Lineare Modelle (Beispiele)

$$y = \beta_0 + \beta_1 x + u$$

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + u$$

$$y = \beta_0 + \beta_1 \log x + u$$

$$y = \beta_0 + \beta_1 \sqrt{x} + u$$

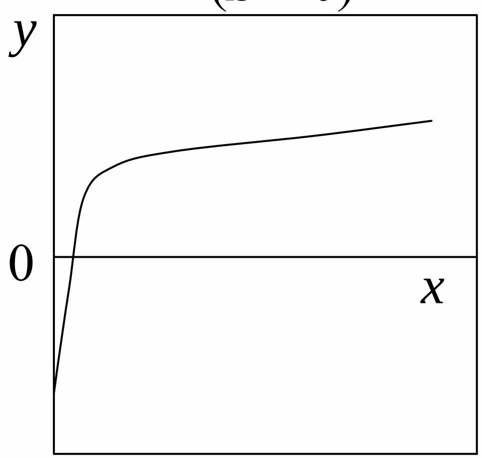
Nichtlineares Modell (Beispiel)

$$y = \beta_0 + e^{\beta_1 x} + u$$

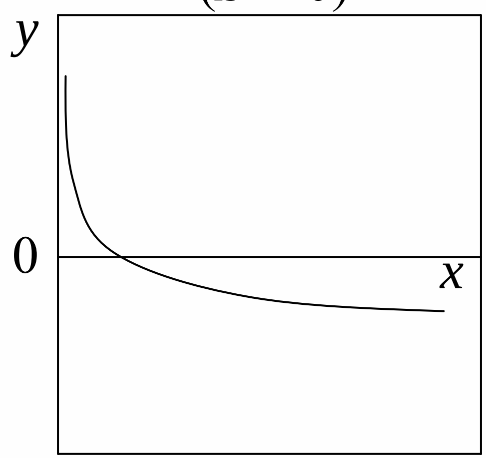
Linearisierbare Funktionen (Beispiel 1)

Funktion	Transformation	Lineare Form
$y = \alpha + \log x^{\beta}$	$x' = \log x$	$y = \alpha + \beta x'$

(ß > 0)



(ß < 0)



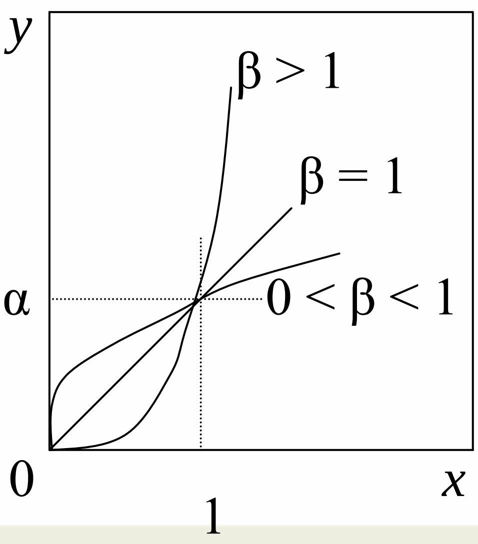
$$y = \alpha + \log x^{\beta}$$

$$y = \alpha + \log\left(x^{\beta}\right) = \alpha + \beta(\log x)$$

Linearisierbare Funktionen (Beispiel 2)

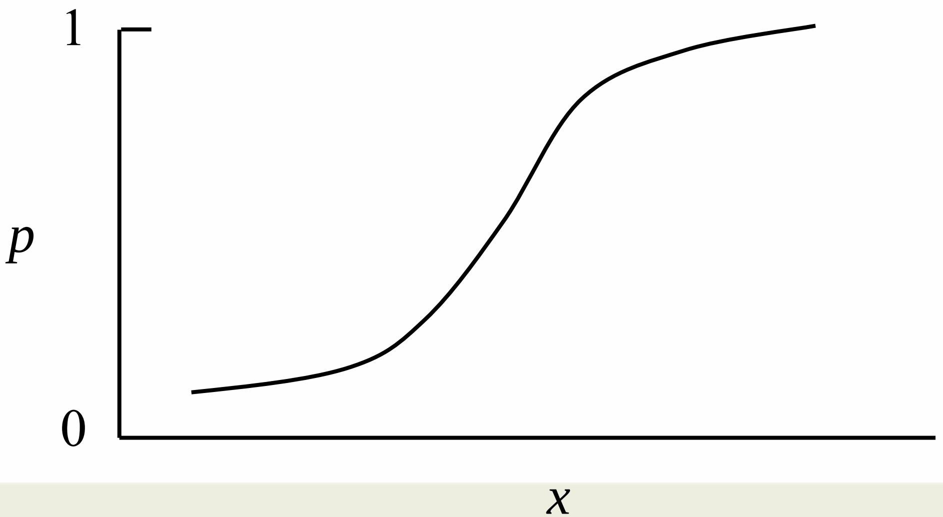
Funktion	Transformation	Lineare Form
$y = \alpha x^{\beta}$	$y' = \log y$ $x' = \log x$	$y' = \log \alpha + \beta x'$

(α, ß, $x > 0$)



Linearisierbare Funktionen (Beispiel 3)

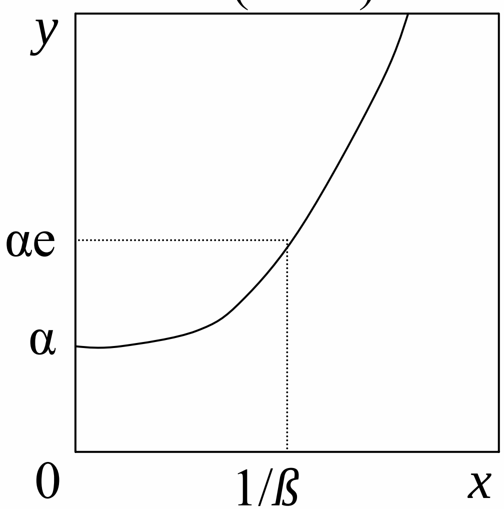
Funktion	Transformation	Lineare Form
$y=\dfrac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$	$y'=\ln\left(\dfrac{y}{1-y}\right)$	$y'=\alpha+\beta x$



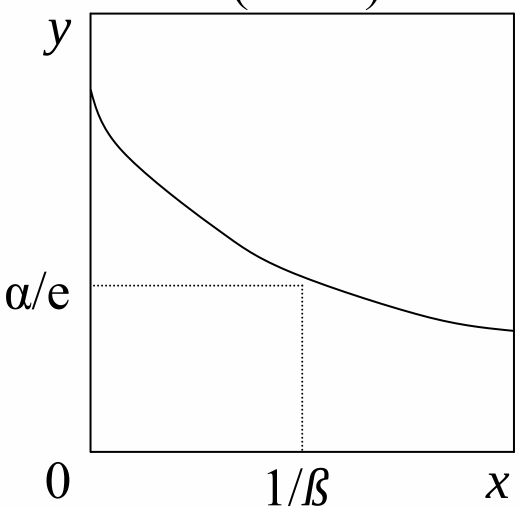
Linearisierbare Funktionen (Beispiel 4)

Funktion	Transformation	Lineare Form
$y = \alpha \cdot e^{\beta x}$	$y' = \ln y$	$y' = \ln \alpha + \beta x$

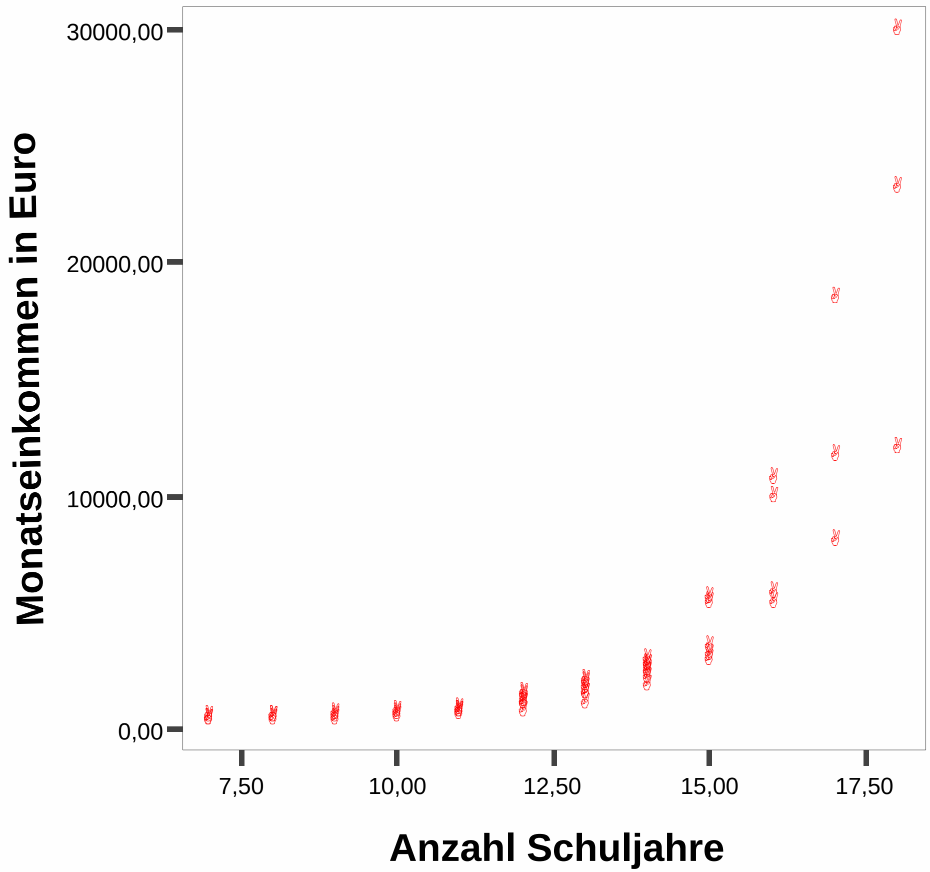
(ß > 0)



(ß < 0)



Ausgangspunkt: Beispiel einer *kurvi*linearen Beziehung zw. Einkommen und Schulbildung



[n=50; fiktive Zahlen]

wenn als **lineare** Einfachregression geschätzt, ..

Modellzusammenfassung[b]

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,708[a]	,501	,491	4277,19534

a. Einflußvariablen : (Konstante), EDUC Anzahl Schuljahre

b. Abhängige Variable: INCOME Monatseinkommen in Euro

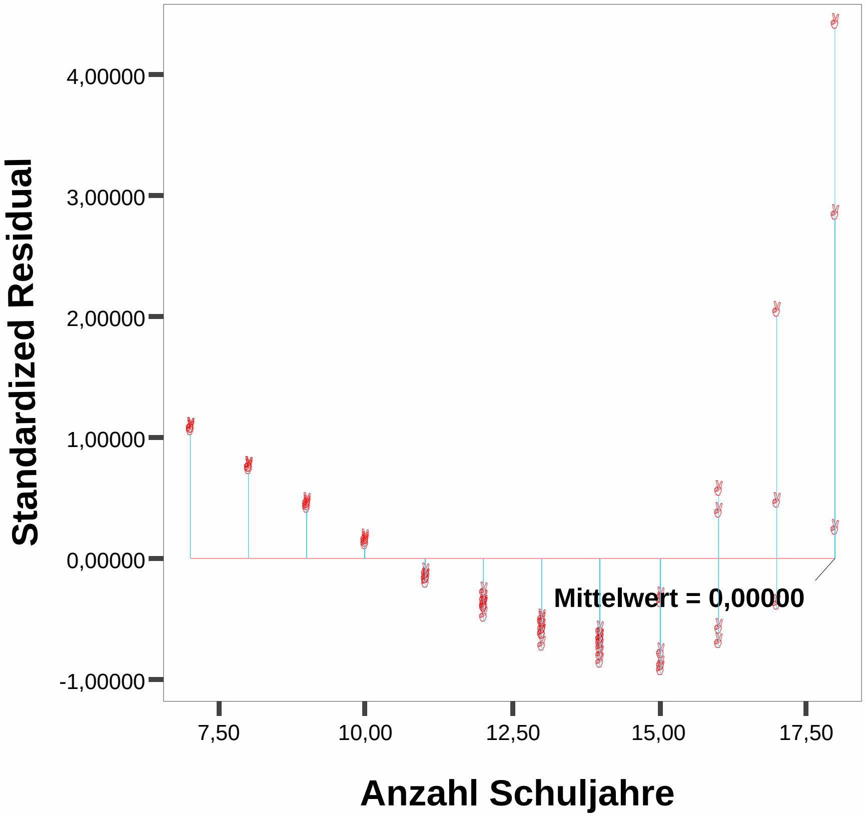
Ergebnis:
50% erklärte Varianz, positiver Effekt der Schulbildung,

aber …. »

Koeffizienten[a]

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	-13780,111	2607,219		-5,285	,000
	EDUC Anzahl Schuljahre	1379,899	198,752	,708	6,943	,000

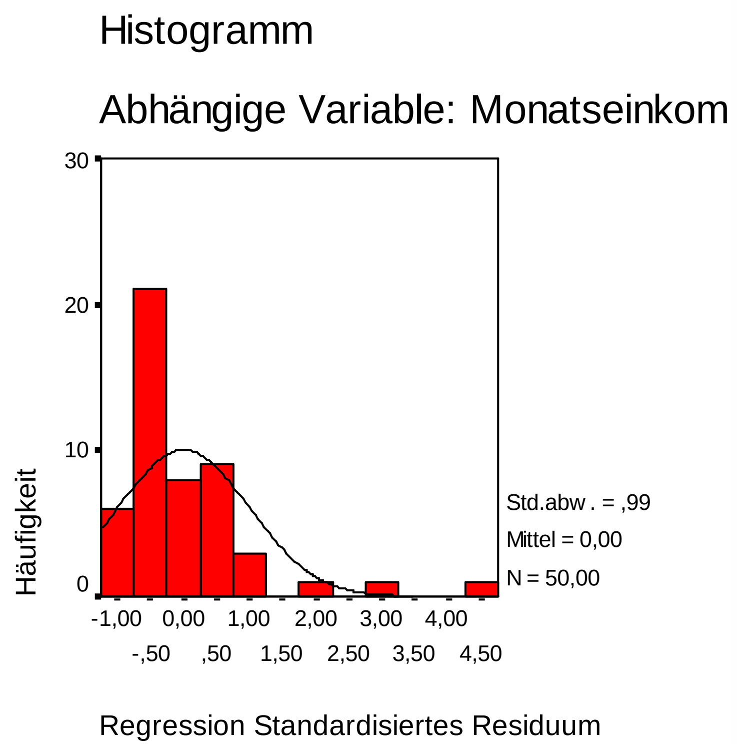
a. Abhängige Variable: INCOME Monatseinkommen in Euro



Mittelwert

.. es werden in Abhängigkeit von X die stand. Residuen erst über-, dann unter-, dann wieder überschätzt [Residuen streuen nicht um Null]

NV Annahme verletzt; Stand. Residuen weichen von Null stärker nach oben als nach unten ab

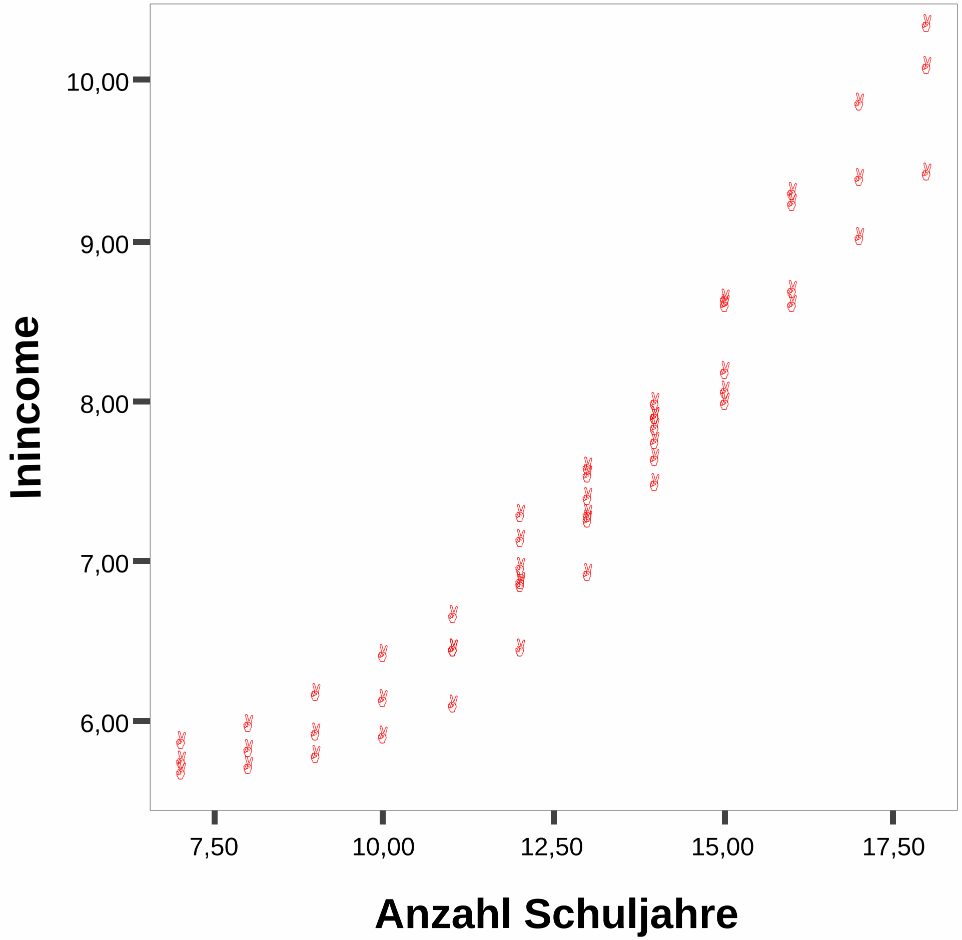


Residuenstatistik[a]

	Minimum	Maximum	Mittelwert	Standardabweichung	N
Nicht standardisierter vorhergesagter Wert	-4120,8179	11058,0703	3827,4000	4242,24759	50
Nicht standardisierte Residuen	-4088,3735	18661,9297	,0000	4233,32551	50
Standardisierter vorhergesagter Wert	-1,874	1,704	,000	1,000	50
Standardisierte Residuen	-,956	4,363	,000	,990	50

a. Abhängige Variable: INCOME Monatseinkommen in Euro

Beziehung zwischen dem *natürlichen Logarithmus* von Einkommen und Anzahl der Schuljahre ..



.. stellt in stärkerem Maße eine lineare Beziehung dar als die Ursprungsbeziehung

Modellzusammenfassung[b]

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,959[a]	,919	,918	,36843

a. Einflußvariablen : (Konstante), EDUC Anzahl Schuljahre

b. Abhängige Variable: LNINCOME

Erklärte Varianz liegt nun bei ca. 92% ….

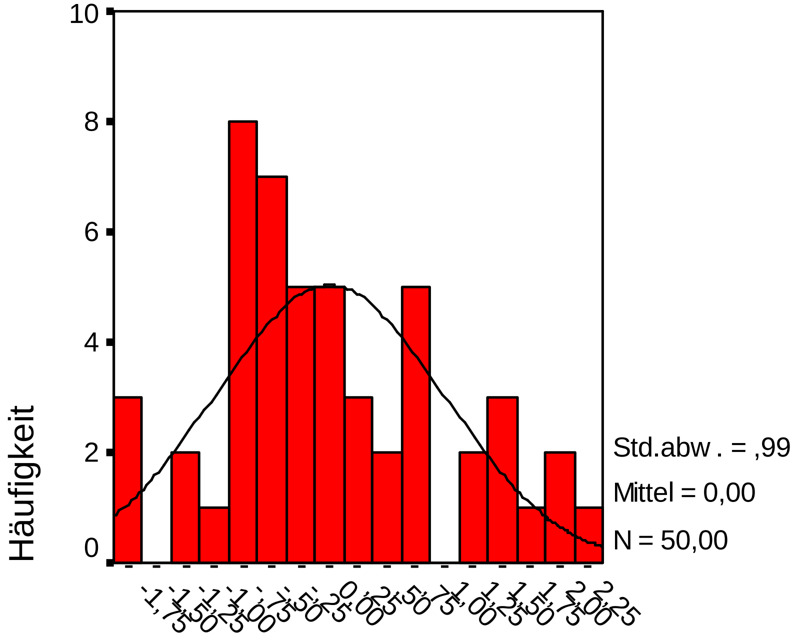
Koeffizienten[a]

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	2,288	,225		10,190	,000
	EDUC Anzahl Schuljahre	,401	,017	,959	23,402	,000

a. Abhängige Variable: LNINCOME

Histogramm

Abhängige Variable: LNINCOME



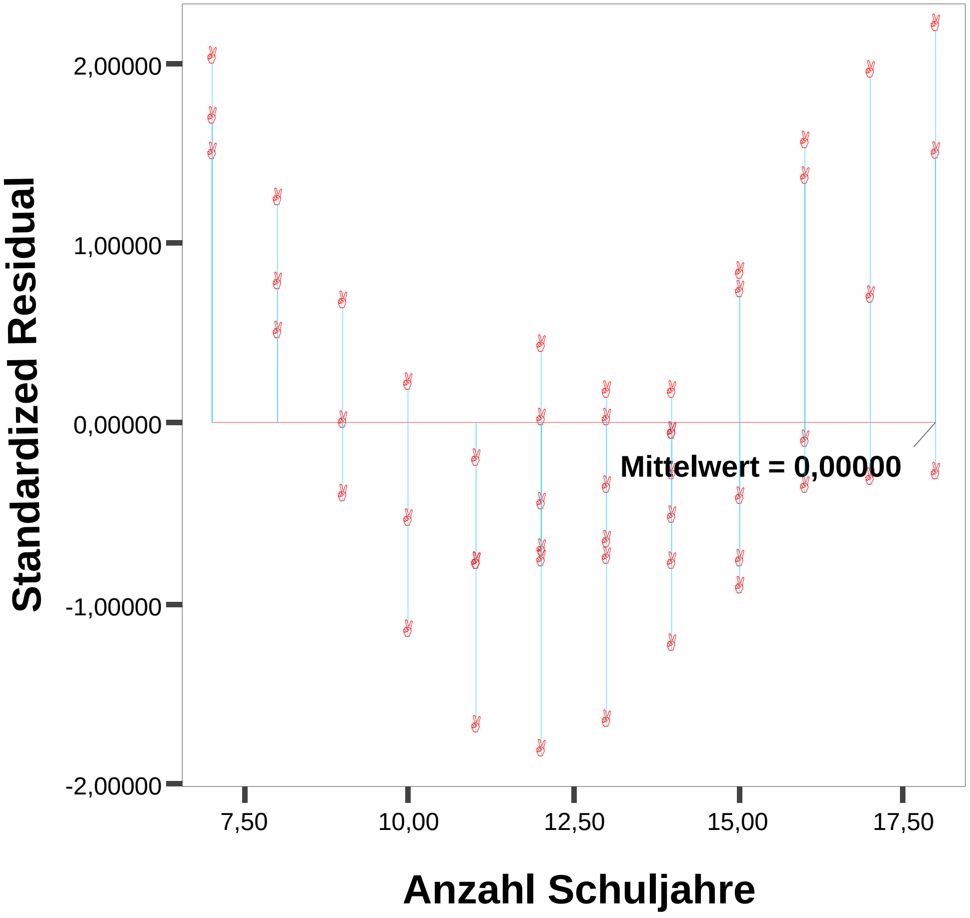
Regression Standardisiertes Residuur

NV Annahme ist weniger stark verletzt; stand. Residuen liegen ausgewogener um Null

Residuenstatistik[a]

	Minimum	Maximum	Mittelwert	Standardabweichung	N
Nicht standardisierter vorhergesagter Wert	5,0930	9,5001	7,4007	1,23171	50
Nicht standardisierte Residuen	-,6827	,7995	,0000	,36465	50
Standardisierter vorhergesagter Wert	-1,874	1,704	,000	1,000	50
Standardisierte Residuen	-1,853	2,170	,000	,990	50

a. Abhängige Variable: LNINCOME



Mittelwert

.. aber die stand. Residuen zeigen noch das Muster, erst über, dann unter, dann wieder über Null zu liegen, wenn X ansteigt

Deshalb: Schätzung einer polynomischen Regression, in der zusätzlich zu „educ" das Quadrat dieser Variablen, „educ_sq" einbezogen wird

Im Beispiel wird die lineare Form damit zu …

$$y' = \ln \alpha + \beta x + \gamma x^2$$

Durch Bildung der Ableitung* von …

$$\beta x + \gamma x^2$$

nach *x* resultiert mit ..

$$\beta + 2 \cdot \gamma x$$

ein Ausdruck, der zeigt, wie sich der Effekt von *x* in Abhängigkeit von den Ausprägungen von *x* verändert

*[$x^n \rightarrow n \cdot x^{n-1}$]

Modellzusammenfassung[b]

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,980[a]	,961	,959	,25984

a. Einflußvariablen : (Konstante), EDUC_SQ, EDUC Anzahl Schuljahre

b. Abhängige Variable: LNINCOME

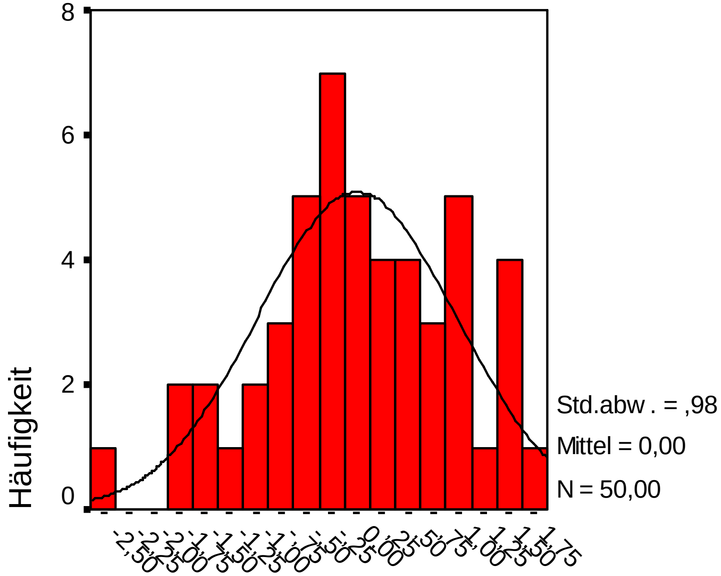
Koeffizienten[a]

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	6,055	,558		10,846	,000
	EDUC Anzahl Schuljahre	-,243	,092	-,581	-2,631	,011
	EDUC_SQ	2,582E-02	,004	1,553	7,036	,000

a. Abhängige Variable: LNINCOME

Histogramm

Abhängige Variable: LNINCOME



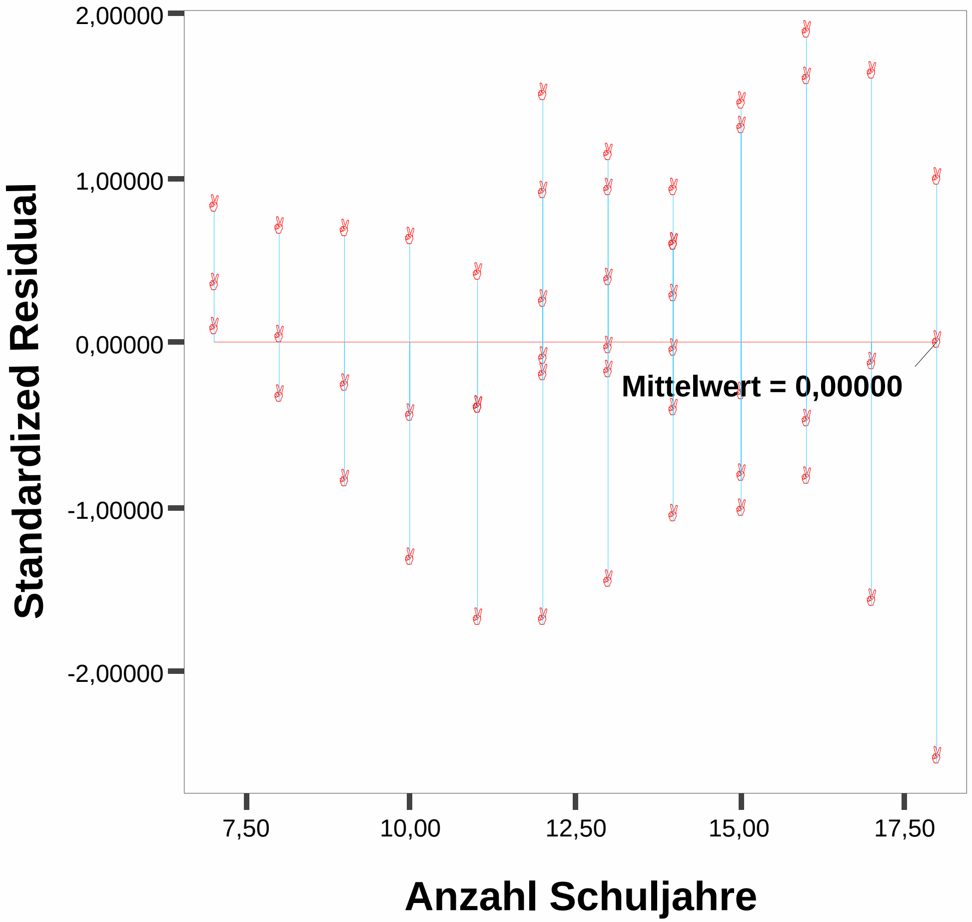
Regression Standardisiertes Residuum

NV Annahme nochmals weniger stark verletzt;

Residuenstatistik[a]

	Minimum	Maximum	Mittelwert	Standardabweichung	N
Nicht standardisierter vorhergesagter Wert	5,6212	10,0510	7,4007	1,25909	50
Nicht standardisierte Residuen	-,6667	,4802	,0000	,25448	50
Standardisierter vorhergesagter Wert	-1,413	2,105	,000	1,000	50
Standardisierte Residuen	-2,566	1,848	,000	,979	50

a. Abhängige Variable: LNINCOME



Mittelwert

.. und die stand. Residuen streuen nun ausgewogener um Null,
wenn X ansteigt

$$y' = \ln\alpha + \beta x + \gamma x^2$$

$$y' = 6{,}055 - 0{,}243x + 0{,}026x^2$$

$$y = \alpha \cdot e^{(\beta x + \gamma x^2)} = \alpha \cdot e^{\beta x} \cdot e^{\gamma x^2}$$

$$y = 426{,}2 \cdot e^{(-0{,}243x + 0{,}026x^2)} = 426{,}2 \cdot e^{-0{,}243x} \cdot e^{0{,}026x^2}$$

Bil-dung	7	8	9	10	11	12	13	14	15	..
Ein-kom.	278	322	393	505	684	976	1465	2319	3866	..
Diff.		44	71	112	179	292	489	854	1547	..