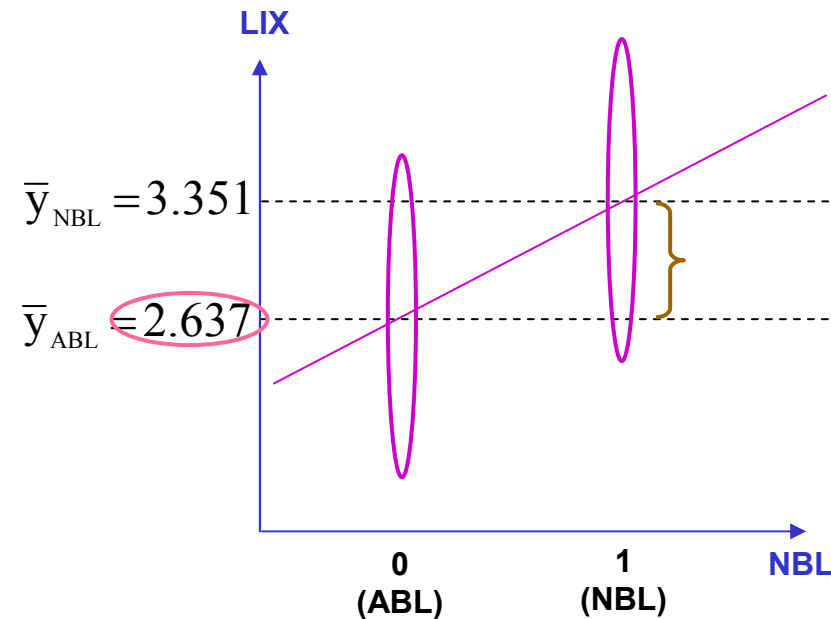
NBL

Modell 1: AV = LIX, UV = NBL

Modellzusammenfassung^b

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,439 ^a	,193	,192	,62236

- a. Einflußvariablen : (Konstante), NBL
- b. Abhängige Variable: lix extrem linke Einstellungen



Koeffizienten^b

Modell	Nicht standardisierte Koeffizienten	Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik			
					B	Standardfehler	Beta	Toleranz
1	(Konstante)	2,637	,015		175,384	,000		
1	NBL	,714	,031	,439	23,166	,000	1,000	1,000

- a. Abhängige Variable: lix extrem linke Einstellungen

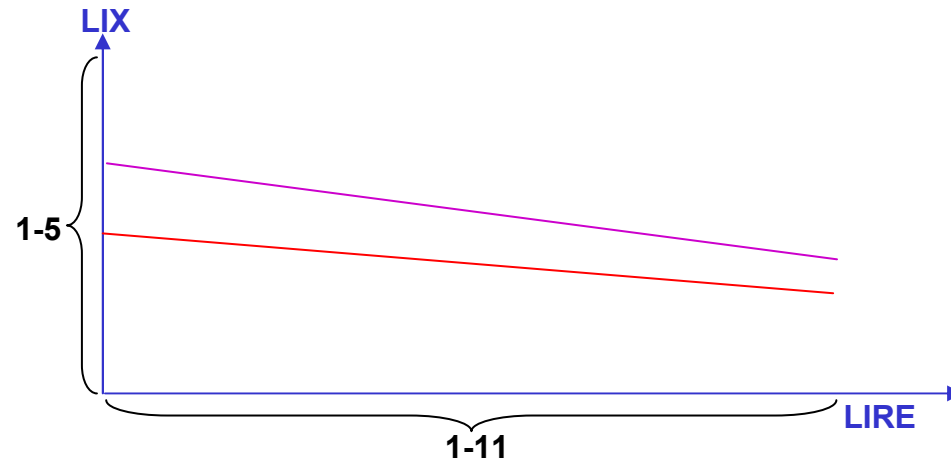
Interpretation?

Modell 2: AV = LIX, UV = LIRE (getrennt: ABL, NBL)



Weiteres Vorgehen:

- Ein Modell?
- leichter Interaktionseffekt?



Koeffizienter^a

Modell	Nicht standardisierte Koeffizienten	Standardfehler	Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
						B	Beta
1	(Konstante)	2,989	,048	62,324	,000	1,000	1,000
	LIRE	-,063	,008	-,188	,000		

R²=.035

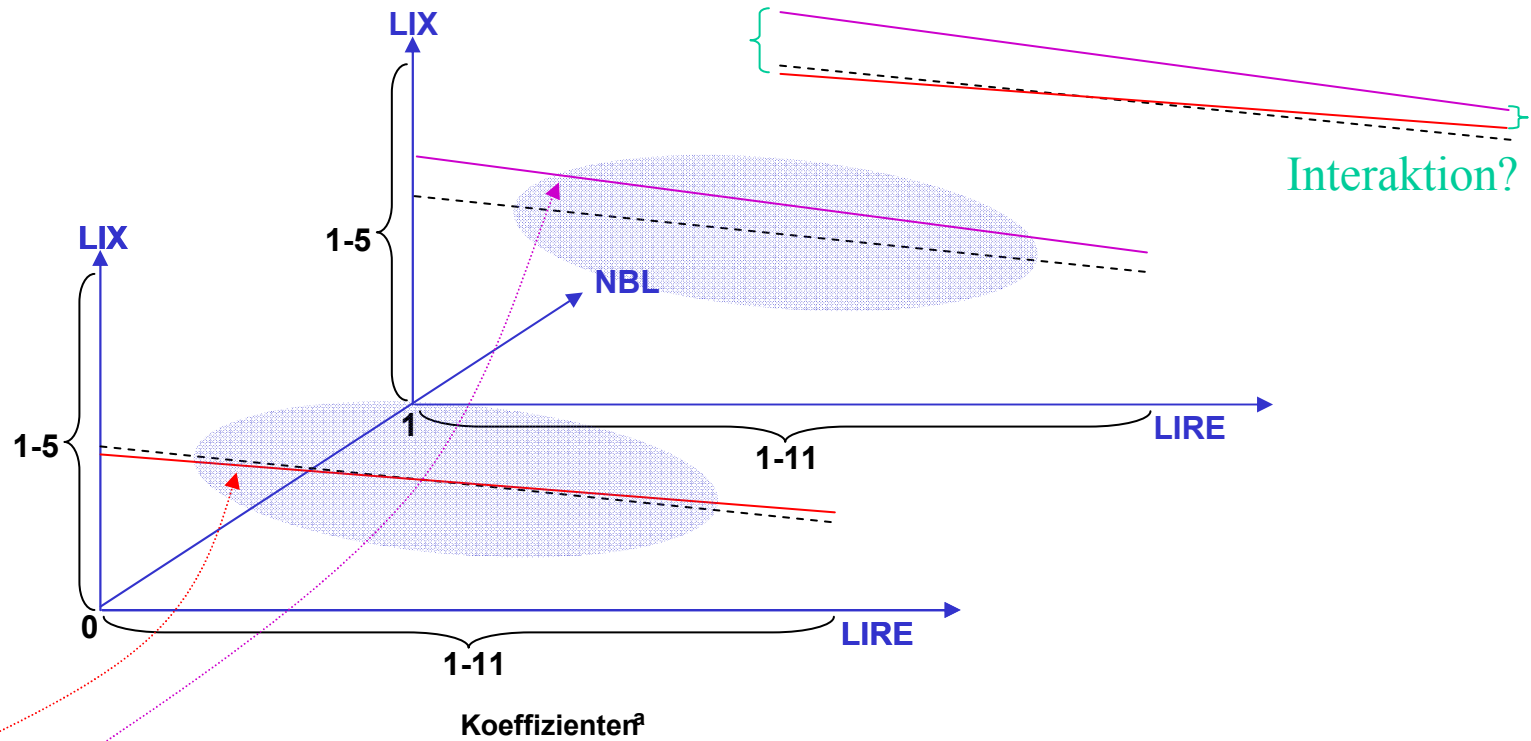
Koeffizienter^a

Modell	Nicht standardisierte Koeffizienten	Standardfehler	Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
						B	Beta
1	(Konstante)	3,825	,078	48,952	,000	1,000	1,000
	LIRE	-,094	,015	-,271	,000		

R²=.073

a. Abhängige Variable: lix extrem linke Einstellungen

Modell 3: AV = LIX, UV = LIRE und NBL



ABL + NBL		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	3,032	,043		70,979	,000		
	NBL	,672	,031	<u>,412</u>	21,819	,000	,976	1,024
	LIRE	-,070	,007	-,189	-9,996	,000	,976	1,024

$R^2 = .035$
+
 $R^2 = .073$
↓
 $R^2 = .230$

a. Abhängige Variable: lix extrem linke Einstellungen

Modell 4: AV = LIX, UV = LIRE und NBL + Interaktion



Korrelationen

		NBL	LIRE	zNBLLIRE
NBL	Korrelation nach Pearson	1	-,154	-,188
	Signifikanz (2-seitig)		,000	,000
	N	2261	2224	2224
LIRE	Korrelation nach Pearson	-,154	1	,022
	Signifikanz (2-seitig)	,000		,303
	N	2224	2254	2224
zNBLLIRE	Korrelation nach Pearson	-,188	,022	1
	Signifikanz (2-seitig)	,000	,303	
	N	2224	2224	2224

Koeffizienter^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	3,031	,043		71,002	,000		
	NBL	,661	,031	,405	21,098	,000	,943	1,061
	LIRE	-,070	,007	-,189	-10,015	,000	,976	1,024
	zNBLLIRE	-,025	,013	-,036	-1,903	,057	,965	1,036

a. Abhängige Variable: lix extrem linke Einstellungen



Dummyvariablen als unabhängige Variablen im linearen Regressionsmodell



frs3 Religion

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig 1 Einer evangelischen Kirche	892	39,6	39,6	39,6
2 Katholische Kirche	716	31,8	31,8	71,5
6 Aus der ev. Kirche ausgetreten	241	10,7	10,7	82,2
7 Aus der kath. Kirche ausgetreten	89	4,0	4,0	86,1
8 Nie Mitglied einer Kirche / Glaubensgemeinschaft gewesen	312	13,9	13,9	100,0
Gesamt	2250	100,0	100,0	

AusgangsvARIABLE

ev

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig ,00	1358	60,4	60,4	60,4
1,00	892	39,6	39,6	100,0
Gesamt	2250	100,0	100,0	

NoKo

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig ,00	1608	71,5	71,5	71,5
1,00	642	28,5	28,5	100,0
Gesamt	2250	100,0	100,0	

kath

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig ,00	1534	68,2	68,2	68,2
1,00	716	31,8	31,8	100,0
Gesamt	2250	100,0	100,0	

nachfolgende Berechnungen:

- Daten aus der Studie 2003
- AV: Sympathie für die CSU (-5 ... +5)



Lineare Abhängigkeit der Dummyvariablen



Katholiken: $kath = 1$ $ev = 0$ $NoKo = 0$ $\rightarrow \Sigma = 1$
Protestanten: $kath = 0$ $ev = 1$ $NoKo = 0$ $\rightarrow \Sigma = 1$
keine Konf.: $kath = 0$ $ev = 0$ $NoKo = 1$ $\rightarrow \Sigma = 1$

Modellzusammenfassung^b

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	1,000 ^a	1,000	1,000	,00000

a. Einflußvariablen : (Konstante), NoKo, ev

b. Abhängige Variable: kath

Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	1,000	,000		9,0E+08	,000		
	ev	-1,000	,000	-1,050	-6,7E+08	,000	,738	1,355
	NoKo	-1,000	,000	-,969	-6,2E+08	,000	,738	1,355

a. Abhängige Variable: kath

Bivariate Regressionen



Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	5,281	,072		73,516	,000		
	kath	1,122	,127	,183	8,813	,000	1,000	1,000

R² = 0.033

Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	5,584	,078		71,955	,000		
	ev	,138	,123	,024	1,120	,263	1,000	1,000

R² = 0.001

Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	6,026	,070		86,414	,000		
	NoKo	-1,355	,130	-,214	-10,386	,000	1,000	1,000

R² = 0.046

a. Abhängige Variable: fr5_3 Sympathie/Antipathie: CSU

AV: Sympathie CSU, UV: 2 der 3 Dummyvariablen



Koeffizienter^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	4,671	,110		42,587	,000		
	ev	1,051	,144	,180	7,303	,000	,694	1,441
	kath	1,733	,151	,283	11,467	,000	,694	1,441

R² = 0.056

Koeffizienter^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	6,403	,104		61,619	,000		
	ev	-,682	,140	-,117	-4,883	,000	,738	1,355
	NoKo	-1,733	,151	-,274	-11,467	,000	,738	1,355

R² = 0.056

Koeffizienter^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	5,722	,093		61,404	,000		
	kath	,682	,140	,111	4,883	,000	,813	1,230
	NoKo	-1,051	,144	-,166	-7,303	,000	,813	1,230

R² = 0.056

a. Abhängige Variable: fr5_3 Sympathie/Antipathie: CSU



Empirisch vorkommende Kombinationen der Variablenwerte

<u>Variable:</u>	Protestanten	Katholiken	Konfessionslose
ev	1	0	0
kath	0	1	0
NoKo	0	0	1

$$ev + kath + NoKo = 1$$

$$\hat{y} = 4.671 + 1.051 \cdot ev + 1.733 \cdot kath \quad | \quad ev = 1 - kath - NoKo$$

$$\hat{y} = 4.671 + 1.051 \cdot (1 - kath - NoKo) + 1.733 \cdot kath$$

$$\hat{y} = 4.671 + 1.051 - 1.051 \cdot kath - 1.051 \cdot NoKo + 1.733 \cdot kath$$

$$\hat{y} = 5.722 + 0.682 \cdot kath - 1.051 \cdot NoKo$$

Empirische Ergebnisse zum Vergleich (Wdh.)



Koeffizienter^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	4,671	,110		42,587	,000		
	ev	1,051	,144	,180	7,303	,000	,694	1,441
	kath	1,733	,151	,283	11,467	,000	,694	1,441

R² = 0.056

Koeffizienter^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	6,403	,104		61,619	,000		
	ev	-,682	,140	-,117	-4,883	,000	,738	1,355
	NoKo	-1,733	,151	-,274	-11,467	,000	,738	1,355

R² = 0.056

Koeffizienter^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	5,722	,093		61,404	,000		
	kath	,682	,140	,111	4,883	,000	,813	1,230
	NoKo	-1,051	,144	-,166	-7,303	,000	,813	1,230

R² = 0.056

a. Abhängige Variable: fr5_3 Sympathie/Antipathie: CSU

Dummyvariablen zentriert



Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	5,639	,059		96,175	,000		
	zentev	1,051	,144	,180	7,303	,000	,694	1,441
	zentkath	1,733	,151	,283	11,467	,000	,694	1,441

R² = 0.056

a. Abhängige Variable: fr5_3 Sympathie/Antipathie: CSU

Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	5,639	,059		96,175	,000		
	zentkath	,682	,140	,111	4,883	,000	,813	1,230
	zentNoKo	-1,051	,144	-,166	-7,303	,000	,813	1,230

R² = 0.056

a. Abhängige Variable: fr5_3 Sympathie/Antipathie: CSU



Empirisch vorkommende Kombinationen der Variablenwerte

<u>Variable:</u>	Protestanten	Katholiken	Konfessionslose	AM
ev_{zent}	0.60355556	-0.39644444	-0.39644444	0.39644444
$kath_{zent}$	-0.31822222	0.68177778	-0.31822222	0.31822222
$NoKo_{zent}$	-0.28533333	-0.28533333	0.71466667	<u>0.28533333</u>
				Σ 1

$$ev_{zent} + kath_{zent} + NoKo_{zent} = 0$$

$$\hat{y} = 5.639 + 1.051 \cdot ev_{zent} + 1.733 \cdot kath_{zent} \quad | \quad ev_{zent} = 0 - kath_{zent} - NoKo_{zent}$$

$$\hat{y} = 5.639 + 1.051 \cdot (0 - kath_{zent} - NoKo_{zent}) + 1.733 \cdot kath_{zent}$$

$$\hat{y} = 5.639 + 0 - 1.051 \cdot kath_{zent} - 1.051 \cdot NoKo_{zent} + 1.733 \cdot kath_{zent}$$

$$\hat{y} = 5.639 + 0.682 \cdot kath_{zent} - 1.051 \cdot NoKo_{zent}$$

Dummyvariablen zentriert (Wdh.)



Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	5,639	,059		96,175	,000		
	zentev	1,051	,144	,180	7,303	,000	,694	1,441
	zentkath	1,733	,151	,283	11,467	,000	,694	1,441

R² = 0.056

a. Abhängige Variable: fr5_3 Sympathie/Antipathie: CSU



Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	5,639	,059		96,175	,000		
	zentkath	,682	,140	,111	4,883	,000	,813	1,230
	zentNoKo	-1,051	,144	-,166	-7,303	,000	,813	1,230

R² = 0.056

a. Abhängige Variable: fr5_3 Sympathie/Antipathie: CSU

Dummyvariablen z-standardisiert



Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	5,639	,059		96,175	,000		
	zev	,514	,070	,180	7,303	,000	,694	1,441
	zkath	,807	,070	,283	11,467	,000	,694	1,441

R² = 0.056

a. Abhängige Variable: fr5_3 Sympathie/Antipathie: CSU

Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	5,639	,059		96,175	,000		
	zkath	,318	,065	,111	4,883	,000	,813	1,230
	zNoKo	-,475	,065	-,166	-7,303	,000	,813	1,230

R² = 0.056

a. Abhängige Variable: fr5_3 Sympathie/Antipathie: CSU



Empirisch vorkommende Kombinationen der Variablenwerte

	Protestanten	Katholiken	Konfessionslose	SD _{Originalvari}
<u>z-stand. Variablen-</u>				
<u>Werte:</u>				
ev_{zstd}	1.23359026	-0.81028166	-0.81028166	0.48926745
$kath_{zstd}$	-0.68304183	1.46338851	-0.68304183	0.46588980
$NoKo_{zstd}$	-0.63172484	-0.63172484	1.58226410	0.45167344
<u>zentrierte Werte</u>				
(z-stand. Variablen-				
werte x SD):				
ev_{zent}	0.60355556	-0.39644444	-0.39644444	0.48926745
$kath_{zent}$	-0.31822222	0.68177778	-0.31822222	0.46588980
$NoKo_{zent}$	-0.28533333	-0.28533333	0.71466667	0.45167344



zentrierte Werte

$$ev_{zstd} \cdot SD_{ev} + kath_{zstd} \cdot SD_{kath} + NoKo_{zstd} \cdot SD_{NoKo} = 0$$

$$ev_{zstd} \cdot SD_{ev} = 0 - kath_{zstd} \cdot SD_{kath} - NoKo_{zstd} \cdot SD_{NoKo} \quad | :SD_{ev}$$

$$ev_{zstd} = 0 - kath_{zstd} \cdot SD_{kath} / SD_{ev} - NoKo_{zstd} \cdot SD_{NoKo} / SD_{ev}$$

$$ev_{zstd} = 0 - 0.952 \cdot kath_{zstd} - 0.923 \cdot NoKo_{zstd}$$

SDs dividieren!

$$\hat{y} = 0.180 \cdot ev_{zstd} + 0.283 \cdot kath_{zstd}$$

$$| ev_{zstd} = 0 - 0.952 \cdot kath_{zstd} - 0.923 \cdot NoKo_{zstd}$$

$$\hat{y} = 0.180 \cdot (0 - 0.952 \cdot kath_{zstd} - 0.923 \cdot NoKo_{zstd}) + 0.283 \cdot kath_{zstd}$$

$$\hat{y} = 0 - 0.171 \cdot kath_{zstd} - 0.166 \cdot NoKo_{zstd} + 0.283 \cdot kath_{zstd}$$

$$\hat{y} = 0.112 \cdot kath_{zstd} - 0.166 \cdot NoKo_{zstd}$$

Dummyvariablen z-standardisiert (Wdh.)



Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	5,639	,059		96,175	,000		
	zev	,514	,070	,180	7,303	,000	,694	1,441
	zkath	,807	,070	,283	11,467	,000	,694	1,441

R² = 0.056

a. Abhängige Variable: fr5_3 Sympathie/Antipathie: CSU

Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz	Kollinearitätsstatistik	
		B	Standardfehler	Beta			Toleranz	VIF
1	(Konstante)	5,639	,059		96,175	,000		
	zkath	,318	,065	,111	4,883	,000	,813	1,230
	zNoKo	-,475	,065	-,166	-7,303	,000	,813	1,230

R² = 0.056

a. Abhängige Variable: fr5_3 Sympathie/Antipathie: CSU

(SPSS) AV: Sympathie CSU; UVs: ev, kath, NoKo



REGRESSION

```
/MISSING LISTWISE  
/STATISTICS COEFF OUTS R ANOVA TOL  
/CRITERIA=PIN(.05) POUT(.10)  
/NOORIGIN  
/DEPENDENT fr5_3 (Sympathie/Antipathie CSU)  
/METHOD=ENTER ev kath NoKo  
/RESIDUALS HIST(ZRESID).
```



Aufgenommene/Entfernte Variablen^b

Modell	Aufgenommene Variablen	Entfernte Variablen	Methode
1	NoKo, kath ^a	.	Eingeben

a. Toleranz = ,000 erreichte Grenzen.

b. Abhängige Variable: fr5_3 Sympathie/Antipathie: CSU



**Vielen Dank für Ihre
Aufmerksamkeit!**