Stationariät und Kointegration ökonomischer Zeitreihen
Schätzung von Fehlerkorrektur-Modellen

1. Einleitung

Die in einer Regressionsgleichung enthaltenen Variablen sollten nach gängiger statistisch/ökonometrischer Theorie ***stationären*** Charakter haben, d.h.sie sollten - locker formuliert - mit einer konstanten Varianz um einen konstanten Mittelwert schwanken. Die meisten ökonomischen Variablen sind aber nicht-stationär, sondern sie unterliegen unterschiedlichen und sich zum Teil über die Zeit ändernden Trends. Das Erscheinen solcher Variablen in ökonometrischen Gleichungen ist problematisch, weil sich die Nicht-Stationarität der Variablen im Störterm der Gleichung niederzuschlagen droht, womit die Annahme einer zeitlich unabhängigen Verteilung des Störterms verletzt ist. Dem Auftreten nicht-stationärer Variablen wird in ökonometrischen Untersuchungen auf unterschiedliche Arten begegnet:

1. Häufig wird der nicht-stationäre Charakter der Variablen einfach ***ignoriert*** (wie bisher in dieser Vorlesung). Bei diesem Vorgehen besteht das Problem der ***Scheinkorrelation***: Trendbehaftete Variablen sind häufig statistisch miteinander korreliert und es können sich in Regressionsgleichungen signifikante t-Werte ergeben - auch wenn die Variablen kausal überhaupt nichts miteinander zu tun haben (z.B. abnehmende Population von Störchen und rückläufige Geburtenzahlen).

2. Nicht-stationäre Variablen (x_t) werden durch ***Differenzenbildung*** ($\Delta x_t = x_t - x_{t-1}$) stationär gemacht. Das Problem bei diesem Vorgehen besteht darin, dass durch die Differenzenbildung die in den Daten enthaltene Information über ***langfristige Zusammenhänge*** verloren geht.

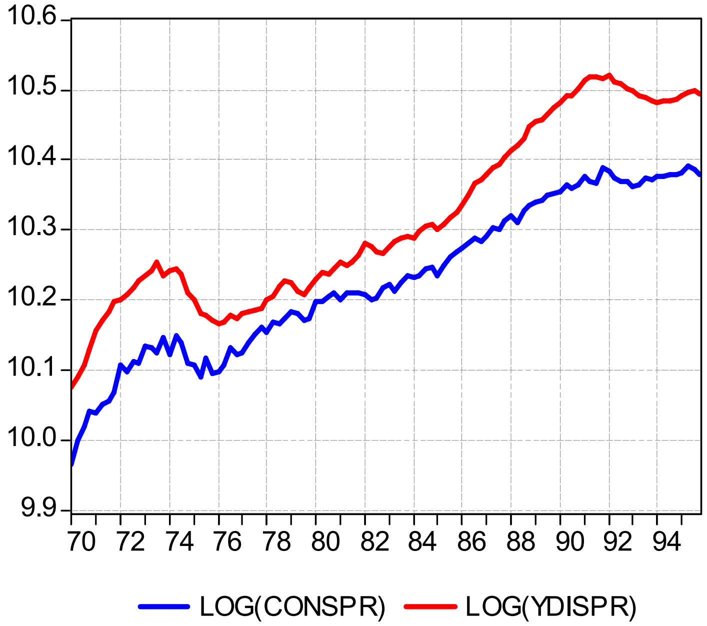
3. Ein Ansatz, der die Probleme von 1 und 2 zu vermeiden sucht, besteht in der Schätzung von ***Fehlerkorrektur-Modellen***. Voraussetzung dazu ist, dass die involvierten Variablen ***kointegriert*** sind. Kointegration ist - wiederum locker formuliert - dann gegeben, wenn die Variablen einzeln nicht-stationär sind, sich aber in einer bestimmten Linearkombination zu einer stationären Variablen zusammenfügen lassen. Diese Linearkombination wird als Gleichgewichtsbeziehung interpretiert, zu der die Variablen nach Störungen mehr oder weniger rasch zurücktendieren.

Beispiel: Konsumfunktion

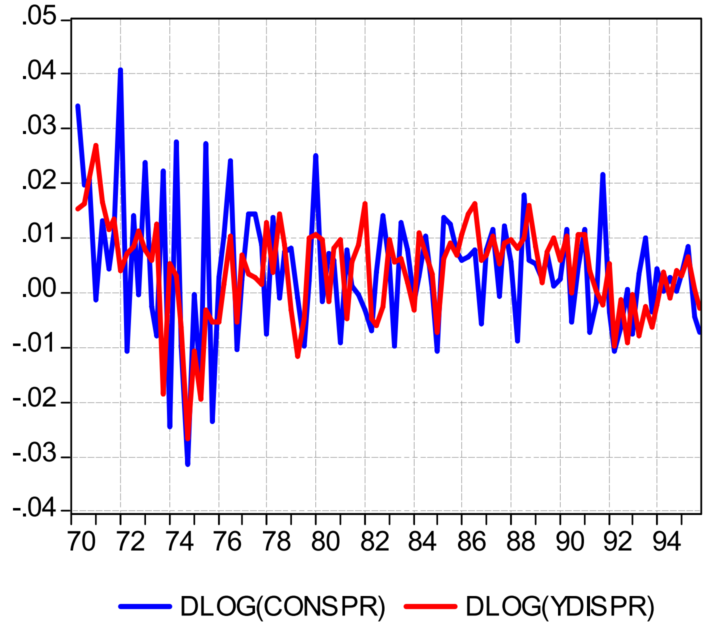
Der gesamtwirtschaftliche reale Konsum (CONSPR) soll mit dem real verfügbaren Einkommen der Haushalte (YDISPR) erklärt werden. Als zusätzlicher Bestimmungsfaktor wird der Anteil der Saisonniers und Grenzgänger an der Bevölkerung (SANT) in Betracht gezogen. Hier kann ein negativer Einfluss vermutet werden, weil die Konsumquote dieser ausländischen Arbeitskräfte unterdurchschnittlich sein dürfte. Die Analyse wird mit Quartalsdaten von von 1970 bis 1995 (alte Volkswirtschaftliche Gesamtrechnung, File KonsumDaten.wf1) in Form eines logarithmischen Gleichungsansatzes durchgeführt.

(In der Übung wird eine ähnliche Fragestellung behandelt. Das aktualisierte Datenset (File Fehlerkorr.wf1) umfasst Quartalsdaten von 1980 bis 2000 gemäss der neuen Volkswirtschaftlichen Gesamtrechnung der Schweiz).

Konsum und Einkommen (real, logarithmiert)



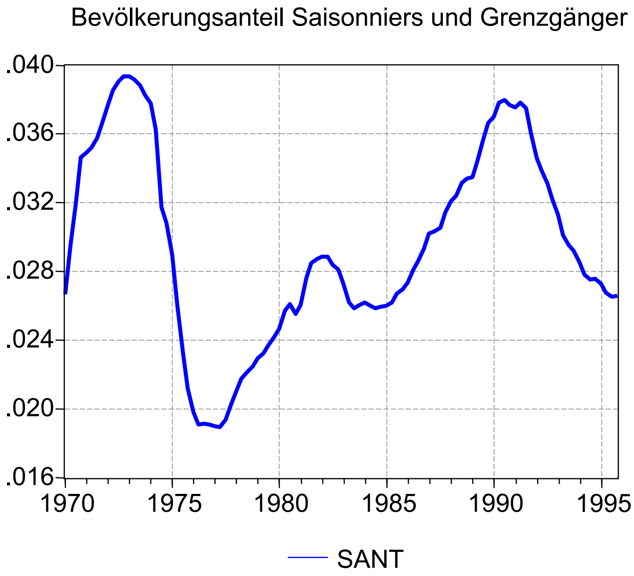
Konsum und Einkommen (real, Veränderungsraten)



Die logarithmierten Variablen LOG(CONSPR) und LOG(YDISPR) sind offensichtlich nicht-stationär. Durch die Bildung von Differenzen, z.B. in Eviews

$$\begin{aligned} \mathrm{DLOG(CONSPR)} &= \Delta \mathrm{LOG}(\mathrm{CONSPR}_t) = \mathrm{LOG}(\mathrm{CONSPR}_t) - \mathrm{LOG}(\mathrm{CONSPR}_{t-1}) \\ &= \mathrm{LOG}(\mathrm{CONSPR}_t/\mathrm{CONSPR}_{t-1}) \approx (\mathrm{CONSPR}_t/\mathrm{CONSPR}_{t-1}) - 1 \end{aligned}$$

erhält man die Vorperioden-Veränderungsraten. In dieser transformierten Form sind die beiden Variablen vermutlich als stationär zu betrachten, d.h. man kann sich vorstellen, dass diese aus Verteilungen mit konstantem Mittelwert und konstanter Varianz stammen.



Ob der Anteil der Saisonniers und Grenzgänger an der Bevölkerung stationär ist oder nicht, kann auf den ersten Blick nicht klar entschieden werden.

1. Wir ignorieren den nicht-stationären Charakter von CONSPR und YDISPR:

Sample: 1970:1 1995:4
Included observations: 104

(1) LOG(CONSPR) = ALFA + BETA*LOG(YDISPR)

	Coefficient	Std. Error	t-Statistic	Prob.
ALFA	1.532199	0.213194	7.186886	0.0000
BETA	0.843062	0.020663	40.80047	0.0000

R-squared 0.942265	Adjusted R-squared 0.941699
S.E. of regression 0.026745	Durbin-Watson stat 0.187587

Die Schätzung der Einkommenselastizität BETA = 0.843 ist hochsignifikant. Ein Warnsignal wird jedoch durch die DW-Statistik von 0.188 gegeben: Positive Autokorrelation der Residuen. Die Voraussetzung eines zufallsverteilten Störterms ist verletzt. Die unter dieser Annahme berechneten statistischen Kennzahlen (Standardfehler, t-Werte) sind nicht aussagekräftig.

2. Wir gehen zu einer Schätzung in ersten Differenzen (Veränderungsraten) über:

Sample(adjusted): 1970:2 1995:4
Included observations: 103 after adjusting endpoints

(2) DLOG(CONSPR) = A + B*DLOG(YDISPR)

	Coefficient	Std. Error	t-Statistic	Prob.
A	0.002148	0.001214	1.768525	0.0800
B	0.457210	0.128155	3.567648	0.0006

R-squared 0.111917	Adjusted R-squared 0.103124
S.E. of regression 0.011137	Durbin-Watson stat 2.694079

Die Schätzung der Einkommenselastizität schrumpft auf B = 0.457 zusammen. Kompensierend besagt die Schätzung A = 0.00215, dass der Konsum bei unverändertem Einkommen pro Quartal um 0.215% zunimmt. Zusammen widerspiegeln diese Ergebnisse die kurzfristige Trägheit der Konsumnachfrage: Die Konsumnachfrage nimmt trendmässig zu (A), reagiert aber eher schwach auf kurzfristige Änderungen im Einkommenswachstum (B). Bei vorübergehend stagnierendem Einkommen nimmt der Konsum aufgrund des positiven A zu. Würde das Einkommen aber tatsächlich über längere Zeit stagnieren, wäre eine Konsumzunahme von fast 1% pro Jahr nicht möglich. Der eher hohe Wert der DW-Statistik von 2.694 deutet im Übrigen auf negative Autokorrelation des Störterms hin, was man als "Überdifferenzierung" der Ausgangsgleichung interpretieren kann.

3. Wir schätzen ein Fehlerkorrektur-Modell. Dieses kann als Erweiterung von Gleichung 2 um den in Gleichung 1 dargestellen langfristigen Zusammenhang zwischen Konsum und Einkommen interpretiert werden. Die Erweiterung besteht darin, dass man die Veränderungsrate des Konsums DLOG(CONSPR) zusätzlich von der Abweichung des Konsums vom langfristigen Normalwert gemäss Gleichung 1 in der Vorperiode abhängig macht. Diese Abweichung ist durch

LOG(CONSPR(-1)) - ALFA - BETA*LOG(YDISPR(-1))

gegeben und entspricht dem um eine Periode verzögerten Residuum von Gleichung 1. Für Parameter GAMMA erwarten wir in der nachfolgenden Gleichung einen negativen Wert, denn wenn der Konsum unter (über) dem langfristigen Normalwert liegt, sollte in der nächsten Periode als Korrekturtendenz - deshalb der Ausdruck “Fehlerkorrektur“-Modell - die Wachstumsrate des Konsums positiv (negativ) beeinflusst werden. Dieser Mechanismus sorgt dafür, dass die Beziehung zwischen Konsum und Einkommen gemäss Gleichung 1 langfristig erfüllt bleibt: LOG(CONSPR) tendiert immer auf ALFA + BETA*LOG(YDISPR) hin, wobei die Stärke dieser Tendenz von Parameter GAMMA gemessen wird. Zusätzlich kann man in die rechte Gleichungsseite - sofern nötig (zur Vermeidung eines autokorrelierten Störterms) - verzögerte Werte der abhängigen Variablen einfügen, z.B. DLOG(CONSPR(-1)):

Sample(adjusted): 1970:3 1995:4
Included observations: 102 after adjusting endpoints
Convergence achieved after 15 iterations

**(3) DLOG(CONSPR) = RHO*DLOG(CONSPR(-1))+B*DLOG(YDISPR)
+GAMMA*(LOG(CONSPR(-1))-ALFA-BETA*LOG(YDISPR(-1)))**

	Coefficient	Std. Error	t-Statistic	Prob.
RHO	-0.332275	0.090564	-3.668963	0.0004
B	0.563708	0.123662	4.558468	0.0000
GAMMA	-0.087519	0.038349	-2.282212	0.0247
ALFA	3.094846	1.159599	2.668895	0.0089
BETA	0.694821	0.111494	6.231923	0.0000

R-squared 0.262055	Adjusted R-squared 0.231624
S.E. of regression 0.010018	Durbin-Watson stat 2.276681

Der Term nach GAMMA ist der sog. ***Fehlerkorrekturterm*** und GAMMA ist der Fehlerkorrekturparameter. Die übrigen Rechthand-Terme stellen die sogenannte ***Kurzfristdynamik*** dar. Die kurz- und langfristigen Einkommenselastizitäten werden auf B = 0.564 bzw. BETA = 0.695 geschätzt. Abweichungen des Konsums vom langfristigen

Normalwert werden in der nächsten Periode (Quartal) gemäss GAMMA jeweils zu 8.75% korrigiert. Diese Korrekturtendenz ist numerisch schwach, erscheint aber statistisch signifikant (t = -2.282).

4. Die vorstehende Gleichung (3) ist nicht-linear in den zu schätzenden Parametern. Die Nicht-Linarität besteht in der multiplikativen Verknüpfung von GAMMA, ALFA und BETA. Steht keine entsprechende Ökonometrie-Software zur Verfügung (Non-Linear Least Squares, NLS), so kann man die Fehlerkorrekturgleichung auch schätzen, indem man zweistufig vorgeht:

- Schätzung der Gleichung (1), Abspeichern der Residuen (z.B unter dem Namen RES1).

- Verwendung der gelagten Residuen RES1(-1) in der Fehlerkorrekturgleichung:

Sample(adjusted): 1970:3 1995:4
Included observations: 102 after adjusting endpoints

(4) DLOG(CONSPR) = A+RHO*DLOG(CONSPR(-1))+B*DLOG(YDISPR) +GAMMA*RES1(-1)

	Coefficient	Std. Error	t-Statistic	Prob.
A	0.002770	0.001122	2.468603	0.0153
RHO	-0.316517	0.090755	-3.487603	0.0007
B	0.580808	0.124194	4.676633	0.0000
GAMMA	-0.085980	0.038645	-2.224875	0.0284

R-squared	0.242412	Adjusted R-squared	0.219221
S.E. of regression	0.010098	Durbin-Watson stat	2.255863

Gemäss dieser Schätzung ist die Einkommenselastizität kurzfristig B = 0.581. Die langfristige Einkommenselastizität beträgt gemäss Gleichung (1) BETA = 0.843.

Wie eingangs erwähnt, sollten die in einer Regressionsgleichung enthaltenen Variablen stationären Charakter haben.

- Diese Voraussetzung ist in Gleichung (1) nicht erfüllt. Deshalb sind die statistischen Testmasse (z.B. t-Werte) nicht interpretierbar. Dies geht auch daraus hervor, dass die Hypothese eines unkorrelierten Störprozesses aufgrund des tiefen DW-Wertes abgelehnt werden muss.

- In Gleichung (2) ist die Voraussetzung stationärer Variablen erfüllt. Durch die Differenzenbildung geht aber Information bezüglich des langfristigen Zusammenhangs zwischen Konsum und Einkommen verloren. Die Gleichung ermöglicht keine Aussage bezüglich des langfristigen Zusammenhangs zwischen Einkommen und Konsum.

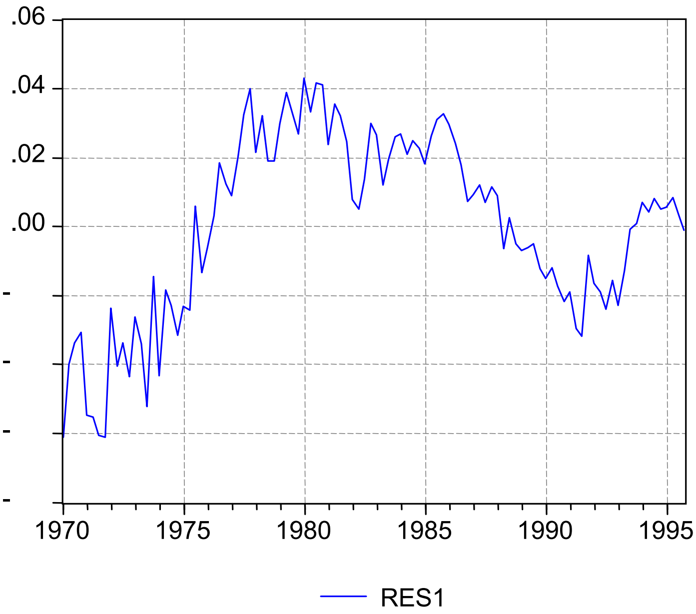
- In Gleichung (3) kommen stationäre Variablen (Veränderungsraten) und nicht-stationäre Variablen (logarithmierte Niveaus) vor. Dies ist auch in Gleichung (4) zu vermuten, denn das in diese Gleichung eingebaute Residuum von Gleichung (1) ist eine Linearkombination nicht-stationärer Variablen und solche Linearkombinationen sind im allgemeinen nicht-stationär. Es ist jedoch nicht ausgeschlossen, dass sich nicht-stationäre Variablen linear zu einer stationären Variablen kombinieren lassen. Dies ist dann der Fall, wenn sich die nicht-stationären Komponenten dieser Variablen gegenseitig neutralisieren. In diesem Fall spricht man von ***Kointegration***.

Konkret stellt sich in unserem Anwendungsbeispiel also die Frage, ob Konsum und Einkommen kointegriert sind, ob also die Linearkombination

LOG(CONSPR) - ALFA - BETA*LOG(YDISPR) = RES1

trotz dem nicht-stationären Charakter von LOG(CONSPR) und LOG(YDISPR) stationär ist. Kann diese Frage bejaht werden, so ist die Schätzung der Fehlerkorrekturgleichungen (3) und (4) zulässig, weil diese Gleichungen dann nur stationäre Variablen enthalten bzw. die Nicht-Stationäritäten sich gegenseitig aufheben. Andernfalls wäre der Störterm der Gleichung als Linearkombination stationärer Variablen mit *einer* nicht-stationären Variablen (dem Fehlerkorrekturterm) zwangläufig nicht-stationär - und damit sicher nicht zeitlich unabhängig normalverteilt. In einer solchen Situation wäre die Schätzung eines Fehlerkorrekturmodells nicht zulässig.

Kointegration von Konsum und Einkommen ist mit anderen Worten dann gegeben, wenn die Residuen von Gleichung 1 als stationär betrachtet werden können.



Ob Stationariät vorliegt, kann visuell nicht eindeutig gesagt werden, scheint aufgrund des grafischen Bildes aber eher fraglich. Im nächsten Abschnitt werden statistische Kriterien bereitgestellt, welche eine genaue Beurteilung dieser Frage zulassen.

2. Stationarität von ökonomischen Zeitreihen (Einheitswurzel-Test)

Variablen bzw. die ihnen zugrundeliegenden stochastischen Prozesse werden als stationär bezeichnet, falls ihre Wahrscheinlichkeitsverteilung zeitinvariant ist. Was allerdings für die meisten Anwendungen erforderlich ist, ist nicht diese sog. strikte Stationarität, sondern schwache Stationarität. Schwache Stationarität (im folgenden einfach Stationarität genannt) liegt vor, falls der Mittelwert und die Varianzen und Kovarianzen endlich und zeitinvariant sind. Stationäre Variablen kehren langfristig immer wieder zu einem wohldefinierten Mittelwert zurück, und die Schwankungen um diesen Mittelwert erfolgen mit einer konstanten Varianz. Bei nicht-stationären Variablen hingegen kann man nicht sinnvollerweise von einem Mittelwert und einer Varianz sprechen (ohne die Aussage auf eine bestimmte Zeitperiode zu beziehen).

Das einfachste Beispiel eines ***nicht-stationären*** stochastischen Prozesses ist der sogenannte “random walk“:

(1) $x_t = x_{t-1} + \varepsilon_t$ mit $\varepsilon_t \sim IN(0, \sigma^2)$

Gemäss dieser Gleichung wird x_t gegenüber dem Vorperiodenwert x_{t-1} immer durch einen unabhängig normalverteilen Störterm ε_t mit Mittelwert 0 nach oben oder unten ausgelenkt. Der in x_t enthaltene Störeinfluss erscheint in der nächsten Periode in Form x_{t-1} auf der rechten Gleichungsseite, bleibt also vollständig erhalten. Dazu kommt wieder ein weiterer Störeinfluss. Weil die ε_t zeitlich unabhängig sind, kann dieser "Zufallsspaziergang" irgendwohin führen. Nimmt man z.B. einen Startwert $x_0 = 0$ an, so gilt

$$x_t = \sum_{i=1}^{t} \varepsilon_i$$

Die Varianz von x_t ist $t\sigma^2$, sie wird also mit fortschreitender Zeit t immer grösser. Deshalb macht es auch keinen Sinn, von einem Mittelwert von x_t zu sprechen, denn es besteht überhaupt keine Tendenz für x_t, auf irgendeinem Wert hin zu tendieren. Die Variable ist somit nicht-stationär.

Ein “random walk“ mit “drift“ stellt sich folgendermassen dar:

(2) $x_t = \delta + x_{t-1} + \varepsilon_t$ mit $\varepsilon_t \sim IN(0, \sigma^2)$

Man spricht in diesem Fall auch von einem ***stochastischen Trend***, weil x_t gegenüber x_{t-1} einerseits systematisch zu- oder abnimmt (je nach dem Vorzeichen des "drift"-Parameters δ), anderseits aber auch rein zufällig ausgelenkt wird und diese Störungen permanent erhalten bleiben.

Ein ***stationärer*** AR(1)-Prozess (autoregressiver Prozess erster Ordnung) unterscheidet sich von einem “random walk“ dadurch, dass der Vorperiodenwert der Variablen mit einem Koeffizienten, der absolut kleiner als Eins ist, in den laufenden Wert der Variablen eingeht:

(3) $x_t = \delta + \rho x_{t-1} + \varepsilon_t$ mit $\varepsilon_t \sim IN(0, \sigma^2)$ und $-1 < \rho < 1$.

Im Folgenden wird positive Autokorrelation angenommen: $0 < \rho < 1$. Der von Gleichung (3) beschriebene Prozess ist stationär, weil die vergangenen Störeinflüsse zwar in Form von x_{t-1} auf der rechten Gleichungsseite wieder erscheinen, dort aber mit ρ $(0 < \rho < 1)$ multipiziert werden. Vergangene Störeinflüsse bleiben somit in x_t nicht vollumfänglich erhalten, sondern werden über die Zeit kontinuierlich abgebaut. Der langfristige Mittelwert dieses Prozesses ist $\mu = \delta/(1-\rho)$ und die Varianz beträgt $var(x) = \sigma^2/(1-\rho^2)$.

(3) lässt sich wie folgt umformulieren:

(4) $\Delta x_t = \delta + (\rho-1)x_{t-1} + \varepsilon_t$ oder

(4') $\Delta x_t = \delta + \theta x_{t-1} + \varepsilon_t$ mit $\theta = \rho-1$.

Von einem ***trendstationären*** Prozess spricht man, falls die Abweichungen von einem ***deterministischen*** Zeittrend stationär sind:

(5) $(x_t - \gamma - \delta\, t) \;=\; \rho(x_{t-1} - \gamma - \delta\,(t-1)) \;+\; \varepsilon_t$ mit $\varepsilon_t \sim IN(0, \sigma^2)$ und $-1 < \rho < 1$.

Wir betrachten wiederum nur den Fall positiver Autokorrelation: $0 < \rho < 1$. Gemäss (5) bleibt x_t immer in der Nähe des linearen Trends $\gamma + \delta\, t$, weil vergangene Störeinfüsse von Periode zu Periode durch Multiplikation mit ρ ($0 < \rho < 1$) abgebaut werden.

(5) lässt sich umformulieren zu:

(6) $\Delta x_t \;=\; \gamma(1-\rho) + \rho\,\delta \;+ \delta\,(1-\rho)\,t \;+ (\rho-1)\,x_{t-1} + \varepsilon_t$ oder

(6') $\Delta x_t \;= a + b\,t + \theta\, x_{t-1} + \varepsilon_t$ mit $\theta = \rho-1$.

Falls die Bedingung $0 < \rho < 1$ nicht gilt, sondern $\rho = 1$ ist, geht (6) über in

$$\Delta x_t \;= \delta \;+\; \varepsilon_t\,,$$

d.h. man erhält einen “random walk mit drift“ bzw. - falls $\delta = 0$ - einen reinen “random walk“, also nicht-stationäre Prozesse. Die Unterscheidung zwischen stationären (bzw. trend-stationären) Prozessen und nicht-stationären Prozessen richtet sich somit danach, ob $0 < \rho < 1$ oder $\rho = 1$ ist. Wenn $\rho = 1$, spricht man von einer Einheitswurzel (*unit root*).

Einheitswurzel-Test

Um abzuklären, ob eine Variable stationär ist, wird getestet, ob man die Null-Hypothese von Nicht-Stationarität ablehnen kann. Als Alternativhypothese kommt entweder Stationarität oder Trend-Stationarität in Frage. Für den logarithmierten realen Konsum LOG(CONSPR) z.B. kommt als Alternativhypothese offensichtlich nur Trend-Stationarität in Frage. Für den Anteil der Saisonniers und Grenzgänger an der Bevölkerung LOG(SANT) hingegen scheint als Alternativhypothese Stationarität sinnvoll.

- Falls die Alternativhypothese auf Stationarität lautet, schätzt man eine Gleichung der Form (4') und testet, ob man die Null-Hypothese einer Eihheitswurzel, d.h. $\theta = 0$ (bzw. $\rho = 1$) zugunsten der Alternativhypothese $\theta < 0$ ablehnen kann (θ signifikant negativ bzw. ρ signifikant kleiner als 1).
- Falls die Alternativhypothese auf Trend-Stationarität lautet, schätzt man eine Gleichung der Form (6') und testet wiederum, ob man die Null-Hypothese eine Einheitswurzel, d.h. $\theta = 0$ (bzw. $\rho = 1$) zugunsten der Alternativhypothese $\theta < 0$ ablehnen kann (θ signifikant negativ bzw. ρ signifikant kleiner als 1).

Die kritischen t-Werte für diese sogenannte “Einheitswurzel-Tests“ entsprechen nicht der üblichen t-Verteilung, da die Variable x unter der Null-Hypothese nicht stationär ist. Die kritischen Werte sind grösser und wurden von Dickey und Fuller tabelliert. Deshalb nennt man den Test auch Dickey-Fuller-Test (DF). Es ist dafür zu sorgen, dass die Störprozesse der Gleichungen (4') bzw. (6') als reine Zufallsprozesse aufgefasst werden können (keine

Autokorrelation). Um dies zu erreichen, müssen unter Umständen gelagte Werte der abhängigen Variablen auf der rechten Gleichungsseite eingebaut werden, also z.B:

(4") $\Delta x_t = \delta + r_1 \Delta x_{t-1} + r_2 \Delta x_{t-2} + \theta x_{t-1} + \varepsilon_t$

In diesem Fall spricht man von einem “Augmented Dickey-Fuller“-Test (ADF-Test).

Integrierte Prozesse und Kointegration

Nicht-stationäre Prozesse werden auch ***integrierte Prozesse*** genannt. Eine Variable x_t kann als Realisation eines sogenannten I(1)-Prozess (integriert vom Grad 1) aufgefasst werden, falls sie durch einmaliges Differenzieren ($\Delta x_t = x_t - x_{t-1}$) stationär wird. Stationarität wird entsprechend auch als I(0)-Prozess bezeichnet. Eine Variable x_t kann als I(d)-Prozess (integriert vom Grad d) aufgefasst werden, falls sie durch d-maliges Differenzieren stationär wird. Die meisten ökonomischen Variablen sind I(1), haben also stationäre erste Differenzen.

Eine Gruppe von I(1)-Variablen wird dann als ***kointegriert*** bezeichnet, falls eine Linearkombination dieser Variablen stationär, also I(0) ist. Ob Kointegration vorliegt, lässt sich anhand der Residuen einer entsprechenden Regression untersuchen. Um z.B. abzuklären, ob zwei Variablen x_t und y_t kointegriert sind, rechnet man die Regression

(7) $y_t = \alpha + \beta x_t + u_t$

und schaut, ob $u_t = y_t - \alpha - \beta x_t$ stationär ist, d.h. ob für die Residuen der Gleichung die Null-Hypothese der Nicht-Stationarität abgelehnt werden kann. Um dies zu testen, führt man für u_t einen Einheitswurzeltest gemäss Gleichung (4') durch, d.h. man rechnet die Regression

(8) $\Delta u_t = \delta + \theta u_{t-1} + \varepsilon_t$

Es ist wiederum dafür zu sorgen, dass der Störprozess dieser Gleichung als reiner Zufallsprozess aufgefasst werden kann (keine Autokorrelation in ε_t). Um dies zu erreichen, muss (8) unter Umständen gemäss (4") erweitert werden. Falls nun θ einen signifikant negativen Wert annimmt, ist die Null-Hypothese von Nicht-Stationarität zugunsten der Alternativhypothese stationärer Residuen abzulehnen. Somit wären x_t und y_t kointegriert, d.h. einzeln I(1) aber in der Linearkombination ($y_t - \alpha - \beta x_t$) I(0).

Die kritischen t-Werte für die Null-Hypothese $\theta = 0$ hängen neben dem Stichprobenumfang von der Anzahl der involvierten Variablen ab (vgl. Tabelle "Critical Values for Cointegration Tests", aus James McKinnon, 1990).

3. Welches ist die adäquate Regressionsgleichung?

Um die adäquate Form einer Regressionsgleichung zu bestimmen, wird für die Variablen zuerst einzeln abgeklärt, ob sie stationär bzw. trend-stationär (I(0)) oder nicht-stationär (I(1)) sind. Danach wird untersucht, ob die I(1)-Variablen kointegriert sind. Letztlich geht es darum, die Regressionsgleichung so zu formulieren, dass sie Annahme eine stationären Störprozesses nicht verletzt wird.

Es gibt drei Fälle mit eindeutigen Implikationen:

1. Falls alle Variablen stationär sind, wird eine Regression mit den Niveauvariablen durchgeführt.
2. Falls neben stationären Variablen auch trendstationäre Variablen auftreten, ist eine Regression in den Niveaus unter Einbezug eines Zeittrends adäquat. Der Zeittrend kompensiert den deterministischen Trend in den trendstationären Variablen.
3. Falls alle Variablen I(1) sind, ist weiter abzuklären ob sie kointegriert sind. Ist dies nicht der Fall (a), so ist eine Regression in ersten Differenzen der Variablen durchzuführen. Liegt Kointegration vor (b), so ist ein Fehlerkorrektur-Modell adäquat.

Schwieriger zu beurteilen sind die folgenden Fälle:

4. Ist eine Variable I(1) und sind alle anderen Variablen stationär oder trendstationär, so ist kein sinnvoller Regressionsansatz möglich, denn die Residuen einer solchen Gleichung wären zwangsläufig nicht-stationär (weil das nicht-stationäre Verhalten der I(1)-Variablen durch die anderen Variablen nicht kompensiert werden kann und sich somit im Residuum niederschlägt). Falls ökonomisch sinnvoll, ist die I(1)-Variable zu differenzieren.
5. Enthält die Regression eine Gruppe x von stationären und trendstationären Variablen sowie eine Gruppe z von nicht-stationären I(1) Variablen, so ist wiederum abzuklären, ob die z-Variablen kointegriert sind. Ist dies nicht der Fall (a), so ist zu überlegen, ob eine Regression mit x und Δz ökonomisch sinnvoll ist. Liegt Kointegration vor (b), so kommt ein Fehlerkorrektur-Modell in Betracht, wobei der Fehlerkorrekturterm nur aus den z-Variablen besteht und die Kurzfristdynamik die Variablen x und Δz enthält.

Beispiel Konsumfunktion

Einheitswurzeltest (DF bzw. ADF-Test)

Privater Konsum, real

Um abzuklären, ob LOG(CONSPR) stationär (I(0)) oder nicht-stationär (I(1)) ist, wird die folgende Regression durchgeführt:

Sample(adjusted): 1970:2 1995:4
Included observations: 103 after adjusting endpoints

DLOG(CONSPR) = A+B*@TREND+THETA*LOG(CONSPR(-1))

	Coefficient	Std. Error	t-Statistic	Prob.
A	2.075897	0.513299	4.044223	0.0001
B	0.000682	0.000189	3.610497	0.0005
THETA	-0.206026	0.051125	-4.029854	0.0001

R-squared 0.163312 S.E. of regression 0.010864
Durbin-Watson stat 2.237349

Der kritische t-Wert (5%) für THETA beträgt: -3.4126 -4.039/103 = -3.452 (Tabelle "Critical Values for Cointegration Tests").
Die Null-Hypothese, THETA = 0 bzw LOG(CONSPR) sei ein I(1)-Prozess, wird somit abgelehnt.
➔ LOG(CONSPR) ist als ***trendstationär*** zu betrachten.

Aufgrund der eher hohen DW-Teststatistik mag aber eine Erweiterung der Gleichung um die gelagte abhängige Variable angebracht erscheinen (ADF-Test):

Sample(adjusted): 1970:3 1995:4
Included observations: 102 after adjusting endpoints

DLOG(CONSPR) = A+R*DLOG(CONSPR(-1))+B*@TREND+THETA*LOG(CONSPR(-1))

	Coefficient	Std. Error	t-Statistic	Prob.
A	1.668603	0.554129	3.011215	0.0033
R	-0.172291	0.093266	-1.847311	0.0677
B	0.000532	0.000202	2.635192	0.0098
THETA	-0.165390	0.055187	-2.996922	0.0035

R-squared 0.145909 S.E. of regression 0.010722
Durbin-Watson stat 1.996586

Gemäss diesem Ansatz kann die Null-Hypothese (THETA = 0) nicht abgelehnt werden.
➔ LOG(CONSPR) ist ***nicht-stationär***.

Das Beispiel zeigt, dass Einheitswurzeltests nicht immer zu eindeutigen Ergebnissen führen. Die Variable LOG(CONSPR) ist sicher nicht stationär; dies zeigt schon ein Blick auf die Grafik. Hingegen lässt sich nicht eindeutig sagen, ob man die Variable als trendstationär oder als nicht-stationär betrachten soll. Schliesst man auf Nicht-Stationarität, so stellt sich weiter die Frage des Integrationsgrades. Nicht-Stationarität bedeutet ja nicht automatisch, dass die Variable als I(1)-Prozess aufzufassen ist. Sie könnte auch I(2) oder noch höher integriert sein. Die Variable ist I(1), falls gezeigt werden kann, dass ihre ersten Differenzen stationär sind. Also muss das gleiche Testverfahren auf DLOG(CONSPR) angewendet und gezeigt werden, dass man die Null-Hypothese THETA = 0 ablehnen kann. Dies ist tatsächlich der Fall (vgl. Übung für das aktualisierte Datenset).

Verfügbares Haushaltseinkommen, real

Analog wird Stationarität von LOG(YDISPR) wie folgt untersucht:

Sample(adjusted): 1970:2 1995:4
Included observations: 103 after adjusting endpoints

DLOG(YDISPR) = A+B*@TREND+THETA*LOG(YDISPR(-1))

	Coefficient	Std. Error	t-Statistic	Prob.
A	0.228169	0.205000	1.113020	0.2684
B	5.30E-05	8.62E-05	0.615489	0.5396
THETA	-0.021994	0.020283	-1.084336	0.2808

R-squared 0.026362 S.E. of regression 0.008575
Durbin-Watson stat 1.018772

Hier ist aufgrund der ausgeprägten Autokorrelation (DW = 1.018772) eindeutig zu einem ADF-Test überzugehen:

Sample(adjusted): 1970:3 1995:4
Included observations: 102 after adjusting endpoints

DLOG(YDISPR) = A+R*DLOG(YDISPR(-1))+B*@TREND+THETA*LOG(YDISPR(-1))

	Coefficient	Std. Error	t-Statistic	Prob.
A	0.301607	0.180270	1.673091	0.0975
R	0.492436	0.087318	5.639596	0.0000
B	0.000104	7.58E-05	1.375359	0.1722
THETA	-0.029578	0.017838	-1.658089	0.1005

R-squared 0.260308 S.E. of regression 0.007484
Durbin-Watson stat 2.235760

Die Null-Hypothese (THETA = 0) kann nicht abgelehnt werden.
➔ LOG(YDISPR) ist ***nicht-stationär.***

Um zu zeigen, dass LOG(YDISPR) einem I(1)-Prozess folgt, ist noch sicherzustellen, dass die ersten Differenzen DLOG(YDISPR) stationär sind.

Bevölkerungsanteil Saisonniers und Grenzgänger

Als Alternativhypothese zur Nullhypothese, dass LOG(SANT) nicht-stationär ist, wird hier Stationarität postuliert (im Unterschied zu Konsum und Einkommen, wo als Alternativhypothese Trendstationarität sinnvoll erschien; vgl. die Grafiken am Anfang dieses Kapitels). Die DF-Testgleichung lautet somit:

Sample(adjusted): 1970:2 1995:4
Included observations: 103 after adjusting endpoints

DLOG(SANT) = DELTA + THETA*LOG(SANT(-1))

	Coefficient	Std. Error	t-Statistic	Prob.
DELTA	-0.054037	0.061566	-0.877720	0.3822
THETA	-0.015247	0.017356	-0.878483	0.3818

R-squared 0.007583 S.E. of regression 0.035090
Durbin-Watson stat 0.394771

Augrund der ausgeprägten Autokorrelation (DW = 0.395) ist wiederum zu einem ADF-Test überzugehen:

Sample(adjusted): 1970:3 1995:4
Included observations: 102 after adjusting endpoints

DLOG(SANT) = DELTA + R*DLOG(SANT(-1)) + THETA*LOG(SANT(-1))

	Coefficient	Std. Error	t-Statistic	Prob.
DELTA	-0.090597	0.035552	-2.548287	0.0124
R	0.776373	0.057208	13.57108	0.0000
THETA	-0.025323	0.010025	-2.525910	0.0131

R-squared 0.652569 S.E. of regression 0.020172
Durbin-Watson stat 2.107505

Die Null-Hypothese (THETA = 0) kann nicht abgelehnt werden.
➔ LOG(SANT) ist nicht-stationär.

LOG(SANT) folgt einem I(1)-Prozess, falls man die Null-Hypothese THETA = 0 für DLOG(SANT) ablehnen kann.

Zusammenfassung der Einheitswurzel-Tests

Aufgrund der vorstehenden Einheitswurzel-Tests ergibt sich die Folgerung, dass LOG(YDISPR) und LOG(SANT) nicht-stationäre Variablen sind. Für LOG(CONSPR) ist nicht ganz eindeutig, ob man auf Nicht-Stationarität oder Trendstationarität schliessen soll. Wir gehen im folgenden von Nicht-Stationarität aus und nehmen weiter an, dass die drei Variablen als Realisationen von I(1)-Prozessen aufzufassen sind, was sich tatsächlich zeigen lässt. Somit stellt sich als nächstes die Frage, ob die Variablen kointegriert sind.

Kointegrationstest

Wir beschränken uns hier auf die Frage, ob LOG(CONSPR) und LOG(YDISPR) kointegriert sind, blenden also die Variable LOG(SANT) aus. Kointegration von LOG(CONSPR) und LOG(YDISPR) ist Voraussetzung für die Zulässigkeit der Fehlerkorrekturgleichungen (3) und (4) auf Seiten 4 und 5. Die Frage lässt sich anhand der Residuen von Gleichung (1) auf Seite 3 beantworten (vgl. Grafik auf Seite 6). Es ist mit einem Einheitswurzel-Test zu zeigen, dass diese Residuen stationär sind. Der ADF-Test ergibt:

Sample(adjusted): 1970:3 1995:4
Included observations: 102 after adjusting endpoints

D(RES1) = DELTA + R*D(RES1(-1)) + THETA*RES1(-1)

	Coefficient	Std. Error	t-Statistic	Prob.
DELTA	0.000650	0.001038	0.626913	0.5322
R	-0.346160	0.090816	-3.811655	0.0002
THETA	-0.082887	0.040297	-2.056907	0.0423

R-squared 0.180761 S.E. of regression 0.010462
Durbin-Watson stat 2.092343

Die Null-Hypothese THETA = 0 (Residuen nicht stationär), kann nicht verworfen werden. Der kritische t-Wert (5%) ist -3.3377 -5.967/102 = -3.396.

➔ Konsum und Einkommen können somit nicht als kointegriert betrachtet werden. Die Schätzung eines Fehlerkorrektur-Modells (Gleichungen 3 und 4 auf Seiten 4 und 5) ist folglich problematisch. Dies zeigt sich dort im numerisch kleinen Wert des Fehlerkorrekturparameters GAMMA.

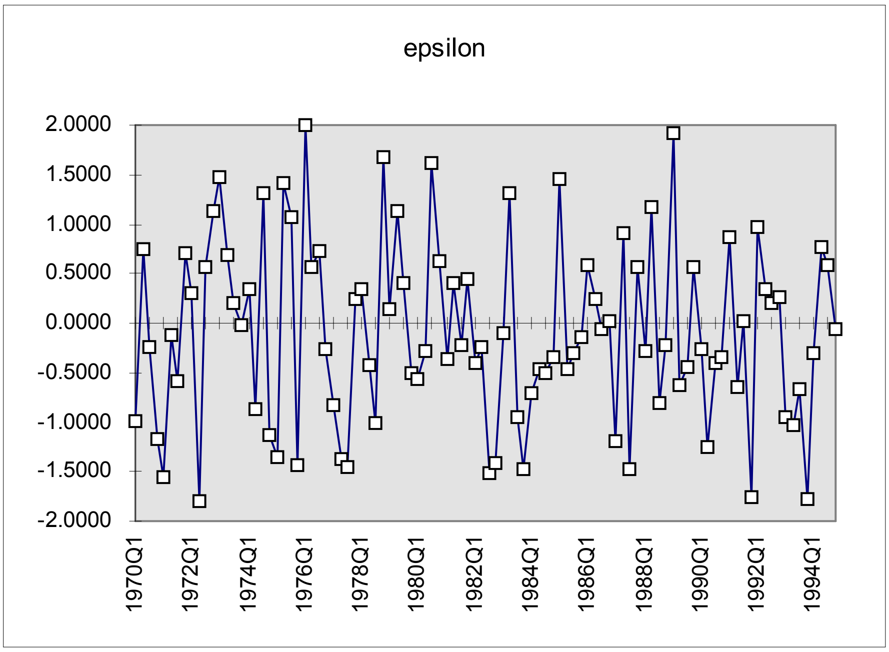
Möglicherweise können die drei Variablen LOG(CONSPR), LOG(YDISPR) und LOG(SANT) als kointegriert betrachtet werden. Sie können dies anhand der im File KonsumDaten.wf1 zur Verfügung gestellten Daten selbst untersuchen.

In der Übung gehen wir anhand eines aktualisierten Datensets (FehlerKorr.wf1, Quartalsdaten 1980-2000) einer ähnlichen Fragestellung nach. Es zeigt sich ebenfalls, dass LOG(CONSPR) und LOG(YDISPR) nicht als kointegriert betrachtet werden können. Als weitere Variablen werden in der Übung die Arbeitslosenrate (URATE) und der Langfristzins (LRATE) in Betracht gezogen. Möglicherweise sind LOG(CONSPR), LOG(YDISPR), URATE und LRATE kointegriert, so dass man mit diesen vier Datenreihen ein Fehlerkorrekturmodell schätzen kann.

Beispiele zu verschiedenen stochastischen Prozessen:

Zuerst generieren wir einen standard-normalverteilten Störterm $\varepsilon_t \sim IN(0, 1)$. Ein mögliches Ergebnis ist in Grafik A dargestellt (Zeitachse willkürlich).

Grafik A: standard-normalverteilter Störterm

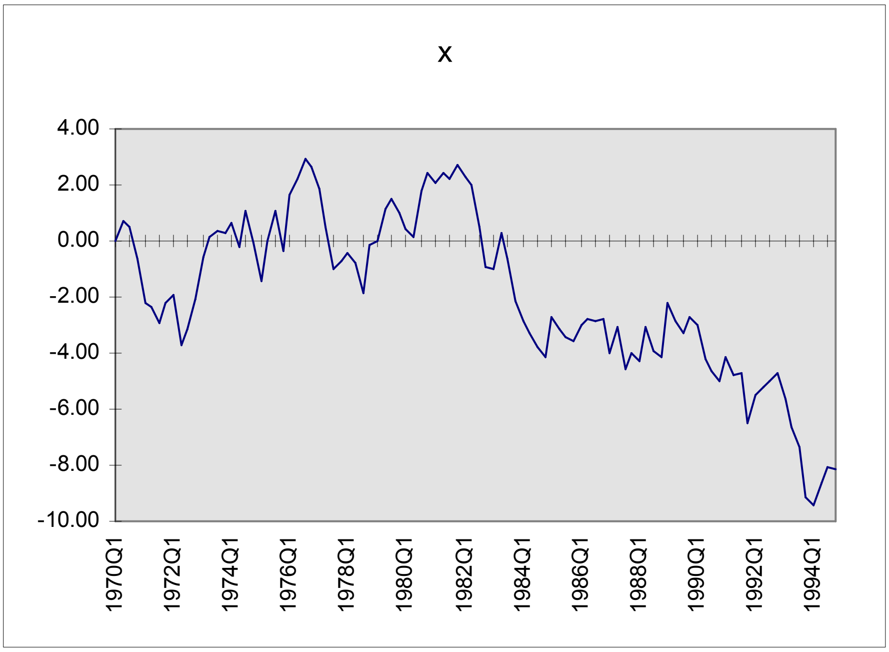


Über die Gleichung

$$x_t = x_{t-1} + \varepsilon_t$$

ergibt sich daraus ein “random walk“, wobei wir als Startwert $x_0 = 0$ wählen (Grafik B).

Grafik B: “random walk“



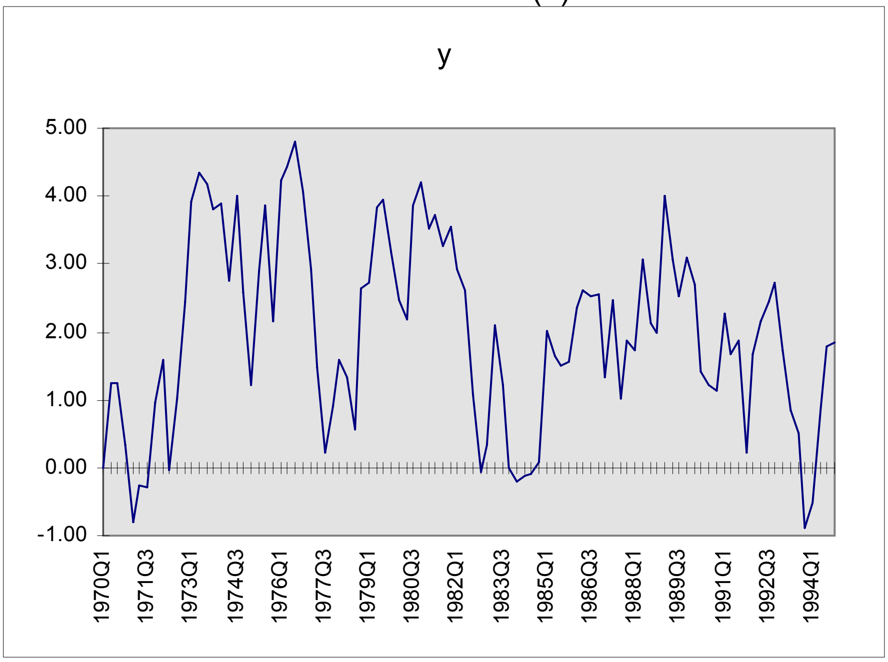
Im weiteren betrachten wir die Prozesse

$$y_t = \delta + \rho y_{t-1} + \varepsilon_t \text{ und}$$

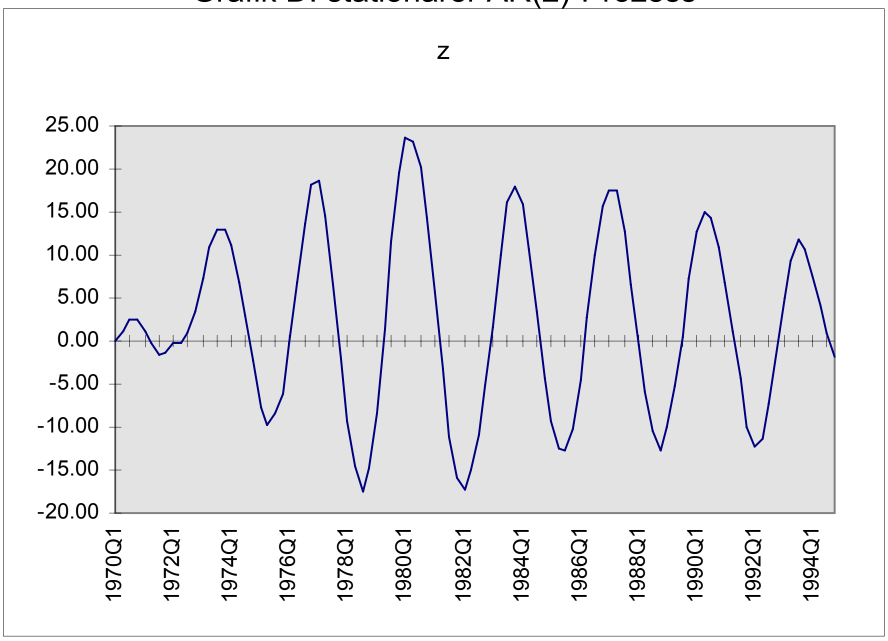
$$z_t = \delta + \rho_1 z_{t-1} + \rho_2 z_{t-2} + \varepsilon_t$$

Für $\delta = 0.5$ und $\rho = 0.8$ bzw. $\rho_1 = 1.8$ und $\rho_2 = -1.0$ sind dies stationäre autoregressive Prozesse erster und zweiter Ordnung (AR(1), AR(2)). Die entsprechenden Verläufe von y_t und z_t sind in Grafiken C und D dargestellt.
Für $\delta = 0.5$ und $\rho = 1.0$ bzw. $\rho_1 = 1.8$ und $\rho_2 = -0.8$ sind dies nicht-stationäre I(1)-Prozesse mit “drift“ δ (vgl. Grafiken E und F).

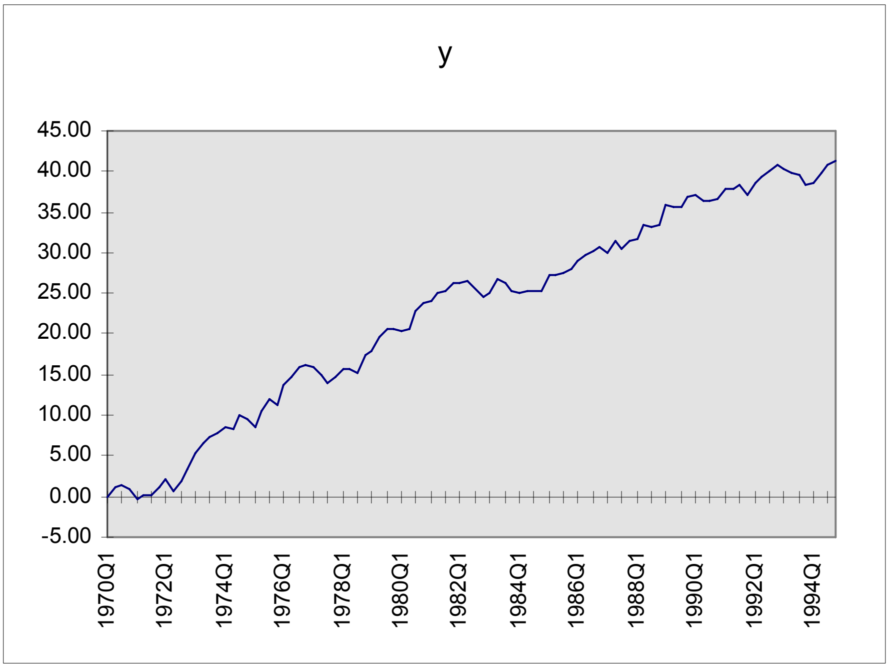
Grafik C: stationärer AR(1)-Prozess



Grafik D: stationärer AR(2)-Prozess



Grafik E: “random walk“ mit “drift“



Grafik F: I(1)-Prozess mit “drift“