

Einführung in das Arbeiten mit EViews

Die folgende Darstellung bietet nur einige Hinweise auf die notwendigsten Operationen, um mit EViews die einfachsten ökonomischen Berechnungen durchführen zu können. Sie beschränkt sich auf mehr oder weniger einführenden Gegenstände, die in der Vorlesung behandelt werden und orientiert sich auch an dem, was üblicherweise an Kenntnissen erforderlich ist, um eine empirische Untersuchung im Rahmen einer Bachelor-, Master- oder Diplomarbeit durchführen zu können.

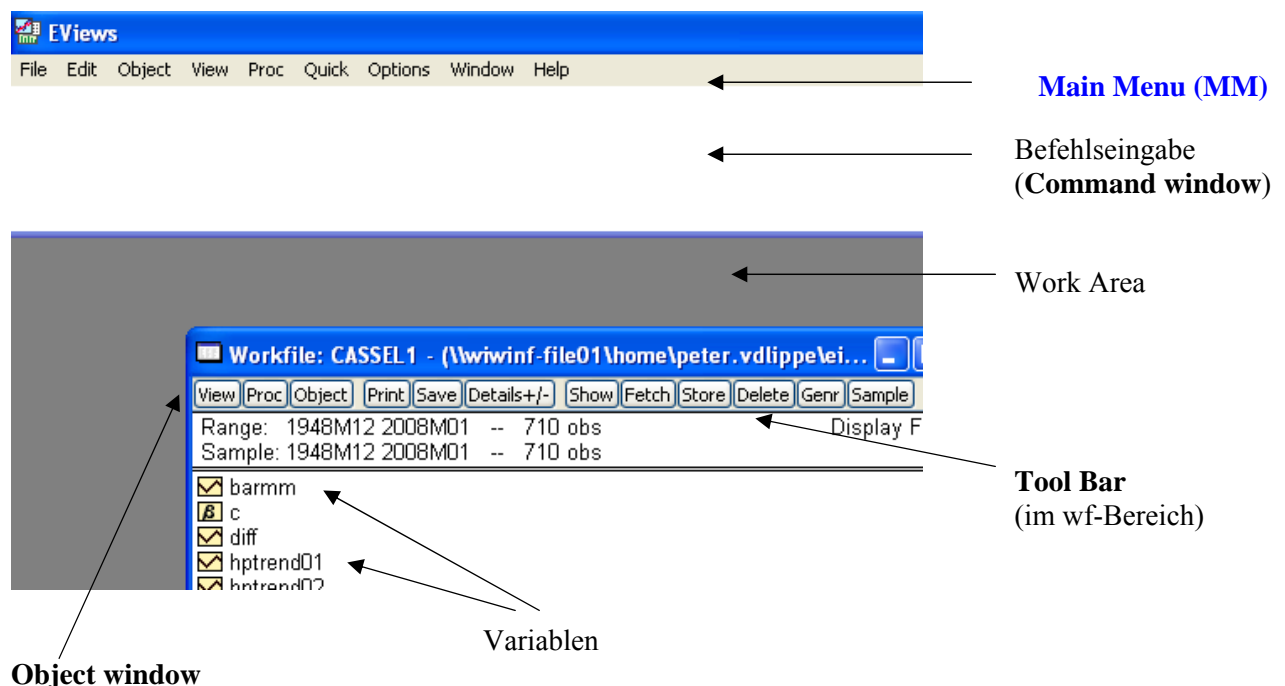
Grundsätzlich gilt, dass man alle Operationen in EViews auf eine ganze Reihe verschiedener Arten durchführen kann. Hier werden meist nur eine oder eine wenige Arten genannt. Es ist durchaus möglich, dass es bei einer Operation eine elegantere Art geben kann als die Methode, die hier dargestellt wird. Einige der hier angesprochenen Fragen sind bereits in meinen Übungsaufgaben (vgl. Downloads A, B, ... auf dieser Website) behandelt und ausführlich beschrieben worden, so dass eine Wiederholung hier nicht notwendig ist. Hinweise auf diese Stellen erfolgen mit dem Buchstaben der betreffenden Übungsblätter und der Seite in Fettdruck und in roter Farbe (vgl. ein Beispiel hierfür in Fußnote 4). Der Text ist im Folgenden an einigen Stellen im Telegrammstil gehalten und wie folgt gegliedert:

1. Wichtigste Grundregeln (Konzepte "workfile" und "object")	Seite 1
2. Daten (Kreieren und importieren, Gruppen, Datenbasis)	4
3. Arbeiten mit den Daten (Steuerungsarten [Maus, Befehle], Notation bei Funktionen)	5
4. Bestimmung einer Regressionsfunktion	6
5. Zugang zu den von EViews bereitgestellten Auswertungsmöglichkeiten	7 (8)
6. Graphiken und Tabellen, Ergänzungen	14

Nicht behandelt werden insbesondere Gegenstände wie Mehrgleichungsmodelle, GARCH Schätzung, logistische Regression (Logit, Probit usw.), ARIMA Theorie, State Space Models, VAR und VEC Modelle einschl. Kointegration und andere Schätzmethoden als OLS (wie etwa NLS, TSLS, GMM). Will man EViews für derartige weiterführende Themen nutzen, so sind die entsprechenden Handbücher zu konsultieren. Weiter gilt: **Learning by doing** ist wichtiger als der vorliegende Text.

1. Wichtigste Grundregeln

1.1. Bildschirm



¹ Für eine kritische Durchsicht und interessante Hinweise danke ich Herrn MA Jens Mehrhoff (Deutsche Bundesbank) und Herrn MSc Claus Christian Breuer.

Hier das Fenster für den workfile "cassel 1" (dieser Name wurde benutzt im Übungstext G), entsprechend Fenster für andere Objekte; dunkle Färbung (der blauen Titelleiste) heißt Fenster ist "aktiv".

Man kann verschiedene Objektfenster in der grauen work area anordnen, verschieben (mit click & drag) und in der Größe verändern. Die Befehlsleiste (**toolbar**) ist abgesehen von Tasten wie **View**, **Proc** oder **Procs** oder **Object** je nach Art des Objekts (Begriff unten Abschn. 1.4) etwas unterschiedlich belegt. Auch bei der gleichen Taste, etwa bei View erhält man ein unterschiedliches drop down menu, je nachdem ob man z.B. im group window oder im equation window (Abschn. 5.2 und 5.3) oder einem anderen object window ist.

1.2. Einige allgemeine Bemerkungen zu EViews

EViews ist weitgehend menügesteuert und somit sehr benutzerfreundlich. Weil die hierfür nötigen Tasten (buttons) und drop down menus auf sehr verschiedenen Wegen erreicht werden können ist es kaum möglich, beim Arbeiten mit EViews alle Wege zu kennen. **Man muss vieles einfach ausprobieren**; dieser Text kann hierzu nur eine Hilfestellung sein und man kann sich mit ihm Ausprobieren (und Fehler machen) nicht ersparen. Auch das Handbuch ist nicht überall hilfreich. Es ist auch so aufgebaut, dass man die Antwort auf eine Frage nicht immer leicht findet.

Einige Operationen sind ausgesprochen umständlich. Als Beispiele seien genannt: Änderung der Frequenz (Periodizität) einer Zeitreihe, Import von Daten, Durchführung von Korrekturen bei einer einmal gemachten Einstellung. Sehr störend ist, dass es keine Rücknahme eines Befehls und Wiederherstellung des bisherigen Zustands gibt, also keine \odot Taste. Häufig kommt man zur "alten" (bisherigen) Situation indem man noch einmal die gleiche Taste drückt (man hat z.B. eine Gleichung geschätzt und will etwas verändern, dann entweder Taste Estimate drücken oder Proc und dann Option "specify/estimate" wählen. Man muss bei dieser Verfahrensweise (gleiche Taste noch einmal drücken für eine \odot Operation) dann jedoch nicht selten jeweils alle Einstellungen neu vornehmen, um den alten Zustand wieder herzustellen.

1.3. Workfile (wf)

Die Arbeit beginnt in der Regel damit, dass man auf dem **main menu** (= MM, obere Befehlsleiste²) anklickt: **File/New/Workfile** → Abfrage nach Art der Daten (Querschnittsdaten [= undated] oder Zeitreihe, dann Frequenz und Beginn/Ende eingeben [bei "undated data" nur Anzahl der Beobachtungen], Name des wf eingeben; erhält dann die Endung *.wfl). Man sollte den wf möglichst schnell benennen (spätestens bei "save as...").

Weiterarbeiten an einem vorhandenen wf **File/Open/Workfile** oder einfach den wf (als einen der letzten 19 benutzten und nach anklicken von **File** unten aufgelisteten) anklicken, an dem man weiterarbeiten möchte.

In einem wf kann nur mit Daten *einer* Frequenz (Periodizität, z.B. nur mit Monatsdaten) und einem Geltungsbereich (= **range**) gerechnet werden³. Will man die Daten in eine andere Frequenz transformieren (etwa von Monatsdaten zu Quartalsdaten) und auch mit diesen transformierten Daten weiterrechnen, muss ein neuer wf gebildet werden,⁴ Beim etwas umständlichen Verfahren der "**frequency conversion**" wird unterschieden "high to low" (z.B. Monatsdaten konvertieren in Jahresdaten) oder umgekehrt "low to high". Eine Möglichkeit, dabei vorzugehen wird beschrieben in **G, 6-8**

Man kann zwar nicht mit unterschiedlicher Frequenz innerhalb eines wf arbeiten, wohl aber mit verschiedenen zeitlichen Abgrenzungen (Teilintervallen einer Zeitreihe, also mit verschiedenen **samples**)

² Wenn eine Taste auf MM anzuklicken ist dann ist sie hier fett und **blau**, sonst (bei Befehlen im **Arbeitsbereich** eines workfiles, genauer im **object's window** [z.B. in der "equation **tool bar**"; eine Gleichung ist ein spezielles Objekt]) fett und **schwarz** gekennzeichnet.

³ Diese Begrenzung gilt nicht bei einer database (db): in ihr kann man Daten von verschiedener Frequenz und mit einem verschiedenen Bereichen sammeln und dann aus der db in einen entsprechend definierten wf einbringen (vgl. unten Abschn. 2.4).

⁴ Das dabei etwas komplizierte Verfahren ist im Einzelnen beschrieben im download G auf Seite 6 (also in **G,6**). Die dort angesprochene **Dated Data Table** erreicht man im Group Window über views (ähnlich wie auch verschiedene Graphiken usw. vgl. Abschn. 5.3).

der gleichen Variable: dazu kann man Daten auf weitere Seiten des wf kopieren und auf diesen das sample anders definieren (beschrieben in **G**).

Wenn ein wf definiert ist gibt EViews ein **workfile directory** (workfile Liste)⁵ mit den Angaben über Range, Sample, alle Variablen und sonstige Objekte [z.B. Gleichungen] an (zu sehen oben auf S. 1 und in **D,1,3,4** sowie in **E,4** und in weiteren Übungsblättern). Diesen **directory nicht löschen**, weil sonst alle Daten verloren gehen.⁶ Man erhält ferner eine zweite Befehlsleiste (tool bar), ähnlich wie das MM mit z.T. den gleichen, aber auch unterschiedlichen Befehlsflächen (buttons).⁷

1.4. Objects

Einheiten, mit denen gearbeitet wird sind "Objekte" (zusammengehörige Informationen). Ein wf besteht aus Objekten. Damit Objekte nicht verloren gehen, im wf integriert sind und wieder aufgerufen werden können, um mit ihnen weiter arbeiten zu können müssen sie *benannt* sein. Objects können sein

eine *Variable* (series), eine *Gruppe* von Variable (group), eine *Gleichung* (equation), eine Festlegung über eine *Stichprobe* (z.B. Angabe von Anfang und Ende eines Teilintervalls) aus dem gesamten Datensatz (sample), eine *Matrix*, ein System (Gleichungssystem), ein Zustandsraummodell (state space), ein VAR-Modell, ein "Pool" (ein Datensatz mit Querschnitts- und Zeitreihendaten)⁸ und vieles mehr (z.B. auch Texte, Graphiken, Tabellen, Koeffizientenvektoren usw.).

Ein Objekt ist mit einem Symbol + Name im workfile directory aufgelistet; etwa $\boxed{=}$ als Symbol für eine Gleichung \boxed{G} für eine Gruppe, $\boxed{\leftrightarrow}$ für ein sample [= smpl]⁹, $\boxed{\#}$ für ein Skalar, ferner \boxed{var} für ein VAR-Modell, \boxed{P} für einen (Daten)pool oder \boxed{TXT} für einen Text usw.

Ein Objekt bilden (= creating objects) mit **Objects/New Object** [oder **Objects/New Object**]... Wenn ein Objekt "geöffnet" wird (Doppelklick oder **Show** [oder **Quick/Show**] anklicken) erhält man ein "**object window**" mit einer Befehlsleiste (tool bar). Dann erfolgt die Definition (Spezifizierung) des Objekts meist menügesteuert. Beispiel: kreiert man z.B. als neues Objekt eine Gleichung, wird man (im equation specification window) aufgefordert in speziellen Feldern die Gleichung anzugeben (vgl. unten Abschn. 3 und 4 für den Syntax), die Abgrenzung der Stichprobe und das Schätzverfahren anzugeben und mit Mausclick bestimmte Optionen einzustellen.

Hat man ein Object kreiert kann man auf der toolbar des object window mit dem dort u.a. gegebenen button \boxed{Object} auch wieder Änderungen vornehmen, oder kopieren, umbenennen, ausdrucken usw. Die toolbar beim Objekt "Variable/Zeitreihe" (series) sieht z.B. wie folgt aus (mehr unten Abschn. 5.1).¹⁰

View	Proc	Object	Properties	Print	Name	Freeze	→
→	Sample	Genr	Sheet	Stats	Ident	Line	Bar

Man kann jederzeit mit Markieren des Namens eines Objekts im wf-directory und **Copy-and-paste** (kopieren und einfügen)¹¹ ganze Objekte innerhalb eines workfiles oder zwischen verschiedenen workfiles hin und her bewegen und so z.B. die Kopie einer Zeitreihe oder einer Gleichung woanders ablegen um dort das duplizierte Objekt weiter zu bearbeiten.

Wegnehmen eines Objekts (aus dem wf entfernen): Objekt (z.B. Reihe/series) markieren, mit rechter Maustaste Delete wählen.

⁵ Abgebildet am Beispiel des wf "Cassel 1" (aus dem Übungstext **G**) im screen shot auf der ersten Seite.

⁶ Siehe dazu auch **D,1**. EViews beenden mit **File/Close**, natürlich nicht bevor alles weggespeichert ist.

⁷ weil das wf-directory ein object window ist und jedes solches Objektfenster (z.B. auch bei einer Gleichung oder einem VAR-Objekt) eine eigene toolbar hat (mit speziellen Befehlen, die dort angeklickt werden können).

⁸ Meist auch "Panel" genannt, z.B. Zeitreihen für eine Gruppe von Ländern (die Unterschiedlichkeit der Länder macht den "Querschnitt" aus), d.h. für jedes Land der gleiche Satz von Zeitreihen.

⁹ Dieses Objekt enthält nur die Verabredung über frequency und die in der Regel verringerte range (hat man z.B. im wf Jahres-Daten einer Zeitreihe von 1980 bis 2008, so könnte ein sample die Verabredung sein, dass nur Werte von 1990 bis 2005 genommen werden, also eine Teilmenge aus der "range" ausgewählt wird).

¹⁰ Wir gehen im Abschnitt 5 in der Regel nur auf die beiden ersten buttons "view" und "procs", weil diese den Zugang zu den Auswertungsmöglichkeiten liefern.

¹¹ Befehle nach Markierung Edit/Copy und Edit/Paste.

Freeze (einfrieren) oder "Object/Freeze Output" bei einem Objekt (Graphik, Tabelle etc) bedeutet, den aktuellen Stand einer Bearbeitung (wie ein Schnappschuss) festhalten und sichern um es später weiterzubearbeiten und z.B. in eine Präsentation oder Textdatei einzubringen. Diese Funktion spielt eine große Rolle bei Graphiken (siehe unten Abschn. 6).

2. Daten

2.1. Manuelle Dateneingabe (keyboard entry → **D,2**)

Object/new object → Abfrage nach Art des Objekts: **series** wählen → spread sheet zur Dateneingabe öffnet sich (auf **Edit +/-** Taste achten (Zahleneingabe nur bei +, wenn beendet – wählen; Edit – ist der "protection mode", die Eingabe ist damit beendet und das Gitternetz [spread sheet] verändert seine Farbe und Layout etwas);¹² es gibt keinen speziellen Befehl zum Wegspeichern. Nach Eingabe ist die Series (Variable) auch in dem wf aufgenommen, sofern sie benannt ist. Es kann dann mit gleicher Befehlsfolge (object/new object...) eine neue series eingegeben werden.

Ergänzung oder Streichung eines Wert innerhalb einer Zeitreihe mit $0 \leq t \leq T$ Werten: im spreadsheet die entsprechende Stelle (t), an der eingefügt werden soll, markieren und dann **InsDel** button anklicken. Darauf achten dass range entsprechend angepasst wird, sonst kann passieren, dass alle Werte eine Periode weiter rücken und der T-te Wert (der an die Stelle T+1 kommen soll) verschwindet.

Will man in einem Arbeitsgang mehrere Variablen (in die Spalten eines spreadsheet) manuell eingeben muss man das **group window** öffnen (zum Begriff "Gruppe" siehe Abschn. 2.3). Oder am einfachsten mit **Quick/Empty Group (Edit Series)**.¹³

Procs (procedures): hier erzeugt EViews selbst neue Daten (eine neue Variable) aufgrund einer Funktion mit generate series. Auch hier (also bei procs) möglich Änderung des samples, import/export von Daten.

2.2. Datenimport (z.B. aus einer Excel Datei)

Über **Procs** (oder auf dem MM mit **Procs**) → Abfrage (Beschreibung des im Wesentlichen recht einfachen Verfahrens, z.B. bei Import aus einer Excel Tabelle¹⁴ in **F,2** [dort Fußnote]). Auch möglich über **File/Open/Foreign Data as Workfile**. und **Procs/Import/Read...** beim Datenimport oder **Procs/Export/Write...** beim Datenexport von der workfile toolbar aus, oder mit **Procs ...** vom main menu aus. Man kann auch viele Daten in eine EViews **Database** (DB) ablegen, um sie in einem oder mehreren workfiles zu benutzen. Die Befehle Store, Update, Fetch usw. beziehen sich hierauf. Mehr dazu → 2.4. Mit copy-and-paste kann man auch direkt Spalten oder Zeilen einer Excel Tabelle in das EViews spreadsheet zur Bildung einer Variable übertragen (vgl. Abschn. 2.4).

2.3. Gruppen (groups)

Die Zusammenfassung von mindestens zwei Variablen zu einer Gruppe ist ein wichtiges Hilfsmittel für alle weiteren Berechnungen, insbesondere auch für die Durchführung einer Regressionsrechnung. Wie unter 5.2 gezeigt gibt es eine Reihe von Auswertungsmöglichkeiten (insbesondere solche, die Zusammenhänge zwischen mehreren Variablen betreffen), die man erhält wenn man eine Gruppe "öffnet".¹⁵ Die group wird definiert durch Angabe der einzelnen Variablen (series), die zu ihr gehören sollen, und die man vorher eingegeben und benannt haben muss. Liste der series angeben mit blancs, ohne Komma.

¹² Eine +/- Taste ist ein toggle Schalter mit zwei Einstellungen (im Falle von **edit**: + edit mode und - protection mode), ähnlich **smpl** (sample + alle Werte in der range des wf werden gezeigt/- nur Werte des samples) oder **wide** (spreadsheet mit + allen/- einer Datenspalte) oder **label +/-** mit oder ohne textliche Zusätze im Verzeichnis der Objekte eines wf.

¹³ In die erste Zeile einer Datenspalte jeweils den Namen eingeben. Jede Variable eine Spalte.

¹⁴ Wohl die häufigste Art des Datenimports. Er ist nur, wie hier gesagt, einfach, wenn man genau weiß, auf welchem Blatt und in welchem Feld einer Excel-Datei die gesuchten Daten stehen. Ansonsten ist der Datenimport eher unkomfortabel.

¹⁵ Man beachte: Für Gruppen ergeben sich andere Auswertungsmöglichkeiten als für die einzelnen Variablen als solche (vgl. unten Abschn. 5.1 und 5.2)

Man kann eine Gruppe (oder z.B. auch Gleichung [vgl. Abschn. 4]) auch kreieren indem man die Variablen, die man einbeziehen will anklickt (evtl. bei gedrückter ctrl Taste) und markiert, dann irgendwo in den markierten Bereich doppelt klickt und im pop up menu "**open/as group**" (andre Optionen: as equation, as VAR, as Multiple series) wählt. Man erhält dann eine spreadsheet Ansicht aller Variablen der Gruppe. Das und die möglichen Multiple-Graph Ansichten kann (und sollte auch) der Prüfung dienen, ob die Daten korrekt sind.

2.4. EViews Databases (*.db)

Hier können Daten (idealerweise große Datenmengen) mit verschiedener frequency und range zusammengestellt werden, die (auch mit copy-and-paste) zwischen workfiles und databases hin und her bewegt werden können. Create db mit: **File/New/Database**. Nach im Dialog erfragten Angaben (Pfad, Name für die db usw.) erhält man das Database window mit einer eigenen toolbar. Dabei gibt es u.a. die buttons "Easy Query" und (Advanced) "Query", mit denen die Datenbasis nach Variablen durchsucht werden können, die im Hinblick auf bestimmte Gemeinsamkeiten (z.B. hinsichtlich des Namens, der verbalen Beschreibung,¹⁶ der Frequenz usw.) ausgefiltert sind. Im Database Registry Window (erreichbar über **Options/ Database Registry**) kann man Inhalte der db überprüfen, ergänzen, Einstellungen für die Suche (query) vornehmen usw. Vom workfile window (abgebildet oben auf S. 1) aus kann man mit den buttons "Fetch" und "Store" Reihen von der db in den wf bringen (= fetch) oder vom wf in die db (= store).

3. Arbeiten mit den Daten

Operationen starten mit Buttons (Anklicken von Befehlsflächen): Viele wichtige Arbeiten, wie die Berechnung einer Regressionsgleichung (stochastische Funktion), oder das Generieren einer Variable als (nicht-stochastische) Funktion anderer Variablen, die Anfertigung einer Graphik usw. kann man starten mit der Befehlsfolge

Object/New Object/... (im Arbeitsbereich) oder

Object/New Object/... (im main menu MM)

Unter Object findet man ferner die Datenbankoperationen (fetch, store, copy) sowie rename und delete.

Commands: statt mit Buttons kann man eine Operation auch mit in das command window (weiß) getippten Befehlen starten. Zu wählende Parameter meist in Klammern hinter dem Kommandowort. Beispiel Berechnung der Hodrick Prescott Filters (d.h. der glatten Komponente) der Zeitreihe xy wobei dann die gefilterte Reihe xy_hp heißen soll

Entweder bei der Variable xy über den button Procs (siehe unten Abschn. 5.1) und Einstellung des Parameters λ als $\lambda = 1500$ und Vergabe des Namens (z.B. der Name xy_hp) im Dialog

Oder mit dem Befehl xy.hp(1500) xy_hp

Weitere Beispiele:

- mit barmm.hp(14400) hptrend01 wird ein HP-Filter mit $\lambda = 14000$ auf die Zeitreihe barmm angewendet und das Ergebnis als Zeitreihe hptrend01 im wf abgelegt (wie oben auf Seite 1 hptrend01 im wf-directory)
- erste Zeile: equation eq1 – dann zweite Zeile: eq1.ls ger c phos1 nit2 (wie in **D,2** und unten in Abschn. 4); das erzeugt eine Gleichung, die eq1 heißt.
- series dm1 = d(m1) erzeugt die Zeitreihe $m1_t - m1_{t-1}$ der ersten Differenzen von m1
- barmm.uroot (4,c) führt einen unit-root-Test (ADF-Test) mit der Zeitreihe barmm durch bei einer Konstanten c und 4 lags der ersten Differenzen.
- Bildung einer Zeitreihe y als Funktion einer anderen Zeitreihe x (etwa als $y = e^x$) im command window mit series y =exp(x) oder (gleichbedeutend) genr y = exp(x).

Zu weiteren Einzelheiten siehe Handbuch.

Funktionen/Operatoren

Autofunktionen (autoseries): Wenn die Zeitreihe y_t beispielsweise (wie auf Seite 1) den Namen barmm (= Bargeldmenge monatlich) hat erhält man die folgenden einfachen Ausdrücke für Differenzen leads und lags usw.

¹⁶ sofern eine solche mit **label** erfolgt ist.

Symbol	Bedeutung	Symbol	Bedeutung
barmm(-1)	y_{t-1}	barmm(2)	y_{t+2}
barmm(0 to -2)	y_t, y_{t-1}, y_{t-2}	d(barmm)	$\Delta y_t = y_t - y_{t-1}$
log(barmm)	$\log y_t$	dlog(barmm)	$\log y_t - \log y_{t-1}$

Höhere Differenzen der Zeitreihe y mit $d(y,n)$ etwa $d(y,2)$ ist die Reihe $\Delta(\Delta y_t) = y_t - y_{t-1} - (y_{t-1} - y_{t-2})$ der zweiten Differenzen.

Die Bedeutung versteht sich von selbst bei $\exp(\text{barmm})$ und $\sqrt{\text{barmm}}$ ¹⁷ sowie $d\log(\text{barmm}(-1)) = \log y_{t-1} - \log y_{t-2}$ und den Operatoren $+$ $*$ $/$ und $^{\wedge}$.

Bestimmte Funktionen werden mit $@$ eingeleitet, wie etwa

Summary functions: $@\text{mean}(\dots)$ bildet einen Mittelwert aus allen Werten einer Variable oder $@\text{var}(\dots)$ die Varianz, mit $@\text{movav}$ wird ein moving average gebildet und mit $@\text{trend}$ die natürliche Zahlenfolge (1, 2, ...) erzeugt, so dass man mit $y = c(1)+c(2)*@\text{trend}$ in einer Regressionsfunktion einen linearen Trend der abhängigen Variable (Zeitreihe) y erhält.

Element functions die jeweils auf einen Wert der Variable wirken, wie $@\text{abs}$ oder der observation identifier $@\text{elem}(x,4)$ wählt aus der Variable x das vierte Element aus; mit $@\text{all}$ ist $\text{sample} = \text{range}$, mit $@\text{first}$ 1990 2007 $@\text{last}$ wird das sample auf einen Teil-Zeitraum [1990 - 2006] der range begrenzt.

Mit $@$ -Funktionen kann man auch bestimmte vom Programm berechnete Größen gezielt absichern und benennen, um mit ihnen weiterrechnen zu können (z.B. als Konstante in einer Gleichung benutzen): Beispiele $@f$ (F-Statistik) $@dw$ (Durbin Watson Koeffizient), $@r2$ ($= R^2$), $@ssr$ (sum of squared residuals) usw. Das gilt nicht nur für Skalare, auch für Matrizen und Vektoren etwa $@\text{coefs}$ (liefert den Vektor der geschätzten Koeffizienten) oder $@\text{coefcov}$ (Matrix der Varianzen und Kovarianzen der geschätzten Koeffizienten, also der Größen $\hat{\sigma}_{\hat{\alpha}}^2$, $\hat{\sigma}_{\hat{\beta}}^2$, $\hat{\sigma}_{\hat{\alpha}\hat{\beta}}$ usw. Weitere Einzelheiten im Handbuch.

4. Bestimmung einer Regressionsfunktion

4.1. Gleichungen spezifizieren

Einfache Möglichkeiten sind bereits beschrieben in **D,2**: wie so oft gibt es verschiedene Wege zum gleichen Ziel. Hat man eine Gruppe definiert, so kann man die Regressionsgleichung berechnen mit

Quick (auf dem MM) dann **Estimate Equation** (Gleichung eingeben wie unter **D,2** [mit dem Beispiel "ger c phos1 nit2"] oder auch unten beschrieben: "create equation by formula")¹⁸ oder mit

Procs/Make Equation (auch hier nur wenn die Gruppe geöffnet ist, also über die tool bar im group window [mehr dazu siehe unter 5.2 unten]) oder auch einfach mit **Object/New Object/Equation** oder **Object/New Object/Equation**.

Gleichungen (Regressionsfunktionen)¹⁹ kann man auf zwei Arten einbringen (create, specify)

by List (etwa $\text{ger} = c(1) + c(2)*\text{phos1} + c(3)*\text{nit2}$ in **D,2** oder ein linearer Trend für die Variable "ger" mit $\text{ger} = c(1) + c(2)*@\text{trend}$)²⁰ oder

by Formula ($\text{ger} = c(1) + c(2)*\text{phos1} + c(3)*\text{nit2}$).

Man kann bei Spezifikation der Regressionsgleichung auch eine Variable angeben, die erst aus einer anderen als auto-series gebildet werden muss. Beispiel: $\log(\text{ger}) = c(1) + c(2)*\text{phos1} + c(3)*\text{nit2}$ als nichtlineare Gleichung statt wie oben $\text{ger} = c(1) + c(2)*\text{phos1} + c(3)*\text{nit2}$.

¹⁷ $\exp(y_t) = e^{y_t}$ und Wurzel aus y_t .

¹⁸ **Quick/Generate Series** (by equation) oder **Procs/Generate Series** ist - anders als in **D,2** beschrieben - mehr gedacht für das Generieren einer Variable durch eine deterministische (nicht-stochastische) Funktion, etwa $\text{diff} = \text{series01} - \text{series02}$ oder $\text{series4} = \text{series4} + \text{series4}(-1)$ bildet eine Summe (kumulierte Werte) für "series4".

¹⁹ Reine Definitionsgleichungen werden vorzugsweise gebildet mit **Generate Series** (per button) oder mit $\text{gen} \dots$ (= generate) per command Zeile im (weißen) command window.

²⁰ Man kann auch im Dialogfenster das Gleichungsfeld einmal anklicken und dann mit der gedrückten CTRL-Taste die Variablen markieren und danach eine der Variablen doppelt anklicken und Open/Equation wählen.

Die folgende Art der Spezifizierung einer (linearen) Gleichung ist besonders komfortabel

die Variablen im wf-directory sukzessive anklicken (auswählen): die erste Variable wird dann automatisch zur endogenen, die folgenden zu den exogenen²¹ und die Gleichung gebildet, wenn man dann über Rechtsklick **open/as equation** wählt (Gleichung wird dann wie folgt geschrieben $y \ x_1 \ x_2 \ \dots \ c$; Konstante c [für intercept] wird automatisch hinzugefügt)

Variablenauswahl mit gedrückter linker Maustaste

zusammenhängende Variablen: erste und letzte markieren bei gedrückter **Shift** (\uparrow) Taste

nicht zusammenhängende ... Variablen einzeln bei gedrückter **Ctrl** (= **Strg**) Taste markieren.

Wenn die Gleichung geschätzt worden ist, ihr am besten gleich mit dem button "Name" einen Namen geben.²² Außerdem hat man ab jetzt interessante Auswertungsmöglichkeiten (mehr unten unter 5.3).

4.2. Schätzmethode

Die Grundeinstellung ist ls (Least squares, also Methode der kleinsten Quadrate, auch OLS = ordinary least squares). Es ist aber auch möglich nls (= nonlinear; eine iterative Schätzung einer nichtlinearen Regressionsfunktion), weighted LS oder GMM (Generalized Method of Moments) usw. Hier ist auch der Zugang zu GARCH Modellen,²³ und Logit (logistische Regression) als Spezialfall der Option BINARY. Bei TSLS (two stage LS) handelt es sich um die Schätzung einer Gleichung mit (einem) mit u_i korrelierten Regressor(en) - also bei einer Verletzung der Annahme C1 - mit der Methode der Instrument(al)variablen.²⁴

4.3. Equation Output

Er ist in den downloads wiederholt abgebildet, sei es als screen shot (**B, D,3, E,4**), sei es durch ein anderes layout etwas überarbeitet (**C,4, E1, F2,**):

Den Equation Output (bei linearer Regression) richtig zu interpretieren ist ein wesentliches Ziel der Veranstaltung und – anders als die hier dargestellten Details zum Umgang mit EViews – auch **klausurrelevant**.

Die t-Werte eines Koeffizienten $\hat{\beta}_i$ für den Regressor x_i im Ergebnisausdruck beziehen sich jeweils auf die Hypothese $H_0: \beta_i = 0$. Sie sind deshalb berechnet als $t = \hat{\beta}_i / \hat{\sigma}_{\hat{\beta}_i}$ und somit leicht aus den mitgeteilten Zahlen $\hat{\beta}_i$ (Spalte "Coefficient") und $\hat{\sigma}_{\hat{\beta}_i}$ (Spalte "Std.Error") zu errechnen, so dass die Spalte "t-Statistic" streng genommen überflüssig ist. Will man die Hypothese $H_0: \beta_i = 2$ (ein konkreter, von Null verschiedener Wert) testen, muss man den **Wald Test** auf Koeffizienten Restriktionen anwenden (vgl. unten 5.3 [der Test ist zugänglich als view im equation object window]).

5. Zugang zu den von EViews bereitgestellten Auswertungsmöglichkeiten

Es wird davon ausgegangen, dass ein Objekt ausgewählt wurde. Das "**Öffnen**" einer Variable, Gruppe usw. (ganz allgemein eines Objekts) um weitere Auswertungen zu starten kann erfolgen mit:

im wf-directory markieren und Doppelklick oder mit rechter Maustaste anklicken und open wählen²⁵ oder mit **Show** auf der toolbar, bzw. **Quick/Show** im MM. Show zeigt Daten im spreadsheet nicht als eine Graphik.

²¹ Die Reihenfolge ist hier anders als bei einer Graphik des Streudiagramms (die erste angeklickte Variable wird nach Open/as group/view/scatter zur x-Achse (unabhängige Variable, Regressor), die zweite zur y-Achse.

²² Die Gleichung [mit Name] erscheint dann mit Symbol = im wf-directory.

²³ Umfasst natürlich ARCH als Spezialfall.

²⁴ Wählt man als Schätzmethode TSLS öffnet sich ein equation specification Fenster, in dem auch ein Feld für "instrument list" ist (man gibt dort die Instrumentvariablen ein).

²⁵ Dann kann man z.B. die unter 5.1 erwähnten Auswertungsmöglichkeiten für eine Variable starten.

5.1. Bei einer einzelnen Variable (Reihe, series)

wenn diese geöffnet ist, erhält man mit dem Anklicken der buttons [= Befehlsflächen] **View** und **Procs**²⁶ auf der "tool bar" im "series window":

View	spread sheet , verschiedene graphische Darstellungen , deskriptive Statistiken (=Stats, z.B. Mittelwerte, Varianzen etc. → E,3) und deren Tests, Correlogramm (ACF, PACF), Einheitswurzeltest (Unit Root Test) auf Stationarität → vgl. unten, BDS ²⁷ Independence Test und Conversion Options (betrifft frequency conversion Methode, die angewendet werden soll bei der entsprechenden Zeitreihe [mehr dazu in G,*])
Procs	Gleichungen und Berechnungen nach eigener Wahl (generate by equation), resample ²⁸ , Verfahren der Trendberechnung , bzw. <i>allgemein</i> Glättung (Hodrick Prescott → G,* , Exponential smoothing ²⁹ [Holt-Winters]), Saisonbereinigung (seasonal adjustment mit Census X12-ARIMA oder Tramo/Seats)

Auswertungsmöglichkeiten **bei einer Variablen** (insbesondere bei einer Zeitreihe)
(die einzelnen Tests kann man auch mit einzutippenden commands statt mit den buttons erreichen)

Deskriptive Statistiken (Häufigkeitsverteilungen)	Zeitreihenoperationen
Stats sind die Mittelwerte, Varianz und andere Momente. Bei Stats+Histogram zusätzlich graphische Darstellung der Häufigkeitsverteilung, Jarque Bera Test auf Normalverteiltheit, ferner Parameter tests gegen hypothetisch angenommene Werte für Mittelwert*, Median, Varianz usw. und Kerndichteschätzung (Anpassung einer stetigen Verteilung)	EViews bietet zahlreiche Möglichkeiten zur Glättung , Saisonbereinigung (Seasonal adjustment nur bei quarterly + monthly data; bei einigen Operationen kann die Meldung kommen, dass die Zeitreihe zu kurz ist) und Schätzung von ARIMA -Modellen. Als Unit Root Tests gibt es ADF- (Augmented Dickey Fuller) und PP-Test (Phillips-Perron)

* Beispiel: Gegeben eine Stichprobe mit $n = \dots$ Werten, einer Häufigkeitsverteilung und einem sich daraus ergebenden Stichprobenmittelwert von 112. Geprüft wird, ob dies z.B. verträglich mit der Nullhypothese $H_0: \mu = 100$ (Zahl muss eingegeben werden) ist oder mit $H_0: \mu = 120$ usw. Entsprechend auch Testen von Varianzen oder des Unterschieds zweier Mittelwerte, zweier Varianzen usw. (aus zwei Stichproben).

Exkurs: Häufigkeitsverteilungen (nicht primär für Ökonometrie von Interesse)

Häufigkeitsverteilungen (**klassierte Verteilungen**) ein- und mehrdimensional als Tabellen laufen unter dem Stichwort (**one-way**) **tabulation** {bei einer series} oder **cross tabulation** {zweidimensional; view im group window}. EViews geht davon aus, dass Querschnittsdaten als Einzelangaben (für jeweils eine Einheit) vorliegen und bietet Optionen (grouping options), wie sie zu Größenklassen oder Gruppen nach Art der Merkmalsausprägung zusammengefasst (to group into bins) werden können. Berechnet werden absolute Häufigkeiten (counts), relative Häufigkeiten (percentages) und Summenhäufigkeiten (cumulatives) sowie Zeilen- und Spaltensummen und Kontingenzmaße bei cross tabulation.

Graphiken (unter Views) hier auch u.a. erreichbar der QQ-plot (vergleicht Quantile der gegebenen empirischen Häufigkeitsverteilung mit denen einer Wahrscheinlichkeitsverteilung, wie der Normal-, Exponential-, logistischen-, Extremwertverteilung usw.). Bei Häufigkeitsverteilungen ist CDF (cumulative distribution function) die Summenhäufigkeitsfunktion $F(x)$; der Ausdruck survivor function (das ist $1-F(x)$) mag etwas ungewohnt erscheinen.

²⁶ Wir beschränken uns weitgehend auf Operationen, die mit View oder Procs gestartet werden. Der Unterschied ist, dass views meist Ansichten der bestehenden (Roh-)Daten sind während bei procs durch Berechnungen an den Daten neue Daten entstehen.

²⁷ Abkürzung steht für die Autoren: Brock, Dechert, Scheinkman. Getestet wird, ob Beobachtungen unabhängig identisch verteilt (iid) sein können.

²⁸ Nicht zu verwechseln mit der Bildung von Teilzeiträumen für die Auswertung (→ **G,4ff**). Resample heißt Einheiten aus den series einer group zufällig auswählen (eher sinnvoll bei Querschnittsdaten, seltener bei Zeitreihen).

²⁹ EViews bietet zahlreiche Möglichkeiten der Prognose und Glättung mit exponential smoothing (worauf hier nicht eingegangen werden kann) und gibt der geglätteten Reihe automatisch einen Namen (hängt SM an den Variablennamen an) und listet sie im wf-directory auf, so dass mit dieser Variable weitergerechnet werden kann.

Weitere Anmerkungen zu Operationen mit Zeitreihen

Vorweg: man sollte immer zuerst eine deskriptive Interpretation der Zeitreihen durchführen, also die Grafik anschauen, deskriptive Statistiken beurteilen etc. Das wird häufig vernachlässigt. Viele fangen sofort mit dem Schätzen und Testen an. Einen Strukturbruch in der Zeitreihe³⁰ beispielsweise sieht man nicht an den Parametern, sehr wohl aber in einer Grafik.

Will man *mehrere Zeitreihen bildlich darstellen* (als Kurvenzug [line] oder mit unterschiedlich hohen Balken [bars]), muss man diese zunächst als eine Gruppe zusammenfassen (wie unter 4.1 beschrieben, markieren) und dann mit **Quick/Graph..** über Views zeichnen lassen.

EViews erlaubt nicht nur die Schätzung autoregressiver (**AR Modelle**) wie $y_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + u_t$, ein ar(3) Prozess³¹ – und die Modellierung einer Störgöße als AR und/oder MA (moving average) Prozess – sondern auch die Schätzung von distributed lag Modellen (verzögerte Variable ist exogen, nicht endogen wie bei AR Prozessen) $y_t = \alpha + \gamma_1 x_{t-1} + \gamma_2 x_{t-2} + \gamma_3 x_{t-3} + \gamma_4 x_{t-4} + u_t$. Weil hier aus verständlichen Gründen schnell Probleme mit Kollinearität entstehen gilt es Restriktionen für die Parameter γ einzuführen (sie etwas als Wahrscheinlichkeiten einer diskreten Wahrscheinlichkeitsverteilung zu konzipieren oder als Funktionswerte eines Polynoms [niedrigerer Ordnung] wie dies im Fall des **polynomial distributed lag**³² (PDL) Modells der Fall ist, das mit EViews geschätzt werden kann.

Einheitswurzeltest (**Unit Root Test** wird auch ausführlich beschrieben in **G,3,7**), d.h. dem Test auf Stationarität³³ einer Zeitreihe: Beim ADF (Augmented Dickey Fuller) Test sind über eine Dialogbox die folgenden Einstellungen (Optionen) möglich

level, 1 st und 2 nd differences	In (Beispiel) $\Delta y_t = \mu + \gamma y_{t-1} + u_t$ liegt ein Test vor für level (also y_t)
mit oder ohne Trend und mit oder ohne intercept	In $\Delta y_t = \mu + \gamma y_{t-1} + u_t$ ist ein intercept (Konstante), mit $\Delta y_t = \mu + \alpha t + \gamma y_{t-1} + u_t$ wäre die Option intercept (μ) und trend (α) gewählt
Anzahl der lags	$\Delta y_t = \mu + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \delta_3 \Delta y_{t-3} + u_t$ number of lags ist 3

Beispiel: beim DF Text wird die Hypothese, dass ϕ in $y_t = \mu + \phi y_{t-1} + u_t$ den Wert 1 einnimmt (= Einheitswurzel = stochastischer Trend = y ist nicht-stationär) wie folgt getestet

$$y_t - y_{t-1} = \mu + (\phi - 1)y_{t-1} + u_t \text{ oder } \Delta y_t = \mu + \gamma y_{t-1} + u_t$$

Statt ϕ gegen 1 zu testen, testet man $\gamma = \phi - 1$ gegen 0. Durch Elemente wie $\delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \delta_3 \Delta y_{t-3}$ ist der DF Test "augmented" zum ADF Test.

Wird H_0 (y_t oder 1st difference Δy_t usw. hat eine Einheitswurzel, ist also nicht-stationär) **abgelehnt**, liegt **Stationarität** vor. Die angegebenen kritischen Werte sind stets negativ (unit-root Test ist immer einseitig), ist ADF Teststatistik größer (z.B. positiv) kann H_0 nicht abgelehnt werden (Nicht-Stationarität, könnte problematisch sein für weitere Auswertungen).

Beim PP-Test wird eine (Newey West) HAC - Korrektur an der std.dev. der t-statistic vorgenommen.

5.2. Bei einer Gruppe

Geht es um Zusammenhänge zwischen mehreren Variablen erhält man die betreffenden Auswertungsmöglichkeiten über buttons auf der tool bar im group window. Es ist somit konsequent, dass man über diesen Weg auch Zugang zum Schätzen von **Regressionsfunktionen** (auch mit mehreren Gleichungen, z.B. VAR-Modelle) und zu verschiedenen plots von **Streuungsdiagrammen** erhält. Neben View und Procs ist auch der **Transform** button nützlich: man kann so die Zahlen für eine Variable y_t in einem spread sheets umstellen auf (transformieren in) Differenzen (Δy_t) oder Wachstumsraten von y_t usw.

³⁰ Strukturbruchtests betreffen eine geschätzte Regressionsfunktion. Sie sind deshalb nur über das Objekt "equation" erreichbar (vgl. unten Abschn. 5.3), nicht über das Objekt "series" (Abschn. 5.1). Ob die Zeitreihe y selbst und nicht die Regressionskoeffizienten in einer Gleichung mit y als abhängiger Variable einen Strukturbruch aufweisen kann man nur an der Graphik erkennen (bei Betrachtung von ihr vermuten).

³¹ AR-Modelle kann man auch für einen Vektor von y -Variablen [y_{1t} y_{2t} ... y_{Gt}] bestimmen – man erhält so ein VAR-Modell – zu schätzen mit procs in der group toolbar (vgl. Abschn. 5.2).

³² Auch (nach dem Autor) Almon Lag genannt.

³³ Mehr zu diesen Tests vgl. meinen Download Nr. 5.

Für eine group kann man per buttons auf der group tool bar (oder über commands) aktivieren:

View	<p>Dated Data Table bietet die Möglichkeit, die Frequenz der Daten umzustellen (etwa Monatsdaten zu Jahresdaten konvertieren) und so automatisch Daten für einen anderen wf zu generieren (im einzelnen beschrieben in → G,6-8). Mit Multiple Graphs und scatter erhält man Zeitreihen und Streuungsdiagramme^{a)} → E,2, F,4, 5,10. Descriptive Statistics liefert Momente und Hypothesentests über diese (Jarque Bera Test → E,5, F,4), eine Kovarianz- und Korrelationstabelle (→ E,2, F,5), und</p> <p>a) bei Querschnittsdaten Hauptkomponenten (auf der Basis der Korrelationstabelle; liefert Eigenwerte und -vektoren → Multivariate Analyse)</p> <p>b) bei Zeitreihen: Correlogram, Correlations (mit Variablen mit lag) und Konintegrations-test (Cointegration Test) sowie Granger Causality (siehe unten)</p>
Procs	<p>Make Equation (Definition einer Regressionsfunktion zwischen den Variablen der Gruppe, EViews schlägt bereits eine vor, die geändert werden kann.³⁴ Man kann auch Ergebnisse wegspeichern^{b)} und in der equation box das sample (den Teilzeitraum) neu definieren. Make Vector Autoregression (VAR)^{c)} und (weniger relevant) Resample.</p>

a) siehe hierzu auch eine Bemerkung unter 5.3.

b) etwa mit **Procs/Make Residual Series** die Residuen benennen und als eigene Variable weiterverarbeiten oder mit "scalar x = eq...@r2 den Wert von r^2 bei der Gleichung "eq..." unter dem Namen x speichern.

c) Man erhält ein Menü (VAR specification window) in dem man das Modell (VAR oder VEC), die endogenen Variablen und die Anzahl der lags eingeben kann.

Regressionsfunktionen und Streuungsdiagramme

Neben dem **Simple Scatter** (Streuungsdiagramm nur mit den Punkten, erste Variable auf der x-Achse, zweite auf der y-Achse) gibt es bei einer *Gruppe* mit **View/Graph/...** unter anderem auch **Scatter with Regression**³⁵ bei dem auch die lineare oder nichtlineare Regressionsfunktion³⁶ eingezeichnet wird (man kann also z.B. die Funktion $\hat{y}_t = \alpha + \beta \ln(x_t)$ oder $\hat{y}_t = \alpha + \beta \sqrt{x_t}$ berechnen oder auch **Box-Cox-Transformation** vornehmen³⁷ (aber immer nur einfache Regression und nur als Graphik, will man die Koeffizienten bestimmen muss die Gleichung als nichtlineare Regression geschätzt werden). Man kann jedoch hier (bei Scatter with Regression) die Variable \hat{y}_t benennen und so in den wf einbringen.³⁸ Andere Möglichkeiten sind etwas umständlicher (vgl. Fußnote 39).

Für diese Graphiken und die Schätzung einer einfachen Regression ist es sinnvoll, **eine ad hoc Gruppe** zu bilden: Bei einem Streuungsdiagramm ist dann die erste Variable die x-Variable (waagrechte Achse, explanatory variable, Regressor), die zweite Var. der Gruppe die y (senkrechte Achse, dependent variable)

Bemerkung zur **Granger Kausalität**: Hat man $v = 2$ Variablen x und y so wird bei zu wählender (dem Programm anzugebender) lag-Länge p geprüft ob im Gleichungssystem

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + \beta_1 x_{t-1} + \dots + \beta_p x_{t-p}$$

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + \dots + \alpha_p x_{t-p} + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p}$$

in der ersten Gleichung $H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$ (x ist nicht Granger kausal für y) und bei der zweiten Gleichung $H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$ (y ist nicht Granger kausal für x). Angegeben werden für alle $v(v-1)/2$ Paare von Variablen in der v -Variablen umfassenden Gruppe die F-Statistik und deren prob-

³⁴ Grundsätze zur Schreibweise: dependent variable, Liste der Regressoren, dann Absolutglied (= c), ohne Kommas nur mit blancs getrennt.

³⁵ Scatter with Nearest Neighbor Fit (und Kernel Fit) ist eine nichtparametrische und lokale (mit jeweils einem wenige Werte umfassenden Stützbereich) Anpassung einer Regressionsfunktion im Unterschied zu einer globalen (alle T Werte benutzenden) Anpassung mit einer expliziten (Parameter enthaltenden) Funktion, wie im Falle der OLS-Schätzung.

³⁶ Besser (weil nicht notwendig ein geschlossener Funktionsausdruck) "fit line".

³⁷ Stellt man bei den Transformationsmöglichkeiten bei der Y- und X-Variable jeweils "none" ein, erhält man die lineare Regression, also die Regressions-gerade im Streuungsdiagramm eingezeichnet.

³⁸ Damit ist es auch möglich, in ein Streuungsdiagramm verschiedene regression lines einzuzeichnen und so zu sehen, welche Art nichtlineare Regression sich der Punktwolke des Streuungsdiagramms am besten anpasst.

value (wenn größer als vorgegebenes Signifikanzniveau, dann H_0 annehmen) sowie die Entscheidung in Worten ausgedrückt. Wird die erste H_0 angenommen, die zweite nicht, dann ist $y \rightarrow x$; es ist auch möglich $x \leftrightarrow y$ und dass weder x für y noch y für x kausal sind. Die Aussagefähigkeit der Betrachtung sollte nicht überbewertet werden.

5.3. Bei einer Gleichung (equation)

Man hat jetzt einige buttons, die spezifisch sind für das object window "equation"; darunter *Name* (auch eine Gleichung sollte benannt werden), *Forecast* und *Resids* (damit erhält man die Möglichkeit, die geschätzten Residuen allein oder zusammen mit den Daten (y_t) und den fitted values (Regresswerte also \hat{y}_t) graphisch darzustellen (man kann das aber auch mit *View* und dann *Actual, Fitted, Residual* wählen; zu diesen Graphiken → **D,2,4** und **F,6,7**). *View* und *Procs* auf der "equation tool bar" liefert

View	man erhält jetzt neben dem üblichen estimation output verschiedene Graphiken (z.B. der Residuen in verschiedenen Varianten, aber <i>nicht</i> Streuungsdiagramme*), ferner einige Tests (→ F,4) im Zusammenhang mit der Beurteilung der Spezifikation, nämlich Coefficient-, Residual- und Stability Tests (siehe <i>Übersicht unten auf dieser Seite</i>)
Procs	mit Make Residual Series kann man die Residuen benennen, um sie von den Residuen anderer Gleichungen zu unterscheiden und getrennt als einzelne Variable (wie unter 5.1) auszuwerten. ³⁹ Weitere Möglichkeiten Forecast (unterscheidet sich nicht von dem entsprechenden button), Make Regressor Group , Make Model (in dem von den geschätzten Koeffizienten in bestimmten Gleichungen Gebrauch gemacht wird) usw.

* Die **Graphiken** bei einer Regression (**Streuungsdiagramme**) sind erreichbar mit dem Views button beim **group window** (siehe Abschn. 5.2), **nicht beim equation window**. Die Graphiken, die man als views im equation window erhält betreffen vor allem die **Störgröße**, allerdings auch die **Konfidenzellipse**.

Übersicht über die mit EViews erhältlichen **Tests bei einer Gleichung**
(die einzelnen Tests kann man auch mit einzutippenden commands statt mit den buttons erreichen)

Coefficient Tests	Residual Tests	(Specification and) Stability Tests
Confidence Ellipse	Correlogram-Q statistic ^{c)}	Chow Breakpoint Test
Wald-Coefficient Restrictions	Correlogram-Squared Residuals ^{d)}	Chow Forecast Test
Omitted Variables-LR Test ^{a)}	Histogram-Normality Test ^{e)}	Ramsey RESET Test (OLS only)
Redundant Variables LR-Test ^{b)}	Serial Correlation LM Test ^{f)}	Recursive Estimates ^{h)} (OLS only)
	ARCH LM Test	
	White Heteroskedasticity ^{g)}	

Fußnoten:

- a) LR = Likelihood Ratio
- b) → **F,7** für eine Anwendung
- c) = **Box Pierce Test** (Ljung Box Q-statistic) → **F,3,4** ,
- d) werden u.a. benötigt für White- und ARCH LM Test
- e) **Jarque Bera Test** (auch als beschreibende Statistik bei Abschn. 5.1; Verletzung der Abnahme B4 = Normalverteilung der Störgröße) → **F,4**,
- f) **Breusch Godfrey Serial Correlation LM Test** (wie c Test auf Autokorrelation) → **F,4,8,9**,
- g) **White Test** auf Homoskedastizität (oder Verletzung der Modellannahmen in einem weiteren Sinne) → **F,3, 9**,
- h) CUSUM Test mit diversen Varianten

³⁹ Im einzelnen auch beschrieben in **F,6**. Ein ähnliches Verfahren (Wegspeichern als eigene Variable) bei den "fitted values" ("y-Dach") ist nicht so bequem. Man erhält sie als Variable im equation window über **Procs/Forecast** oder mit dem **Forecast** button (EViews macht ein "in-sample-forecast", wenn das gesamte Sample zum Schätzen der Regression benutzt wurde), dann einen Namen eingeben, so dass dies als Variable in den wf aufgenommen wird und als Zeitreihe zeichnen lassen (i.d.R. mit eine Standardabweichung der fitted values). Änderungen vornehmen (quasi die ↻ Taste), indem man noch einmal forecast anklickt.

Anders als man vielleicht vermuten könnte erhält man **Streuungsdiagramme** mit und ohne eingezeichnete Regressionsfunktion nicht hier sondern bei den **Gruppen Views**.

Die **Konfidenzellipse** (confidence ellipse, oder confidence set), d.h. die Menge der Kombinationen $\hat{\alpha}, \hat{\beta}$ die bei einem Signifikanzniveau von % zur Annahme der $H_0: \alpha = \beta = 0$ führen.⁴⁰ Das ist einer der Coefficient Tests und wird demonstriert in **F,2**.

Ebenfalls im Übungstext **F** ist mehr gesagt über folgende **Coefficient Tests**

Redundant Variables Test (ausführlich beschrieben in **F,7**): Test ob L bestimmte hinzugekommene Regressoren $x_{K+1}, x_{K+2}, \dots, x_{K+L}$ irrelevant sind (also Test der $H_0: \beta_{K+1} = \beta_{K+2} = \dots = \beta_{K+L} = 0$ mit L Restriktionen) und somit das Modell over-fitted (redundant) machen.

Die Prüfgröße ist die in verschiedenen downloads⁴¹ ausführlich kommentierte $F_{L,T-K-1}$ verteilte Größe

be
$$F = \frac{(S_{\hat{u}\hat{u}}^0 - S_{\hat{u}\hat{u}})/L}{S_{\hat{u}\hat{u}}/(T-K-1)}$$
. Der Vorteil des Tests ist, dass man nicht zwei Regressionen mit zwei

"sum squared resid" ($S_{\hat{u}\hat{u}}^0$ und $S_{\hat{u}\hat{u}}$) berechnen muss nämlich die restringierte (mit $S_{\hat{u}\hat{u}}^0$) und die nicht restringierte (mit $S_{\hat{u}\hat{u}}$), sondern dies alles "in einem Aufwasch" bewerkstelligen kann.

Auf der gleichen Ebene steht der **Omitted Variables Test**, bei dem geprüft wird, ob potenzielle Regressoren $x_{K+1}, x_{K+2}, \dots, x_{K+L}$ fehlen und so das Modell underfitted machen (mit der Konsequenz dass die geschätzten Regressionskoeffizienten verzerrt und inkonsistent sind). EViews fordert auf, eine Liste von $q \geq 1$ potenziellen omitted variables anzugeben ("one or more test series") und liefert eine F verteilte F-statistic und eine χ_q^2 verteilte LR-statistic⁴² und deren prob-values (p-values).

Sind diese kleiner als das vorgegebene Signifikanzniveau wird H_0 (additional set of regressors not jointly significant) abgelehnt (die q Variablen sind nicht "omitted").

Der **Wald Test** ist geboten, wenn man einzelne Koeffizienten nicht gegen 0 sondern einen konkreten Wert oder lineare Restriktionen testet. Das Programm verlangt deshalb Angaben wie $c(2)=0.5$ oder $c(2)=c(3)$.

Tests auf Annahmeverletzungen als Residual Tests (Annahmen B2, B3 bei v.Auer)

Heteroskedastizität (B2) Beim **White Heteroskedasticity Test** wird eine Gleichung für die quadrierten Residuen geschätzt ($H_0 =$ Homoskedastizität). Hierauf wird im download F (Seite **F,1** und **F,9**) ausführlich eingegangen, so dass hier eine weitere Darstellung nicht nötig ist.

Autokorrelation⁴³ (B3) Der **Durbin-Watson** Test wird weitgehend abgelehnt, die DW Statistik aber berechnet. Statt dessen bietet EViews zwei Tests:

1. **Box-Pierce** Test auf der Basis des Korrelograms mit der Ljung Box Q-Statistik mit der Prüfgröße

$$Q = T(T+2) \sum_{j=1}^J \frac{\rho_j^2}{T-J}, \text{ die } \chi^2 \text{ verteilt ist mit } J \text{ Freiheitsgraden, also } \chi_J^2 \text{ verteilt } (\rho_j \text{ ist die Auto-}$$

korrelation j-ter Ordnung) mit Angabe des prob-values (H_0 heißt keine Autokorrelation, liegt vor wenn bis zu dem betreffenden lag die Funktionen ACF und PACF Werte um Null annehmen und der p-Wert [= prob value] groß ist). Kleine p-Werte bei allen lags heißt Autokorrelation (H_0 wird abgelehnt). Das gilt auch für den folgenden Test.

⁴⁰ im Unterschied zur isolierten Betrachtung von einzelnen Konfidenzintervallen (für α und β getrennt).

⁴¹ Nr. 2 und 7 sowie in den Übungsblättern F.

⁴² Log likelihood ratio $LR = -2(l_r - l_u)$ mit l als maximalem Wert der Likelihoodfunktion im restricted (r) und unrestricted (u) Modell. Die F-Statistik beruht wie beim "redundant variables test" auf der Differenz der Summe der quadrierten Residuen.

⁴³ = serial correlation. Die H_0 lautet "keine Autokorrelation". Im Unterschied zum unit root test (oder auch vielen andren Tests) ist hier Bestätigung, nicht Verwerfung von H_0 das gewünschte Ergebnis.

2. **Serial correlation LM** (= Lagrange multiplier) Test; das ist der Test von **Breusch** und **Godfrey**: Mehr hierzu in **F,3,4,8,9** und in der Vorlesung. Auch hier wie beim White Test eine Hilfsregression für die Residuen u_t geschätzt (beim **White** Test oder auch beim **Breusch Pagan** Test von H_0 : Homoskedastizität oder beim **ARCH LM**-Test für die **quadrirten** Residuen).

Korrekturen bei Annahmeverletzungen (B2, B3) nach **White** und **Newey-West** (HAC-Korrektur)

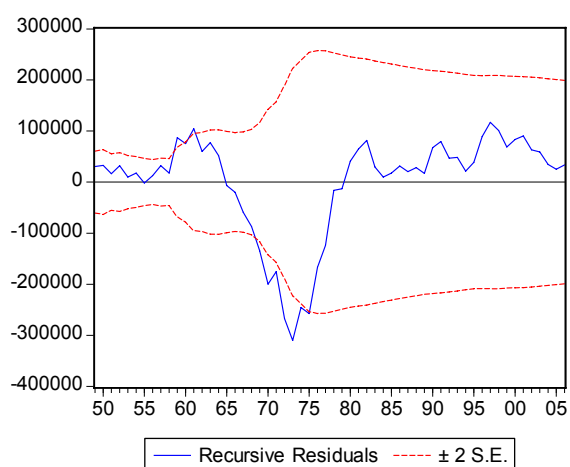
Erreichbar mit **Estimate** Taste auf der Equation Output Seite, dann erhält man eine "Equation Estimation Box" mit zwei Registern "Specification" (erlaubt es, sample neu zu definieren, Gleichung zu verändern, eine andere Schätzmethode als LS zu wählen usw.) und "**Options**". Diese anklicken dann wählen "Heteroskedasticity consistent coefficients covariances". Dann gibt es zwei Möglichkeiten, White (berücksichtigt Verletzung von B2) oder Newey West (berücksichtigt Verletzung von B2 und B3, daher "heteroskedasticity and autocorrelation consistent" = **HAC**). Es ändern sich dann (gegenüber der ersten Schätzung mit OLS) nur die Standardabweichungen und t-Werte der Regressionskoeffizienten, nicht aber die Schätzwerte für die Regressionskoeffizienten, R^2 , DW und andere Größen.

Bei **Stability Tests** (auf Strukturbruch) ist zu beachten, dass es hier um die Annahme A3 (Die Parameter einer Regressionsfunktion sind für alle T Beobachtungen ($y_t, x_{1t}, \dots, x_{kt}$) konstant) im Lehrbuch von v. Auer geht, nicht um Strukturbrüche bei einer Zeitreihe (vgl. Bemerkungen hierzu auf S. 9 oben und Fußnote 30). Es ist daher konsequent, dass dieses Tests nur erreichbar sind über view im equation window, nicht im series window (Abschn. 5.1).

Chow's Breakpoint Test prüft ob sich die Regressionskoeffizienten in m Teilintervallen nicht signifikant unterscheiden (keine "breaks" = H_0). Man gibt $m-1 \geq 1$ Zeitpunkte an, etwa die Jahre 1995 2000 2005 ($m-1 = 3$) bei einer Zeitreihe von 1990 bis 2008 (range) und es geht somit um die 4 Intervalle (1990,1995), (1995,1999), (2000,2004), (2005,2008). Angegeben wird eine F Statistik und eine $\chi^2_{(m-1)k}$ verteilte LR Statistik ($k = K+1$ bei K Regressoren und dem Absolutglied). Man muss die Zeitpunkte potenzieller Strukturbrüche und damit die zeitliche Abgrenzung der m subsamples kennen, anders als beim

Chow's Forecast Test:⁴⁴ gibt man ein Jahr an, etwa 2006 im Beispiel, dann wird geprüft ob die Prognose für den Zeitraum T_{II} (ab 2006 bis 2008) sich signifikant von dem unterscheidet, was man mit den Koeffizienten prognostiziert, die man mit den Daten des Zeitraums T_I (bis 2005) geschätzt hat. Gibt man 2 und mehr Daten an, wird automatisch der Breakpoint Test berechnet. Eine Sequenz von Forecast Tests liegt vor bei der Option N-Step Forecast Test im Rahmen der folgenden Gruppe von Stabilitätstests:

Recursive Estimates liefert rekursive Residuen einer Gleichung (vgl. Bild)

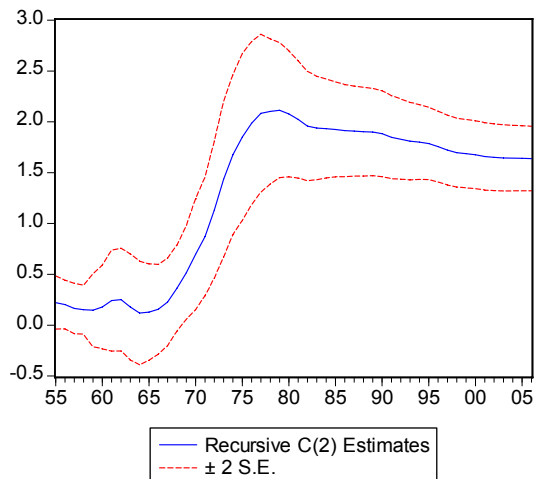


Auf denen der CUSUM (cumulative sum of recursive residuals) und CUSUM of Squares Test beruht. Das nebenstehende Bild bezieht sich auf die Residuen der im Übungsblatt **F** behandelten Schätzung der Anzahl der Geburten, speziell der Gleichung
geb c ehen negeb

also im Unterschied zu **F,6** ohne den Regressor "ueber". Das Bild lässt vermuten, dass Ende der 60er Jahre ein Strukturbruch vorliegen könnte, was sich jedoch bei der Graphik des CUSUM Tests nicht bestätigt.

Man kann auch sehen, wie die rekursiv (mit weniger als T Daten) berechneten Regressionskoeffizienten aussehen (vgl. nächstes Bild für c(2), also den Koeffizienten von "ehen").

⁴⁴ Er heißt "Prognostischer Chow Test" in dem Lehrbuch von L. v. Auer.



Spezifikationstests:¹

Hierzu gehört vor allem der Ramsey **RESET Test**,¹ der – wie in der Vorlesung eingeführt – primär der Entdeckung einer möglichen Nichtlinearität dient.¹ Er beruht auf Einbeziehung von Regresswerten $\hat{y}^2, \dots, \hat{y}^m$ (\hat{y} ist nicht dabei; Kollinearität!) der ursprünglichen Regression in einer Hilfsregression und testet ob alle zugehörigen Regressionskoeffizienten gemeinsam nichtsignifikant sind. H_0 annehmen heißt für Linearität entscheiden. Die Teststatistik ist eine LR Statistik und F nach Art der oben eingerahmten Formel. Das Programm verlangt Angabe der number of fitted items: bei 4 wird gerechnet mit $\hat{y}^2, \dots, \hat{y}^5$.

6. Graphiken und Tabellen, Ergänzungen

Graphiken werden i.d.R. mit View in der entsprechenden tool bar (einer Variable oder Gruppe⁴⁵) erzeugt – worauf in den meisten Fällen bereits hingewiesen wurde – oder mit **Quick/Graph...** (dann series oder group angeben) auf dem main menu. Man kann so schnell Graphiken der verschiedensten Art (Zeitreihen, Kreisdiagramme [pie charts], Streudiagramme [scatter plots] usw. herstellen. Im Folgenden geht es nur um die Gestaltung einer einmal durch eine view-Taste aufgerufenen Graphik.

6.1. Gestaltung von Graphiken

Formatieren

Es empfiehlt sich die Bearbeitung einer zunächst mit view erzeugten Graphik zu beginnen. Man hat dann die folgende tool bar⁴⁶

View	Proc	Object	Print	Name	Freeze	Sample	Sheet*	Stats	Spec
------	------	--------	-------	------	--------	--------	--------	-------	------

* liefert spread sheet Tabelle der Daten

Mit Freeze ("Einfrieren" durch den *freeze* button im object window) entsteht ein *graph object*⁴⁷ (ihm einen Namen geben mit "Name") aus einer graphical view eines Objekts (wie z.B. einer Zeitreihe oder Häufigkeitsverteilung, einer Gleichung oder einer Gruppe) und man erhält auch ein anderes **Object window** mit einer anderen toolbar

View	Proc	Object	Print	Name	Add Text	Line/Shade	Remove	Template*	Options	Zoom
------	------	--------	-------	------	----------	------------	--------	-----------	---------	------

* damit werden die Einstellungen als Muster für andere Graphiken übernommen.

Über Object kann man erneut freeze erreichen. Mit dem button "Options" (oder auch durch Klick an irgendeiner Stelle in der Zeichnungsfläche und Wahl "options" mit rechter Maustaste) erhält man Zugang zum **Graph Options window** mit weiteren Registern:⁴⁸


⁴⁵ Siehe die erste der drei gelb markierten einzeiligen Tabellen mit den aufgelisteten buttons.

⁴⁶ Die obige tool bar betrifft eine Gruppe als Objekt. Sie ist nicht bei jedem Objekt gleich. Bei Zeitreihen gibt es z.B. auch Tasten für Lines und Bars

⁴⁷ Erst mit freeze entsteht ein graph object. Mit Freeze wird nicht eine neue Graphik erzeugt, sondern eine existierende weiter bearbeitet. Man kann sie wieder öffnen indem man auf den für sie vergebenen Namen klickt und **Quick/Show/...** wählen (man kann so auch mehrere Graphiken öffnen und Elemente verschiedener Graphiken miteinander kombinieren).

⁴⁸ Auch über die Object Taste kann view options einstellen und darüber hinaus auch z.B. eine Graphik kopieren,

Type	Frame*	Axes & Scaling	Legend	Lines & Symbols	Fill Areas	Objects	Template
------	--------	----------------	--------	-----------------	------------	---------	----------

* Taste hieß früher "General". Mit frame erhält man weitere Optionen u.a. mit frame attributes und frame size sowie graph position in frame (mit Zahleneingabe in Horizontal indent und Vertical indent kann man Graphik mehr in den Ursprung des Koordinatensystems oder von ihm weg bewegen). Man kann ferner für die frame border verschiedene Formate wählen, wie z.B. nur zwei Achsen  oder vier Achsen.

Wie man sieht erhält man so auch Optionen zur Formatierung der Achsen (und deren Skalierung), der Linien und Symbole (Strickstärke, Farbe), Legende usw.

Kürzerer Zugang zum Graph options window:

Man kann auch in einer Graphik die betreffenden Elemente (eine Achse, Linie, Legende oder generell in der plot area) anklicken und erhält so die passenden Optionen zur Veränderung des Elements.

Mit click and drag kann man auch jederzeit Teile einer Graphik verschieben. Beseitigen eines Elements in einer Graphik: dieses anklicken und Remove button drücken.

Bei Formatierung einer **Multiple Graph** (z.B. mehrere Zeitreihen wie in **F,5**) kann man Optionen finden mit der procs Taste. Man erreicht diese "Multiple Graph Options" auch mit Doppelclick an irgendeiner Stelle der Multiple Graph. Bei den Optionen ist zu unterscheiden: "Options on all graphs" und "Options for selected" (markierte und blau umrandete) graphs.

Drucken und Kopieren (gilt auch für Tabellen)

Der print button im graph view Modus oder der print button nach freezing in der toolbar vom graph object window⁴⁹ erzeugt einen Druck auf Papier. Ausdrucken einer Graphik und z.B. Einbringen in einen Word Text auch möglich mit object/view options/copy. Kopieren (z.B. in andere Windows Programme, etwa Word) bei dem geöffneten Graphik (Tabellen) Objekt auch mit **Edit/Copy** (und dann **Edit/Paste** an der Stelle, an der graph/table object hin soll)

Mit **Files/Print Setup** kann man Einstellungen für das Drucken vornehmen.

Wegspeichern (gilt auch für Tabellen)

Mit **Proc/Save Graph to disk** (dann Pfad angeben) kann man eine Graphik auch als eigene Datei – unabhängig vom wf – ablegen. Das geht auch über proc auf der toolbar des (Graph) object windows (nach freeze).

6.2. Gestaltung von Tabellen

Auch hier ist "table view" (vor freezing) und nach Drücken der **freeze** Taste "table object" zu unterscheiden. Im zweiten Fall hat man mehr Gestaltungsoptionen als bei einer table view. Beim "table object" kann man die Optionen mit procs wählen. Man kann in der Tabelle mit edit +/- Zahlen und Texte ändern (+ wieder edit mode, - protection mode).

Will man z.B. aus dem Ergebnis output einer Gleichung daraus eine Tabelle erzeugen wählt man in der object (hier equation) toolbar Object/Freeze Output dann erscheint ein Tabellen Fenster mit

View	Proc	Object	Print	Name	Edit +/-	CellFmt	InsDel	Grid +/-	Title	Comments +/-
------	------	--------	-------	------	----------	---------	--------	----------	-------	--------------

Die Tabelle ist nach Art einer Excel Tabelle aufgebaut (Spalten A, B, ..., Zeilen 1,2 ...). InsDel erlaubt einfügen und streichen von Feldern, Zeilen, Spalten. Geht nur im Edit mode (edit +). Mit CellFmt erhält man zahlreiche Optionen zu Format (Art der Zahlendarstellung z.B. mit ... Nachkommastellen) Font/Clor (z.B. Schriftgröße) und Borderlines.

6.3. Ergänzungen und Nachträge

Vorgesehen für spätere Versionen dieses Papiers.

⁴⁹ oder einfach zu kopierenden Bereich markieren und mit rechter Maustaste copy wählen (also per copy-and-paste).