

Einführung in die angewandte Ökonometrie

Herbert Stocker

SS 2011



Blick von meinem Schreibtischsessel (Canon Ixus, 18.03.2008)

Dieses Manuskript ist 'work in progress', d.h. unvollständig und nicht komplett korrigiert! Hinweise auf Fehler, Ungereimtheiten etc. sowie Verbesserungsvorschläge aller Art sind willkommen!

Kapitel 8

Nichtlineare Funktionsformen

“Alle guten Dinge haben etwas Lässiges und liegen wie Kühe auf der Wiese.”
(Friedrich Nietzsche)

Die Gauss Markov Annahme der *Linearität* ist nicht ganz so restriktiv wie es auf den ersten Blick scheinen mag, denn sie bezieht sich nur auf Linearität in den Parametern, aber *nicht* auf Linearität in den Variablen. Modelle, die nicht-linear in den Variablen sind, können ganz normal mit OLS geschätzt werden.

So ist z.B. das Modell

$$y = \beta_0 + \beta_1 x_1^2 + \beta_2 \log(x_2) + \varepsilon$$

zwar nicht-linear in den Variablen x_1 und x_2 , aber linear in den Parametern β_0, β_1 und β_2 , und erfüllt deshalb die Gauss Markov Annahme der Linearität. Deshalb kann dieses Modell problemlos mit OLS geschätzt werden.¹

Hingegen benötigt man für Modelle, die *nicht-linear in den Parametern* sind, wie z.B.

$$y = \beta_0 + \beta_1^2 x_1 + \log(\beta_2) x_2 + \beta_1 \beta_2 x_1 x_2 + \varepsilon$$

andere Methoden, auf die wir hier nicht eingehen werden.²

Hier beschränken wir uns auf Modelle, die *nicht-linear in den Variablen* sind, denn solche Modelle werden in der Praxis sehr häufig angewandt. Das Problem bei solchen Modellen ist weniger die Schätzung – die funktioniert vollkommen gleich wie bisher – sondern die korrekte Interpretation der Parameter.

8.1 Logarithmische Transformationen

Die in der angewandten Ökonometrie am häufigsten verwendete Funktionsform ist vermutlich die logarithmische Transformation von Variablen. Wenn die abhängige

¹Manchmal ergeben sich Multikollinearitätsprobleme, da z.B. x und x^2 über einen begrenzten Bereich häufig korreliert sind, aber Multikollinearität verletzt ebenfalls keine Gauss-Markov Annahme, weshalb die Effizienz davon nicht berührt wird.

²Modelle, die nicht-linear in den Parametern sind, können in EViews trotzdem sehr einfach geschätzt werden, z.B. mit `ls y = c(1) + c(2)^2 * x1 + @log(c(3)) * x2`, aber die Theorie dahinter und die Eigenschaften unterscheiden sich deutlich vom OLS Modell.

und die erklärende Variable logarithmiert wird spricht man von einem log-linearen oder log-log Modell. Wenn hingegen nur die abhängige *oder* die erklärende Variable logarithmiert wird spricht man von einem semi-log Modell, oder manchmal auch von einem log-lin bzw. lin-log Modell.

Ein wesentlicher Grund für die Beliebtheit logarithmischer Modelle liegt in der Tatsache, dass Differenzen logarithmierter Variablen näherungsweise den *relativen* Änderungen der ursprünglichen Variablen entsprechen. Wir erinnern uns an die Ableitungsregel für den natürlichen Logarithmus

$$\frac{d \ln x}{dx} = \frac{1}{x}$$

ist. Intuitiv können wir dies umgeschrieben denken als

$$d \ln x = \frac{dx}{x}$$

d.h. eine infinitesimal kleine Änderung des logarithmierten x ist gleich der *relativen* Änderung von x , da dx eine infinitesimal kleine Änderung von x bezeichnet.

In Analogie dazu würden wir erwarten, dass für diskrete Fälle *näherungsweise* (\approx) gilt

$$\ln(x + \Delta x) - \ln(x) \approx \frac{\Delta x}{x}$$

Dieser Zusammenhang gilt tatsächlich, wenn $\Delta x/x$ ‘relativ’ klein ist, d.h. die logarithmische Differenz einer Variable misst tatsächlich näherungsweise die relative Änderung der ursprünglichen Variable.

8.1.1 Log-lineare (log-log) Modelle

Das Modell

$$y_i = A x_i^\beta \exp(\varepsilon_i)$$

lässt sich einfach durch logarithmieren linearisieren:³

$$\ln y_i = \alpha + \beta \ln x_i + \varepsilon_i \quad (\alpha := \ln A)$$

Dieses logarithmierte Modell ist *linear in den Parametern* und kann deshalb ganz normal mit OLS geschätzt werden. Abbildung 8.1 zeigt mögliche Verläufe für positive (links) und negative β (rechts).

Interpretation: Auch die Koeffizienten log-linearer Modelle messen wie üblich den *marginalen Effekt*

$$\beta = \frac{d \ln y}{d \ln x}$$

das heißt, der Koeffizient gibt an, um wieviele Einheiten sich $\ln y$ ceteris paribus ändert, wenn $\ln x$ um eine Einheit zunimmt. Das Problem dabei ist, dass kaum

³ $\exp(\varepsilon_i) := e^{\varepsilon_i}$, wobei e die Eulersche Zahl 2.71828 ist. Um Verwechslungen mit den Stichprobenresiduen e zu vermeiden werden wir meist die Funktion $\exp(\varepsilon_i)$ verwenden.

Exkurs: Logarithmus

Logarithmieren zu einer Basis ist einfach die Umkehrung des Potenzieren einer Basis, d.h. wenn

$$a = b^x \quad \text{dann ist} \quad \log_b(a) = x$$

oder in anderen Worten, der Logarithmus zur Basis b ist die Umkehrfunktion der allgemeinen Exponentialfunktion zur Basis b .

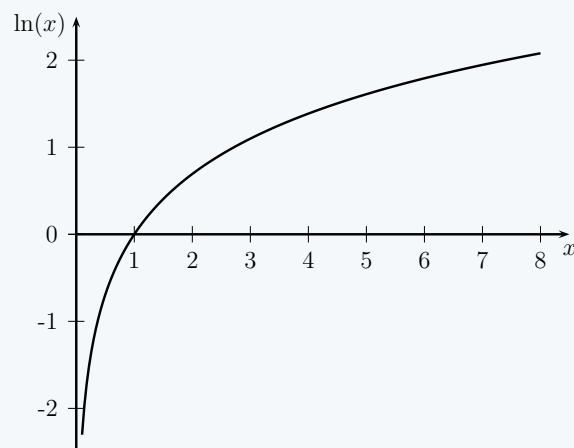
In der Ökonometrie wird fast ausschließlich der natürliche Logarithmus verwendet, d.h. der Logarithmus zur Basis $e = 2.71828$. Für den natürlichen Logarithmus wird häufig das Funktionszeichen \ln verwendet.

$$\ln(x) = a \quad \text{bedeutet} \quad e^a := \exp(a) = x$$

Zum Beispiel ist $e^2 = 2.71828^2 = 7.389$, deshalb ist $\ln(7.389) = 2$. Der Logarithmus einer negativen Zahl ist deshalb nicht definiert.

Hier ein paar Logarithmen positiver Zahlen:

$$\begin{aligned} \ln(0) &= -\infty \\ \ln(0.1) &= -2.303 \\ \ln(0.5) &= -0.693 \\ \ln(1) &= 0 \\ \ln(10) &= 2.303 \\ \ln(1000) &= 6.908 \\ \ln(1000000) &= 13.816 \end{aligned}$$



Aus den Potenzgesetzen lassen sich die üblichen Rechenregeln für Logarithmen ableiten, z.B.

$$\begin{aligned} \ln(a \cdot b) &= \ln(a) + \ln(b) \\ \ln\left(\frac{a}{b}\right) &= \ln(a) - \ln(b) \\ \ln(a^x) &= x \ln(a) \\ \ln(1/x) &= -\ln(x) \end{aligned}$$

Achtung: Viele Ökonometrieprogramme verwenden für den natürlichen Logarithmus die Funktion `log` (EViews z.B. `@log`).



Exkurs: Logarithmische Differenz und relative Änderungsraten

Wir haben im Text behauptet, dass die logarithmische Differenz einer Variable ungefähr gleich der relativen Änderung dieser Variable ist

$$\ln(x + \Delta x) - \ln(x) \approx \frac{\Delta x}{x} \quad \text{für kleine } \frac{\Delta x}{x}$$

Dies kann allgemein mit Hilfe einer Taylor Expansion gezeigt werden. Mit einer Taylor Reihenentwicklung können nichtlineare (differenzierbare ...) Funktionen in der Umgebung bestimmter Punkte durch Potenzreihen dargestellt werden. Insbesondere gilt

$$\ln(1 + x) = \sum_{n=1}^{\infty} (-1)^{(n+1)} \frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \quad \text{für } |x| < 1$$

bzw. für $x = \frac{\Delta y}{y}$

$$\ln\left(1 + \frac{\Delta y}{y}\right) = \frac{\Delta y}{y} - \frac{1}{2} \left(\frac{\Delta y}{y}\right)^2 + \frac{1}{3} \left(\frac{\Delta y}{y}\right)^3 - \dots$$

Die linke Seite kann auch als logarithmische Differenz geschrieben werden, da

$$\ln\left(1 + \frac{\Delta y}{y}\right) = \ln\left(\frac{y + \Delta y}{y}\right) = \ln(y + \Delta y) - \ln(y)$$

Daraus folgt

$$\ln(y + \Delta y) - \ln(y) \approx \frac{\Delta y}{y}$$

da für kleine $\Delta y/y$ die Folgeterme $-1/2 (\Delta y/y)^2 + 1/3 (\Delta y/y)^3 - \dots$ der Reihe sehr klein werden und oft vernachlässigbar sind.



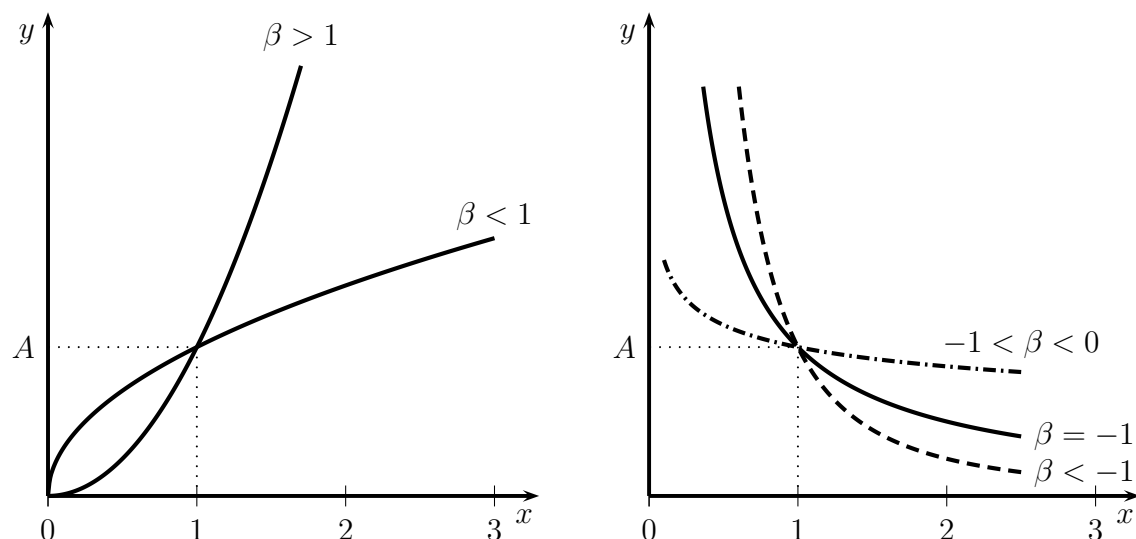


Abbildung 8.1: Log-lineare (bzw. log-log) Modelle: $y = Ax^\beta$

jemand für die Veränderungen der Logarithmen interessieren wird, wir wollen meist wissen, wie sich y selbst verändert, wenn x um eine Einheit zunimmt.

Hier zeigt sich ein erster großer Vorteil log-linearer Modelle, denn die Koeffizienten haben in diesem Fall auch in Bezug auf die ursprünglichen Variablen x und y eine sehr einfache Interpretation, sie können nämlich direkt als *Elastizitäten* interpretiert werden.

Erinnern wir uns, die Elastizität E zwischen y und x ist definiert als

$$E_{y,x} := \frac{dy}{dx} \frac{x}{y} := \frac{\frac{dy}{y}}{\frac{dx}{x}}$$

und gibt (näherungsweise) an, um wieviel *Prozent* sich die abhängige Variable y ändert, wenn die erklärende Variable x ceteris paribus um *ein Prozent* zunimmt.

Man kann sich dies intuitiv anhand der Ableitungsregel für Logarithmen veranschaulichen:

$$\frac{d \ln(y)}{dy} = \frac{1}{y} \Rightarrow d \ln(y) = \frac{dy}{y}$$

Ebenso gilt $d \ln(x) = dx/x$. Daraus folgt unmittelbar die Elastizität:

$$E_{y,x} = \frac{d \ln(y)}{d \ln(x)} = \frac{\frac{dy}{y}}{\frac{dx}{x}} = \frac{dy}{dx} \frac{x}{y} = b_1$$

Beispiel: Cobb-Douglas Funktionen der Art

$$y = b_0 x_1^{b_1} x_2^{b_2} \exp(e)$$

spielen in den Wirtschaftswissenschaften eine große Rolle und können durch Logarithmieren einfach Linearisiert werden

$$\ln y = \ln(b_0) + b_1 \ln(x_1) + b_2 \ln(x_2) + e$$

Es ist bekannt, dass die Koeffizienten der Cobb-Douglas Funktion direkt die Elastizitäten angeben, wie einfach gezeigt werden kann: die Ableitung von $y = b_0 x_1^{b_1} x_2^{b_2} \exp(e)$ nach x_1 ist⁴

$$\begin{aligned}\frac{\partial y}{\partial x_1} &= b_1 (b_0 x_1^{b_1-1} x_2^{b_2} \exp(e)) \\ &= b_1 (b_0 x_1^{b_1} x_2^{b_2} \exp(e)) x_1^{-1} \\ &= b_1 \frac{y}{x_1}\end{aligned}$$

bzw.

$$b_1 = \frac{\partial y}{\partial x_1} \frac{x_1}{y} := E_{y,x_1}$$

wobei E_{y,x_1} die übliche Elastizität bezeichnet.

Es ist offensichtlich, dass man den gleichen Koeffizienten viel einfacher unmittelbar aus der Ableitung der log-log Funktion $\ln(y) = b_0 + b_1 \ln(x_1) + b_2 \ln(x_2) + e$ erhält,

$$b_1 = \frac{\partial \ln(y)}{\partial \ln(x_1)} = E_{y,x_1}$$

Achtung: Die Beziehung $y = \alpha x^\beta$ kann ökonometrisch auf verschiedene Arten modelliert werden, z.B.

$$\begin{array}{llll} 1) & y_i & = & \alpha x_i^\beta \exp(\varepsilon_i) \quad \Rightarrow \quad \ln y_i = \ln \alpha + \beta \ln x_i + \varepsilon_i \\ 2) & y_i & = & \alpha x_i^\beta \varepsilon_i \quad \Rightarrow \quad \ln y_i = \ln \alpha + \beta \ln x_i + \ln \varepsilon_i \\ 3) & y_i & = & \alpha x_i^\beta + \varepsilon_i \quad \Rightarrow \quad \ln y_i = \ln(\alpha x_i^\beta + \varepsilon_i) \end{array}$$

Nur die erste Gleichung kann unmittelbar mittels OLS geschätzt werden, sofern die Gauss-Markov Annahmen erfüllt sind.

Wenn die Residuen der zweiten Gleichung normalverteilt sind, d.h. $\ln(\varepsilon_i) \sim \mathcal{N}(0, \sigma^2)$, so ist der Störterm ε_i log-normalverteilt mit Mittelwert $\exp(\sigma^2/2)$ (siehe Exkurs Seite 250).

Die dritte Gleichung $\ln y_i = \ln(\alpha x_i^\beta + \varepsilon_i)$ ist schließlich nicht linear in den Parametern, da $\ln(A + B) \neq \ln A + \ln B$, und kann nicht mit OLS geschätzt werden.

8.1.2 Das Log-Lin Modell

Das log-lin Modell⁵

$$\ln y = \alpha + \beta x + \varepsilon$$

ist in Abbildung 8.2 dargestellt. Die linke Abbildung zeigt den Funktionsverlauf für einen positiven Steigungskoeffizienten β , die rechte Grafik für einen negativen Steigungskoeffizienten.

Für die Interpretation des Koeffizienten muss man sich den marginalen Effekt ansehen, der aus der (partiellen) Ableitung $d \ln y / dx = \beta$ folgt

⁴Wir verwenden hier das partielle Ableitungszeichen ∂ , weil die Funktion mehr als eine erklärende Variable enthält.

⁵Das *log-lin* Modell ist nicht zu verwechseln mit dem *log-linearen* Modell $\ln y = \alpha + \beta \ln(x) + u$.

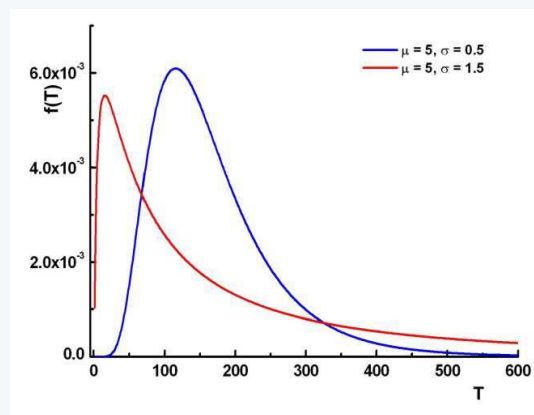
Exkurs: Die Log-Normalverteilung (Logarithmische Normalverteilung)

Eine Zufallsvariable, deren natürlicher Logarithmus normalverteilt ist, ist log-normalverteilt.

Das heißt, wenn $\ln(X) \sim \mathcal{N}(\mu, \sigma^2)$, dann ist X log-normalverteilt.

Anders herum, wenn eine Zufallsvariable Y normalverteilt ist, dann ist $X = \exp(Y)$ log-normalverteilt.

Die Log-Normalverteilung ist rechtsschief und kann nur positive Werte annehmen.



Eine log-normale Verteilung wird häufig zur Modellierung von Zufallsvariablen herangezogen, die man sich als das (multiplikative) Produkt vieler kleiner unabhängiger Faktoren vorstellen kann.

Erwartungswert und Varianz sind

$$E(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right)$$

$$\text{Var}(X) = \exp(2\mu + 2\sigma^2)[1 - \exp(-\sigma^2)]$$

$$\text{Median}(X) = \exp(\mu)$$

Für $\mu = 0$ ist der Mittelwert $\exp(\sigma^2/2)$ und die Varianz $\exp(\sigma^2)(\exp(\sigma^2) - 1)$.



$$\begin{aligned}
 \beta &= \frac{d \ln y}{dx} = \frac{d \ln y}{dy} \frac{dy}{dx} \\
 &= \frac{1}{y} \frac{dy}{dx} = \frac{\frac{dy}{y}}{dx} \\
 &\approx \frac{\Delta y / y}{\Delta x}
 \end{aligned}$$

d.h. der geschätzte Koeffizient für β gibt *näherungsweise* (\approx) die *relative* Änderung von y an, die durch eine Zunahme von x um *eine Einheit* ausgelöst wird. Wenn man β mit 100 multipliziert erhält man daher näherungsweise die prozentuelle Änderung von y , wenn x um eine Einheit zunimmt

$$100\beta \approx \frac{\frac{\Delta y}{y} \times 100}{\Delta x} = \frac{\text{prozentuelle Änderung von } y}{\text{Änderung von } x \text{ um eine Einheit}}$$

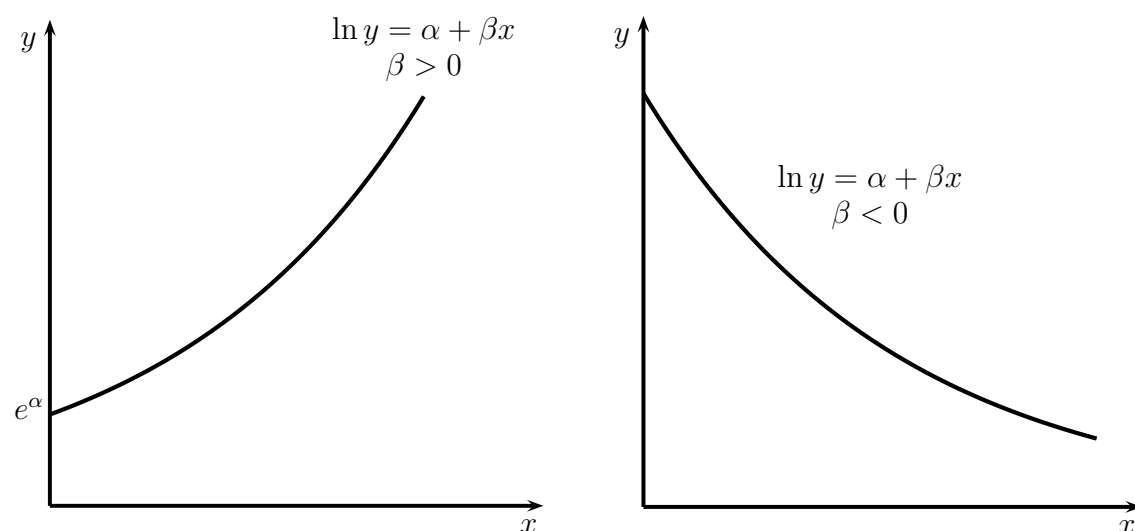


Abbildung 8.2: Log-lin Modelle: $\ln y = \alpha + \beta x$

Stetige und diskrete Änderungen

Der Grund, warum wir oben das Zeichen \approx für *näherungsweise* verwendet haben liegt in der nicht-linearen Funktionsform. Abbildung 8.3 zeigt den Funktionsverlauf. Wenn x um *eine Einheit* zunimmt (d.h. $\Delta x = 1$), ändert sich y um Δy Einheiten.

Wenn die Änderung von x sehr klein gewählt wird ($\Delta x \rightarrow 0$) nähert sich dies der Ableitung, d.h. der Steigung der Tangente. Für infinitesimal kleine Änderungen misst die Steigung der Tangente den marginalen Effekt genau, aber für diskrete Änderungen von x gibt $\frac{dy}{dx} \Delta x$ nur näherungsweise die Auswirkung auf y wieder.

Man beachte, dass wir uns in der Regel für die diskreten Differenzenquotienten $\frac{\Delta y}{\Delta x}$ interessieren, denn dieser gibt uns an, um wieviele Einheiten sich y ändert, wenn x um eine diskrete Einheit zu- oder abnimmt. Wenn die Kurve wenig gekrümmt ist und Δx nicht allzu groß ist, wird dieser Effekt durch den stetigen Differentialquotient

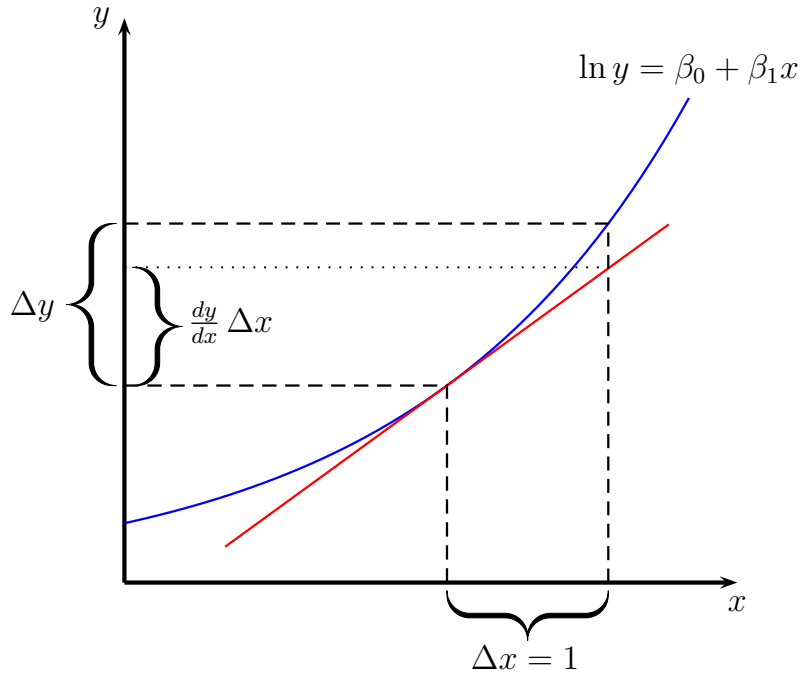


Abbildung 8.3: Auswirkung einer diskreten Änderung von x (d.h. Δx) auf y .

$\frac{dy}{dx}$ relativ gut approximiert. Wie man aus Abbildung 8.3 erkennen kann, ist diese Approximation für stark gekrümmte Kurven und große Δx weniger gut.

Glücklicherweise kann aus den Parametern eines log-lin Modells die genaue *prozentuelle Änderung* von y sehr einfach berechnet werden. Dies sei an folgendem Modell gezeigt

$$\ln y_t = b_0 + b_1 \ln x_{t1} + b_2 x_{t2}$$

(da diese Modelle häufig für Zeitreihen angewandt werden, verwenden wir als Index für die Beobachtungen t)

Dieses Modell ist ein log-log (bzw. log-lineares) Modell in x_1 und ein log-lin Modell in x_2 . Der Koeffizient b_1 ist – wie im vorhergehenden Abschnitt gezeigt – eine gewöhnliche Elastizität; wenn z.B. $b_1 = 0.5$ würde eine Zunahme von x_1 um ein Prozent eine Zunahme des Erwartungswertes von y um 0.5 Prozent bewirken.

Der Koeffizient b_2 gibt *näherungsweise* die relative Änderung von y an, wenn sich x_2 um eine (physische) Einheit ändert. Die *genaue Änderung* erhält man, wenn man x_1 konstant hält ($\Delta x_1 = 0$) und die Differenzen bildet

$$\begin{aligned} \Delta \ln y_t &:= \ln y_t - \ln y_{t-1} = b_1 0 + b_2 (x_{2t} - x_{2t-1}) \\ \ln \left(\frac{y_t}{y_{t-1}} \right) &= b_2 (\Delta x_2) \\ \exp \left[\ln \left(\frac{y_t}{y_{t-1}} \right) \right] &= \frac{y_t}{y_{t-1}} = \exp(b_2 \Delta x_2) \\ \left(\frac{y_t}{y_{t-1}} - 1 \right) &= \exp(b_2 \Delta x_2) - 1 \\ \left(\frac{y_t - y_{t-1}}{y_{t-1}} \right) \times 100 &= (\exp(b_2 \Delta x_2) - 1) \times 100 \end{aligned}$$

bzw. für $\Delta x_2 = 1$

$$\% \Delta y := \left(\frac{y_t - y_{t-1}}{y_{t-1}} \right) \times 100 = [\exp(b_2) - 1] \times 100$$

d.h. im log-lin Modell ist die prozentuelle Änderung von y , die durch eine Änderung von x_2 um eine Einheit ausgelöst wird, gleich $(\exp(b_2) - 1) \times 100$.

Wenn b_2 klein ist (die Funktion also nur wenig gekrümmt ist) macht es keinen großen Unterschied, ob man b_2 oder $(\exp(b_2) - 1) \times 100$ verwendet, aber wenn der geschätzte Koeffizient b_2 z.B. größer als 0.05 ist sollte für obige Interpretation der Wert $(\exp(b_2) - 1) \times 100$ herangezogen werden.

Die folgende Tabelle soll dies verdeutlichen (die Werte müssen mit 100 multipliziert werden um Prozentwerte zu erhalten):

| b_2 | $\exp(b_2) - 1$ |
|-------|-----------------|
| 0.01 | 0.010 |
| 0.05 | 0.051 |
| 0.10 | 0.105 |
| 0.30 | 0.350 |
| 0.40 | 0.492 |
| 0.50 | 0.649 |

Berechnung von durchschnittlichen Wachstumsraten mittels OLS Die vorhin geschilderten Möglichkeiten zur Berechnung von durchschnittlichen Wachstumsraten haben einen Nachteil, sie berücksichtigen nur die Ausgangs- und Endperiode, ohne die Entwicklung zwischen diesen beiden Perioden zu berücksichtigen. Es zeigt sich, dass mit Hilfe einer einfachen log-lin Regression auf den Trend⁶ eine durchschnittliche Wachstumsrate berechnet werden kann, die alle Beobachtungen berücksichtigt.

Wenn i die diskrete Wachstumsrate einer Variable y ist gilt

$$y_t = y_0(1 + i)^t \Rightarrow \ln y_t = \ln y_0 + \ln(1 + i) \times t$$

Dieser Zusammenhang sollte für jede Periode gelten. Um die diskrete Wachstumsrate i zu schätzen können wir deshalb t durch eine Trendvariable $\text{Trend} = 1, 2, 3, \dots, T$ ersetzen. Wenn wir mit y_0 den Wert von y in der Ausgangsperiode bezeichnen ist

$$\ln y_t = \underbrace{\ln y_0}_{b_0} + \underbrace{\ln(1 + i)}_{b_1} \times \text{Trend}$$

Wir können also einfach

$$\ln y_t = b_0 + b_1 \text{Trend} + e_t$$

⁶Eine Trendvariable nimmt mit jeder Beobachtung um eine Einheit zu, z.B. $\text{Trend} = 0, 1, 2, 3, \dots, T$. In EViews kann eine Trendvariable mit der Funktion `@trend` erzeugt werden, z.B. `series Trend = @trend`.

schätzen und aus $b_1 = \ln(1 + i)$ die durchschnittliche diskrete Wachstumsrate i berechnen, denn aus

$$b_1 = \ln(1 + i) \quad \text{folgt} \quad i = \exp(b_1) - 1$$

Die prozentuelle durchschnittliche diskrete Wachstumsrate ist deshalb

$$i\% := i \times 100 = [\exp(b_1) - 1] \times 100$$

Wenn b_1 sehr klein ist (z.B. kleiner als 0.05) wird sich b_1 nur geringfügig von i unterscheiden, bei größeren Werten sollte aber die Korrektur $\exp(b_1) - 1$ angewandt werden.

8.1.3 Das Lin-Log Modell

Die grafische Abbildung des lin-log Modells

$$y_i = \beta_0 + \beta_1 \ln x_i$$

findet sich in Abbildung 8.4.

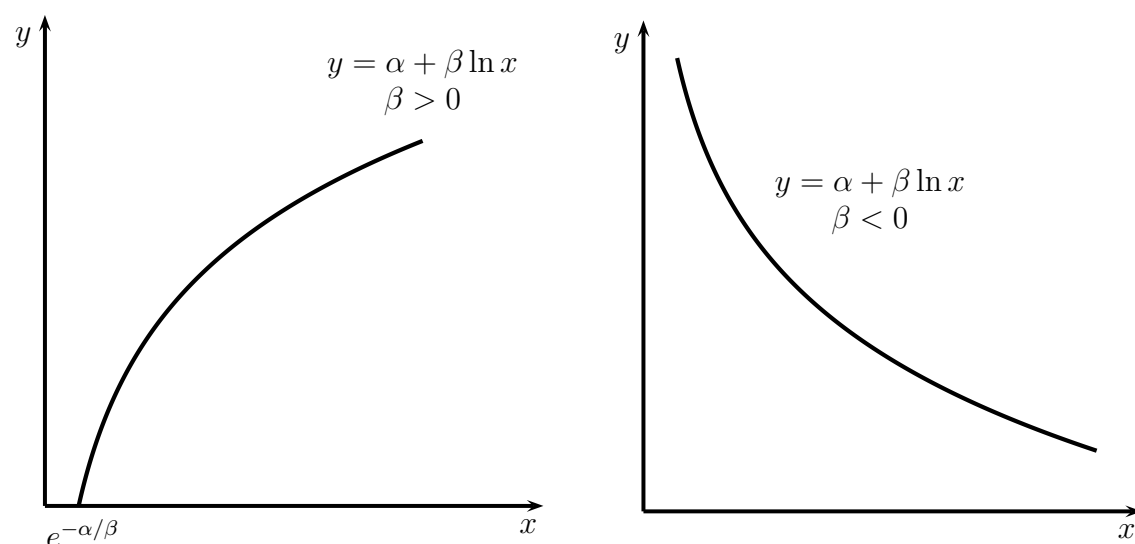


Abbildung 8.4: Lin-log Modell: $y = \alpha + \beta \ln x$

Die Interpretation des Koeffizienten erhält man wieder, indem man die (partielle) Ableitung bildet. Aus der Kettenregel folgt $dy/dx = dy/d(\ln x) \cdot d(\ln x)/dx$ und somit

$$\frac{dy}{dx} = b_1 \frac{1}{x}$$

bzw.

$$b_1 = \frac{dy}{dx/x} = \frac{\text{Änderung in } y}{\text{relative Änderung in } x}$$

Zur Interpretation gilt ähnliches wie im vorhergehenden Fall, d.h. eine Änderung von x um 1% führt zu einer Änderung von y um $0.01 \times b_1$ Einheiten, da

$$b_1 \approx \frac{\Delta y}{\frac{\Delta x}{x}}, \quad \Rightarrow \quad 0.01 b_1 = \frac{\Delta y}{\frac{\Delta x}{x} \times 100} = \frac{\text{Änderung von } y}{\text{prozentuelle Änderung von } x}$$

8.1.4 Wann logarithmieren?

Durch Logarithmierung sind die Steigungskoeffizienten unabhängig von der gewählten Maßeinheit und deshalb einfach zu interpretieren. Darüber hinaus bietet die Logarithmierung Vorteile, wenn die bedingte Verteilung von y schief ist oder die Fehlervarianzen von ε nicht konstant sind (Heteroskedastizität), denn häufig erfüllt die Verteilung von $\ln(y)$ die Gauss-Markov Annahmen besser als y . Da durch die Logarithmierung die Verteilung ‘gestaucht’ wird, sind logarithmische Gleichungen häufig auch weniger anfällig gegen Ausreißer (*outliers*).

Sehr häufig werden v.a. Variablen logarithmiert, die im Niveau (Level) gemessen werden, und von denen man annehmen kann, dass sie im Zeitablauf wachsen (z.B. Geldbeträge, Bevölkerung, ...). Variablen, die eine Zeitdimension haben, werden seltener logarithmiert.

Sehr häufig werden makroökonomische Variablen wie z.B. das BIP logarithmiert, von denen man plausibel annehmen kann, dass sie langfristig ungefähr exponentiell wachsen. Wenn diese Variablen ungefähr mit einer bestimmten durchschnittlichen Wachstumsrate Jahr für Jahr wachsen, dann nimmt der Logarithmus dieser Variablen linear zu.

Ein weiterer Grund besteht darin, dass die Standardabweichung vieler ökonomischer Variablen ungefähr proportional zum Niveau dieser Variablen ist, deshalb ist die Standardabweichung der logarithmierten Variablen näherungsweise konstant. In anderen Worten, das Logarithmieren ökonomischer Variablen führt häufig zu einer Art ‘Stabilisierung’ der Standardabweichung.

Achtung: Es ist darauf zu achten, dass keine Variablen logarithmiert werden, die negative Werte annehmen können, da der Logarithmus einer negativen Zahl nicht definiert ist. Wird eine Variable mit negativen Werten logarithmiert, so werden die negativen Werte in den meisten Programmen ohne Benutzerwarnung mit dem Code für ‘nicht definiert’ (z.B. `na` für *not available*) überschrieben, gehen also verloren. Besondere Vorsicht ist bei Änderungsraten oder Wachstumsraten geboten, die – wenn überhaupt – nur in ganz speziell begründeten Fällen logarithmiert werden sollten.

Das entscheidende Argument für die Wahl der Funktionsform sollte sein, welches Modell mit der Theorie konsistent ist und die Daten am besten abbildet. Wir werden in dem späteren Kapitel *Spezifikation* einen Test von MacKinnon, White & Davidson kennen lernen, der diese Entscheidung erleichtern sollte. Oft reicht allerdings schon ein Blick auf das Histogramm der geschätzten Residuen um zu erkennen, welche Funktionsform am ehesten angebracht ist.

Beispiel: Die folgende Regression erklärt den Gehalt von CEO’s (*Salary*, gemessen in 1000 \$) in Abhängigkeit vom Verkaufsvolumen der Firmen (*Sales*, gemessen in Mio. \$) und dem Return on Equity (*ROE*, alle Daten aus Wooldridge 2000).

ohne Log:

$$\text{SALARY} = 830.631 + 0.016 \text{ SALES} + 19.631 \text{ ROE}$$

$$(3.71) \quad (1.842) \quad (1.772)$$

$$R^2 = 0.029, \quad s = 1358.728, \quad N = 209$$

(*t*-Statistiken in Klammern)

mit Log:

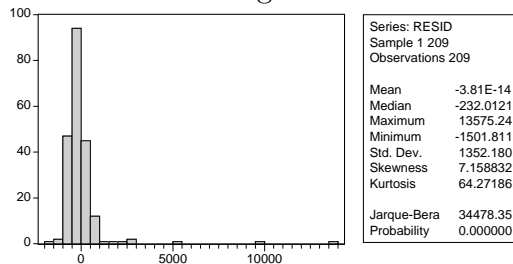
$$\text{@LOG(SALARY)} = 4.362 + 0.275 \text{@LOG(SALES)} + 0.018 \text{ROE}$$

(14.843) (8.272) (4.519)

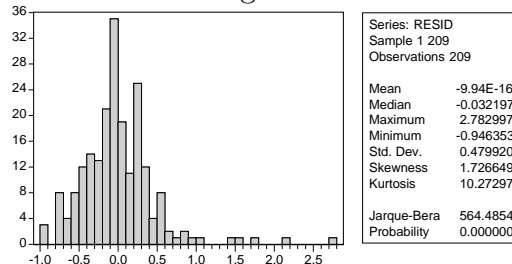
$$R^2 = 0.282, \quad s = 0.482, \quad N = 209$$

(t-Statistiken in Klammern)

Residuen *ohne* Log:



Residuen *mit* Log:



Zusammenfassung: Interpretation der Koeffizienten in Log-Modellen

Logarithmen können verwendet werden für die Transformation der abhängigen Variable y , der erklärenden Variable(n) x , oder von beiden, solange die Variablen positiv sind (der Logarithmus einer negativen Variable ist nicht definiert!)

| Spezifikation | Interpretation |
|--|--|
| I. $y_i = \beta_0 + \beta_1 \ln(x_i) + \varepsilon_i$ | Eine Änderung von x um 1% geht mit einer Änderung von y um $0.01 \times \beta_1$ Einheiten einher. |
| II. $\ln(y_i) = \beta_0 + \beta_1 x_i + \varepsilon_i$ | Eine Änderung von x um eine Einheit (d.h. $\Delta x = 1$) geht ungefähr mit einer Änderung von y um $100 \times \beta_1\%$ einher, oder genauer, mit einer Änderung von $(\exp(\beta_1) - 1) \times 100$ Prozent. |
| III. $\ln(y_i) = \beta_0 + \beta_1 \ln(x_i) + \varepsilon_i$ | Eine Änderung von x um 1% geht einher mit einer Änderung von y um $\beta_1\%$, d.h. β_1 kann als Elastizität interpretiert werden. |



8.2 Reziproke Transformationen

Eine andere Funktionsform, die z.B. häufig für die Schätzung von Phillips-Kurven herangezogen wird, sind reziproke Transformationen (man verwendet einfach den Kehrwert der Variablen, siehe Abbildung 8.5)

$$y = b_0 + b_1 \frac{1}{x} + e$$

Um zum Beispiel die Verkaufsumsätze in Abhängigkeit von den Werbeausgaben zu erklären wird manchmal auch ein log-inverses Modell

$$\ln(y) = b_0 - \frac{b_1}{x} + e$$

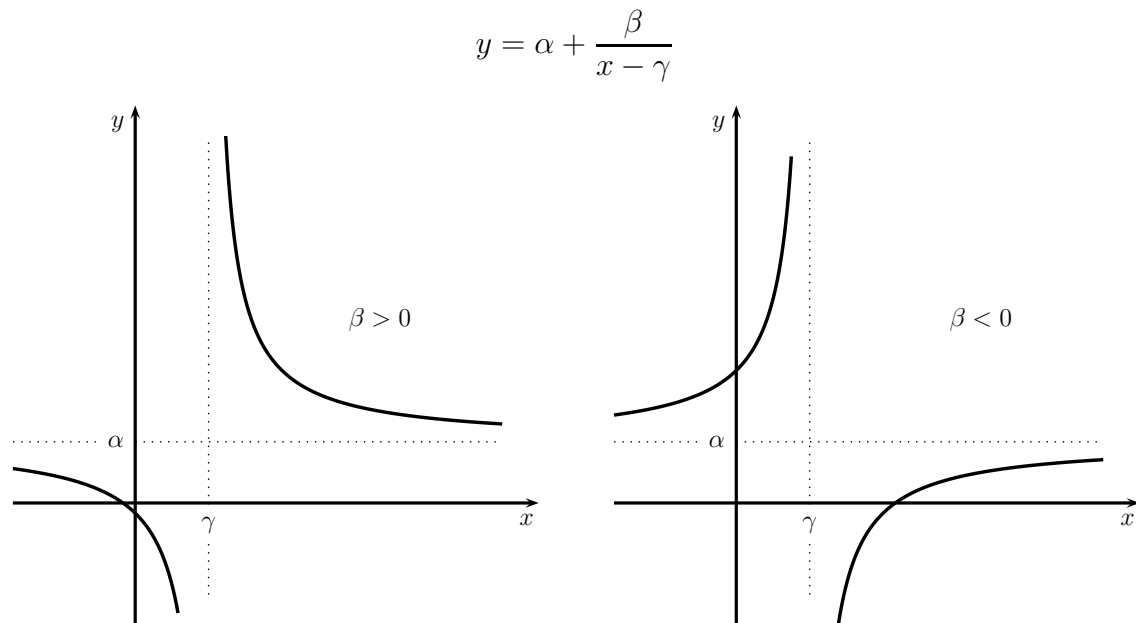


Abbildung 8.5: Reziproke Transformationen

geschätzt. Wie man aus Abbildung 8.6 erkennen kann erlaubt diese S-förmige Funktionsform zuerst zunehmende Grenzerträge von Werbeausgaben, und ab dem Wendepunkt bei $b_1/2$ abnehmende Grenzerträge, und nähert sich asymptotisch einem horizontalen Verlauf an.

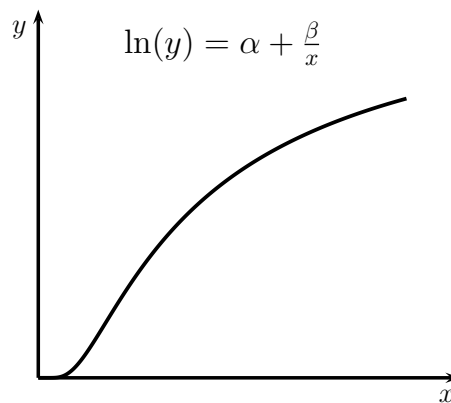


Abbildung 8.6: Log-Reziproke Transformationen

8.3 Quadratische Modelle und höhere Polynome

$$y_i = b_0 + b_1 x_i + b_2 x_i^2 + e_i$$

mit

$$\frac{\partial y}{\partial x} = b_1 + 2b_2 x$$

Der marginale Effekt ist also nicht konstant, sondern hängt von x ab.

Achtung: Die einzelnen Koeffizienten b_1 und b_2 haben keine direkte Interpretation, da sich x und x^2 nicht unabhängig voneinander ändern können!

Für quadratische Funktionen kann man immer den Wert von x berechnen, bei denen y ein Maximum oder Minimum erreicht.

Um das Maximum oder Minimum⁷ x^* zu bestimmen muss wie üblich die Ableitung nach x gleich Null gesetzt und nach x gelöst:

$$x^* = \frac{-b_1}{2b_2}$$

Dies soll anhand einer einfachen Mincer-Lohnleichung demonstriert werden:

$$\text{Wage} = b_0 + b_1 \text{Education} + b_2 \text{Experience} + b_3 \text{Experience}^2 + e$$

wobei der Lohn (*Wage*) der Stundenlohn in Dollar ist, *Education* die Ausbildungsdauer in Jahren und *Experience* die Jahre an Berufserfahrung sind. Die empirische Schätzung ist

$$\text{WAGE} = -3.965 + 0.595 \text{ EDUC} + 0.268 \text{ EXPER} - 0.005 \text{ EXPER}^2$$

(0.752) (0.053) (0.037) (0.001)

$$R^2 = 0.269, \quad s = 3.166, \quad N = 526$$

(Standardfehler in Klammern)

Offensichtlich erhöht ein zusätzliches Ausbildungsjahr den Erwartungswert des Stundenlohns um beinahe 60 Cent.

Die Auswirkung der Berufserfahrung auf den Erwartungswert des Einkommens hängt bei dieser Spezifikation vom Ausmaß (*level*) der Berufserfahrung ab. Zum Beispiel würden wir nach dieser Schätzung für jemanden mit 10 Jahren Berufserfahrung erwarten, dass im elften Jahr der Lohn *ceteris paribus* um 0.168 Dollar zunimmt, denn

$$\frac{\partial \text{WAGE}}{\partial \text{EXPER}} = 0.268 + 2 \times (-0.005) \times 10 = 0.168$$

Für jemanden mit 40 Berufsjahren ($\text{EXPER} = 40$) würden wir hingegen für ein weiteres Jahr Berufserfahrung einen Einkommensrückgang um ca. 13 Cent erwarten (warum?).

Nach dieser Schätzung würde der Lohn nach ca. 27 Berufsjahren ein Maximum erreichen, denn aus $\partial \text{Wage} / \partial \text{Exper} = 0.268 - 2 \times 0.005 \text{ Exper} \stackrel{!}{=} 0$ folgt

$$\text{EXPER}^{\max} = \frac{-0.268}{2 \times (-0.005)} = 26.8$$

⁷Für ein Minimum muss das Vorzeichen der 2. Ableitung positiv sein, für ein Maximum negativ.

8.4 Interaktions-Modelle

Als erklärende Variablen können auch *Produkte* einzelner Variablen verwendet werden, z.B. im folgenden Modell das Produkt von x_1 und x_2

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i1} x_{i2} + \varepsilon_i$$

In diesen Modellen werden x_{i1} und x_{i2} Hauptterme (*'main terms'*) genannt und das Produkt $x_{i1} x_{i2}$ wird als Interaktionsterm (*'interaction term'*) bezeichnet.

Wenn in einem Modell Interaktionsterme berücksichtigt werden sollten auf jeden Fall auch die Hauptterme berücksichtigt werden, da sonst die Gefahr eine Fehlspezifikation aufgrund fehlender relevanter Variablen extrem groß ist (der Interaktionsterm $x_1 x_2$ ist fast immer mit x_1 und x_2 korreliert) (Ozer-Balli and Sorensen, 2010).

Der marginale Effekte von x_1 hängt vom Wert von x_2 ab.

$$\frac{\partial y}{\partial x_1} = \beta_1 + \beta_3 x_2$$

d.h. die ceteris paribus Auswirkung einer Änderung von x_1 auf y hängt auch vom Wert von x_2 ab. Wenn getestet werden soll, ob x_1 einen Effekt auf y hat, darf deshalb nicht der einfache t -Test für den Koeffizienten von x_1 herangezogen werden, sondern es muss z.B. mit einem F -Test die gemeinsame Nullhypothese $\beta_1 = 0$ und $\beta_3 = 0$ getestet werden. Ebenso ändert sich die Interpretation der Koeffizienten. Wenn im Modell Interaktionseffekte berücksichtigt werden, dürfen die Koeffizienten der Hauptterme nicht mehr unabhängig von den Interaktionseffekten interpretiert werden!⁸

Analog gilt für eine Änderung von x_2 , wenn x_1 konstant gehalten wird,

$$\frac{\partial y}{\partial x_2} = \beta_2 + \beta_3 x_1$$

Der Koeffizient des Interaktionsterms ist einfach die zweite Ableitung

$$\beta_3 = \frac{\frac{\partial y}{\partial x_1}}{\partial x_2} = \frac{\frac{\partial y}{\partial x_2}}{\partial x_1} = \frac{\partial^2 y}{\partial x_1 \partial x_2}$$

(das zweite '=' Zeichen folgt aus Young's Theorem).

Da zweite Ableitungen nicht besonders anschaulich sind nähern wir uns der Interpretation auf anderem Weg.

Zur Interpretation Wir erinnern uns, dass für jede nichtlineare Gleichung $y = f(x_1, x_2)$ die Änderung von y , d.h. Δy , *näherungsweise* gleich ist

$$\Delta y \approx f(x_1 + \Delta x_1, x_2 + \Delta x_2) - f(x_1, x_2)$$

In unserem Fall mit Interaktionsterm ist die Funktion

$$y = f(x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon$$

⁸Der Koeffizient β_1 ist nur dann der marginale Effekt von x_1 , wenn $x_2 = 0$!

und die Änderung Δy deshalb

$$\begin{aligned}\Delta y &= \beta_0 + \beta_1(x_1 + \Delta x_1) + \beta_2(x_2 + \Delta x_2) + \beta_3(x_1 + \Delta x_1)(x_2 + \Delta x_2) + \varepsilon - \\ &\quad - \beta_0 - \beta_1 x_1 - \beta_2 x_2 - \beta_3 x_1 x_2 - \varepsilon \\ &= \beta_1 \Delta x_1 + \beta_2 \Delta x_2 + \beta_3 [x_2 \Delta x_1 + x_1 \Delta x_2 + \Delta x_1 \Delta x_2] \\ &= \underbrace{(\beta_1 + \beta_3 x_2) \Delta x_1}_{\left. \frac{\partial y}{\partial x_1} \right|_{\Delta x_2=0}} + \underbrace{(\beta_2 + \beta_3 x_1) \Delta x_2}_{\left. \frac{\partial y}{\partial x_2} \right|_{\Delta x_1=0}} + \beta_3 \Delta x_1 \Delta x_2\end{aligned}$$

Die ersten zwei Terme erfassen die ceteris paribus Effekte einer Änderung von x_1 und x_2 , und der dritte Term erfasst den eigentlichen nicht-linearen Interaktionseffekt.

Das heißt, die Veränderung von y setzt sich zusammen aus dem Effekt einer Änderung von x_1 wenn x_2 konstant gehalten wird ($\Delta x_2 = 0$), bzw. $\partial y / \partial x_1 = \beta_1 + \beta_3 x_2$, aus dem ceteris paribus Effekt einer Änderung von x_2 , $\partial y / \partial x_2 = \beta_2 + \beta_3 x_1$, sowie aus dem *zusätzlichen* Effekt wenn sowohl x_1 als auch x_2 geändert werden.

Der Koeffizient β_3 des Interaktionsterms $x_1 \times x_2$ misst also den *zusätzlichen* Effekt einer simultanen Änderung von x_1 und x_2 , der über den Effekt einer Änderung von x_1 alleine und x_2 alleine hinausgeht! In gewissem Sinne deutet ein signifikanter Interaktionseffekt also darauf hin, dass der Gesamteffekt stärker (oder schwächer) ist als die Summe der einzelnen Effekte.

Zum Beispiel:

$$\text{LOG(WAGE)} = 0.455 + 0.08 \text{ EDUC} + 0.001 \text{ EDUC*EXPER}$$

(4.726) (10.947) (6.272)

$$R^2 = 0.243, \quad s = 0.463, \quad N = 526$$

(t-Statistiken in Klammern)

also

$$\frac{\partial \ln(\text{WAGE})}{\partial (\text{EDUC})} = 0.08 + 0.001 \times \text{EXPER}$$

Nach dieser Schätzung steigt der Wert der Bildung (EDUC) mit der Berufserfahrung (EXPER). Bei einer Berufserfahrung von 5 Jahren ist

$$\frac{\partial \ln(\text{WAGE})}{\partial (\text{EDUC})} = 0.08 + 0.001 \times 5 = 0.085$$

das heißt, für jemanden mit 5 Jahren Berufserfahrung würde ein zusätzliches Ausbildungsjahr den Lohn ungefähr um 8.5% erhöhen, bzw. genauer um $[\exp(0.085) - 1] \times 100 = 8.87\%$.

Für jemanden mit 20 Jahren Berufserfahrung gilt

$$\frac{\partial \ln(\text{WAGE})}{\partial (\text{EDUC})} = 0.08 + 0.001 \times 20 = 0.1$$

d.h. ein zusätzliches Ausbildungsjahr erhöht den Erwartungswert des Einkommens nach 20 Jahren Berufserfahrung um ungefähr 10% (bzw. genauer um $[\exp(0.1) - 1] \times 100 = 10.52\%$).

Der hochsignifikante Koeffizient des Interaktionsterms EDUC*EXPER deutet darauf hin, dass der ‘Wert’ der Bildung mit der Berufserfahrung stärker steigt (bzw., dass der ‘Wert’ der Berufserfahrung mit der Bildung steigt), als dies durch die individuellen Beiträge von Bildung und Berufserfahrung alleine erklärbar ist.

In Interaktionsmodellen sind die interagierten Variablen häufig hoch korreliert, also multikollinear. Das muss aber nicht unbedingt ein Problem sein, da diese Funktion selbst nichtlinear ist und deshalb die Regressionsfläche trotzdem gut durch die Beobachtungspunkte ‘abgestützt’ sein kann.

Übersicht:

| Modell | Gleichung | Steigung ($= \frac{dy}{dx}$) | Elastizität ($= \frac{dy}{dx} \frac{x}{y}$) |
|------------|--------------------------------|--------------------------------|---|
| Linear | $y = \alpha + \beta x$ | β | $\beta(x/y)$ |
| Log-linear | $\ln y = \alpha + \beta \ln x$ | $\beta(y/x)$ | β |
| Log-lin | $\ln y = \alpha + \beta x$ | $\beta(y)$ | $\beta(x)$ |
| Lin-log | $y = \alpha + \beta \ln x$ | $\beta(1/x)$ | $\beta(1/y)$ |
| Reziprok | $y = \alpha + \beta(1/x)$ | $-\beta(1/x^2)$ | $-\beta(1/xy)$ |

Achtung:

- Mit polynomischen Modellen (z.B. quadratischen oder kubischen Modellen) kann man zwar manchmal einen sehr guten Fit in der Stichprobe erreichen, aber für Prognosen sind sie meistens ziemlich unbrauchbar, da die Funktionsform ‘*out of sample*’ häufig extreme Verläufe erzwingt.
- Das Bestimmtheitsmaß R^2 darf nur für den Vergleich von Modellen verwendet werden, in denen die abhängige Variable y nicht transformiert wurde *und* wenn beide Modelle die gleiche Anzahl erklärender Variablen haben.

Wenn die abhängige Variable y nicht transformiert wurde und beide Modelle eine *unterschiedliche* Anzahl erklärender Variablen haben kann das korrigierte Bestimmtheitsmaß \bar{R}^2 verwendet werden.

- Grundsätzlich sollte weder das normale R^2 noch das korrigierte Bestimmtheitsmaß \bar{R}^2 für die Einschätzung der Qualität einer Schätzung überinterpretiert werden, da sie nur die Anpassung in der Stichprobe beschreiben. Viel wichtiger ist z.B., ob die geschätzten Koeffizienten die theoretischen Erwartungen erfüllen und signifikant sind. Im Kapitel über *Spezifikation* werden wir einige weitere Kriterien und Tests kennen lernen.

Literaturverzeichnis

- Duan, N. (1983), ‘Smearing estimate: A nonparametric retransformation method’, *Journal of the American Statistical Association* **78**(383), 605–610.
URL: <http://www.jstor.org/stable/2288126>

Ozer-Balli, H. and Sorensen, B. E. (2010), ‘Interaction Effects in Econometrics’, *CEPR Discussion Paper no. 7929*.

Wooldridge, J. (2005), *Introductory Econometrics: A Modern Approach*, 3 edn, South-Western College Pub.

8.A Appendix

8.A.1 Berechnung von durchschnittlichen Wachstumsraten

Bei der Berechnung durchschnittlicher Wachstumsraten werden häufig Fehler gemacht, obwohl dies gerade in den Wirtschaftswissenschaften ziemlich wichtig ist, deshalb werden wir darauf etwas näher eingehen.

Diskrete Wachstumsraten (i)

Unter der diskreten Wachstumsrate verstehen wir die relative Änderung einer Größe zwischen zwei Perioden

$$i = \frac{y_t - y_{t-1}}{y_{t-1}} = \frac{\Delta y_t}{y_{t-1}} = \frac{y_t}{y_{t-1}} - 1$$

Der Quotient y_t/y_{t-1} wird auch als Wachstumsfaktor bezeichnet.

Wenn eine Variable y mit einer konstanten diskreten Wachstumsrate i wächst nimmt sie in jeder Periode um iy_{t-1} Einheiten zu. Für $t = 1, 2, \dots, T$

$$\begin{aligned} y_1 &= y_0 + iy_0 = y_0(1 + i) \\ y_2 &= y_1(1 + i) = y_0(1 + i)(1 + i) = y_0(1 + i)^2 \\ &\vdots \\ y_T &= y_0(1 + i)^T \end{aligned}$$

Sollte die durchschnittliche Wachstumsrate zwischen den Perioden 0 und T berechnet werden, so darf dazu *nicht* das arithmetische Mittel herangezogen werden!

Man kann sich einfach fragen, welche Wachstumsrate i führt vom Wert y_0 zu y_T . Dazu logarithmieren wir $y_T = y_0(1 + i)^T$ und lösen nach i

$$\begin{aligned} \ln y_T &= \ln y_0 + T \ln(1 + i) \\ \frac{1}{T} (\ln y_T - \ln y_0) &= \ln(1 + i) \\ \exp\left(\frac{1}{T} (\Delta \ln y)\right) &= 1 + i \end{aligned}$$

$$i_{0,T} = \exp\left(\frac{1}{T} \ln \left[\frac{y_T}{y_0}\right]\right) - 1$$

T bezeichnet dabei die Anzahl der Perioden; sollte z.B. die durchschnittliche Wachstumsrate des BIP von 2005 – 2010 berechnet werden, und wird darunter die Periode vom 1.1.2005 – 31.12.2010 verstanden, so ist $T = 6$.

Stetiges Wachstum (r)

Der enge Zusammenhang zwischen Wachstumsraten und der Exponentialfunktion (bzw. dem Logarithmus) wird sofort klar, wenn man mehrere Verzinsungen pro Periode zulässt, und die Anzahl der Verzinsungen pro Periode gegen Unendlich gehen lässt.

Wenn die Verzinsung m Mal pro Periode erfolgt

$$\begin{aligned} y_t &= y_0 \left(1 + \frac{r}{m}\right)^{mt} = y_0 \left[\left(1 + \frac{r}{m}\right)^{m/r}\right]^{rt} \\ &= y_0 \left[\left(1 + \frac{1}{w}\right)^w\right]^{rt} \quad \text{mit } w := \frac{m}{r} \end{aligned}$$

Wir erinnern uns, dass die Eulersche Zahl als Grenzwert einer Folge dargestellt werden kann

$$\lim_{m \rightarrow \infty} \left(1 + \frac{1}{w}\right)^w = e \approx 2.7182819$$

Also

$$\boxed{y_t = y_0 e^{rt} := y_0 \exp(rt)}$$

Man beachte, dass in $y_t = y_0(1 + r/m)^{mt}$ die Zeit noch eine diskrete Variable ist, erst durch die Grenzwertbildung wird die Zeit zu einer stetigen Variable. Außerdem wird hier r als konstant angenommen, aber natürlich müssen tatsächliche Wachstumsraten nicht immer konstant sein.

Mit einer stetigen Wachstumsrate kann die Veränderung über die Zeit einfach als Ableitung nach der Zeit berechnet werden

$$\frac{dy_t}{dt} = \frac{d(y_0 e^{rt})}{dt} = r y_0 e^{rt}$$

Die relative Änderung ist deshalb

$$\frac{\frac{dy_t}{dt}}{y_t} = \frac{r y_0 e^{rt}}{y_0 e^{rt}} = r$$

Für sehr kurze Zeitperioden konvergiert die diskrete Wachstumsrate i gegen die stetige Wachstumsrate r

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} i &= \lim_{\Delta t \rightarrow 0} \left\{ \frac{y_{t+\Delta t} - y_t}{\Delta t} \cdot \frac{1}{y_t} \right\} \\ &= \frac{dy}{dt} \frac{1}{y} = r \end{aligned}$$

Wenn eine Variable y exponentiell mit konstanter Wachstumsrate r wächst gilt $y_t = y_0 e^{rt}$, bzw. $\ln y_t = \ln y_0 + rt$. Daraus folgt durch Umschreiben

$$r = \frac{\ln y_t - \ln y_0}{t} = \frac{\Delta \ln y}{t}$$

Daraus folgt, dass die stetige Wachstumsrate zwischen zwei aneinandergrenzenden Perioden, d.h. wenn $\Delta t = 1$, einfach die logarithmische Differenz ist $r_{t,t+1} = \Delta \ln y_t$. Ebenso ist aus $\ln y_t = \ln y_0 + rt$ auch einfach ersichtlich, dass

$$\frac{d \ln y_t}{dt} = r$$

wenn y_0 konstant ist.

Diese Zusammenhänge werden in der empirischen Wirtschaftsforschung ausgiebig genutzt, auch weil bei 'kleinen' Wachstumsraten (z.B. $r < 0.05$) die stetigen Wachstumsraten eine gute Annäherung an die diskreten Wachstumsraten darstellen.

Umrechnen zwischen stetigen und diskreten Wachstumsraten

Da zu jedem Zeitpunkt für eine beliebige diskrete Wachstumsrate i genau eine stetige Wachstumsrate r existiert, die zum gleichen Betrag y_t führt, kann einfach zwischen diskreten und stetigen Wachstumsraten umgerechnet werden

$$y_0(1+i)^t = y_0 \exp(rt)$$

folgt $\ln y_0 + t \ln(1+i) = \ln y_0 + rt$ bzw.

$$\boxed{r = \ln(1+i) \quad \text{bzw.} \quad i = \exp(r) - 1}$$

Wir haben vorhin gezeigt, dass die stetige Wachstumsrate zwischen zwei Perioden als logarithmische Differenz berechnet werden kann $r = \Delta \ln y / t$. Wenn man diese in eine diskrete Wachstumsrate umrechnet erhält man wieder $i = \exp(\Delta \ln y / t) - 1$. Die prozentuellen Wachstumsraten erhält man wie üblich, indem man diese Wachstumsraten mit 100 multipliziert.

Beispiel: Eine Regression des österreichischen BIPs (verkettete Volumenindizes, 2005 = 100, Statistik Austria⁹) auf den Trend liefert folgendes Ergebnis

$$\begin{array}{rcl} \text{@LOG(BIP)} & = & 3.949333 \quad + \quad 0.022791 \text{ @TREND} \\ & & (0.005735) \quad \quad (0.000299) \end{array}$$

$$\begin{array}{l} R^2 = 0.994532, \quad s = 0.017091, \quad F\text{-Stat} = 5819.7113, \quad DW = 0.998612, \\ N = 34 \text{ (1976 - 2009, Standardfehler in Klammern)} \end{array}$$

Das folgende kleine EViews Programm berechnet die diversen Wachstumsraten für die Periode 1976 – 2009 (34 Perioden).

```
wfcreate(wf=BIPAUT) a 1976 2009
read(t=xls,c7) VGR.xls BIP
table(2,3) WR
WR(1,1) = "durchschnittl. stetige Wachstumsrate"
```

⁹http://www.statistik.at/web_de/static/volkswirtschaftliche_gesamtrechnung_hauptgroessen_019505.xls, 09.05.2010

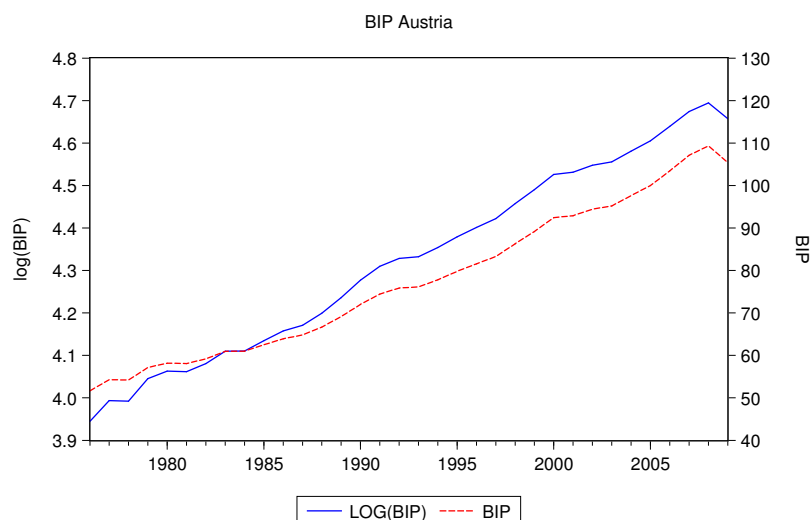


Abbildung 8.7: Die Entwicklung des österreichischen BIPs (1976 – 2009); linke Skala logarithmisch, rechte Skala linear.

```

WR(1,2) = 1/34*( log(@last(BIP)) - @log(@first(BIP)) )*100
WR(2,1) = "durchschnittl. diskrete Wachstumsrate"
WR(2,2) = (@exp(1/34*( log(@last(BIP)) - @log(@first(BIP)) )) - 1)*100
WR(3,1) = "durchschnittl. diskrete Wachstumsrate, OLS"
equation eqlog.ls @log(BIP) c @trend
WR(3,2) = (@exp(eqlog.@coefs(2)) - 1)*100
show WR

```

| | | |
|--|-----------------------------------|--------|
| durchschnittl. stetige Wachstumsrate (%) | $r = 1/t(\ln y_t - \ln y_0)100 =$ | 2.0961 |
| durchschnittl. diskrete Wachstumsrate (%) | $i = (\exp(r) - 1)100 =$ | 2.1182 |
| durchschnittl. diskrete Wachstumsr., OLS (%) | $i^* = (\exp(b_1) - 1)100 =$ | 2.3053 |

Wie man auch aus Abbildung 8.7 ersehen kann ist aufgrund des Einbruchs 2009 die durchschnittliche Wachstumsrate, die nur Ausgangs- und Endperiode berücksichtigt, kleiner als die mittels OLS berechnete durchschnittliche Wachstumsrate, die alle Perioden berücksichtigt.

8.A.2 Prognose mit einer logarithmierten abhängigen Variable

Spezielle Probleme mit der Prognose ergeben sich, wenn ein Modell mit einer logarithmierten abhängigen Variable ($\ln y$) geschätzt wird, z.B.

$$\ln y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

Falls die Gauss-Markov Annahmen erfüllt sind kann $\ln y$ wie üblich prognostiziert werden, aber in den seltensten Fällen sind wir an einer Prognose des Logarithmus von y interessiert, fast immer interessieren wir uns für die Prognose von y selbst, d.h. von $\hat{y} = E(y_t|x_t)$.

Die vordergründige Lösung, einfach $\ln y$ zu prognostizieren und die Exponentialfunktion davon zu nehmen, funktioniert leider nicht, da $\exp[E(\ln y)] \neq E(y)$ (ähnlich wie z.B. $\sqrt{E(y^2)} \neq E(y)$; erinnern Sie sich, dass der Erwartungswert als eine gewichtete Summe verstanden werden kann).

Es gilt zwar

$$y_t = \exp(\beta_0 + \beta_1 x_t + \varepsilon_t) = \exp(\beta_0 + \beta_1 x_t) \exp(\varepsilon_t)$$

und deshalb

$$E(y_t | x_t) = \exp(\beta_0 + \beta_1 x_t) E[\exp(\varepsilon_t)]$$

aber aus $E(\varepsilon_t) = 0$ folgt im allgemeinen nicht $E[\exp(\varepsilon_t)] = 1$, deshalb würde $\tilde{y} = \exp(b_0 + b_1 x_t)$ verzerrte und nicht konsistente Schätzungen für \hat{y} liefern!

Für den Fall normalverteilter Störterme, oder genauer $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$, kann man zeigen, dass $E[\exp(\varepsilon_t)] = \exp(\sigma^2/2) \neq 1$ (vgl. Exkurs Log-Normalverteilung, Seite 250).

Für diesen Fall mit normalverteilten Störtermen erhält man zwar keine erwartungstreue, aber immerhin konsistente Prognose

$$\hat{y} = \exp[b_0 + b_1 \text{Trend} + s^2/2] = \exp[\widehat{\ln y}] \exp(0.5s^2)$$

wobei s^2 die Schätzung aus der Stichprobe für σ^2 ist ($s^2 = \mathbf{e}'\mathbf{e}/(N - K)$).

Man beachte, dass $\exp(0.5s^2) > 1$, da $s^2 > 0$, deshalb würde $\exp[\widehat{\ln y}]$ das wahre y systematisch unterschätzen. Diese Korrektur liefert häufig gute Ergebnisse, beruht aber auf der Normalverteilungsannahme.

Für die schwächere Annahme unabhängig und identisch verteilter Störterme $\varepsilon_t \sim iid(0, \sigma^2)$ hat Duan (1983) gezeigt, dass der Mittelwert der Residuen $N^{-1} \sum_j \exp(e_j)$ ein konsistenter Schätzer für $E[\exp(\varepsilon_t)]$ ist.

Eine alternative Möglichkeit wird in Wooldridge (2005, 218f) vorgeschlagen.

Angenommen, wir möchten wieder das Modell $\ln y = \beta_0 + \beta_1 x + u$ für Prognosen nützen. Dann ist

$$E(y|x) = \exp(\beta_0 + \beta_1 x) \alpha$$

mit $\alpha = \exp(u)$. Eine konsistente Schätzung für α kann berechnet werden, indem

1. aus $\ln y = \beta_0 + \beta_1 x + u$ die gefitteten Werte $\widehat{\ln y}$ berechnet werden,
2. eine neue Datenreihe $z = \exp(\widehat{\ln y})$ gebildet wird, und
3. eine Hilfsregression $y = az$ (ohne Interzept!) gerechnet wird. Die Schätzung von a dieser Hilfsregression ist ein konsistenter Schätzer für $\alpha = \exp(u)$, deshalb können mit a konsistente Prognosen berechnet werden

$$\hat{y} = a \exp(\widehat{\ln y})$$

8.A.3 Vergleich des Bestimmtheitsmasses von log und lin Modellen

Das Bestimmtheitsmaß $R^2 = 1 - \sum_i e_i^2 / (\sum_i y_i - \bar{y})^2$ kann nur für den Vergleich von Modellen verwendet werden, bei denen die abhängige Variable identisch ist und die die gleiche Anzahl erklärender Variablen haben, zum Beispiel also nicht für einen Vergleich eines log-lin Modells mit einem lin-lin Modell (d.h. $\ln y = b_0 + b_1 x + e$ mit $y = b_0 + b_1 x + e$).

Erinnern wir uns, dass das Bestimmtheitsmaß R^2 auch als das Quadrat des Stichproben-Korrelationskoeffizienten zwischen y und \hat{y} interpretiert werden kann. Deshalb wird in Fällen, in denen das R^2 nicht direkt vergleichbar ist, manchmal das Quadrat des Korrelationskoeffizienten zwischen y und \hat{y} angegeben.

Beispiel 1: Wir greifen wieder auf das Beispiel mit dem österreichischen BIP (Berechnung von Wachstumsraten) zurück.¹⁰

Das folgende EViews Programm berechnet die unterschiedlichen Prognosen:

```
wfcreate(wf=BIPAUT) a 1976 2009
' Datei VGR.xls wurde vorher ins Defaultverzeichnis heruntergeladen
read(t=xls,c7) VGR.xls BIP

equation eqlog.ls @log(BIP) c @trend
eqlog.fit(d) lnYhat
series Yhatwrong = @exp(lnYhat)
scalar R2wrong = @cor(BIP,Yhatwrong)^2

series Yhatnormal = @exp(lnYhat) * @exp(0.5*eqlog.@se^2)
scalar R2normal = @cor(BIP,Yhatnormal)^2

eqlog.makesresids uhat
series YhatDuan = @exp(lnYhat) * @mean(@exp(uhat))
scalar R2Duan = @cor(BIP,YhatDuan)^2

equation eqaux.ls BIP @exp(lnYhat)
series YhatWoold = @exp(lnYhat) * eqaux.@coefs(1)
scalar R2Woold = @cor(BIP,YhatWoold)^2

equation eqlevels.ls BIP c @trend
eqlevels.fit Yhatlevels
scalar R2levels = eqlevels.@r2

stats BIP Yhatwrong Yhatnormal YhatDuan YhatWoold Yhatlevels
```

Achtung: falls in EViews als abhängige Variable eine Funktion angegeben wird, z.B. `equation eq1.ls log(y) c x1 x2`, dann berechnet `eq1.fit yhat` nicht (!) $\widehat{\ln y}$ (d.h. die gefitteten Logarithmen von y), sondern die gefitteten Werte von $\hat{y} = \exp(\widehat{\ln y})$.

¹⁰Statistik Austria, BIP, verkettete Volumenindizes, 2005 = 100,
http://www.statistik.at/web_de/static/volkswirtschaftliche_gesamtrechnung_hauptgroessen_019505.xls, 09.05.2010.

Tabelle 8.1: Verschiedene BIP Prognosen mit einem log-lin Modell (siehe EViews Programm)

| | Mean | Median | Maximum | Minimum | Std.Dev. | R^2 |
|------------|---------|---------|----------|---------|----------|--------|
| BIP | 77.5116 | 75.9659 | 109.3240 | 51.6604 | 17.5674 | 1.0000 |
| YhatWrong | 77.4991 | 75.6000 | 110.1067 | 51.9007 | 17.5019 | 0.9937 |
| YhatNormal | 77.5104 | 75.6110 | 110.1227 | 51.9083 | 17.5045 | 0.9937 |
| YhatDuan | 77.5097 | 75.6104 | 110.1218 | 51.9079 | 17.5043 | 0.9937 |
| YhatWoold | 77.5132 | 75.6138 | 110.1267 | 51.9102 | 17.5051 | 0.9937 |
| YhatLevels | 77.5116 | 77.5116 | 106.3924 | 48.6308 | 17.4304 | 0.9845 |

Tabelle 8.2: Verschiedene Einkommen Prognosen mit einem log-lin Modell

| | Mean | Median | Maximum | Minimum | Std. Dev. | R^2 |
|------------|------------|------------|-------------|------------|------------|--------|
| INC | 32811.8500 | 28093.5000 | 408215.9000 | 338.0000 | 22861.6200 | 1.0000 |
| YhatWrong | 28408.7900 | 27134.0200 | 95732.9200 | 14492.7500 | 8901.9450 | 0.1750 |
| YhatNormal | 34081.8000 | 32552.4600 | 114850.0000 | 17386.8300 | 10679.5900 | 0.1750 |
| YhatDuan | 32824.7300 | 31351.8100 | 110613.9000 | 16745.5400 | 10285.6900 | 0.1750 |
| YhatWoold | 32606.8400 | 31143.6900 | 109879.6000 | 16634.3800 | 10217.4100 | 0.1750 |
| YhatLevels | 32811.8500 | 32192.6900 | 85134.2900 | 12620.0000 | 9888.5420 | 0.1871 |

Möchte man die gefitten logarithmierten Werte kann man dies durch die Option `d` erreichen, d.h. `eq1.fit(d) lnyhat`.

Tabelle 8.1 zeigt das Ergebnis dieses Programms. Zu Vergleichszwecken wurde auch eine Prognose mit einem lin-lin Modell berechnet (Yhatlevels). Da eine OLS Regression immer durch den Mittelwert der Stichprobe geht stimmen die Mittelwerte der Level-Prognose und der wahren Stichprobenwerte (BIP) überein, aber der Fit ist schlechter als der log-Prognosen, wie das R^2 zeigt. Außerdem würde man schon aus theoretischen Gründen kein lin-lin Modell schätzen (es ist unplausibel, dass das BIP jährlich um den gleichen Absolutbetrag zunehmen wird, eine relative Änderung ist weit plausibler).

Beispiel 2: Auf Grundlage der EU-SILC Daten 2007 wurde eine übliche Lohn-gleichung geschätzt

$$\text{LOG(INC)} = \underset{(0.05)}{9.153} + \underset{(0.004)}{0.058 \text{ EDUC}} + \underset{(0.001)}{0.022 \text{ EXPER}}$$

$$R^2 = 0.193, \quad s = 0.603, \quad F\text{-Stat} = 259.326, \quad N = 2165$$

mit INC: Einkommen aus unselbständiger Erwerbstätigkeit (Jahresbetrag Brutto 2006, Variable `py010g`); EDUC: Alter bei Beginn erster Erwerbstätigkeit minus sieben (Variable `P032000 - 7`); und EXPER Zahl der erwerbstätigen Jahre (Variable `P033000`).

Tabelle 8.2 zeigt wieder die verschiedenen Prognosen mit Hilfe des log-lin Modells (bzw. des lin-lin Modells für YHATLEVELS). Wieder sieht man, dass $\widehat{\text{YHATWRONG}} = \exp(\ln y)$ das Einkommen systematisch unterschätzt, wenngleich die Unterschiede keinesfalls dramatisch sind.

Quelle:

<http://www.uibk.ac.at/econometrics/>