

UNIVERSITÄT BREMEN

Skripten zur empirischen Wirtschaftsforschung
und angewandten Statistik

GRETLManual

Eine Einführung in die Ökonometriesoftware *gretl*

Martin Missong

Sina Wittenberg



Inhaltsverzeichnis

1 Grundlagen

Das Programmpaket *gretl* ist eine freie, plattformunabhängige Statistiksoftware mit dem Schwerpunkt ökonometrische Methoden. *gretl* steht für „Gnu Regression, Econometrics and Time-series Library“ und ist unter der Internet-Adresse <http://gretl.sourceforge.net/> erhältlich.

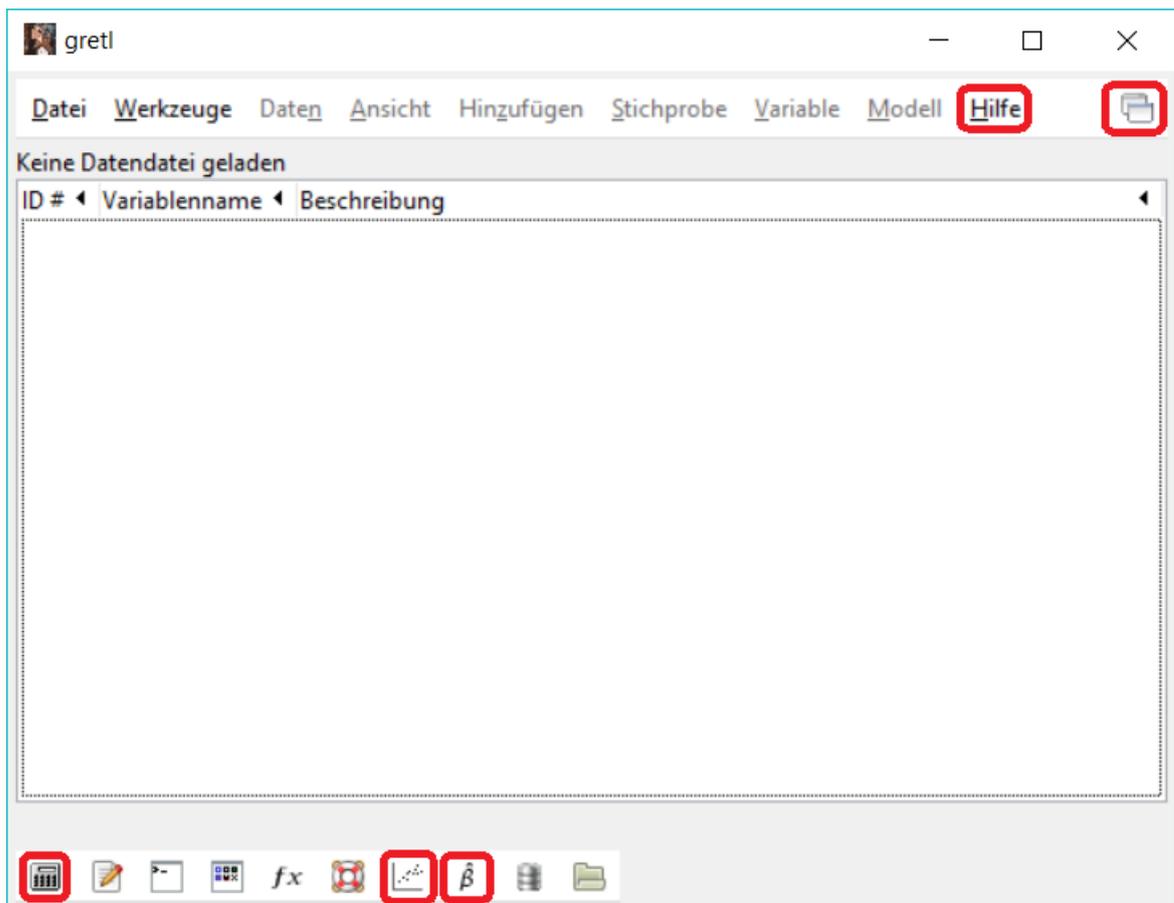
Das vorliegende Manual gibt eine kurze Einführung in *gretl*. Dabei wird nicht nach einer strengen Systematik vorgegangen und nicht auf ein vollständiges „Abarbeiten“ der Programmfunktionen gezielt. Die vorliegende Dokumentation soll vielmehr eine anschauliche Einführung als Begleitmaterial zum Bachelormodul „Einführung in die Ökonometrie“ in den Bachelorstudiengängen der Universität Bremen bieten. Deshalb beschränkt sich die Darstellung zunächst auf die Umsetzung der Beispiele aus dem Vorlesungsskript, und zwar zu den Datensätzen „Konsumfunktion“ und „EDP-Daten“. Beide Datensätze finden Sie im Excel-Format auf den Stud.IP-Seiten zur Vorlesung. Die „EDP-Daten“ sind dem sehr empfehlenswerten Wirtschafts-Lehrbuch „Managerial Economics“ von Mark Hirschey und Erik Bentzen (Cengage, 2016, 14. Auflage, S. 185) entnommen.

Ferner beschränkt sich die folgende Darstellung auf eine menüorientierte Steuerung der Bearbeitungsschritte. Im Schlusskapitel gehen wir in Form eines Ausblicks auf die Speicherung des Programmcodes bzw. die Erstellung von Skriptdateien ein und deuten an, wie mit *gretl* Syntax-orientiert gearbeitet werden kann. Damit lassen sich auch komplexere Problemstellungen bearbeiten und *gretl* stellt eine leistungsfähige Alternative zu Programmiersprachen wie R, Octave oder MATLAB dar.

Unter dem Menüpunkt „Hilfe“ im *gretl*-Fenster ist u. a. eine ausführliche Bedienungsanleitung für die Software abrufbar. Zusätzlich findet man dort eine reichhaltige Befehlsdokumentation und – für weitergehend Interessierte – ein Handbuch zur Programmiersprache *hansl* („Hansl’s A Neat Scripting Language“), in der *gretl* programmiert ist.

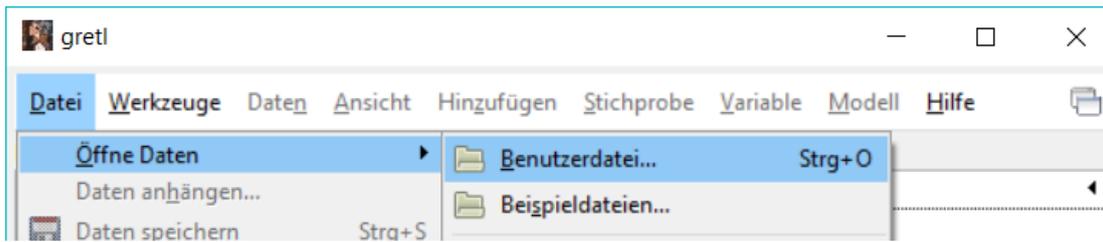
1.1 *gretl* Oberfläche

Die Oberfläche von *gretl* ist sehr übersichtlich gehalten: In der oberen Leiste erlauben verschiedene Menüs den Zugriff auf die einzelnen Funktionen. Über das Menü *Hilfe* lässt sich das Benutzerhandbuch aufrufen. Am unteren Rand des Fensters finden sich einige Schnellzugriffe, darunter  zum Aufrufen des Taschenrechners,  zur Erstellung von Plots und $\hat{\beta}$ zur Durchführung einer KQ-Schätzung. Über  lässt sich zwischen den verschiedenen *gretl*-Fenstern wechseln.

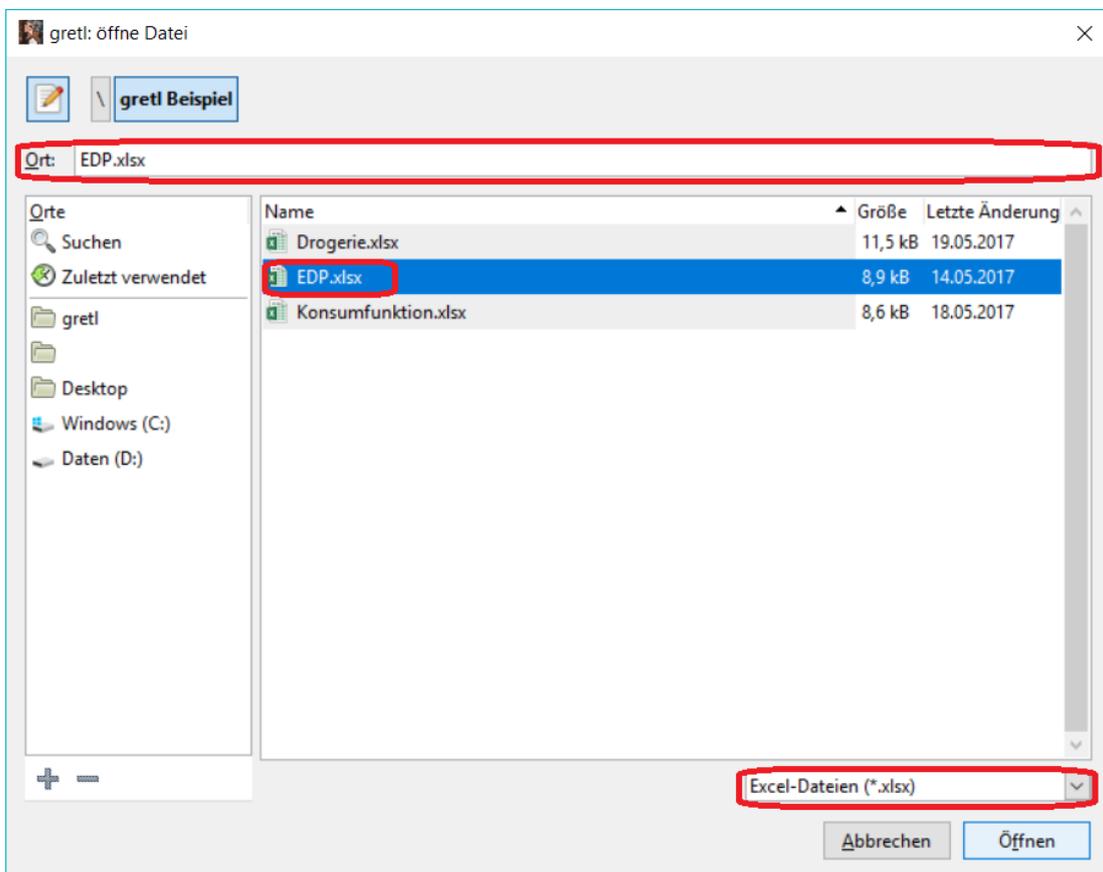


1.2 Daten öffnen

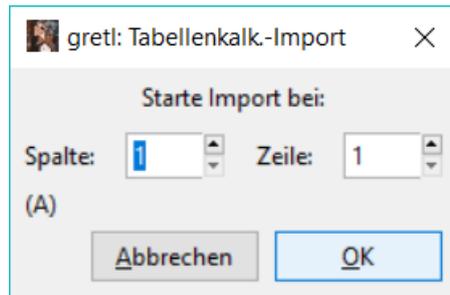
Zunächst muss der relevante Datensatz in *gretl* geöffnet werden.



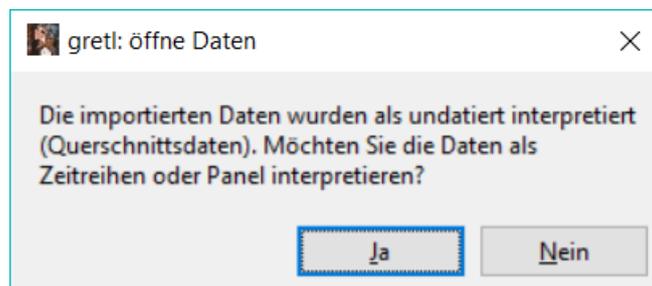
Durch einen Klick auf *Datei* → *Öffne Daten* → *Benutzerdatei* öffnet sich folgendes Menü. Zunächst müssen Ort und Dateityp ausgewählt werden, daraufhin erscheinen alle Dateien des ausgewählten Typs im Auswahlbereich. Die gewünschte Datei kann hier nun ausgewählt und durch einen Klick auf *Öffnen* importiert werden.



Anschließend wird festgelegt, mit welcher Zelle der Tabelle der Import gestartet werden soll. So könnten beispielsweise Erklärungen und überflüssige Überschriften ausgeschlossen werden.



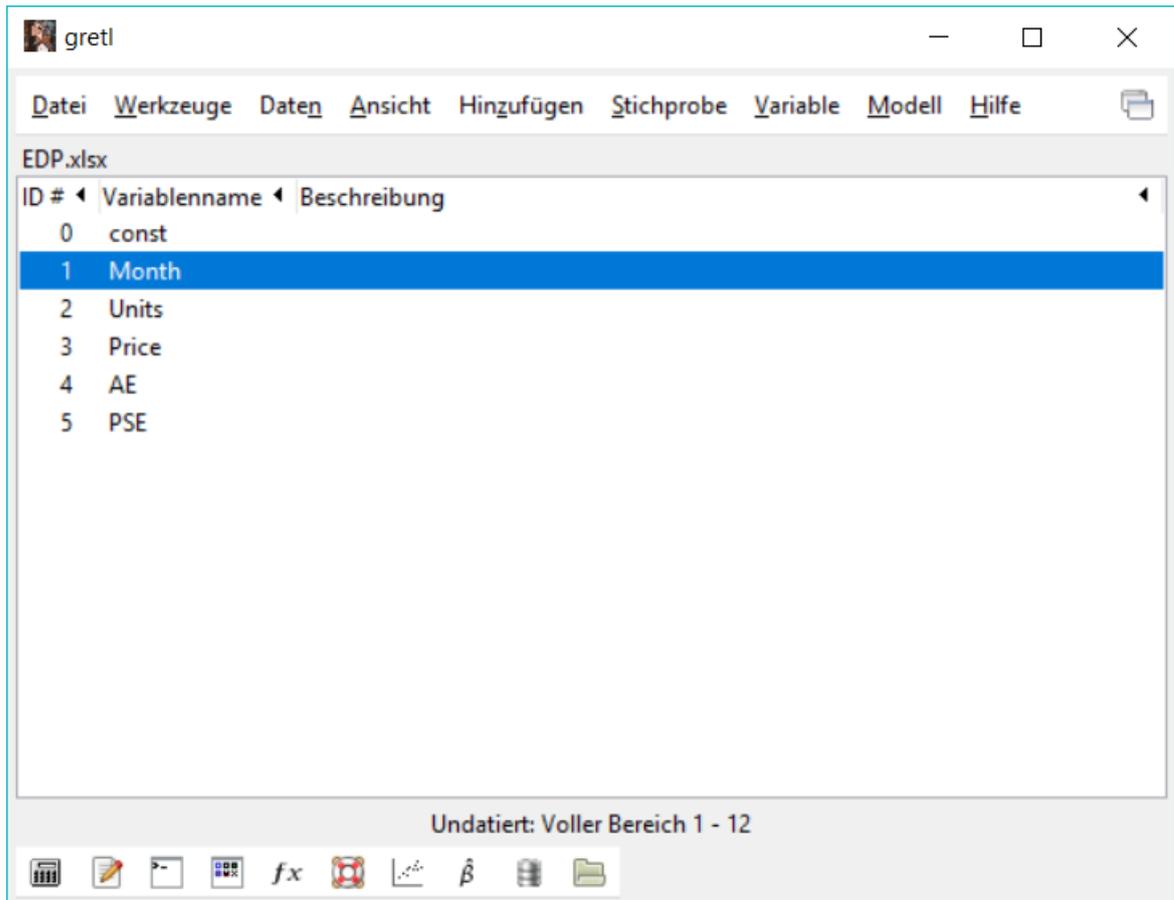
Bei der darauffolgenden Frage, ob die Daten als Zeitreihen oder Panel interpretiert werden sollen, ist *Ja* zu wählen,



und in den darauf folgenden Abfragen der Zeitreihencharakter mit monatlicher Frequenz anzugeben. Der Startzeitpunkt kann beispielhaft auf 2018:1, also Januar 2018 für die erste Beobachtung gesetzt werden.

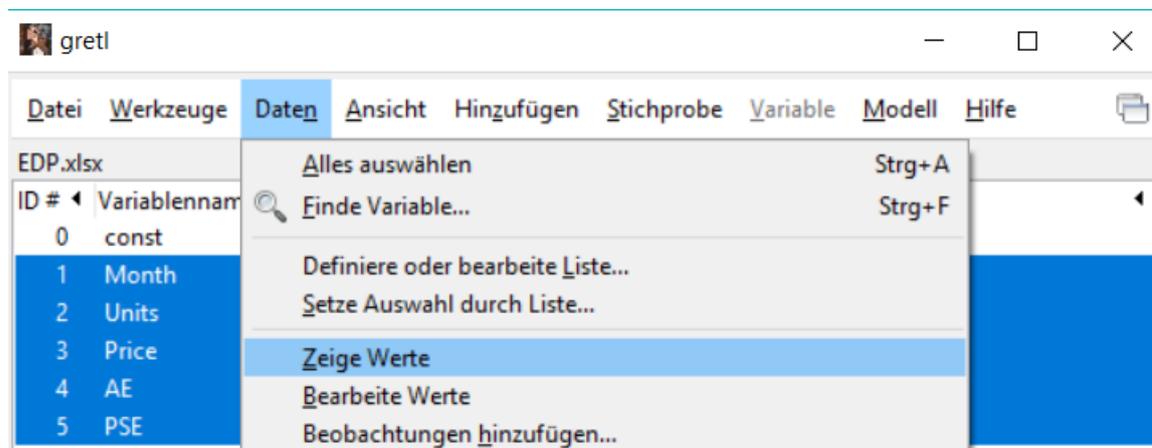


Daraufhin werden die Variablen des importierten Datensatzes in *gretl* angezeigt.

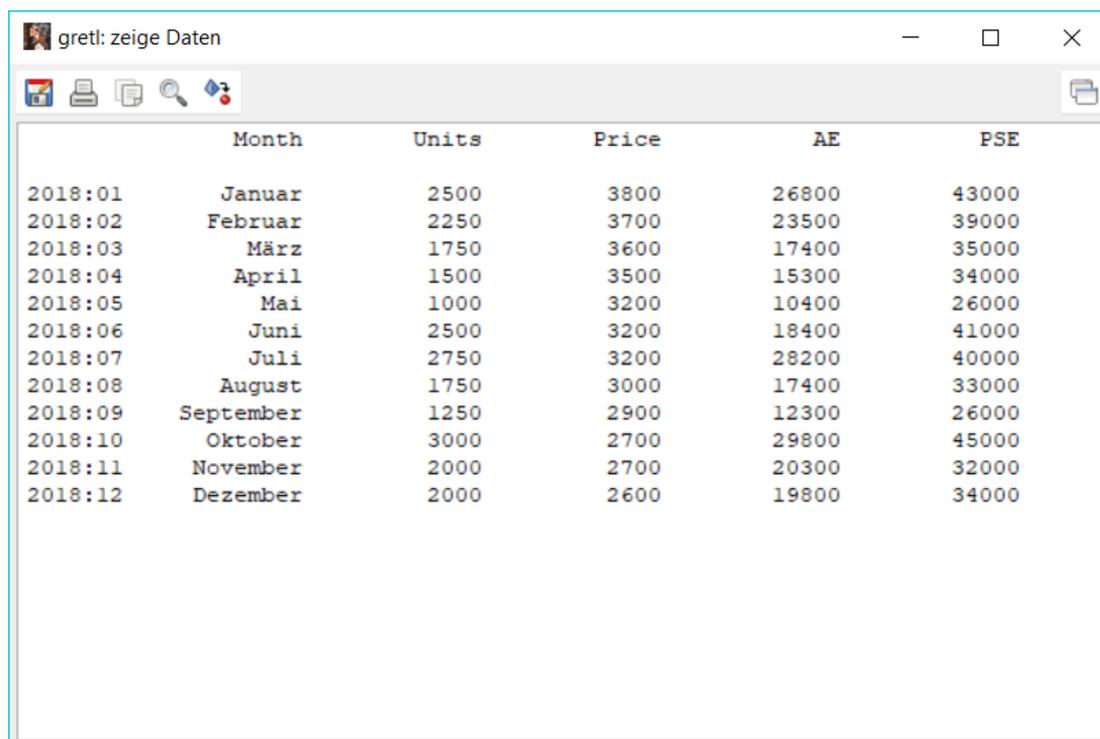


1.3 Daten anzeigen

Durch Auswählen der relevanten Variablen per Mausklick (oder über *Daten* → *Alle auswählen*) und einen Klick auf *Daten* → *Zeige Werte* lassen sich die Beobachtungswerte der einzelnen Variablen in anzeigen.



Ein neues Fenster öffnet sich, die Werte sind tabellarisch dargestellt.



The screenshot shows a window titled 'gretl: zeige Daten' displaying a table of data for the year 2018. The table has six columns: 'Month', 'Units', 'Price', 'AE', and 'PSE'. The data is as follows:

	Month	Units	Price	AE	PSE
2018:01	Januar	2500	3800	26800	43000
2018:02	Februar	2250	3700	23500	39000
2018:03	März	1750	3600	17400	35000
2018:04	April	1500	3500	15300	34000
2018:05	Mai	1000	3200	10400	26000
2018:06	Juni	2500	3200	18400	41000
2018:07	Juli	2750	3200	28200	40000
2018:08	August	1750	3000	17400	33000
2018:09	September	1250	2900	12300	26000
2018:10	Oktober	3000	2700	29800	45000
2018:11	November	2000	2700	20300	32000
2018:12	Dezember	2000	2600	19800	34000

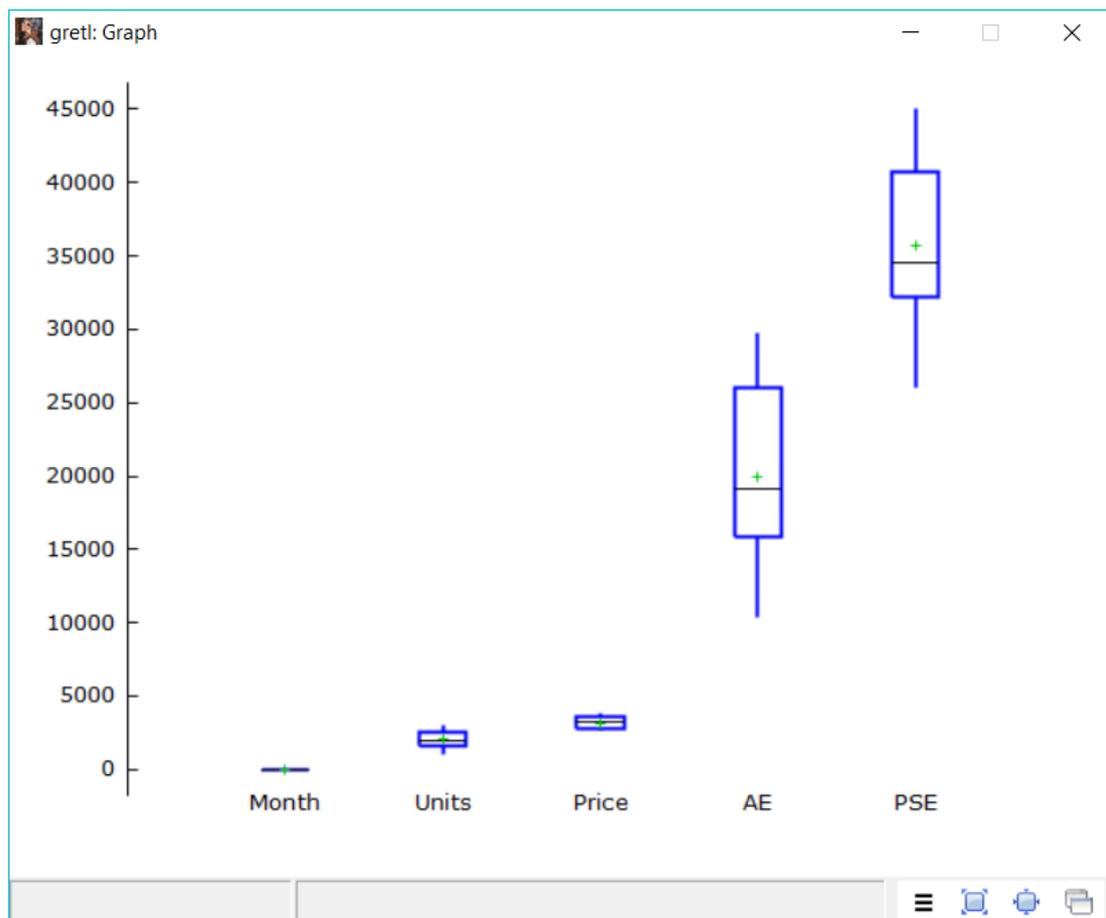
Mit *Units* ist der Absatz des Unternehmens Electronic Data Processing (EDP) in Stück bezeichnet, während *Price* den Produktpreis in US-\$ angibt. *AE* bezeichnet die Werbeausgaben (advertising expenditures) und *PSE* die Vertriebsausgaben (personal selling expenditures), beide werden in US-\$ gemessen.

1.4 Grafische Darstellungen der Daten

Über *Ansicht* → *Plotte spezifizierte Variablen* lassen sich ausgewählte Datenreihen grafisch darstellen. Beispielfhaft werden hier Boxplots, Zeitreihen und Streuungsdiagramme vorgestellt.

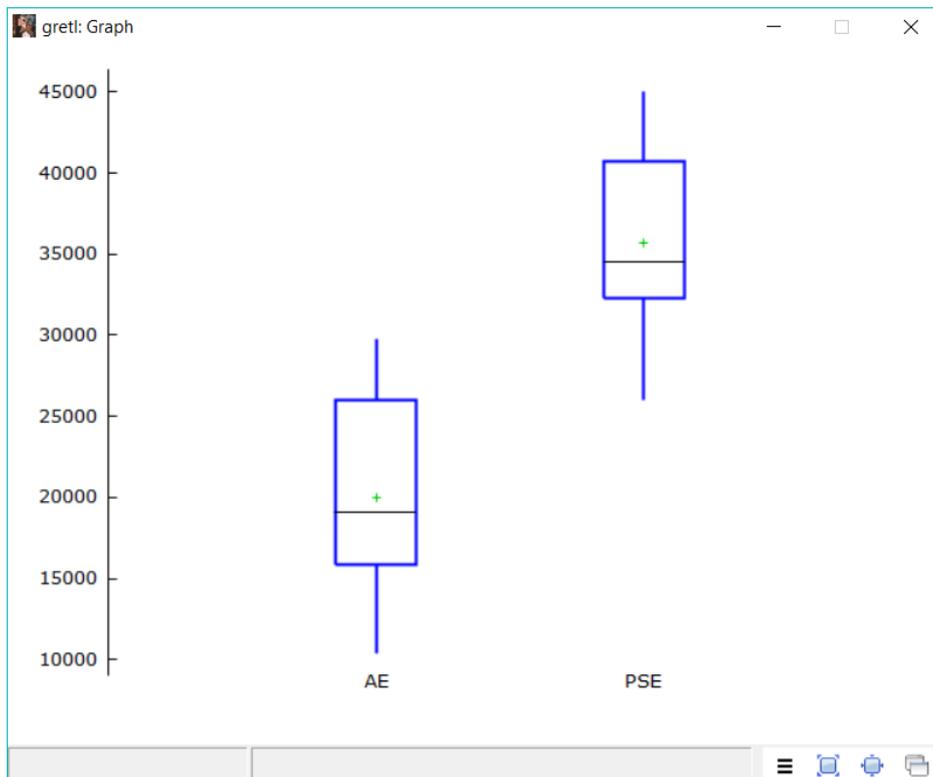
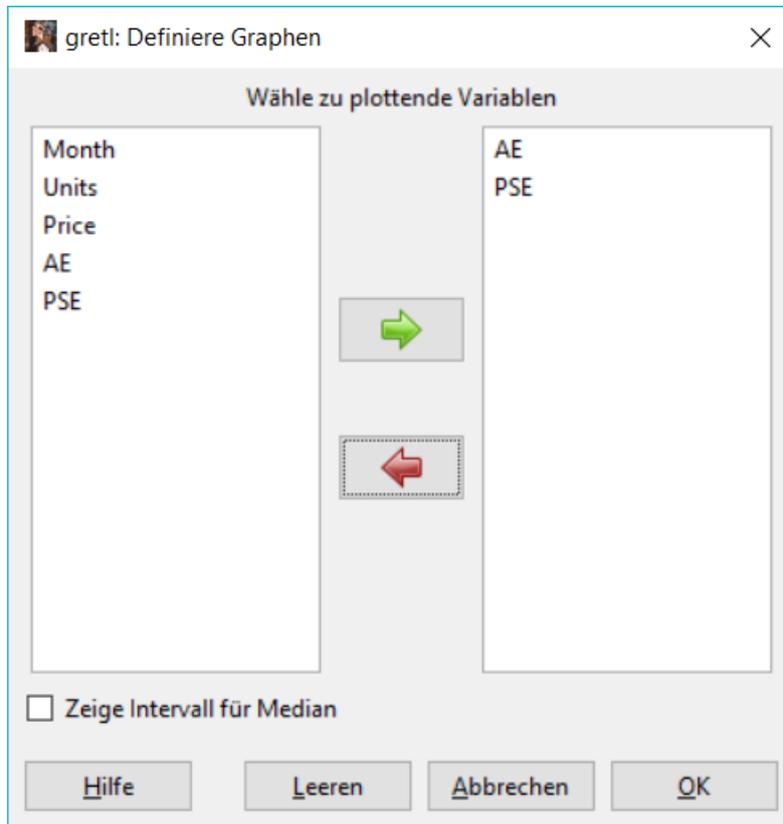
1.4.1 Boxplots

Markiert man zunächst alle Variablen des Datensatzes über *Daten* → *Alle auswählen* und lässt unter *Ansicht* → *Plotte spezifizierte Variablen* → *Boxplots* im folgenden Abfragefenster die Variablenauswahl unverändert, so erhält man die folgende Grafik:



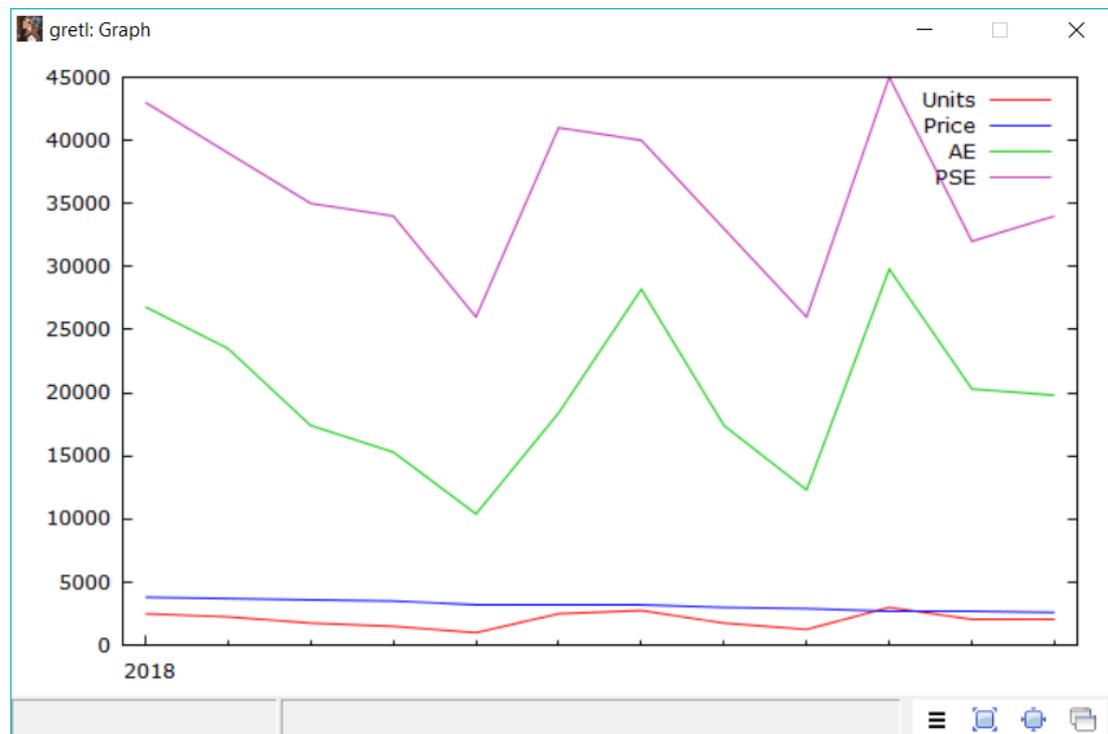
Es wird zum einen deutlich, dass *gretl* die Monatsnamen mit den Zahlen 1 bis 12 codiert hat; so kommt für die Reihe *Month* ein Boxplot zustande, der aber nicht sinnvoll interpretierbar ist. Zum anderen lässt die Darstellung einen Vergleich von Werbeausgaben *AE* und Vertriebsausgaben *PSE* zu: Beide Verteilungen sind linkssteil und besitzen eine vergleichbare Streuung, die Vertriebsausgaben liegen jedoch deutlich über den Werbeausgaben.

Will man lediglich diese beiden Verteilungen vergleichen, so ist der Plot-Befehl erneut aufzurufen und im Dialogfenster können die übrigen Variablen entfernt werden:



1.4.2 Zeitreihen

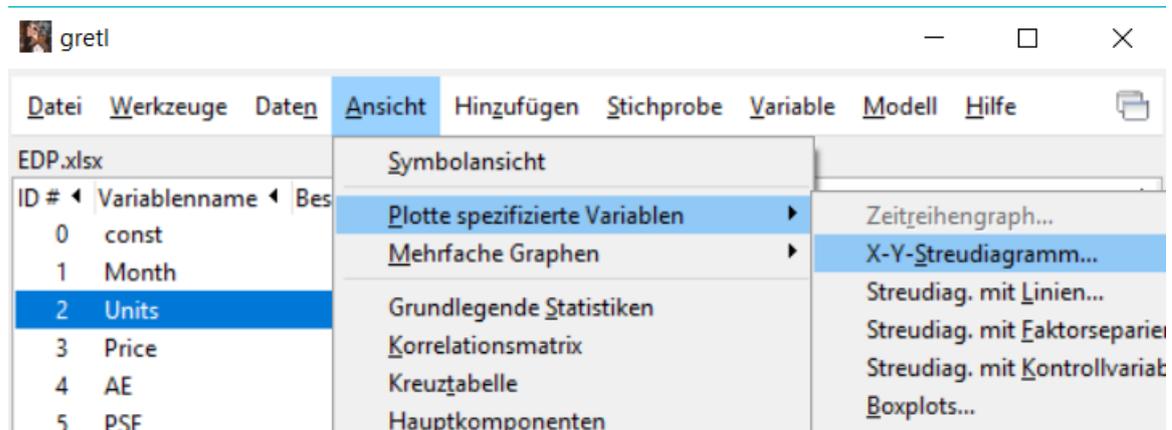
Ansicht → *Plote spezifizierte Variablen* → *Zeitreihengraph* liefert die Zeitreihendarstellung des Datensatzes (die Variable „Monate“ wurde im Dialogfenster ausgeschlossen):



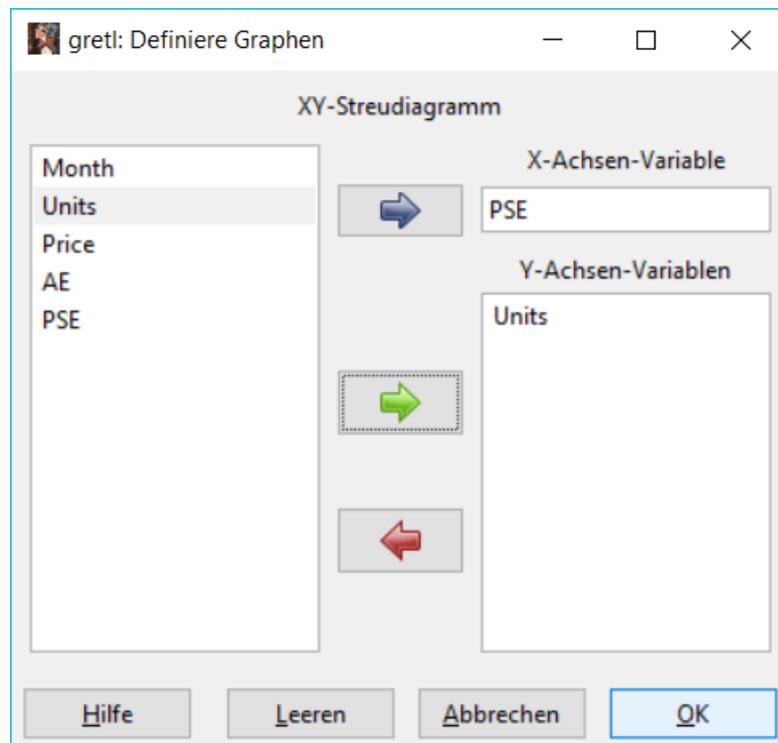
Die gleichgerichteten Änderungen von AE und PSE lassen bereits erahnen, dass es schwierig wird, mit Hilfe eines Linearen Modells die Einflüsse von AE und PSE auf den Absatz zu trennen, hier besteht deutliche (Multi-)Kollinearität.

1.4.3 Streudiagramme

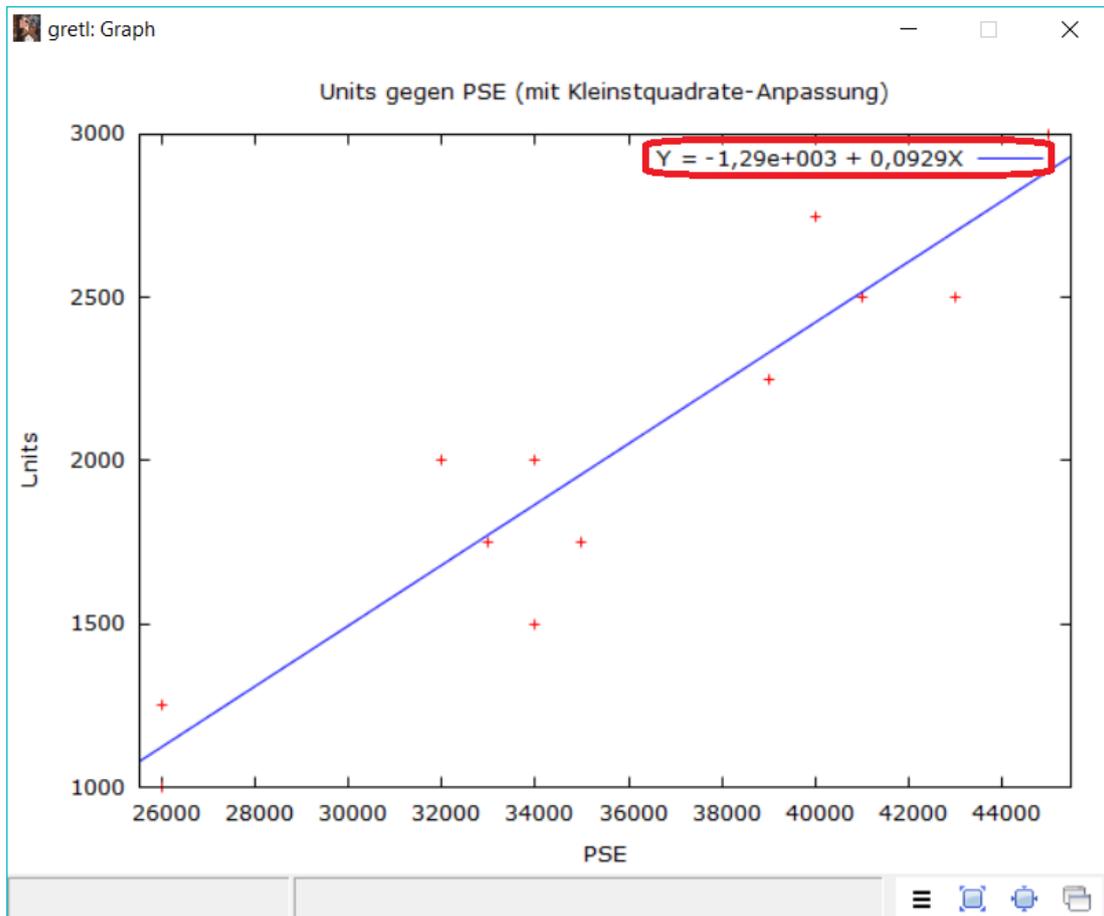
Über *Ansicht* → *Plotte spezifizierte Variablen* → *X-Y-Streudiagramm* lassen sich in *gretl* außerdem Streudiagramme erstellen.



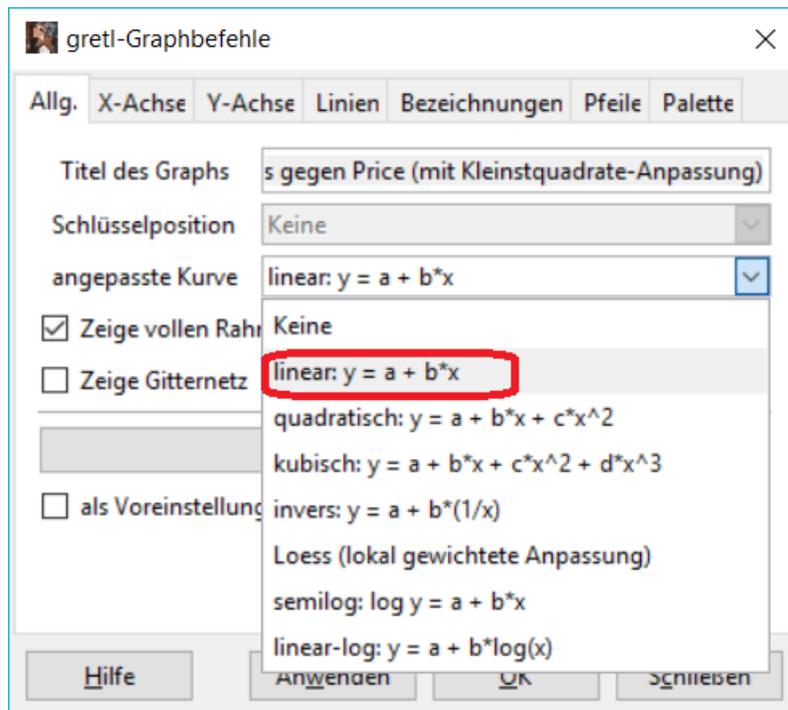
Im darauffolgenden Fenster können X- und Y-Achsen-Variable festgelegt werden. Dazu wird die entsprechende Variable im linken Bereich ausgewählt und mit dem blauen bzw. grünen Pfeil in das entsprechende Feld verschoben. Hier sollen beispielhaft die Vertriebsausgaben (PSE) auf der X-Achse und die Einheiten (Units) auf der Y-Achse aufgetragen werden. Die fertige Auswahl wird mit *OK* bestätigt.



Das Streudiagramm öffnet sich in einem neuen Fenster. Bei einem ausreichend großen Korrelationskoeffizienten wird automatisch die ausgleichende Regressionsgerade eingezeichnet. Die zugehörige Gleichung wird oben rechts in der Grafik angezeigt. Ein ausgeprägter, positiver linearer Zusammenhang zwischen den Vertriebsausgaben und den verkauften Einheiten wird deutlich.

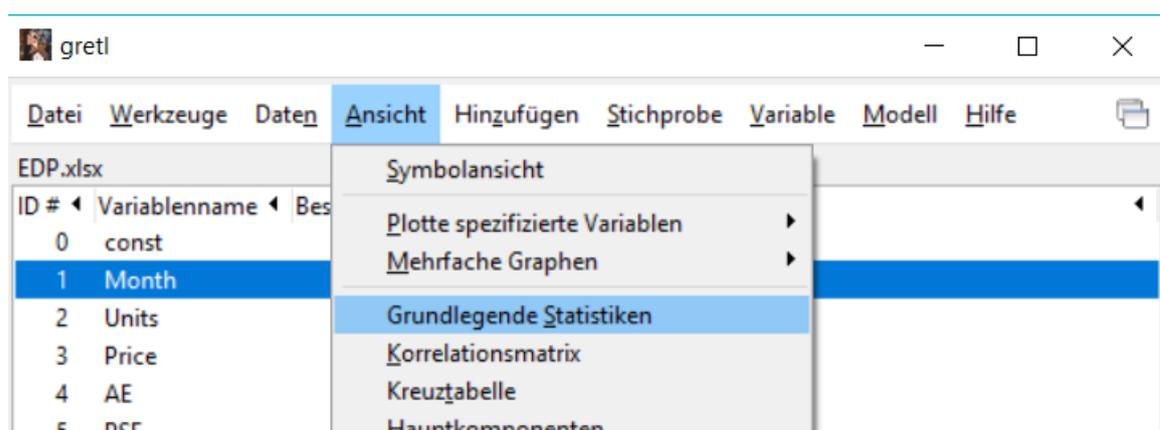


Im *Graph*-Fenster lässt sich mit einem Klick auf  das Menü öffnen. Über *Bearbeiten* öffnet sich ein neues Fenster mit vielen Einstellungsmöglichkeiten für das Diagramm. Hier lässt sich beispielsweise die angepasste Kurve (Regressionsgerade) entfernen oder, sofern sie nicht automatisch eingezeichnet wurde, hinzufügen.

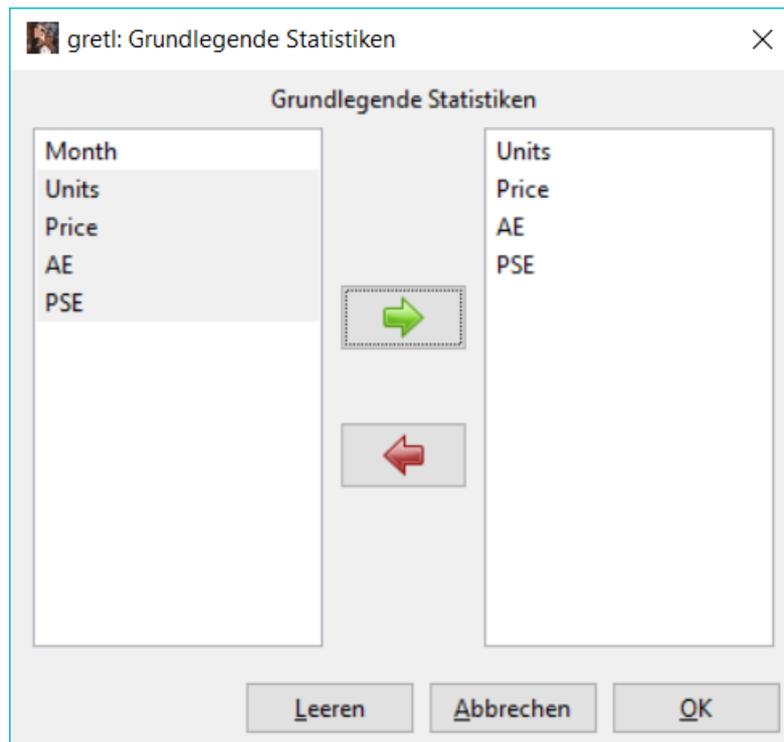


1.5 Deskriptive Statistiken

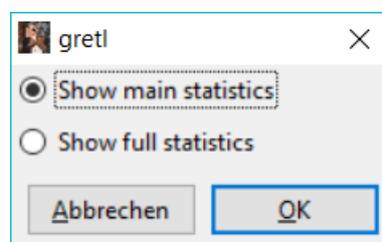
Über *Ansicht* → *Grundlegende Statistiken* lassen sich deskriptive Statistiken zu ausgewählten Variablen anzeigen.



Wieder müssen zunächst die Variablen ausgewählt werden, für die deskriptive Statistiken angezeigt werden sollen.



Anschließend kann gewählt werden, ob nur grundlegende Statistiken (*Show main statistics*) oder alle Statistiken (*Show full statistics*) angezeigt werden sollen. Häufig sind die grundlegenden Statistiken im Rahmen einer ersten Datenanalyse ausreichend.



Die Statistiken öffnen sich in einem neuen Fenster.

	arith. Mittel	Median	S.D.	Min	Max
Units	2021	2000	607,3	1000	3000
Price	3175	3200	409,3	2600	3800
AE	19967	19100	6116	10400	29800
PSE	35667	34500	6125	26000	45000

Über *Ansicht* → *Korrelationsmatrix* lassen sich aus dem Hauptfenster heraus außerdem die paarweisen Korrelationen aller oder ausgewählter Variablen anzeigen.

Korrelationskoeffizienten, benutze die Beobachtungen 1 - 12
5% kritischer Wert (zweiseitig) = 0,5760 für n = 12

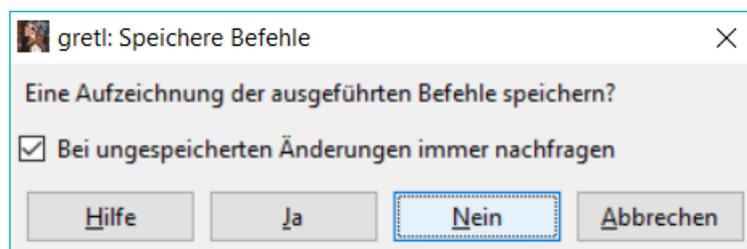
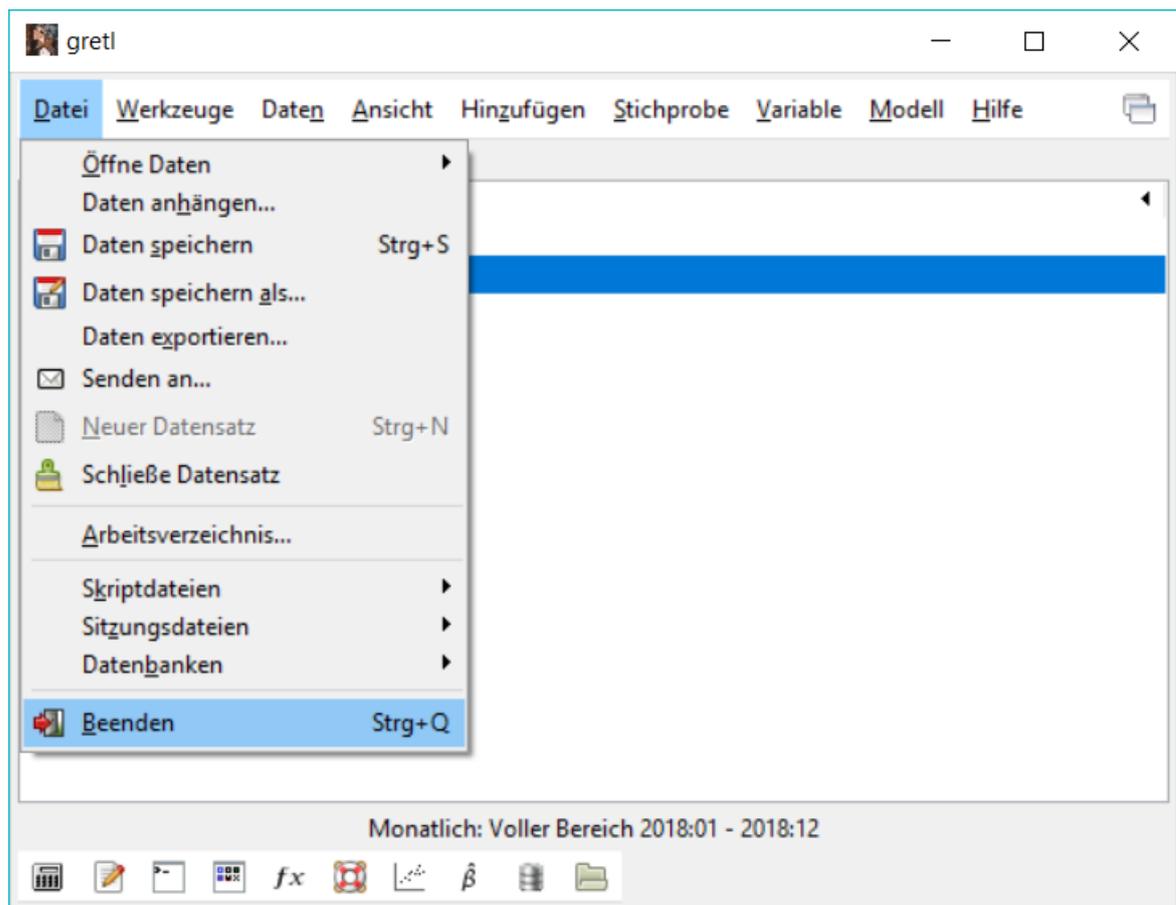
Units	Price	AE	PSE	
1,0000	-0,0160	0,9412	0,9368	Units
	1,0000	0,0461	0,2502	Price
		1,0000	0,8808	AE
			1,0000	PSE

2 Das Lineare Modell mit einem Regressor

2.1 Neuen Datensatz laden

Als Einstieg in das Arbeiten mit dem linearen Modell in *gretl* soll zunächst das Lineare Modell mit einem Regressor am Beispiel der Konsumfunktion gezeigt werden. Anschließend wird das multiple Modell anhand des EDP-Beispiels untersucht.

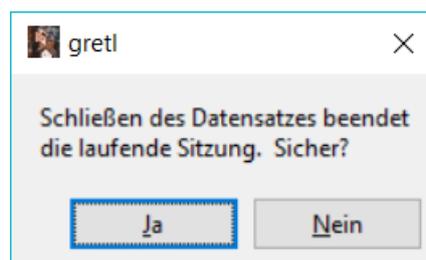
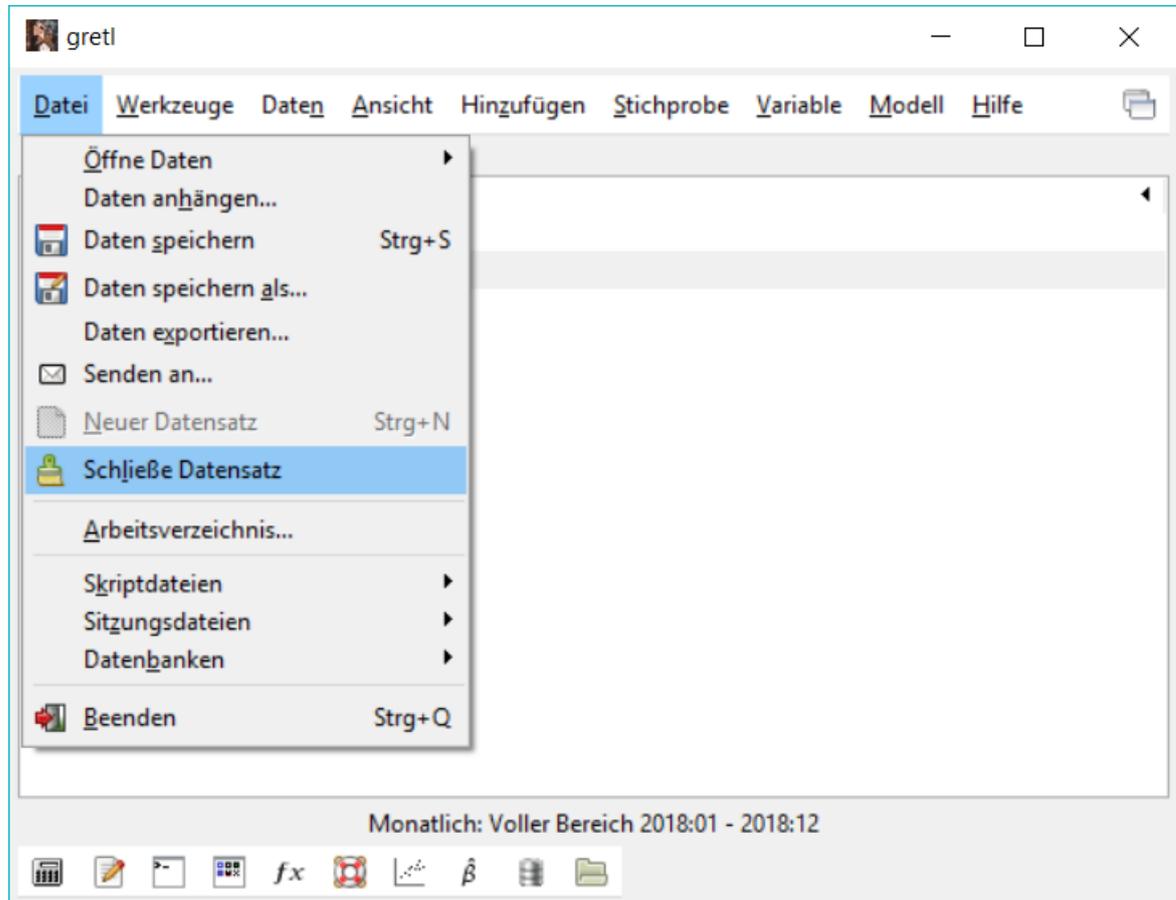
Um die Konsumfunktionsdaten zu laden können Sie entweder eine neue *gretl*-Sitzung starten oder den Datensatz schließen. Im ersten Fall werden Sie nach Aufrufen von *Datei* → *Beenden* gefragt, ob die Befehle gespeichert werden sollen. Diese Möglichkeit wird ausführlich im Schlussabschnitt ?? thematisiert.



An dieser Stelle soll *nein* gewählt und eine neue Sitzung gestartet werden, in der unter

Datei → *Öffne Daten* → *Benutzerdatei...* die Datei *konsumfunktion.xlsx* zu laden ist. In diesem Fall handelt es sich um Querschnittsdaten, weshalb bei der Frage nach Zeitreihen- oder Paneldaten *nein* gewählt werden kann.

Alternativ kann auch unter *Datei* → *Schließe Datensatz* der Datensatz geschlossen und der Hinweis, dass dies die Sitzung beendet, bestätigt werden:



Nun kann über *Datei* → *Öffne Daten* → *Benutzerdatei...* die Datei *konsumfunktion.xlsx* geladen werden.

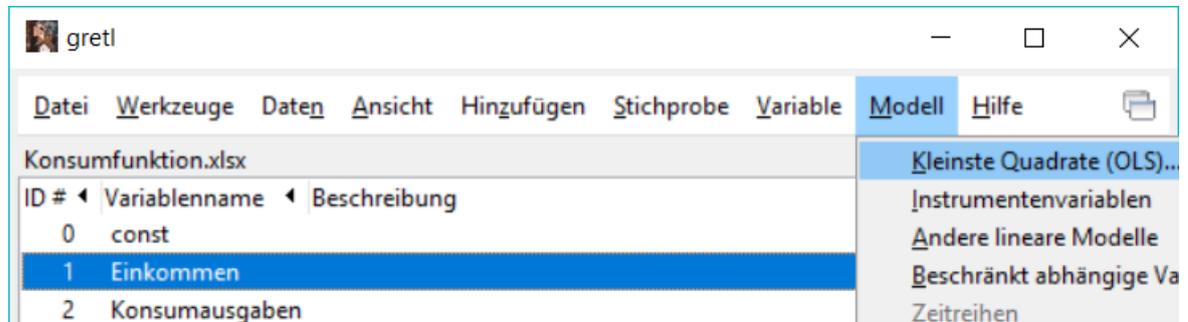
Der Datensatz enthält für 10 Haushalte Daten zum monatlichen Einkommen und zum monatlichen Konsum, jeweils gemessen in Euro.

2.2 Modell schätzen

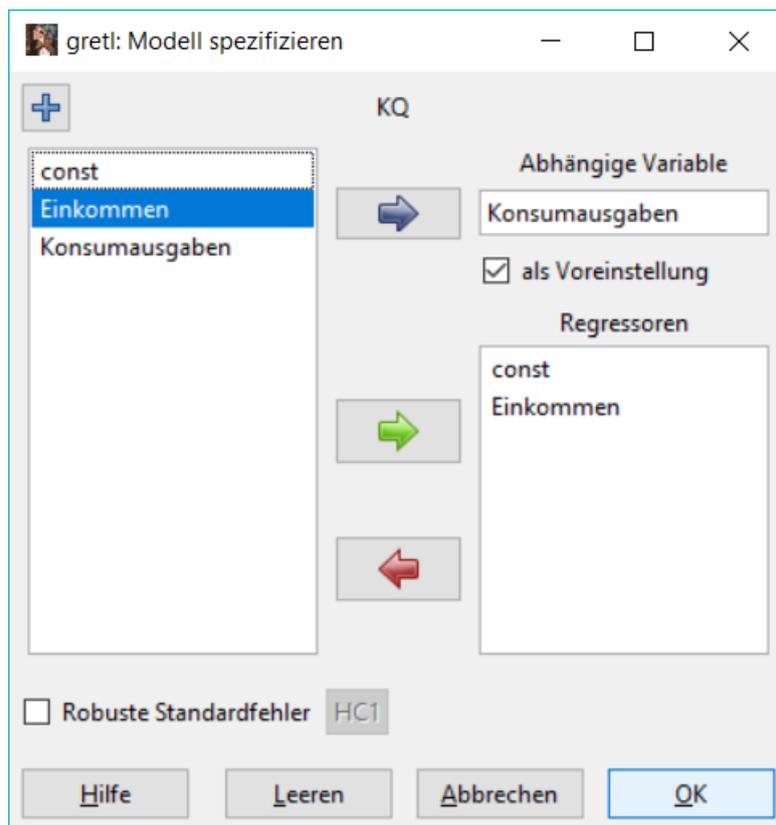
Nachdem die Daten geladen sind, lässt sich das Lineare Modell der Form

$$\text{Konsum}_i = \beta_0 + \beta_1 \text{Einkommen}_i + u_i$$

über *Modell* → *Kleinste Quadrate (OLS)* mit der KQ-Methode schätzen.



Dazu müssen zunächst die abhängige Variable (Y) und die unabhängige Variable (X) festgelegt werden. In diesem Beispiel sollen die Konsumausgaben als abhängige Variable durch das Einkommen als Regressor bzw. unabhängige Variable erklärt werden.



Es ergibt sich folgender Regressionsoutput. Hier sind die Schätzkoeffizienten für den Einfluss der unabhängigen Variable und das Absolutglied sowie deren Standardfehler, empirische t-Werte und p-Werte zu sehen. Zudem finden sich modellbezogene Größen wie das Bestimmtheitsmaß R^2 und die Summe der quadrierten Residuen (RSS).

Modell 1: KQ, benutze die Beobachtungen 1-10
Abhängige Variable: Konsumausgaben

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	-103,920	381,231	-0,2726	0,7921
Einkommen	0,740740	0,163613	4,527	0,0019 ***

Mittel d. abh. Var.	1504,300	Stdabw. d. abh. Var.	778,8235
Summe d. quad. Res.	1532520	Stdfehler d. Regress.	437,6815
R-Quadrat	0,719272	Korrigiertes R-Quadrat	0,684181
F(1, 8)	20,49734	P-Wert (F)	0,001931
Log-Likelihood	-73,88858	Akaike-Kriterium	151,7772
Schwarz-Kriterium	152,3823	Hannan-Quinn-Kriterium	151,1133

2.3 Konfidenzintervalle

Nach der Schätzung des Linearen Modells können aus dem Output-Fenster über *Analyse* → *Konfidenzintervalle für Koeffizienten* Konfidenzintervalle berechnet werden.

Modell 1: KQ, benutze die Beobachtungen 1-10
Abhängige Variable: Konsumausgaben

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	-103,920	381,231	-0,2726	0,7921
Einkommen	0,740740	0,163613	4,527	0,0019 ***

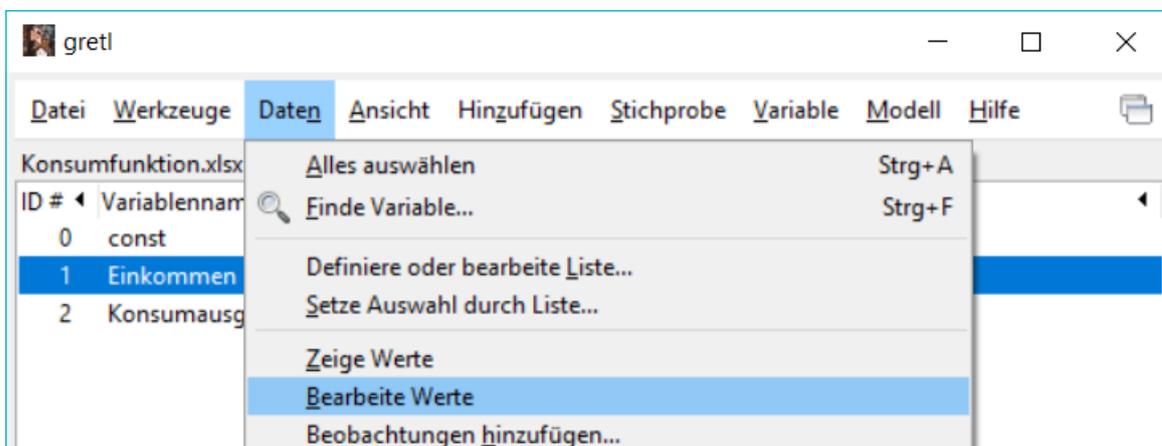
Mittel d. abh. Var.	1504,300	Stdabw. d. abh. Var.	778,8235
Summe d. quad. Res.	1532520	Stdfehler d. Regress.	437,6815
R-Quadrat	0,719272	Korrigiertes R-Quadrat	0,684181
F(1, 8)	20,49734	P-Wert (F)	0,001931
Log-Likelihood	-73,88858	Akaike-Kriterium	151,7772
Schwarz-Kriterium	152,3823	Hannan-Quinn-Kriterium	151,1133

Daraufhin öffnet sich das Output Fenster mit den oberen und unteren Grenzen der Konfidenzintervalle für alle Schätzkoeffizienten. Über α lässt sich das Konfidenzniveau anpassen.

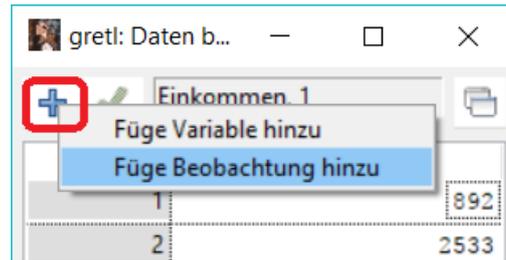
VARIABLE	KOEFFIZIENT	95% KONFIDENZ-INTERVALL	
const	-103,920	-983,041	775,201
Einkommen	0,740740	0,363448	1,11803

2.4 KQ-Prognose

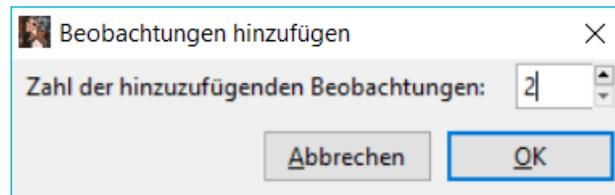
Auf Grundlage des geschätzten Modells können die Konsumausgaben für weitere, über die Beobachtungen hinausgehende Einkommenswerte prognostiziert werden. Zunächst muss dafür im Hauptfenster die unabhängige Variable ausgewählt werden, in diesem Fall das Einkommen. Über *Daten* → *Bearbeite Werte* können dann die zusätzlichen Einkommenswerte eingetragen werden.



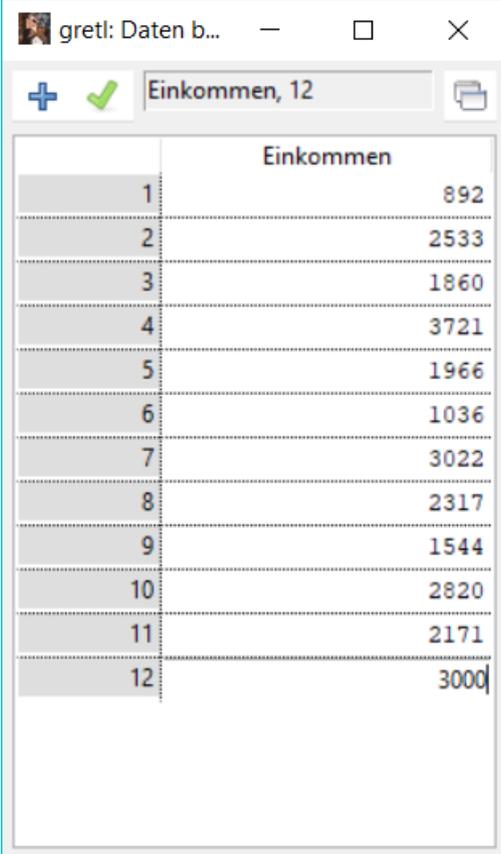
Es öffnet sich ein neues Fenster in dem die einzelnen Werte bearbeitet werden können. Neue Beobachtungen lassen sich über **+** → *Füge Beobachtungen hinzu* eintragen.



Im nächsten Fenster muss die Zahl der hinzuzufügenden Beobachtungen ausgewählt werden.

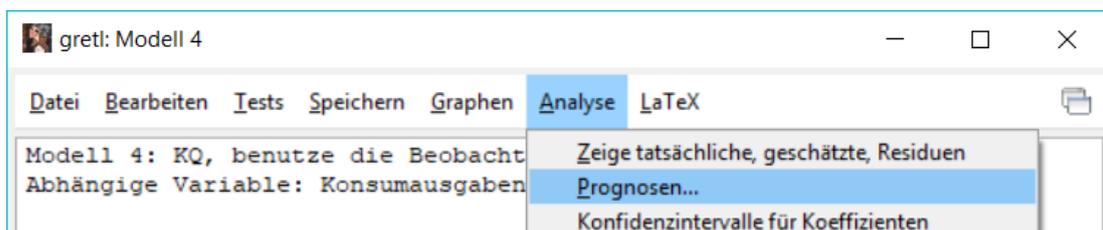


Es erscheinen zwei neue freie Felder, in die neue Werte eingetragen werden können. In diesem Beispiel sollen die Konsumausgaben für ein monatliches Einkommen von 2171 Euro bzw. 3000 Euro prognostiziert werden (vgl. die zugehörige Aufgabe im Skript zur „Einführung in die Ökonometrie“). Über  wird die Eingabe gespeichert.

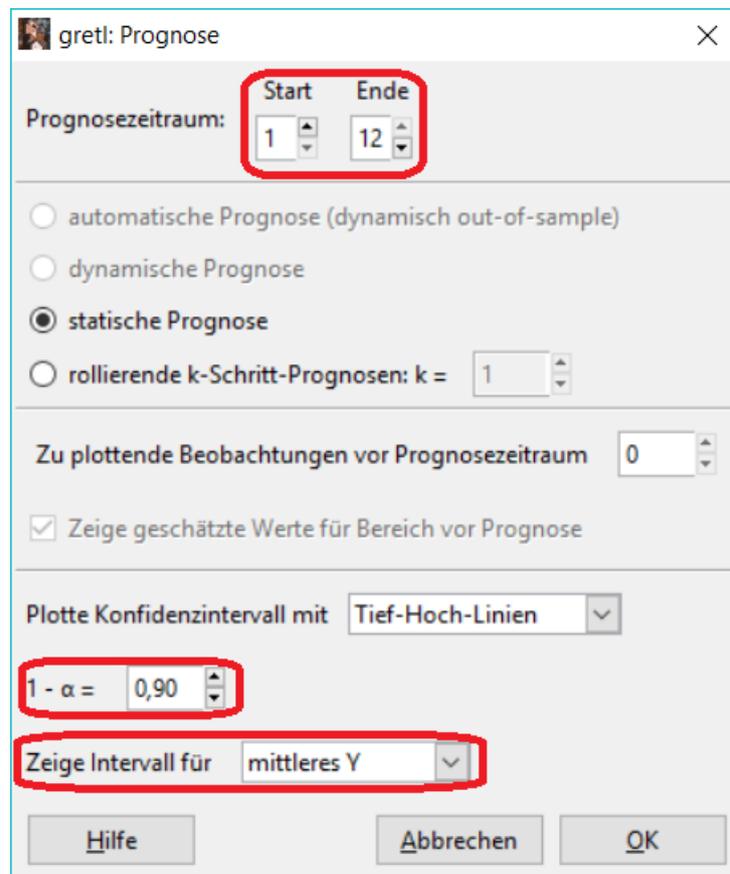


	Einkommen
1	892
2	2533
3	1860
4	3721
5	1966
6	1036
7	3022
8	2317
9	1544
10	2820
11	2171
12	3000

Anschließend muss das Lineare Modell wie in Abschnitt ?? erneut geschätzt werden. Im Fenster mit dem Regressionsoutput kann über *Analyse* → *Prognosen* das Prognosemenü aufgerufen werden.



Für die graphische Darstellung ist es empfehlenswert, den *Prognosezeitraum* über alle Beobachtungen zu strecken und im Feld *Zeige Intervall für* die Option *mittleres Y* auszuwählen. Unter $1 - \alpha$ lässt sich das Konfidenzniveau festlegen, hier soll ein 90%-Prognoseintervall berechnet werden.



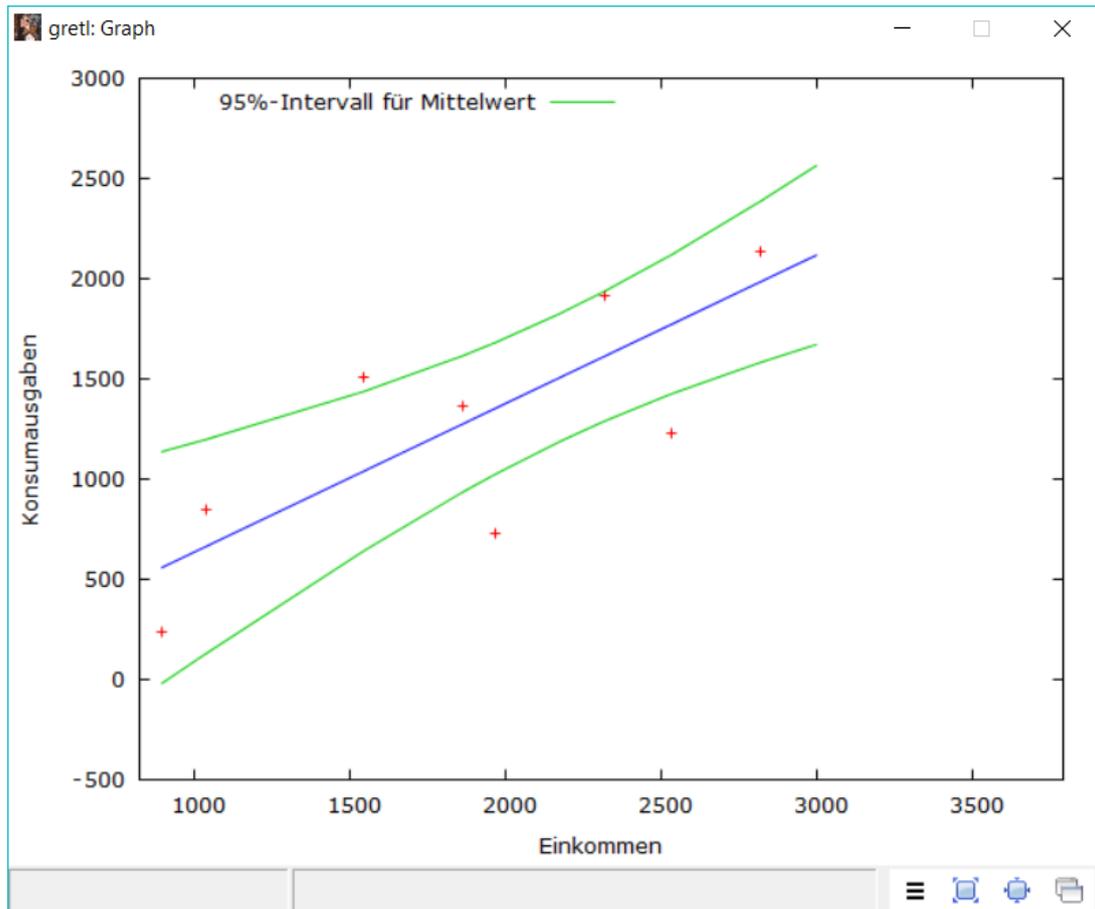
Es öffnen sich zwei neue Fenster. Das erste Fenster enthält die Prognosen, dazugehörige Standardfehler und Konfidenzintervalle.

The screenshot shows the 'gretl: Prognosen' window with the following data:

Für 90%-Konfidenzintervalle, $t(8, 0,05) = 1,860$

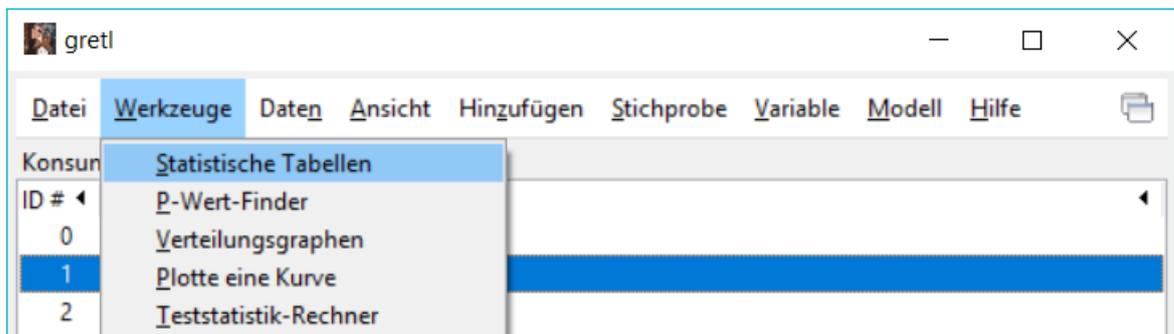
	Konsumausgaben	Prognose	Std.-fehler	90%-Intervall
1	239,00	556,82	250,905	90,25 - 1023,39
2	1227,00	1772,37	150,541	1492,44 - 2052,31
3	1365,00	1273,86	147,470	999,63 - 1548,08
4	2389,00	2652,37	288,896	2115,16 - 3189,59
5	730,00	1352,37	142,417	1087,54 - 1617,21
6	849,00	663,49	231,619	232,78 - 1094,19
7	2679,00	2134,60	196,311	1769,54 - 2499,65
8	1918,00	1612,37	140,450	1351,20 - 1873,55
9	1511,00	1039,78	172,289	719,40 - 1360,16
10	2136,00	1984,97	174,437	1660,59 - 2309,34
11		1504,23	138,407	1246,85 - 1761,60
12		2118,30	193,775	1757,96 - 2478,63

Das zweite Fenster zeigt die Regressionsgerade, beobachtete Werte und den Prognostriecher.

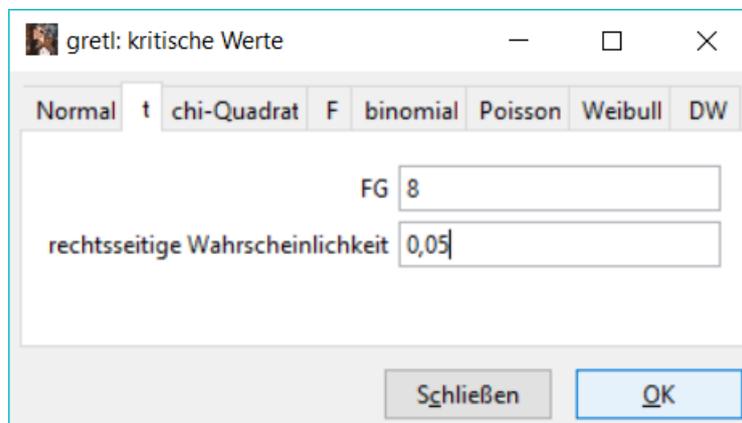


2.5 t-Test

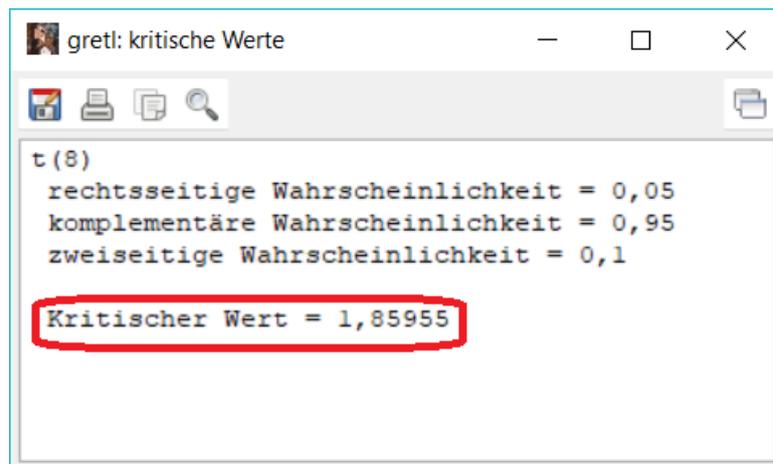
Neben Konfidenzintervallen und Prognosen spielen besonders Hypothesentests eine wichtige Rolle in der Arbeit mit dem Linearen Modell. Mit einem t-Test lässt sich beispielsweise überprüfen ob β_1 , die marginale Konsumquote, den Wert 0,8 unterschreitet. Im Skript findet sich dazu die Aufforderung, auf Basis der Modellschätzung $H_0 : \beta_1 \geq 0,8$ gegen $H_1 : \beta_1 < 0,8$ mit $\alpha = 0,05$ zu überprüfen. Leider kann dieser Test nicht unmittelbar in *gretl* durchgeführt werden. Stattdessen muss der Wert der Teststatistik zunächst wie gewohnt berechnet werden. Er beträgt $t = (\hat{\beta}_1 - 0,8) / \hat{\sigma}_{\hat{\beta}_1} = (0,741 - 0,8) / 0,164 = -0,362$. Der kritische Wert k kann nun über das *pyqrs*-Programm oder in *gretl* über *Werkzeuge* → *Statistische Tabellen* bestimmt werden.



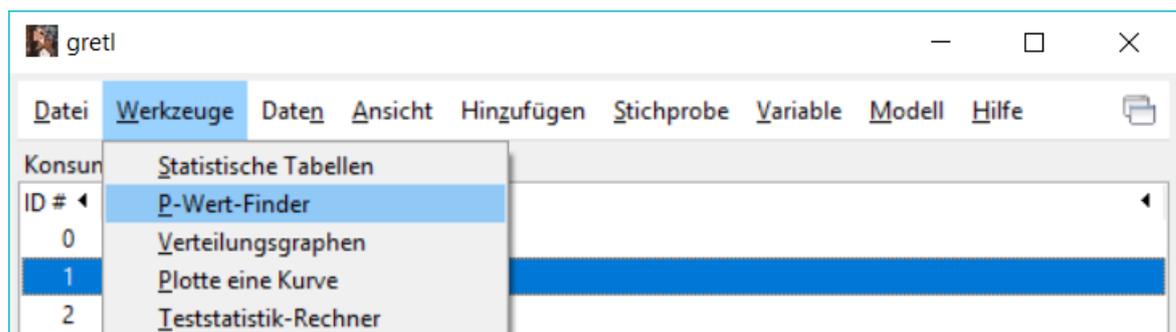
Dazu wird die t-Verteilung ausgewählt und die Anzahl der Freiheitsgrade sowie die rechtsseitige Wahrscheinlichkeit eingetragen.



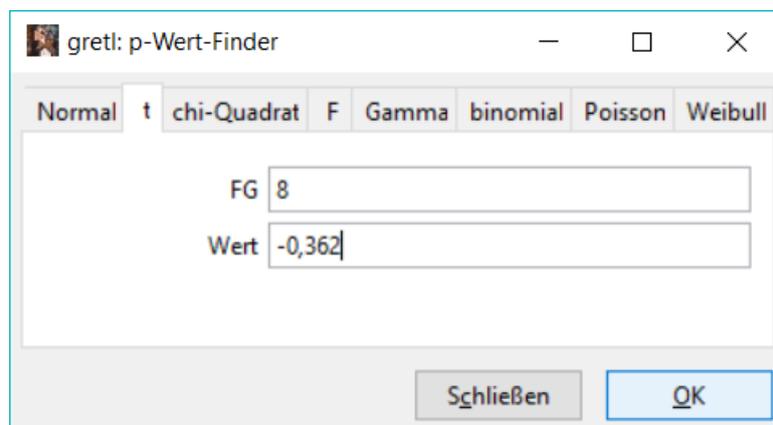
Der kritische Wert $k = 1,86$ wird ausgegeben. Ein Vergleich mit dem empirischen t-Wert $t = -0,362$ ergibt $|t| < k$. Die Nullhypothese kann nicht verworfen werden, die marginale Konsumquote ist nicht signifikant kleiner als 0,8.



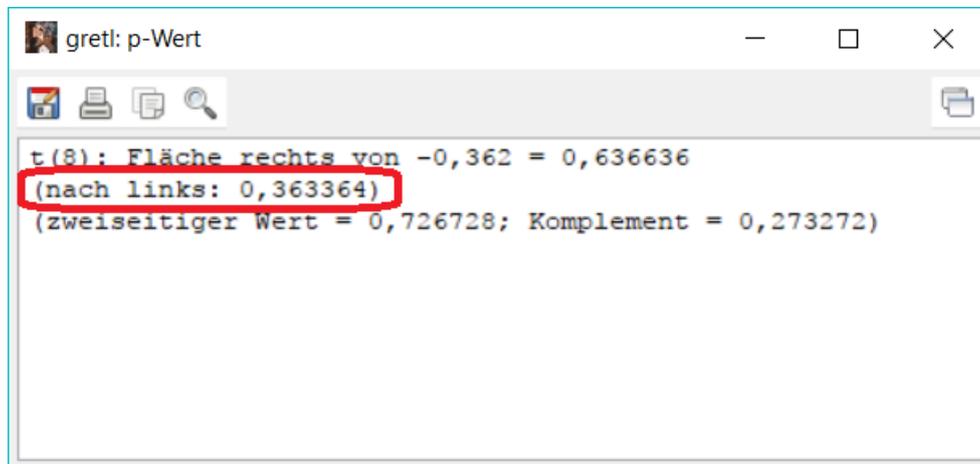
Zusätzlich kann auch das marginale Signifikanzniveau p über *Werkzeuge* → *P-Wert-Finder* bestimmt werden (wiederum als Alternative zur Verwendung des pyqrs-Programms).



Dafür wird zunächst wieder die t-Verteilung gewählt und die Freiheitsgrade sowie der empirische t-Wert eingetragen.



Da der beschriebene Test linksseitig ist, ist der Wert *nach links* der richtige. Der p-Wert beträgt $p = 0,363$, d. h. der beobachtete t-Wert (bzw. die beobachtete Abweichung des Schätzwerts $\hat{\beta}_1 = 0,741$ vom hypothetischen Wert 0,8) ist nicht außergewöhnlich: Wäre $\beta_1 = 0,8$, so würde man den beobachteten t-Wert oder eine noch deutlich negativere Teststatistik mit einer Wahrscheinlichkeit von 36,3% beobachten. Die Daten sprechen also nicht offensichtlich gegen die Annahme, die marginale Konsumquote sei 0,8. Eine ausführliche Darstellung dieses t-Tests findet sich im Vorlesungsskript zur „Einführung in die Ökonometrie“.



```
t(8): Fläche rechts von -0,362 = 0,636636
(nach links: 0,363364)
(zweiseitiger Wert = 0,726728; Komplement = 0,273272)
```

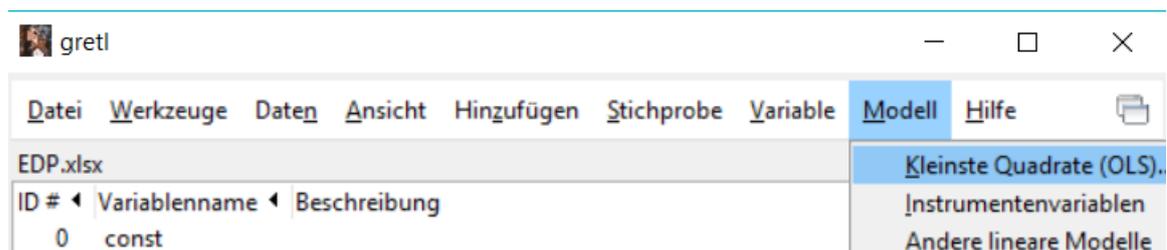
Das Vorgehen zum t-Test erscheint recht umständlich. In Abschnitt ?? wird jedoch gezeigt, wie sich mit Hilfe des F-Tests indirekt auch ein einseitiger t-Test durchführen lässt.

3 Das multiple Lineare Modell

In diesem Abschnitt wird das Vorgehen aus Abschnitt ?? erweitert und auf das multiple Modell übertragen. Außerdem werden weitere Funktionen von *gretl* vorgestellt. Grundlage sind wieder die Daten des EDP-Beispiels, die gemäß den in Abschnitt ?? vorgestellten Optionen zunächst geladen werden müssen.

3.1 Modell schätzen

Über *Modell* → *Kleinste Quadrate (OLS)* lässt sich auch das multiple Lineare Modell mit der KQ-Methode schätzen.



Dazu müssen, im Gegensatz zum Linearen Modell mit einem Regressor, neben der abhängigen Variable mehrere unabhängige Variablen festgelegt werden. In diesem Fall ist die Zahl der verkauften Einheiten *Units* die abhängige Variable, die durch den Preis *Price*, Werbeausgaben *AE* und Vertriebsausgaben *PSE* erklärt werden.

Das Lineare Modell lautet also

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \beta_3 x_{3,t} + u_t$$

mit $y_t = Units_t$, $x_{1,t} = Price_t$, $x_{2,t} = AE_t$, $x_{3,t} = PSE_t$, und u_t als Störgröße.

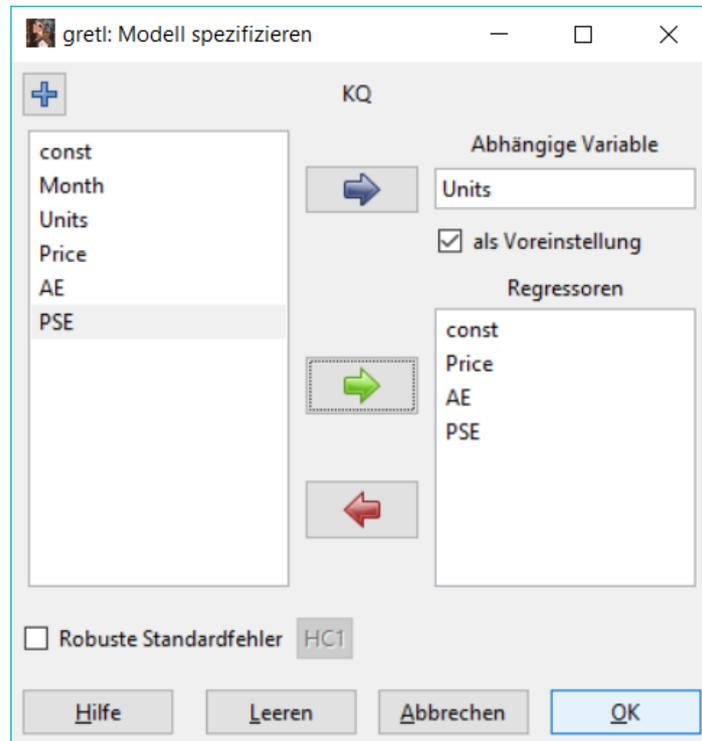
In der *gretl*-Notation bedeutet dies

$$Units_t = const + \beta_{Price} Price_t + \beta_{AE} AE_t + \beta_{PSE} PSE_t + u_t$$

bzw.

$$Units_t = \beta_{[1]} + \beta_{[2]} Price_t + \beta_{[3]} AE_t + \beta_{[4]} PSE_t + u_t \quad .$$

Beachten Sie, dass die Nummerierung der Koeffizienten in *gretl* mit dem Startwert [1] für das Absolutglied beginnt, es gilt also in Bezug zur „konventionellen“ Notation: $\beta_0 = \beta_{[1]}$, $\beta_1 = \beta_{[2]}$ etc.!



Im Regressionsoutput werden wieder die Schätzkoeffizienten der einzelnen Variablen, deren Standardfehler, empirische t-Werte und p-Werte angegeben. Weiterhin finden sich modellbezogene Größen wie das Bestimmtheitsmaß R^2 , die Summe der quadrierten Residuen (RSS) und der F-Wert des „Test of overall significance“. (Das ist ein F-Test der Hypothese, dass alle Steigungskoeffizienten des Modells null sind: $H_0 : \beta_{Price} = \beta_{AE} = \beta_{PSE} = 0$. Der F-Test wird im folgenden Abschnitt ?? beschrieben.)

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	-117,513	333,526	-0,3523	0,7337
Price	-0,296478	0,101950	-2,908	0,0196 **
AE	0,0359758	0,0139494	2,579	0,0327 **
PSE	0,0662060	0,0143722	4,607	0,0017 ***
Mittel d. abh. Var.	2020,833	Stdabw. d. abh. Var.	607,3259	
Summe d. quad. Res.	122854,1	Stdfehler d. Regress.	123,9224	
R-Quadrat	0,969720	Korrigiertes R-Quadrat	0,958365	
F(3, 8)	85,40075	P-Wert (F)	2,04e-06	
Log-Likelihood	-72,43034	Akaike-Kriterium	152,8607	
Schwarz-Kriterium	154,8003	Hannan-Quinn-Kriterium	152,1426	

Wie oben am Beispiel des Konsummodells beschrieben lassen sich nun über *Analyse* → *Konfidenzintervalle für Koeffizienten* Konfidenzintervalle berechnen:

The screenshot shows the main window of gretl titled 'gretl: Modell 1'. The 'Analyse' menu is open, and the option 'Konfidenzintervalle für Koeffizienten' is highlighted. The background displays regression results for a model with dependent variable 'Units'.

	Koeffizient	Std.-f		
const	-117,513	333,52	Stdfehler d. Regress.	123,9224
Price	-0,296478	0,10	Korrigiertes R-Quadrat	0,958365
AE	0,0359758	0,01	P-Wert (F)	2,04e-06
PSE	0,0662060	0,01	Akaike-Kriterium	152,8607
Mittel d. abh. Var.	2020,833		Hannan-Quinn-Kriterium	152,1426
Summe d. quad. Res.	122854,1			
R-Quadrat	0,969720			
F(3, 8)	85,40075			
Log-Likelihood	-72,43034			
Schwarz-Kriterium	154,8003			

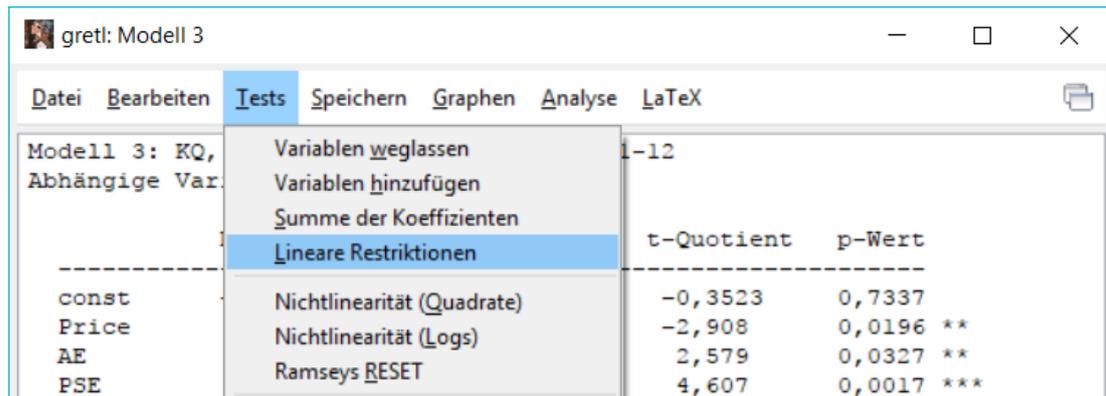
The screenshot shows the 'gretl: Konfidenzintervalle der Koeffizienten' window. The significance level is set to $\alpha = 0,025$. The critical value is $t(8, 0,025) = 2,306$. The 95% confidence interval for the Price coefficient is highlighted with a red box.

VARIABLE	KOEFFIZIENT	95% KONFIDENZ-INTERVALL	
const	-117,513	-886,626	651,599
Price	-0,296478	-0,531575	-0,0613806
AE	0,0359758	0,00380837	0,0681433
PSE	0,0662060	0,0330635	0,0993484

untere Grenze obere Grenze

3.2 F-Test

Ein F-Test lässt sich in *gretl* über *Tests* → *Lineare Restriktionen* durchführen.



The screenshot shows the 'gretl: Modell 3' window. The 'Tests' menu is open, highlighting 'Lineare Restriktionen'. Below the menu, a table displays the results of several tests. The table has two columns: 't-Quotient' and 'p-Wert'. The rows are: 'const' (t-Quotient: -0,3523, p-Wert: 0,7337), 'Price' (t-Quotient: -2,908, p-Wert: 0,0196 **), 'AE' (t-Quotient: 2,579, p-Wert: 0,0327 **), and 'PSE' (t-Quotient: 4,607, p-Wert: 0,0017 ***).

	t-Quotient	p-Wert
const	-0,3523	0,7337
Price	-2,908	0,0196 **
AE	2,579	0,0327 **
PSE	4,607	0,0017 ***

Dabei ist die Überprüfung mehrerer Restriktionen die Regel. Im Fall nur einer Restriktion entspricht der F-Test einem zweiseitigen t-Test.

3.2.1 F-Test mit einer Restriktion

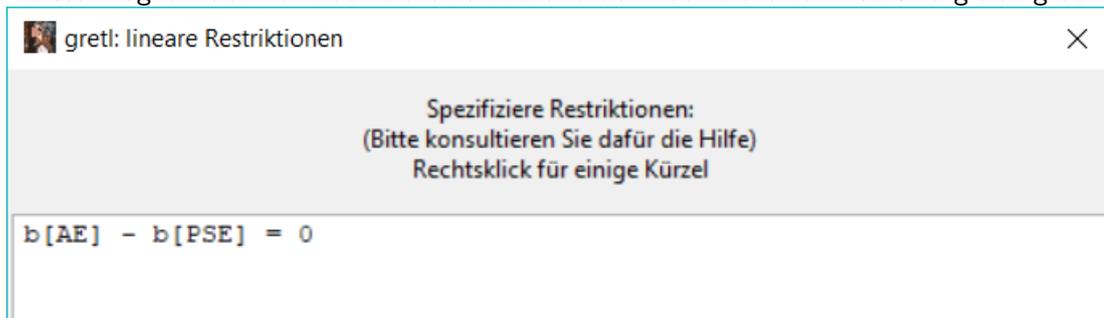
Wird der F-Test wie oben beschrieben aufgerufen, so müssen im sich öffnenden Fenster die Restriktionen spezifiziert werden. Hier soll getestet werden, ob der Einfluss von Werbeausgaben (AE) und Vertriebsaufwendungen (PSE) auf die verkauften Einheiten (Units) gleich groß ist. Die Nullhypothese lautet daher $H_0 : \beta_{AE} = \beta_{PSE}$, die Gegenhypothese ist beim F-Test stets festgelegt auf $H_1 : H_0$ gilt nicht.

Allerdings erlaubt *gretl* nur einen numerischen Wert auf der rechten Seite des Gleichheitszeichens, daher muss die Nullhypothese umformuliert werden zu:

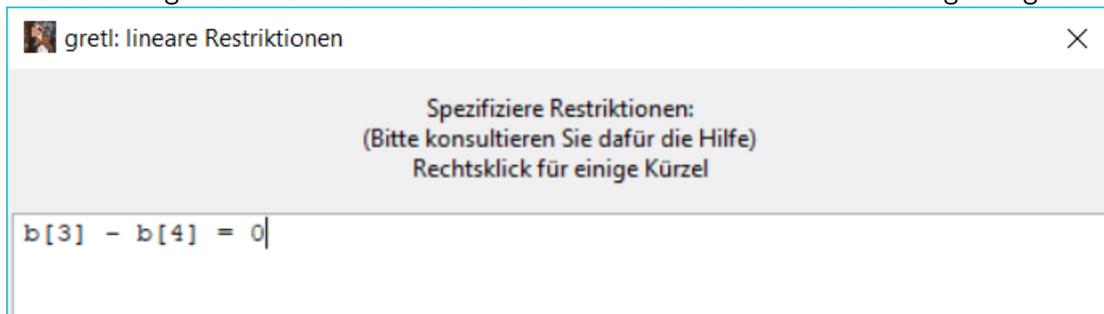
$$H_0 : \beta_{AE} - \beta_{PSE} = 0 \quad .$$

In *gretl* kann diese Hypothese auf zwei verschiedene Arten eingetragen werden, beide sind in den folgenden Abbildungen zu sehen. Nachdem eine dieser Alternativen eingetragen ist, wird die Restriktion mit *OK* bestätigt.

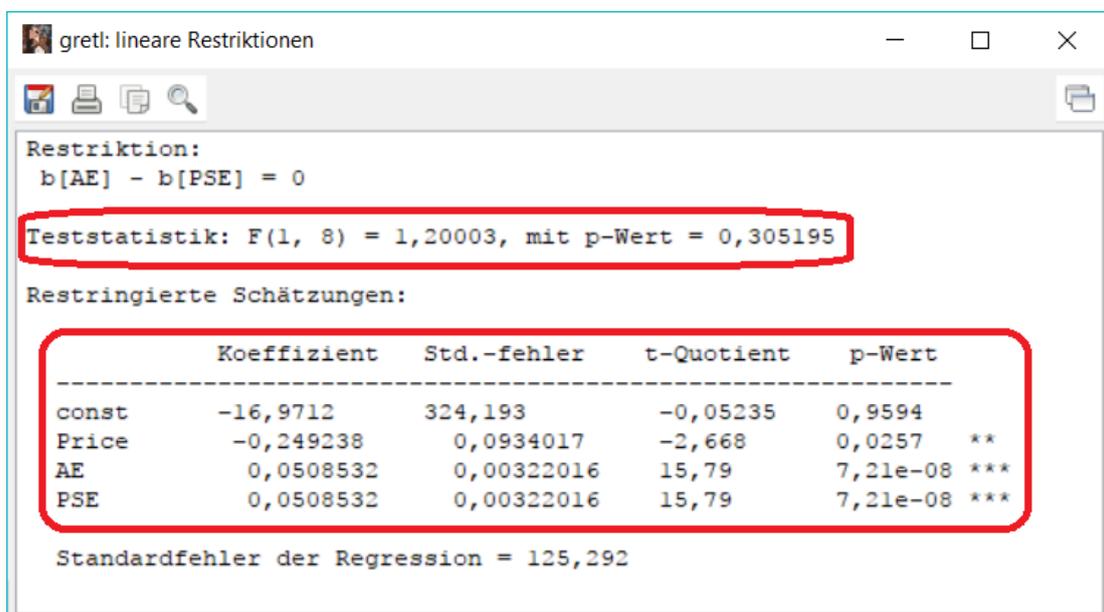
Erste Möglichkeit: Der Koeffizient von AE und der Koeffizient von PSE sind gleich groß:



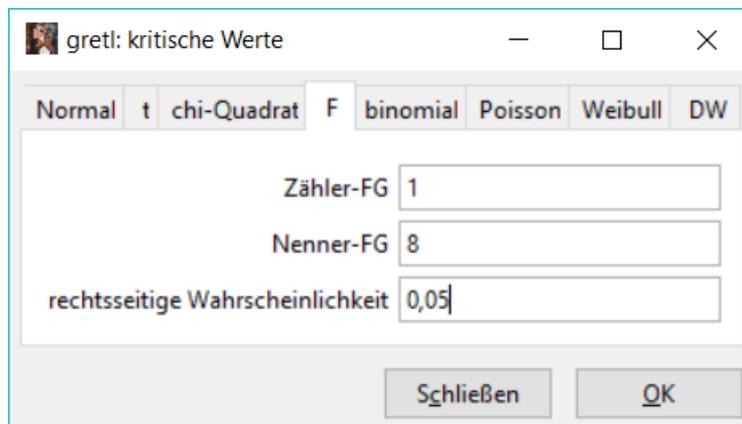
Zweite Möglichkeit: Der dritte und der vierte Koeffizient des Modells sind gleich groß:



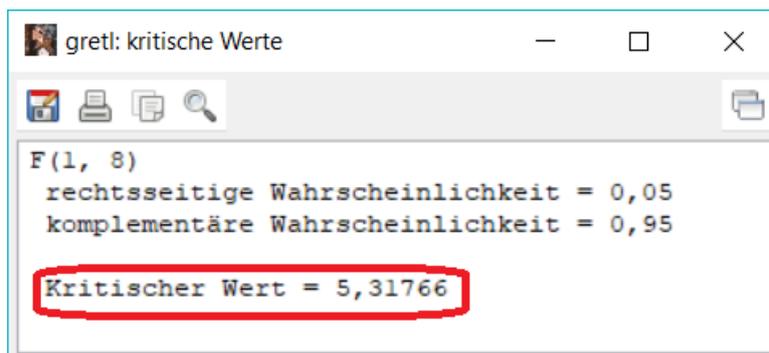
Es öffnet sich ein neues Fenster mit dem Output des F-Tests. Hier lassen sich die Teststatistik und der zugehörige p-Wert ablesen. Zusätzlich ist das Regressionsergebnis des restringierten Modells gezeigt.



Ein kritischer Wert kann ähnlich wie beim t-Test im Hauptfenster über *Werkzeuge* → *Statistische Tabellen* ermittelt werden. (Allerdings zeigt der p-Wert bereits, dass die Nullhypothese zu den üblichen Testniveaus nicht verworfen werden kann.)
 Im Dialogfenster wird die F-Verteilung ausgewählt und Zähler- und Nennerfreiheitsgrade sowie das Testniveau α angegeben.

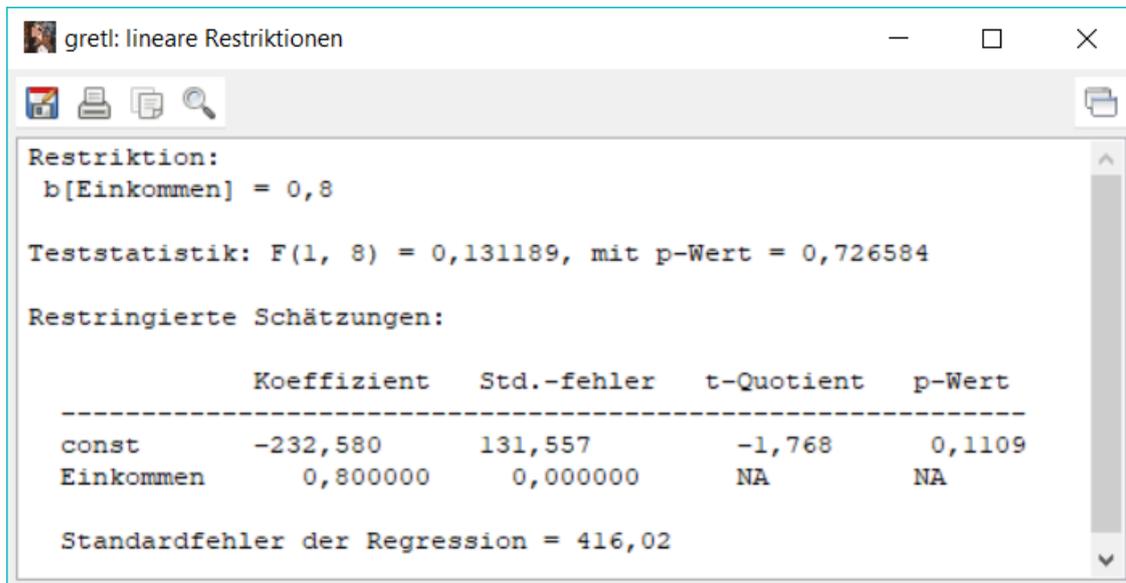


Der kritische Wert wird ausgegeben. Da $F = 1,20003 < k = 5,31766$ kann die Nullhypothese nicht abgelehnt werden.



Im Konsumbeispiel hätte der t-Test aus Abschnitt ??, $H_0 : \beta_1 \geq 0,8$ gegen $H_1 : \beta_1 < 0,8$ demnach durchgeführt werden können, indem man den F-Test aufruft und die Restriktion überprüft (ACHTUNG: Bei Eingabe der Restriktion einen Dezimalpunkt an Stelle des Dezimalkommas verwenden, sonst erhält man eine Fehlermeldung!)

Im Output wird ein p-Wert von $p=0,7266$ ausgewiesen:



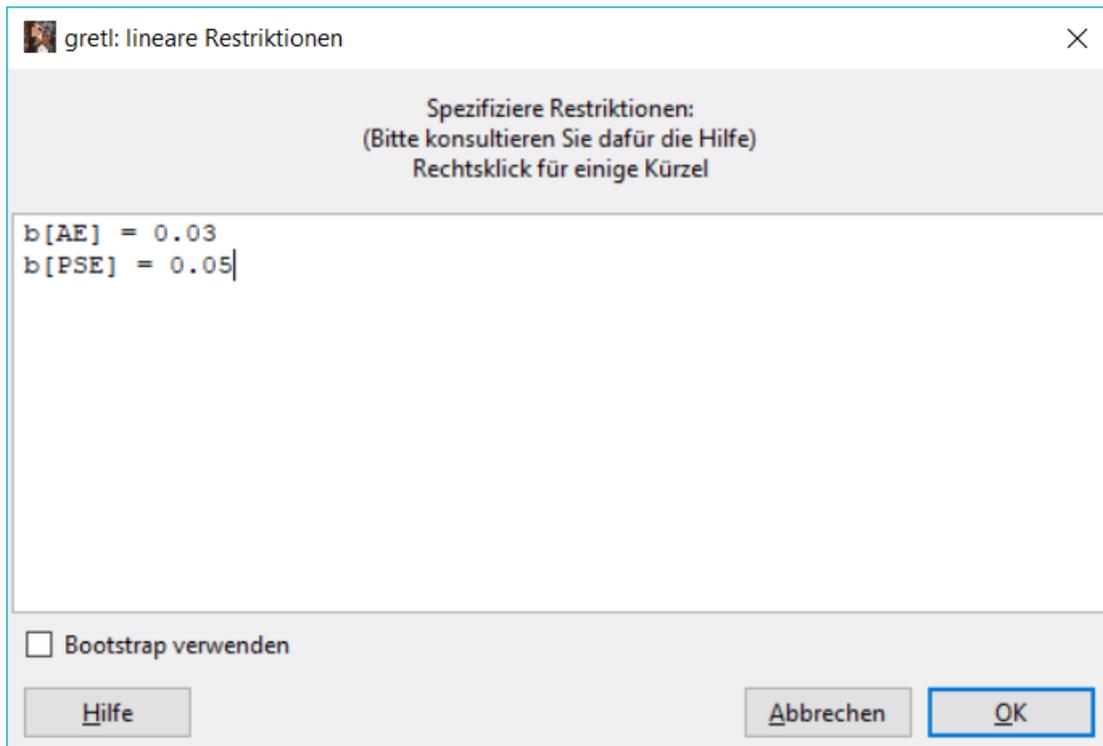
Dieser p-Wert bezieht sich auf die automatisch generierte, zweiseitige Alternativhypothese $H_1 : H_0 \text{ gilt nicht}$ bzw. $H_1 : \beta_1 \neq 0,8$. Will man dagegen nur einseitig testen (nämlich gegen $H_1 : \beta_1 < 0,8$), so ist der p-Wert in diesem Beispiel zu halbieren. $p/2 = 0,7266/2 = 0,363$ deckt sich mit dem in Abschnitt ?? ermittelten Ergebnis.

3.2.2 F-Test mit zwei und mehr Restriktionen

Der F-Test erlaubt, im Gegensatz zum t-Test, auch das gleichzeitige Testen von mehreren Restriktionen. Als Beispiel sollen hier die zwei Restriktionen $\beta_{AE} = 0,03$ und $\beta_{PSE} = 0,05$ gleichzeitig getestet werden. Die Hypothesen lauten daher $H_0 : \beta_{AE} = 0,03 \text{ und } \beta_{PSE} = 0,05$ und $H_1 : H_0 \text{ gilt nicht}$. Die Gegenhypothese umfasst also folgende Fälle:

$$\beta_{AE} \neq 0,03 \text{ oder } \beta_{PSE} \neq 0,03 \text{ (oder } \beta_{AE} \neq 0,03 \text{ und } \beta_{AE} \neq 0,03).$$

Um den Test durchzuführen, werden nun beide Restriktionen unter *Tests* → *Lineare Restriktionen* eingetragen.



Anschließend wird das Modell geschätzt. Die Koeffizienten für AE und PSE zeigen, dass die Restriktionen umgesetzt wurden. Ein Blick auf den p-Wert zur F-Statistik macht deutlich, dass die Nullhypothese, die die beiden Parameter gleichzeitig auf die angegebenen Werte restringiert, abgelehnt werden kann.

```

gretl: lineare Restriktionen
-----
Restriktion angewendet
1: b[AE] = 0,03
2: b[PSE] = 0,05

Teststatistik: F(2, 8) = 6,04421, mit p-Wert = 0,0251523

Restringierte Schätzungen:

-----
                Koeffizient   Std.-fehler   t-Quotient   p-Wert
-----
const          374,049        413,947       0,9036       0,3874
Price          -0,231669                0,129395     -1,790       0,1037
AE              0,0300000                0,000000       NA          NA
PSE              0,0500000                0,000000       NA          NA

Standardfehler der Regression = 175,64
  
```

Das gleichzeitige Testen von mehreren Restriktionen wird häufig zum Ausschluss von Variablen benötigt, man spricht dabei von „Nullrestriktionen“. Auch wenn ein Test der Hypothese $H_0 : \beta_{[AE]} = \beta_{[PSE]} = 0$ im vorliegenden Beispiel nicht sinnvoll ist, soll er beispielhaft durchgeführt werden, um das Prinzip der *gretl*-Optionen zu veranschaulichen. Die beiden Restriktionen können wie zuvor eingegeben und das restringiert Modell geschätzt werden:

```

gretl: lineare Restriktionen
-----
Spezifiziere Restriktionen:
(Bitte konsultieren Sie dafür die Hilfe)
Rechtsklick für einige Kürzel

b[AE] = 0
b[PSE] = 0
  
```

Der p-Wert ist mit $8 \cdot 10^{-7}$ bzw. 0,0000008 sehr klein, daher kann die Nullhypothese abgelehnt werden, der Einfluss von Vertriebs- und Werbeausgaben ist nicht gemeinsam gleich Null. (Gemäß dem p-Wert wäre das Schätzergebnis bei Gültigkeit der Nullhypothese nahezu unmöglich.).

gretl: lineare Restriktionen

Restriktion angewendet
 1: b[AE] = 0
 2: b[PSE] = 0

Teststatistik: $F(2, 8) = 128,067$, mit p-Wert = $8,41509e-007$

Restringierte Schätzungen:

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	2096,22	1501,01	1,397	0,1928
Price	-0,0237449	0,469201	-0,05061	0,9606
AE	0,000000	0,000000	NA	NA
PSE	0,000000	0,000000	NA	NA

Standardfehler der Regression = 636,887

Da Nullrestriktionen im Zuge der Modellspezifikation regelmäßig überprüft werden müssen, bietet *gretl* über *Tests* → *Variablen weglassen* eine weitere Möglichkeit, diese Tests schnell durchzuführen.

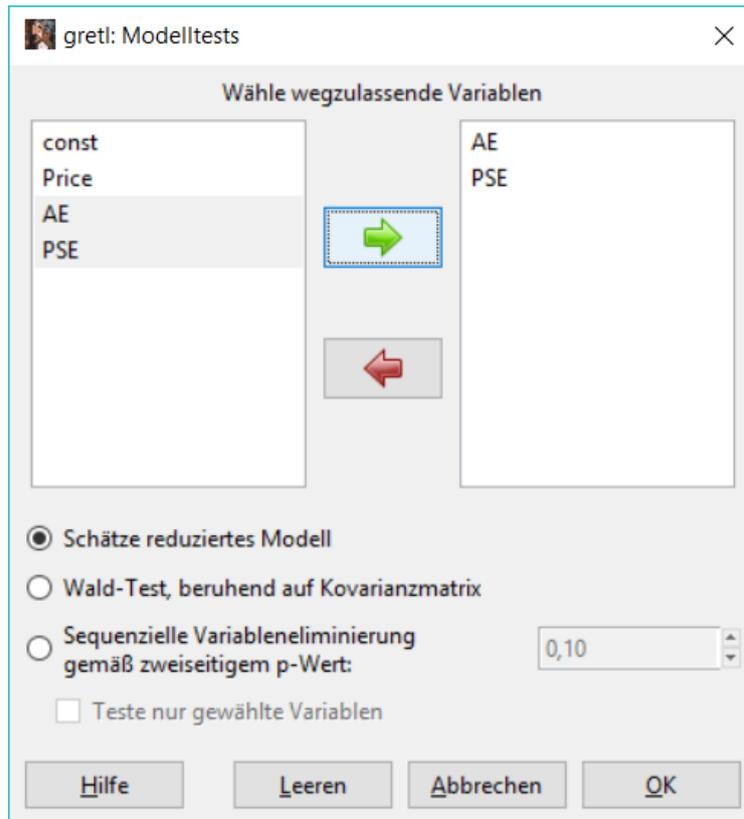
gretl: Modell 1

Datei Bearbeiten **Tests** Speichern Graphen Analyse LaTeX

Modell 1: KQ, Abhängige Var: 1-12

	t-Quotient	p-Wert

Hier müssen die beiden wegzulassenden Variablen ausgewählt werden, in diesem Fall AE und PSE.



Der daraufhin entstehende Output ist etwas anders aufgebaut, die Ergebnisse sind aber identisch:

```

gretl: Modell 2
Datei Bearbeiten Tests Speichern Graphen Analyse LaTeX
Test von Modell 1:
Nullhypothese: Die Regressionskoeffizienten sind Null für die Variablen
AE, PSE
Teststatistik: F(2, 8) = 128,067, p-Wert 8,41509e-007
Das Weglassen von Variablen verbesserte 0 von 3 Informationskriterien.
Modell 2: KQ, benutze die Beobachtungen 1-12
Abhängige Variable: Units
-----
Koeffizient   Std.-fehler   t-Quotient   p-Wert
-----
const         2096,22      1501,01      1,397        0,1928
Price        -0,0237449   0,469201    -0,05061     0,9606
Mittel d. abh. Var.   2020,833   Stdabw. d. abh. Var.   607,3259
Summe d. quad. Res.   4056253   Stdfehler d. Regress.   636,8872
R-Quadrat          0,000256   Korrigiertes R-Quadrat -0,099718
F(1, 10)           0,002561   P-Wert (F)           0,960635
Log-Likelihood     -93,41244   Akaike-Kriterium     190,8249
Schwarz-Kriterium   191,7947   Hannan-Quinn-Kriterium 190,4658

```

3.3 Kreuzeffekte der Regressoren

Es ist plausibel anzunehmen, dass die Preissensitivität der Kunden nicht konstant ist, sondern von den Werbeausgaben abhängt. Je größer die Werbeausgaben, desto geringer die Preissensitivität.

Das ursprüngliche Modell lautet

$$Units_t = const + \beta_{Price}Price_t + \beta_{AE}AE_t + \beta_{PSE}PSE_t + u_t \quad .$$

Ist nun die Preissensitivität abhängig von den Werbeausgaben, bedeutet das

$$\beta_{Price,t} = \gamma_0 + \gamma_1AE_t \quad .$$

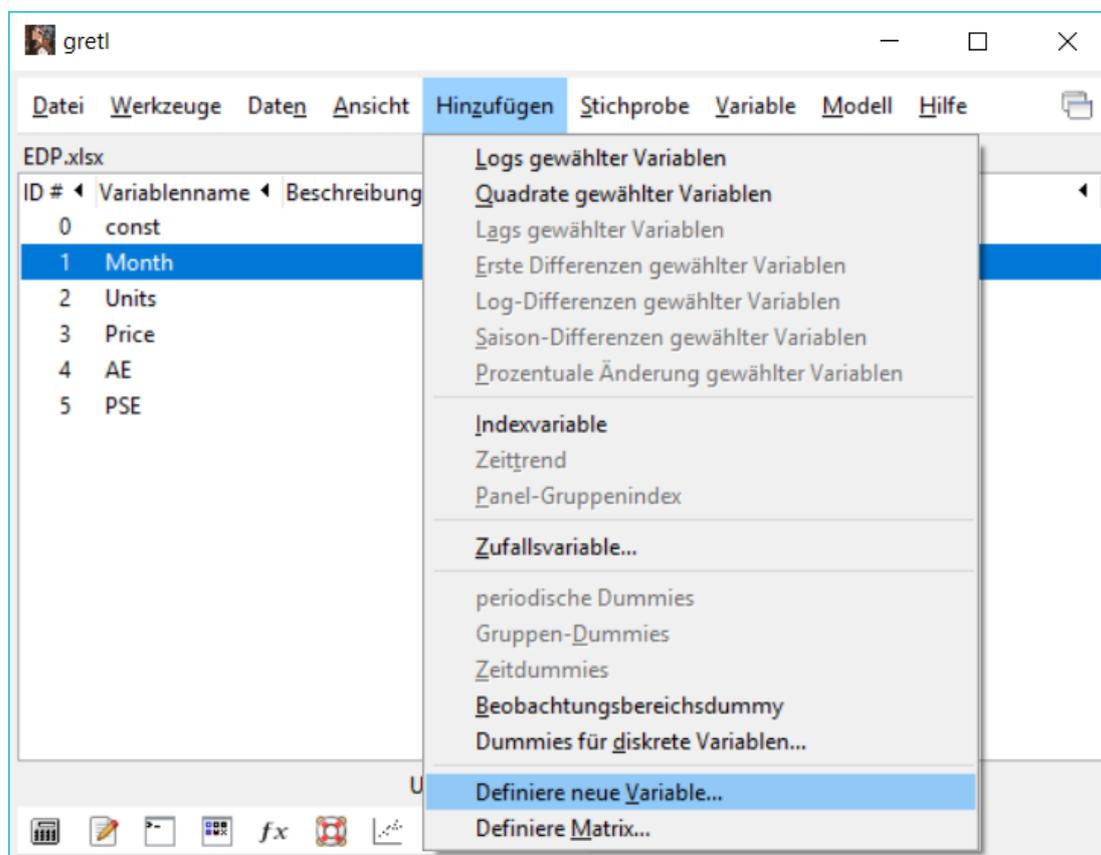
Ersetzen des konstanten Koeffizienten β_{Price} im ursprünglichen Linearen Modell durch $\beta_{Price,t} = \gamma_0 + \gamma_1AE_t$ ergibt

$$Units_t = const + (\gamma_0 + \gamma_1AE_t)Price_t + \beta_{AE}AE_t + \beta_{PSE}PSE_t + u_t \quad .$$

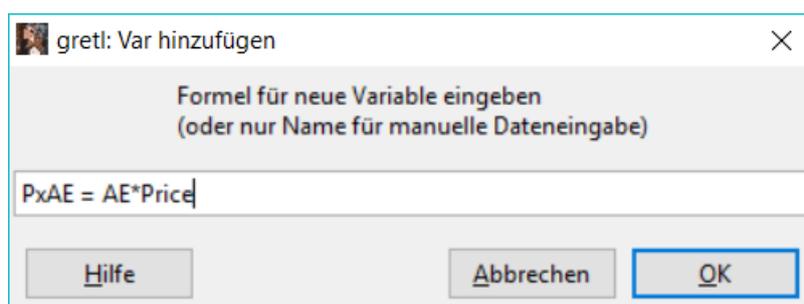
bzw.

$$Units_t = const + \gamma_0Price_t + \beta_{AE}AE_t + \beta_{PSE}PSE_t + \gamma_1P_t \times AE_t u_t + u_t$$

mit $Price_t \cdot AE_t = P_t \times AE_t$ als neuem Regressor. Um ein solches Modell mit Kreuzeffekt zu schätzen, muss über *Hinzufügen* → *Definiere neue Variable* der neue Regressor hinzugefügt werden.



Im dem sich nun öffnenden Fenster muss die Formel für den neuen Regressor eingetragen werden.



Anschließend wird das Modell mit der neuen Variable geschätzt. Es ergibt sich folgender Regressionsoutput:

The screenshot shows the gretl software interface with a window titled 'gretl: Modell 2'. The main content area displays the following regression output:

```

Modell 2: KQ, benutze die Beobachtungen 1-12
Abhängige Variable: Units

-----
                Koeffizient      Std.-fehler      t-Quotient      p-Wert
-----
const          1375,75           1455,29           0,9453           0,3760
Price          -0,811685                    0,499299          -1,626           0,1481
AE             -0,0392996                   0,0727663         -0,5401          0,6059
PSE            0,0729079                    0,0156269          4,666           0,0023 ***
PxAE           2,21494e-05                   2,10194e-05         1,054           0,3270

Mittel d. abh. Var.      2020,833      Stdabw. d. abh. Var.      607,3259
Summe d. quad. Res.     106034,0      Stdfehler d. Regress.     123,0761
R-Quadrat                0,973866      Korrigiertes R-Quadrat    0,958932
F(4, 7)                  65,21209      P-Wert (F)                 0,000013
Log-Likelihood           -71,54691     Akaike-Kriterium          153,0938
Schwarz-Kriterium        155,5184      Hannan-Quinn-Kriterium    152,1962

Abgesehen von Konstante war p-Wert am höchsten für Variable 4 (AE)

```

Der Kreuzeffekt hat keinen signifikanten Einfluss, es kann weiterhin von einer konstanten Preissensitivität β_{Price} ausgegangen werden.

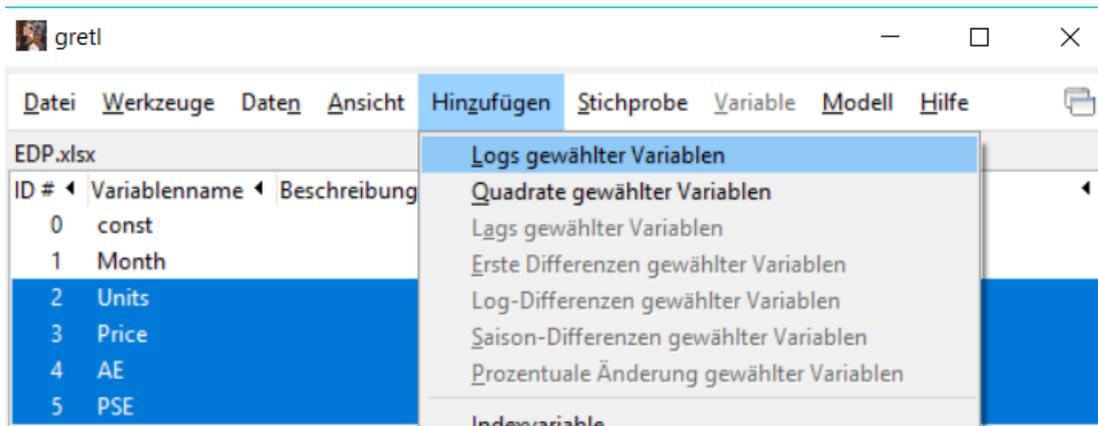
3.4 Alternative Kurvenformen

3.4.1 Logarithmische Spezifikation

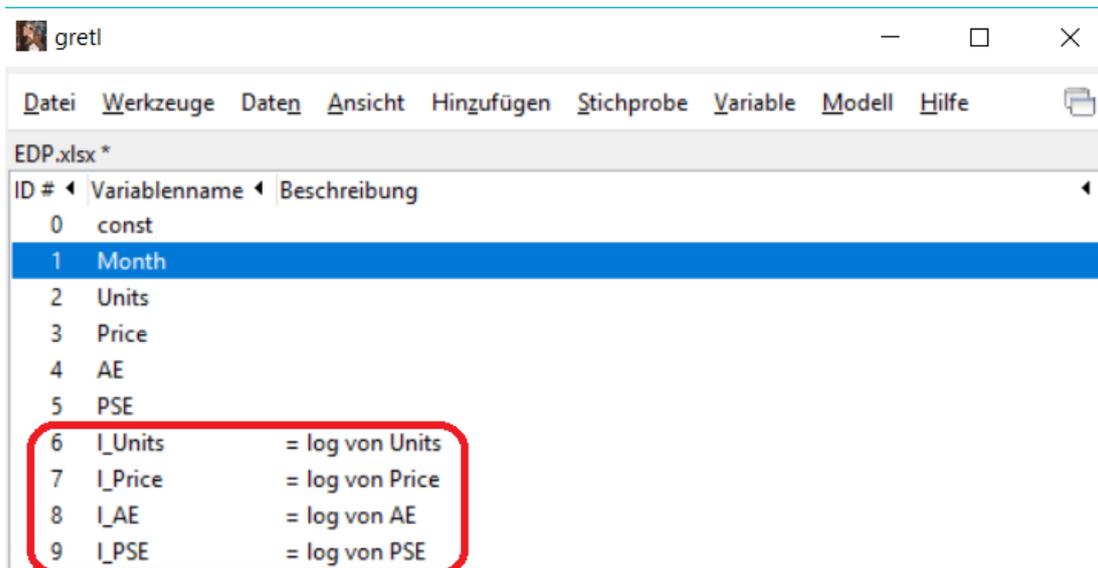
Mit *gretl* ist es leicht möglich, das EDP-Beispiel auch in der doppelt-logarithmischen Form

$$\ln Units_t = const + \beta_{Price} \ln Price_t + \beta_{AE} \ln AE_t + \beta_{PSE} \ln PSE_t + u_t$$

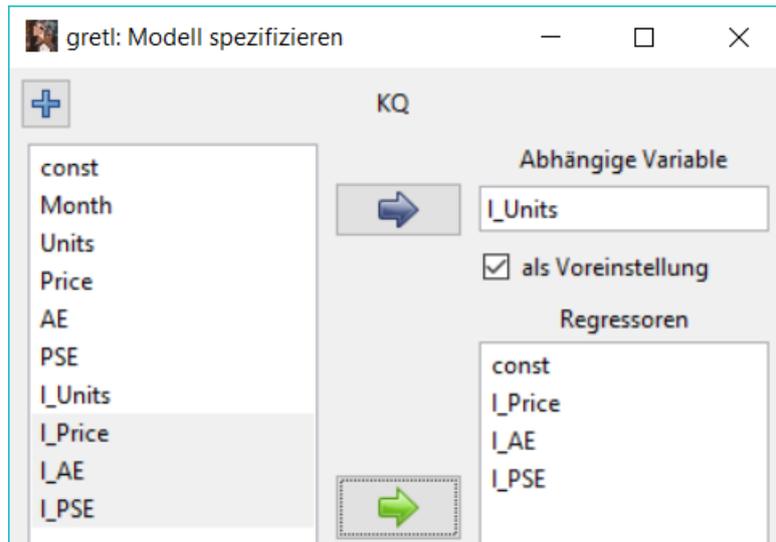
zu schätzen. Dazu müssen zunächst alle relevanten Variablen ausgewählt und über *Hinzufügen* → *Logs gewählter Variablen* logarithmiert werden.



Die logarithmierten Variablen tauchen nun auch in der Übersicht auf und sind am vorangestellten *l_* zu erkennen.



Das Modell kann nun wie gewohnt, diesmal mit den logarithmierten Variablen, geschätzt werden.



Es ergibt sich folgender Regressionsoutput:

The screenshot shows the 'gret! Modell 1' output window. The title bar reads 'gret! Modell 1'. The menu bar includes 'Datei', 'Bearbeiten', 'Tests', 'Speichern', 'Graphen', 'Analyse', and 'LaTeX'. The main text reads 'Modell 1: KQ, benutze die Beobachtungen 1-12' and 'Abhängige Variable: l_Units'. Below this is a table of regression coefficients and statistics.

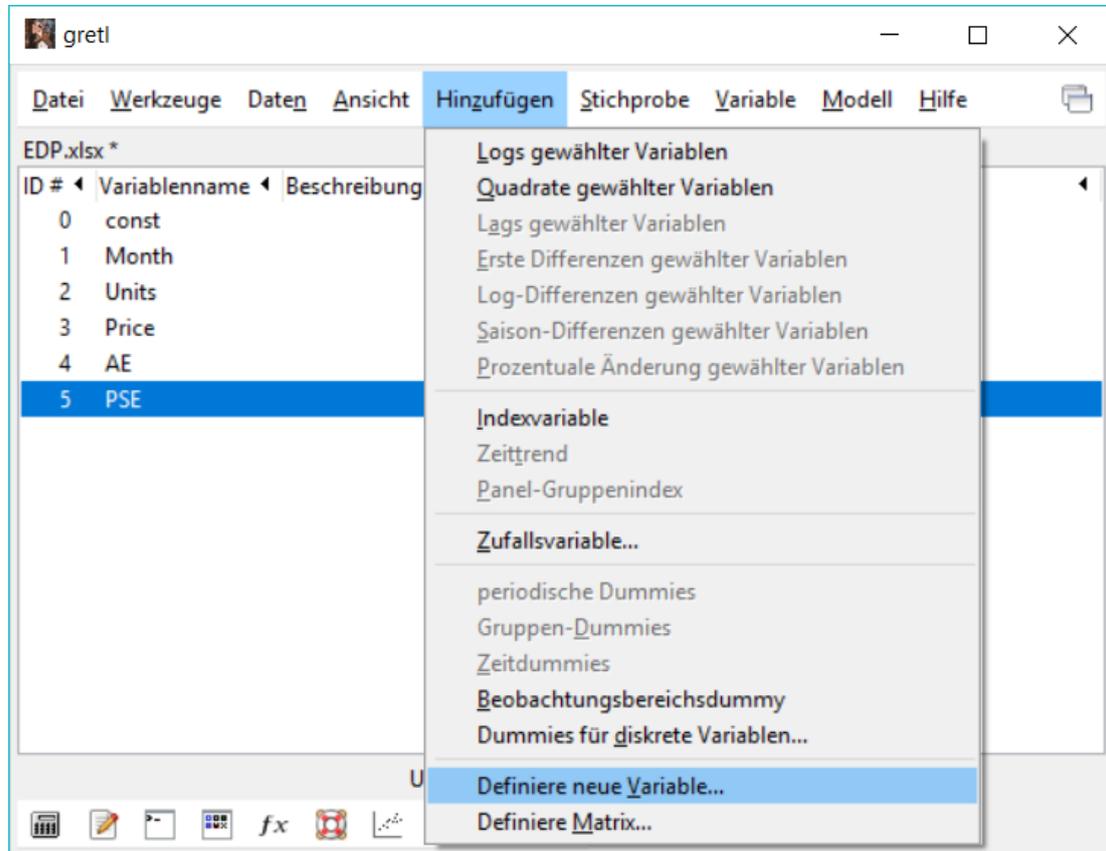
	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	-4,66416	1,73790	-2,684	0,0278 **
l_Price	-0,424409	0,184923	-2,295	0,0509 *
l_AE	0,443414	0,162807	2,724	0,0261 **
l_PSE	1,07737	0,303144	3,554	0,0075 ***

Mittel d. abh. Var.	7,565077	Stdabw. d. abh. Var.	0,328031
Summe d. quad. Res.	0,038552	Stdfehler d. Regress.	0,069419
R-Quadrat	0,967430	Korrigiertes R-Quadrat	0,955216
F(3, 8)	79,20704	P-Wert (F)	2,73e-06
Log-Likelihood	17,41665	Akaike-Kriterium	-26,83330
Schwarz-Kriterium	-24,89367	Hannan-Quinn-Kriterium	-27,55142

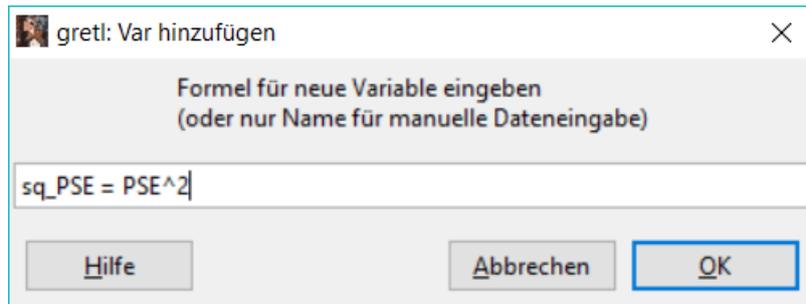
Die Schätzkoeffizienten sind nun als Elastizitäten interpretierbar. Der Koeffizient des logarithmierten Preises stellt die Preiselastizität der Nachfrage dar und besagt, dass eine 1%-ige Preiserhöhung gemäß der Schätzung einen Absatzrückgang von 0,42% erwarten lässt.

3.4.2 Quadratische Spezifikation

Auch ein quadratischer Zusammenhang, beispielsweise im Fall der Vertriebsausgaben, ist denkbar. Analog zur Logarithmierung lassen sich Variablen über *Hinzufügen* → *Quadrate gewählter Variablen* quadrieren. Alternativ ist das auch über *Hinzufügen* → *Definiere neue Variable* möglich. Letzteres ist die allgemeiner einsetzbare Variante und soll deshalb hier vorgestellt werden.

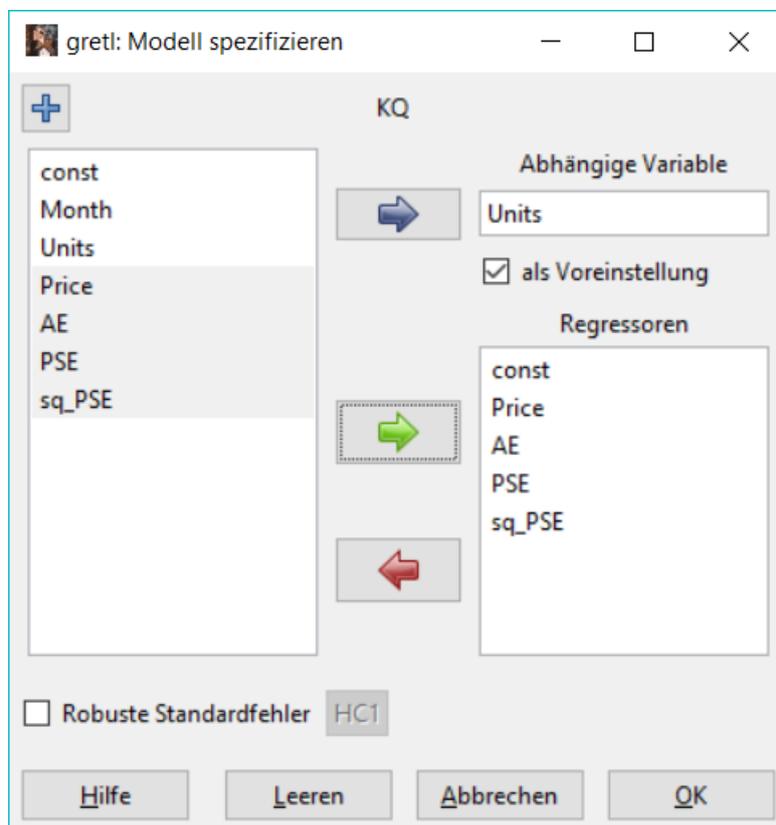


Es öffnet sich ein neues Fenster, in dem die Formel für die neue Variable eingetragen werden muss.

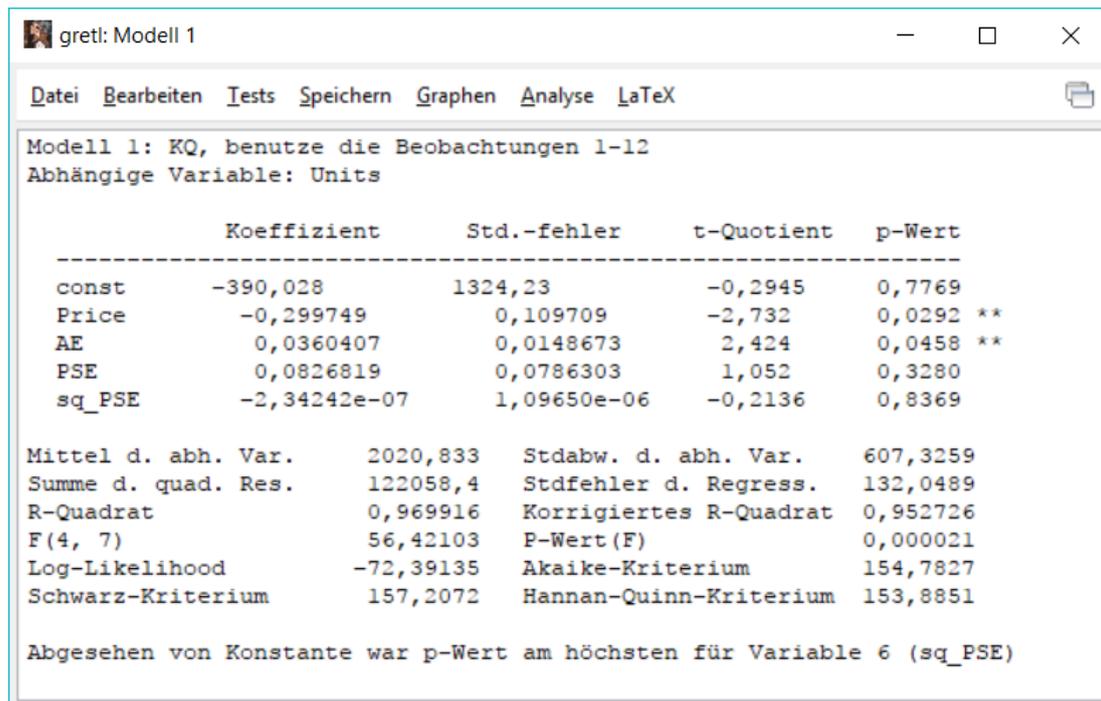


Dabei ist zu unbedingt zu beachten, dass *gret!* bei der Eingabe von \wedge das nächste Zeichen automatisch höher stellt (PSE^2), diese Darstellung aber nicht interpretieren kann und dann eine Fehlermeldung ausgibt aus. Um das zu vermeiden, muss nach dem \wedge die Leertaste gedrückt und dann wie im Screenshot gezeigt fortgefahren werden.

Anschließend muss das Modell mit der neuen Variable geschätzt werden.



Die Schätzung des Modells liefert folgenden Output:



```
gretl: Modell 1
Datei Bearbeiten Tests Speichern Graphen Analyse LaTeX
Modell 1: KQ, benutze die Beobachtungen 1-12
Abhängige Variable: Units

-----
                Koeffizient      Std.-fehler      t-Quotient      p-Wert
-----
const          -390,028          1324,23          -0,2945          0,7769
Price          -0,299749                    0,109709          -2,732           0,0292 **
AE             0,0360407                    0,0148673          2,424            0,0458 **
PSE            0,0826819                    0,0786303          1,052            0,3280
sq_PSE        -2,34242e-07                  1,09650e-06          -0,2136          0,8369

Mittel d. abh. Var.      2020,833      Stdabw. d. abh. Var.      607,3259
Summe d. quad. Res.     122058,4      Stdfehler d. Regress.     132,0489
R-Quadrat               0,969916      Korrigiertes R-Quadrat    0,952726
F(4, 7)                 56,42103      P-Wert (F)                 0,000021
Log-Likelihood          -72,39135      Akaike-Kriterium          154,7827
Schwarz-Kriterium       157,2072      Hannan-Quinn-Kriterium    153,8851

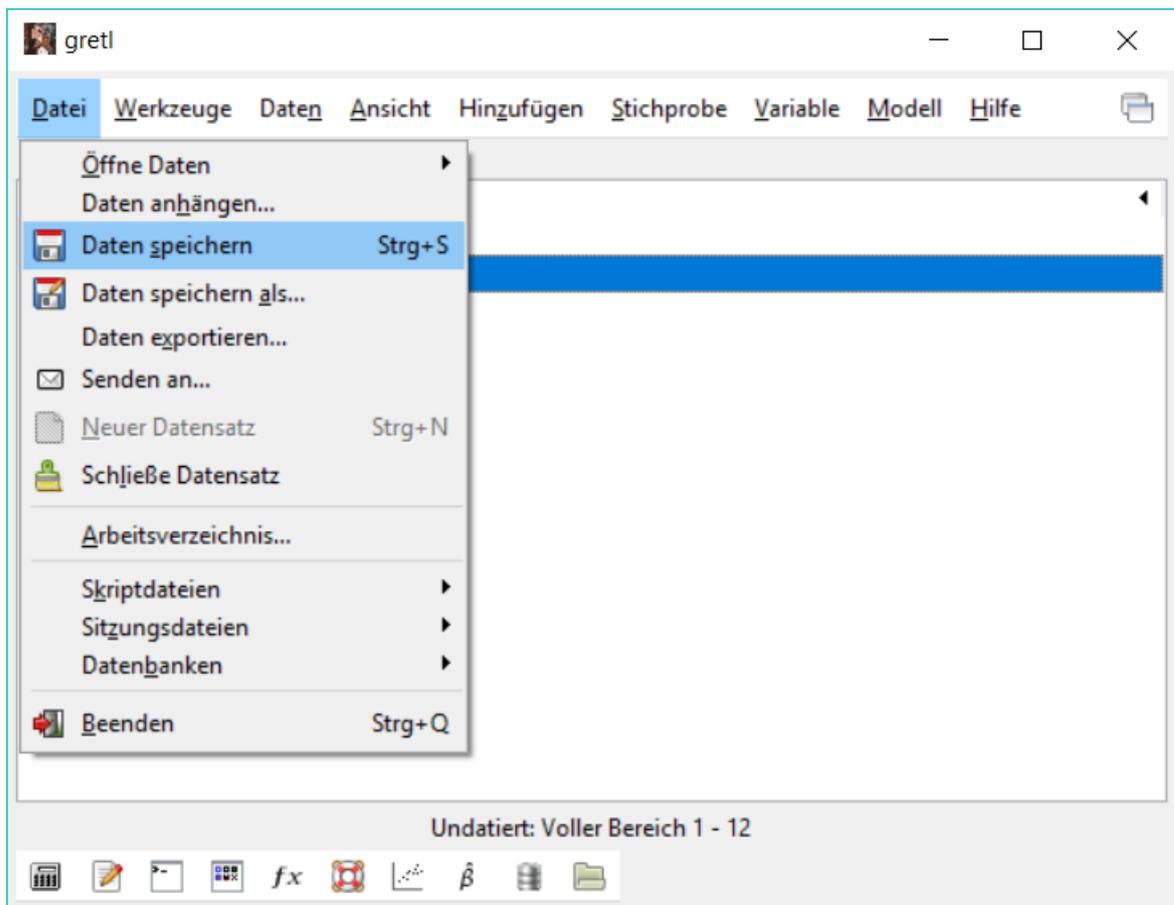
Abgesehen von Konstante war p-Wert am höchsten für Variable 6 (sq_PSE)
```

Ein quadratischer Einfluss der Vertriebsausgaben kann somit nicht festgestellt werden.

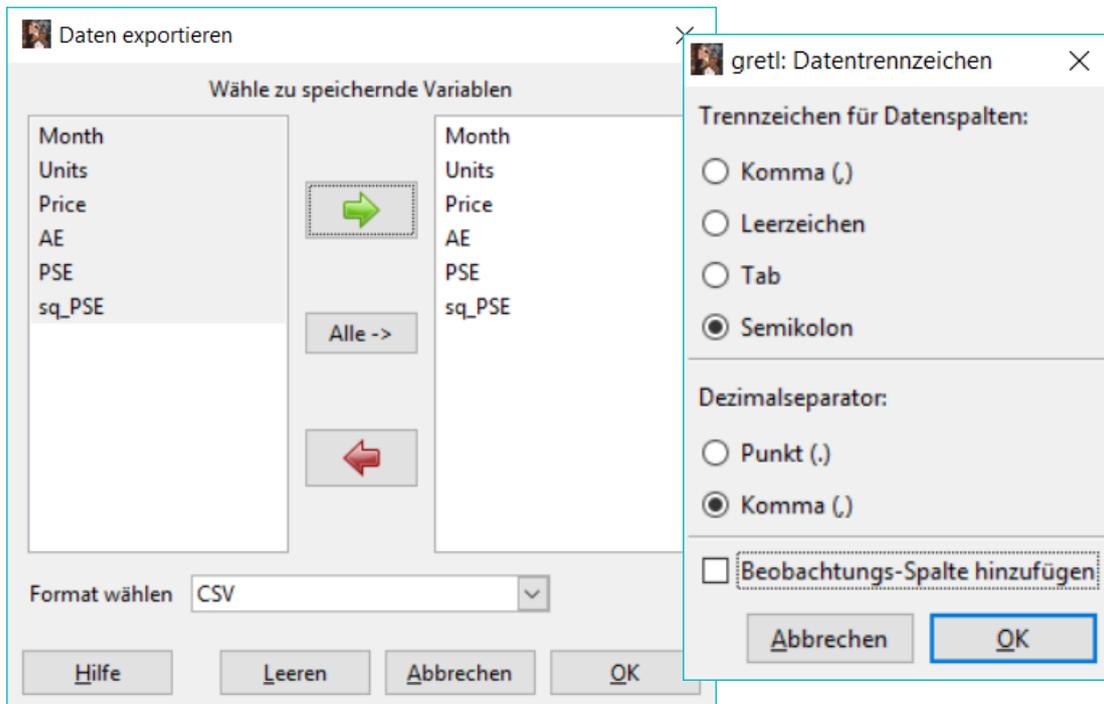
3.5 Daten speichern

Im Zuge der letzten Modellerweiterungen wurden im vorangegangenen Abschnitt neue Variablen definiert und erzeugt. Diese lassen sich im *gretl*-eigenen Datenformat *.gdt* als *gretl*-Datendatei speichern oder in anderen Datenformaten exportieren:

Wählt man unter *Datei* → *Daten speichern*, so erstellt das Programm die *gretl*-Datendatei EDP-Daten.gdt



Die Daten können jedoch auch als Excel-Datei bzw. im CSV-Format gespeichert werden. Hierzu ist *Datei* → *Daten exportieren* aufzurufen, die gewünschten Reihen zu selektieren und das CSV-Format (Voreinstellung) zu wählen. Nun müssen im Dialogfenster noch Spaltenseparator und Dezimaltrennzeichen angegeben werden. Da in der deutschen Notation das Komma als Dezimalzeichen dient, sollte als Spaltenseparator das Semikolon gewählt werden.



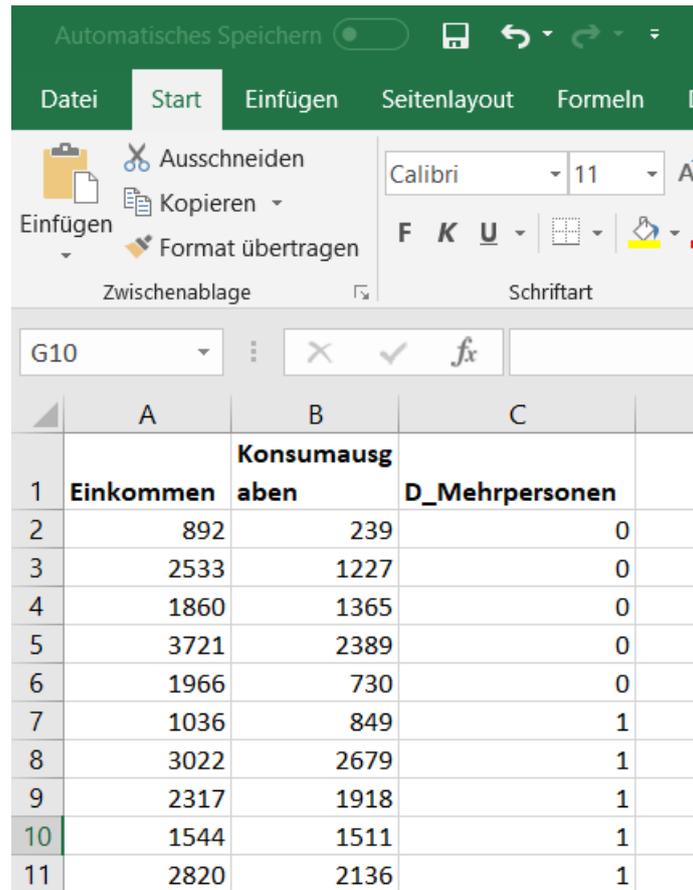
Der folgende screenshot zeigt eine mögliche Datei im CSV-Format, die mit Excel geöffnet werden und dann als xlsx-Datei gespeichert werden kann.

	A	B	C	D	E	F
1	Month	Units	Price	AE	PSE	sq_PSE
2	Januar	2500	3800	26800	43000	1849000000
3	Februar	2250	3700	23500	39000	1521000000
4	MÄrzt	1750	3600	17400	35000	1225000000
5	April	1500	3500	15300	34000	1156000000
6	Mai	1000	3200	10400	26000	676000000
7	Juni	2500	3200	18400	41000	1681000000
8	Juli	2750	3200	28200	40000	1600000000
9	August	1750	3000	17400	33000	1089000000
10	September	1250	2900	12300	26000	676000000
11	Oktober	3000	2700	29800	45000	2025000000
12	November	2000	2700	20300	32000	1024000000
13	Dezember	2000	2600	19800	34000	1156000000
14						

3.6 Dummy-Variablen

Abschließend soll der Gebrauch von Dummy-Variablen in *gretl* illustriert werden. Auch wenn *gretl* die Möglichkeit bietet, programmintern neue Variablen zu definieren, ist es häufig einfacher, diese zunächst in Excel zu erstellen und anschließend in *gretl* einzulesen.

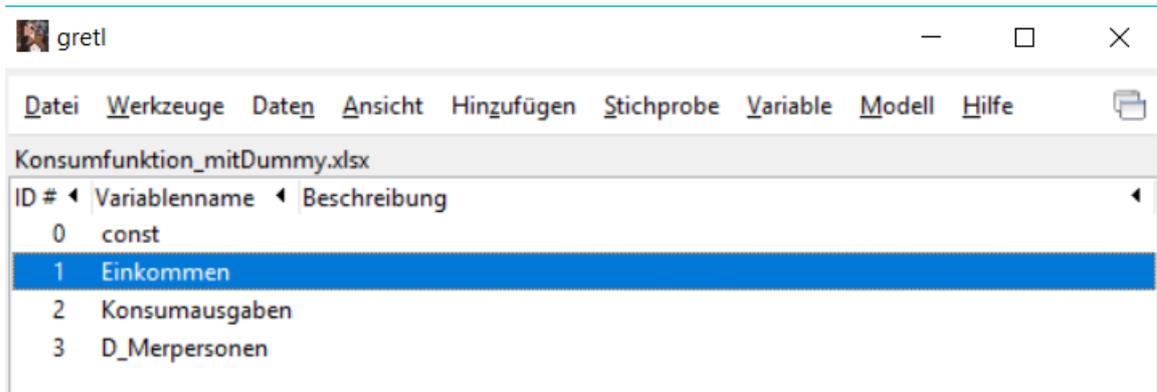
Schließen Sie deshalb die *gretl*-Sitzung mit *Datei* → *Beenden* und speichern Sie die Befehle nicht. Öffnen Sie dann die Datei *Konsumausgaben.xlsx* und ergänzen Sie diese wie folgt: Es wird angenommen, dass die ersten fünf Haushalte Einpersonen-, die übrigen fünf Haushalte dagegen Mehrpersonenhaushalte sind. Definieren Sie deshalb eine Mehrpersonendummy wie im folgenden Bildschirmfenster gezeigt:



The screenshot shows an Excel spreadsheet with the following data:

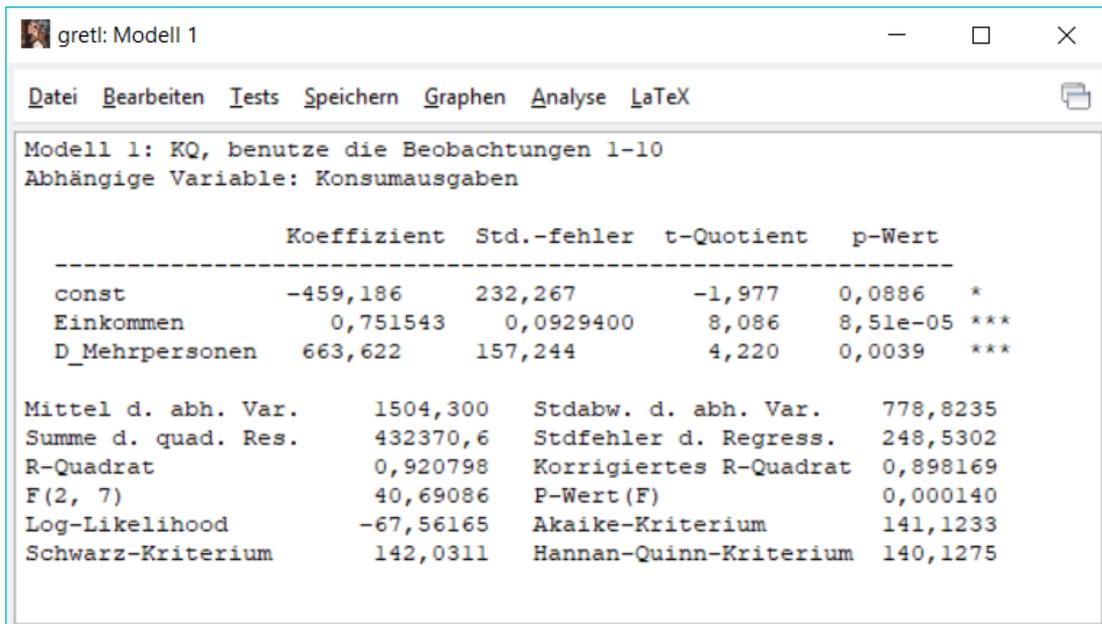
	A	B	C
1	Einkommen	Konsumausgaben	D_Mehrpersonen
2	892	239	0
3	2533	1227	0
4	1860	1365	0
5	3721	2389	0
6	1966	730	0
7	1036	849	1
8	3022	2679	1
9	2317	1918	1
10	1544	1511	1
11	2820	2136	1

Öffnen Sie dann die modifizierte Exceldatei in *gretl*:

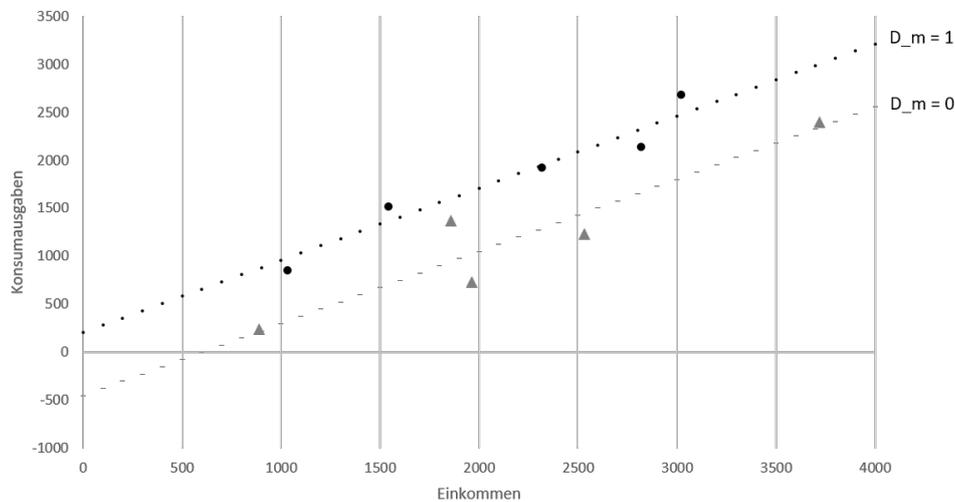


3.6.1 Variables Absolutglied

Das Modell kann in gretl wie gewohnt geschätzt werden, die Dummyvariable $D_{Mehrpersen}$ wird lediglich als zusätzliche unabhängige Variable in das Modell aufgenommen. Daraufhin ergibt sich der folgende Regressionsoutput.



Offensichtlich liegen die Konsumausgaben in Mehrpersonenhaushalten signifikant über denen von Einpersonenhaushalten. Das Absolutglied liegt im Fall $D_i^{MP} = 1$ um 663,60 Euro über dem eines Einpersonenhaushaltes, die Konsumfunktion der Mehrpersonenhaushalte ist gegenüber jener der Single-Haushalte um diesen Betrag nach oben verschoben. Die folgende Grafik wurde zur Illustration in Excel erstellt, sie gehört **nicht** zum *gretl*-Output. Einpersonenhaushalte sind durch Dreiecke, Mehrpersonenhaushalte durch Punkte dargestellt.

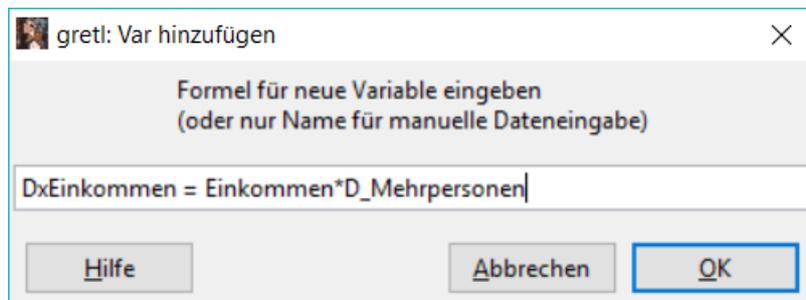


3.6.2 Variable Steigungsparameter

Neben dem Einfluss auf das Absolutglied ist denkbar, dass die Zugehörigkeit zur Gruppe der Mehrperson- bzw. der Einpersonenhaushalte auch einen Einfluss auf den Steigungsparameter hat. Es gilt dann

$$\beta_{1,i} = \gamma_0 + \gamma_1 D_{\text{Mehrpersonen},i}$$

und der Kreuzeffekt $\text{Einkommen} \cdot D_{\text{Mehrpersonen}}$ muss wie in Abschnitt ?? gezeigt im Modell berücksichtigt werden. Dafür ist eine neue Variable zu definieren und hinzuzufügen:



Anschließend wird das Modell mit der zusätzlichen Variable neu geschätzt. Der Regressionsoutput zeigt hier, dass die marginale Konsumquote (Steigung der Schätzgeraden) in einem Mehrpersonenhaushalt nicht signifikant verschieden ist von der eines Einpersonenhaushalts.

gretl: Modell 1

Modell 1: KQ, benutze die Beobachtungen 1-10
Abhängige Variable: Konsumausgaben

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	-404,323	305,817	-1,322	0,2343
Einkommen	0,726542	0,128358	5,660	0,0013 ***
DxEinkommen	0,0628312	0,203483	0,3088	0,7679
D_Mehrpersonen	527,508	471,925	1,118	0,3064

Mittel d. abh. Var.	1504,300	Stdabw. d. abh. Var.	778,8235
Summe d. quad. Res.	425607,5	Stdfehler d. Regress.	266,3355
R-Quadrat	0,922037	Korrigiertes R-Quadrat	0,883055
F(3, 6)	23,65319	P-Wert (F)	0,001006
Log-Likelihood	-67,48282	Akaike-Kriterium	142,9656
Schwarz-Kriterium	144,1760	Hannan-Quinn-Kriterium	141,6379

Abgesehen von Konstante war p-Wert am höchsten für Variable 4 (DxEinkommen)

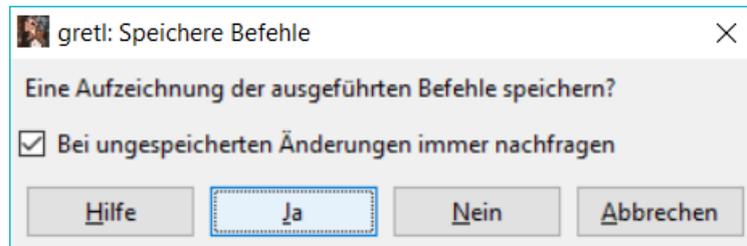
(Hinweis: Die letztgenante Schätzung unterscheidet sich von der getrennten Schätzung zweier Konsumfunktion für Ein- und Mehrpersonenhaushalte allein dadurch, dass im obigen Fall eine gleich große Störgrößenvarianz σ^2 für beide Gruppen angenommen wird.)

Lassen Sie die *gretl*-Sitzung geöffnet und gehen Sie zum nächsten Abschnitt über.

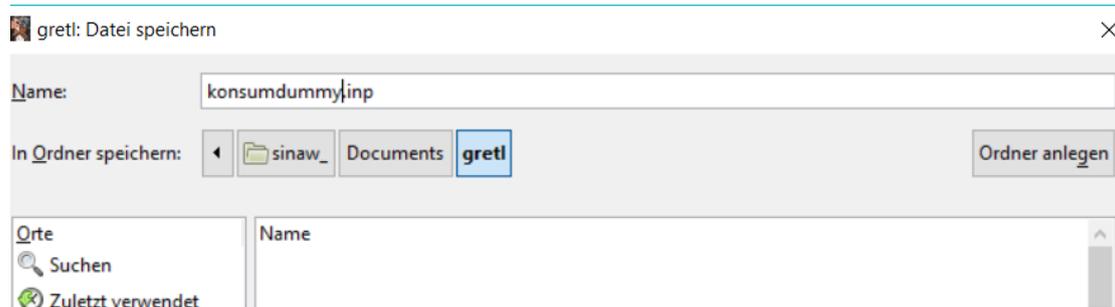
4 Skriptdateien erstellen und modifizieren

In diesem Schlusskapitel wird angedeutet, wie *gretl* als Syntax-orientierte Programmiersprache genutzt werden kann, womit auch die Bearbeitung komplexerer Aufgaben möglich wird.

Schließen Sie die *gretl*-Sitzung aus dem vorangehenden Abschnitt und wählen Sie diesmal die Option, die ausgeführten Befehle zu speichern:

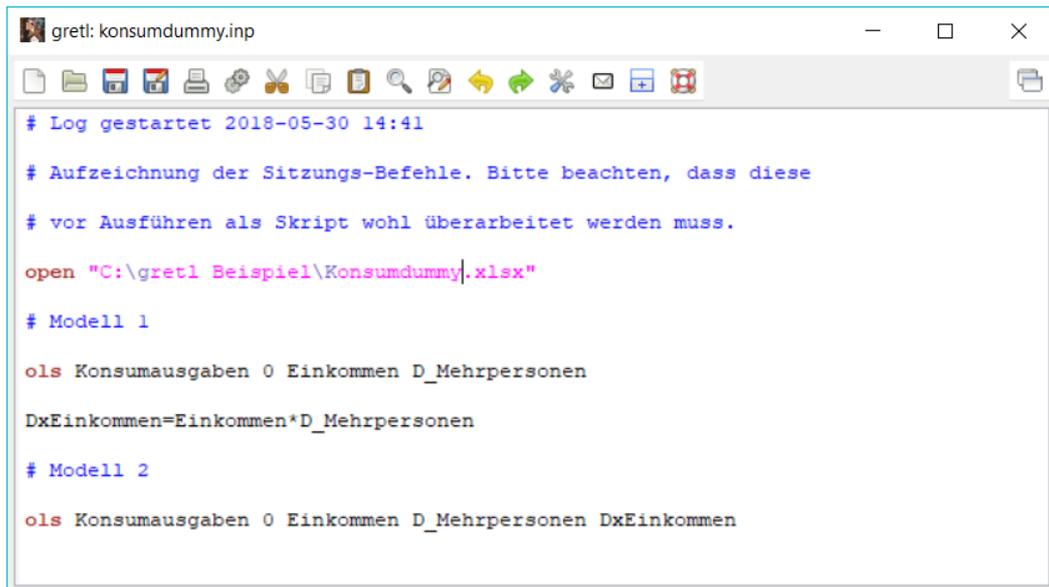


Geben Sie der „Befehlsdatei“ einen geeigneten Namen und wählen Sie einen geeigneten Speicherort:



Dieser Skriptdatei wird automatisch das Suffix `.inp` für „Input“ angehängt. Sie enthält sämtliche Befehle, die in der betreffenden Sitzung ausgeführt wurden.

Doppelklick auf die Datei `konsumdummy.inp` im betreffenden Verzeichnis verknüpft die Datei standardmäßig mit `gretl` und öffnet den `gretl`-internen Skripteditor. Im vorliegenden Beispiel ist die Syntax selbsterklärend:

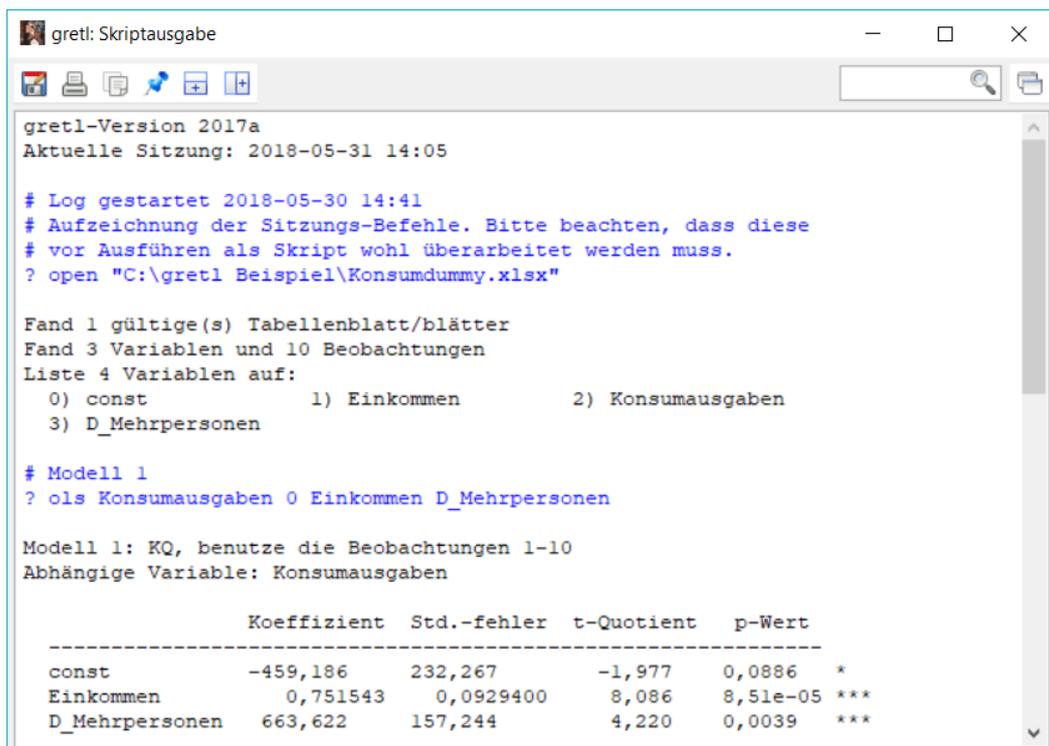


```

gretl: konsumdummy.inp
# Log gestartet 2018-05-30 14:41
# Aufzeichnung der Sitzungs-Befehle. Bitte beachten, dass diese
# vor Ausführen als Skript wohl überarbeitet werden muss.
open "C:\gretl Beispiel\Konsumdummy.xlsx"
# Modell 1
ols Konsumausgaben 0 Einkommen D_Mehrpersonen
DxEinkommen=Einkommen*D_Mehrpersonen
# Modell 2
ols Konsumausgaben 0 Einkommen D_Mehrpersonen DxEinkommen

```

Über die „Zahnrad-Schaltfläche“ *Ausführen* werden diese Befehle in chronologischer Reihenfolge erneut ausgeführt und die Ergebnisse in einem Ausgabefenster dargestellt:



```

gretl: Skriptausgabe
gretl-Version 2017a
Aktuelle Sitzung: 2018-05-31 14:05
# Log gestartet 2018-05-30 14:41
# Aufzeichnung der Sitzungs-Befehle. Bitte beachten, dass diese
# vor Ausführen als Skript wohl überarbeitet werden muss.
? open "C:\gretl Beispiel\Konsumdummy.xlsx"

Fand 1 gültige(s) Tabellenblatt/blätter
Fand 3 Variablen und 10 Beobachtungen
Liste 4 Variablen auf:
  0) const          1) Einkommen          2) Konsumausgaben
  3) D_Mehrpersonen

# Modell 1
? ols Konsumausgaben 0 Einkommen D_Mehrpersonen

Modell 1: KQ, benutze die Beobachtungen 1-10
Abhängige Variable: Konsumausgaben

-----
                Koeffizient  Std.-fehler  t-Quotient  p-Wert
-----
const           -459,186      232,267      -1,977      0,0886  *
Einkommen         0,751543      0,0929400     8,086      8,51e-05 ***
D_Mehrpersonen   663,622       157,244       4,220      0,0039  ***

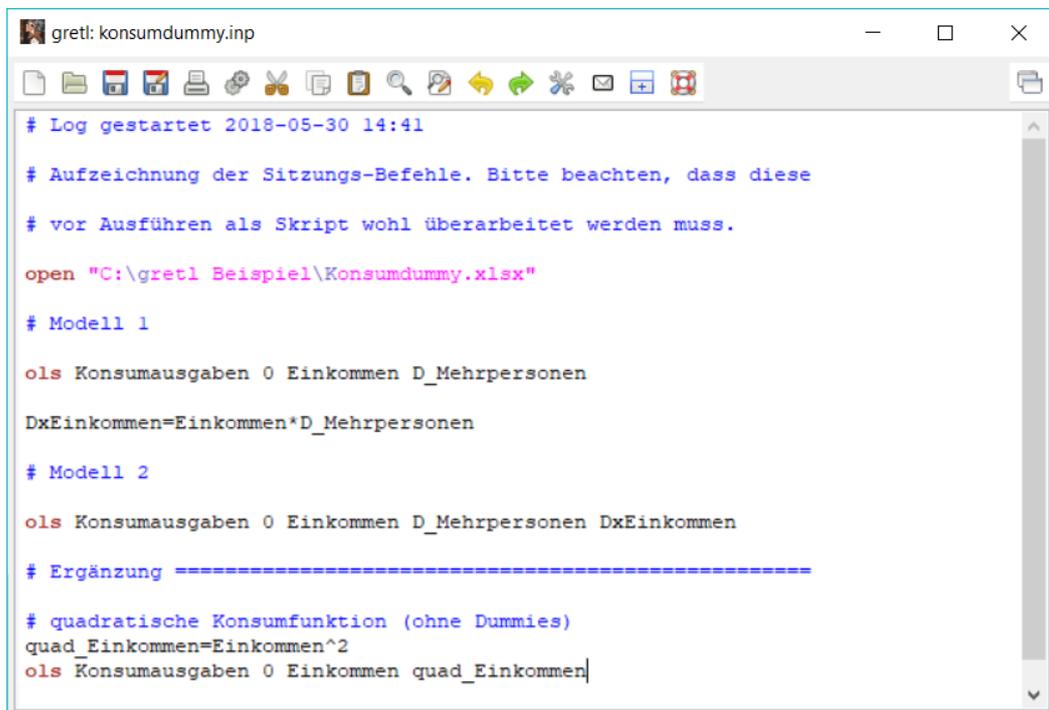
```

Um an die letzte Sitzung anknüpfend weiterzuarbeiten, kann die Skriptdatei bearbeitet und dabei z. B. um weitere Befehle ergänzt werden. Im vorliegenden Beispiel soll – ohne Verwendung der Mehrpersonendummy – eine quadratische Konsumgleichung der

Form

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + u_i$$

geschätzt werden. Dafür wurde die Befehlsdatei wie folgt ergänzt (Kommentarzeilen werden mit dem Rautensymbol # eingeleitet):



```
gretl: konsumdummy.inp
# Log gestartet 2018-05-30 14:41
# Aufzeichnung der Sitzungs-Befehle. Bitte beachten, dass diese
# vor Ausführen als Skript wohl überarbeitet werden muss.
open "C:\gretl Beispiel\Konsumdummy.xlsx"
# Modell 1
ols Konsumausgaben 0 Einkommen D_Mehrpersonen
DxEinkommen=Einkommen*D_Mehrpersonen
# Modell 2
ols Konsumausgaben 0 Einkommen D_Mehrpersonen DxEinkommen
# Ergänzung =====
# quadratische Konsumfunktion (ohne Dummies)
quad_Einkommen=Einkommen^2
ols Konsumausgaben 0 Einkommen quad_Einkommen
```

Ausführen des modifizierten Skripts und Ersetzen des „alten“ Outputs liefert das zusätzliche Ergebnis:

```

gretl: Skriptausgabe 2
Abgesehen von Konstante war p-Wert am höchsten für Variable 4 (DxEinkommen)

# Ergänzung =====
# quadratische Konsumfunktion (ohne Dummies)
? quad_Einkommen=Einkommen^2
Erzeugte Reihe quad_Einkommen (ID 5)
? ols Konsumausgaben 0 Einkommen quad_Einkommen

Modell 3: KQ, benutze die Beobachtungen 1-10
Abhängige Variable: Konsumausgaben

-----
                Koeffizient      Std.-fehler    t-Quotient    p-Wert
-----
const            -363,359         915,374        -0,3970       0,7032
Einkommen         1,01298           0,878903       1,153         0,2869
quad_Einkommen   -6,10798e-05         0,000193303   -0,3160       0,7612

Mittel d. abh. Var.    1504,300    Stdabw. d. abh. Var.    778,8235
Summe d. quad. Res.   1510969    Stdfehler d. Regress.   464,5995
R-Quadrat             0,723220    Korrigiertes R-Quadrat  0,644140
F(2, 7)              9,145414    P-Wert (F)              0,011155
Log-Likelihood        -73,81777    Akaike-Kriterium        153,6355
Schwarz-Kriterium     154,5433    Hannan-Quinn-Kriterium  152,6397

```

Im *gretl*-Hauptfenster können Sie unter *Hilfe* eine interaktive Befehlsdokumentation öffnen, die alle *gretl*-internen Befehle, die bei der Bearbeitung des Skripts verwendet werden können, beschreibt.

```

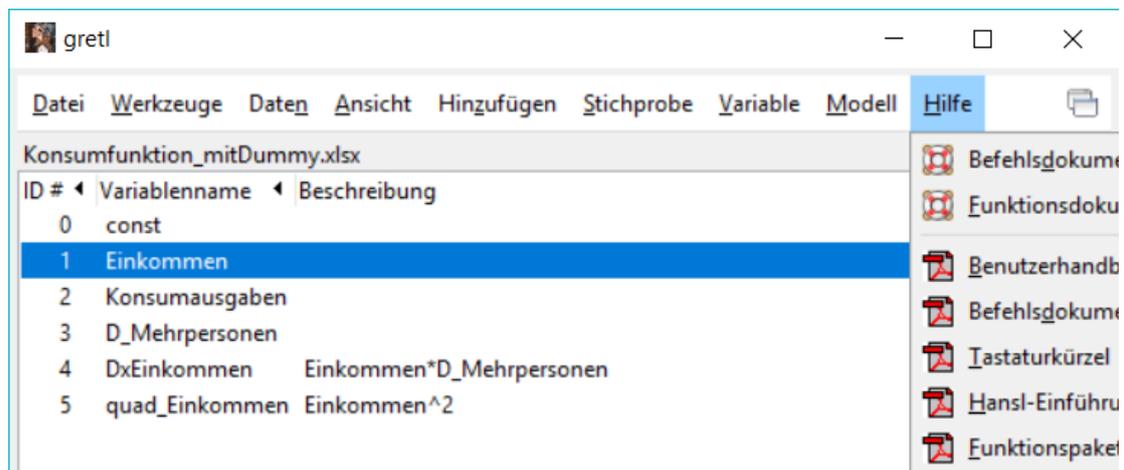
gretl: Befehlsdokumentation
-----
Index
+ Tests
+ Statistiken
+ Datensatz
+ Schätzung
+ Graphen
+ Programmieren

Gretl-Befehlsdokumentation
add      adf      anova     append    ar        ar1
arbond   arch     arima     biprobit  boxplot   break
catch    chow     clear     coeffsum  coint     coint2
corr     corrgm   cusum     data      dataset   debug
delete   diff     difftest  discrete  dpanel    dummify
duration elif     else      end       endif     endloop
eqnprint equation estimate eval       fcast     flush

```

Für die tiefergehende Beschäftigung mit *gretl* finden Sie unter *Hilfe* auch drei sehr hilfreiche Dokumente im PDF-Format:

- Ein Handbuch zu den *gretl*-Befehlen,
- eine *gretl*-Dokumentation,
- eine Einführung in *hansl*, die Programmiersprache, in der *gretl* geschrieben ist.



Viel Spass beim Entdecken der vielfältigen Möglichkeiten
von ökonometrischen Analysen im Allgemeinen
und von *gretl* im Besonderen!