

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (<http://www.ub.tuwien.ac.at>).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (<http://www.ub.tuwien.ac.at/englweb/>).



D I P L O M A R B E I T

# Kointegration: Eine mögliche Beziehung zwischen Rohstoff- und Aktienmärkten

ausgeführt am Institut für  
Wirtschaftsmathematik  
Forschungsgruppe Ökonometrie & Systemtheorie  
der Technischen Universität Wien

unter Anleitung von  
Ao.Univ.Prof. Dipl.-Ing. Mag.rer.nat. Dr.techn. Wolfgang Scherrer

durch  
Christian Kletzmayr  
Franz-Lehar-Str. 10  
4540 Bad Hall

---

Datum

---

Unterschrift

# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>1</b>
<b>2. Mehrdimensionale Prozesse</b>	<b>2</b>
2.1. Annahmen & Definitionen . . . . .	2
2.2. Vektor-Autoregressive Prozesse . . . . .	4
2.2.1. Schätzung von VAR-Prozessen . . . . .	5
2.2.2. Spezifikationstests . . . . .	7
2.2.3. Beispiel . . . . .	9
2.3. Nichtstationäre Prozesse . . . . .	14
<b>3. Kointegration</b>	<b>18</b>
3.1. „Spurious“ Regression . . . . .	18
3.2. Definition . . . . .	19
3.3. „Granger Representation Theorem“ . . . . .	23
3.4. Schätzung von kointegrierten Systemen . . . . .	27
3.4.1. Zweistufiges Verfahren von Engle-Granger (1987) . . . . .	27
3.4.2. Johansen Verfahren (Maximum Likelihood) . . . . .	28
3.4.3. Andere Verfahren . . . . .	34
3.5. Kointegrationstests . . . . .	34
3.5.1. Erweiterte „Unit Root“ Tests . . . . .	34
3.5.2. Johansen . . . . .	35
3.5.3. Phillips und Ouliaris (1990) . . . . .	39
3.6. Prognose in mehrdimensionalen Systemen . . . . .	41
<b>4. Eine empirische Untersuchung</b>	<b>44</b>
4.1. Daten . . . . .	44
4.2. Programm . . . . .	48
4.3. Modellbildung . . . . .	48
4.4. Auswertung . . . . .	52
4.5. Prognose . . . . .	59
4.6. Zusammenfassung der Untersuchungen . . . . .	59
<b>A. Stationäre ARMA Prozesse</b>	<b>61</b>
A.1. Stochastische Prozesse, Stationarität, ... . . . . .	61
A.2. „White Noise“ - Weißes Rauschen . . . . .	63

A.3. „Moving Average“ Prozesse . . . . .	64
A.4. Autoregressive Prozesse . . . . .	70
A.5. Autoregressive „Moving Average“ Prozesse . . . . .	74
<b>B. „Unit Root“- Integrationstests</b>	<b>78</b>
B.1. Dickey - Fuller Tests . . . . .	79
B.1.1. Dickey - Fuller t-Test . . . . .	79
B.1.2. „Augmented“ Dickey - Fuller Test . . . . .	82
B.1.3. Dickey - Fuller F-Test . . . . .	84
B.2. Kwiatkowski, Phillips, Schmidt und Shin Test . . . . .	84
B.3. Phillips - Perron Test . . . . .	87
B.4. Zivot - Andrews Test . . . . .	90
<b>C. Ergänzende Tabellen</b>	<b>92</b>
<b>D. „Source Code“</b>	<b>103</b>
<b>E. Zeichenerklärung</b>	<b>108</b>
E.1. Griechische Buchstaben . . . . .	108
E.2. Andere Zeichen . . . . .	108
E.3. Mathematische Symbole . . . . .	109
<b>Literaturverzeichnis</b>	<b>110</b>

# Kapitel 1.

## Einleitung

Seit der Einführung der Kointegration im Jahre 1983 von C.W.J. Granger und der genauen mathematischen Formulierung von Engle und Granger (1987) haben sich zahlreiche Autoren mit dieser Thematik beschäftigt. Neue Modelltypen, vor allem hinsichtlich der deterministischen Komponenten, wurden entwickelt und dazu Teststrategien entworfen, um die anfänglich theoretischen Varianten auch in der Praxis mit echten Daten benützen zu können. In den letzten Jahren wurde vermehrt der Einfluss von Strukturbrüchen auf die Kointegrationsbeziehung von Prozessen untersucht. Diese neueren Forschungsergebnisse werden hier jedoch nicht behandelt und angewendet und wären somit eine logische Fortsetzung der empirischen Analyse.

Die Idee dieser Arbeit ist folgende: Angenommen ein Unternehmen ist sehr stark von einem oder auch mehreren Rohstoffen abhängig, die Entwicklung des Unternehmens ist vollständig durch den Aktienkurs gegeben und die Entwicklung der Rohstoffe wird durch den Weltmarktpreis repräsentiert, dann kann man die Frage stellen, in welcher Beziehung die beiden Zeitreihen zueinander stehen. Da im Allgemeinen die besprochenen Zeitreihen nicht stationär sind, würde eine direkte Regression keine sinnvollen Ergebnisse liefern. Mit Hilfe der Kointegrationstheorie ist es jedoch möglich Beziehungen zwischen nichtstationären Prozessen zu beschreiben.

Für die Analyse der Daten werden in Kapitel 2 die notwendigen, allgemeinen Definitionen für mehrdimensionale Prozesse eingeführt und beschrieben. In Kapitel 3 wird die Kointegration definiert, einer der wichtigsten Sätze („Granger Representation Theorem“) bewiesen und diverse Modelle und Teststrategien vorgestellt. Im letzten Kapitel werden diese Tests angewendet um Kointegrationsbeziehungen für zwei verschiedene Datenzusammenstellung zu finden. Der Anhang bietet im ersten Kapitel eine kurze Einführung in die ARMA-Modelltheorie, enthält im nächsten Kapitel eine Beschreibung von speziell ausgewählten „unit root“ Tests und listet in den restlichen Kapiteln weitere Ergebnisse der empirischen Untersuchungen und dort verwendete Programme auf.

## Kapitel 2.

# Mehrdimensionale Prozesse

Dieses Kapitel bietet einen kurzen Überblick über die mehrdimensionale Zeitreihenanalyse, die als Grundlage für das nächste Kapitel „Kointegration“ dient.

Einführende Worte über Zeitreihen, stochastische Prozesse und deren Realisationen findet man im Anhang A. Dort werden die wichtigsten Prozesse für den eindimensionalen Fall beschrieben und deren Eigenschaften analysiert.

Wie im Anhang wird auch im Hauptteil dieser Arbeit die Bezeichnung

$$\mathbf{y}_t = \begin{pmatrix} y_{1,t} \\ \vdots \\ y_{n,t} \end{pmatrix}$$

sowohl für die beobachtete Zeitreihe als auch für den stochastischen Prozess verwendet.

### 2.1. Annahmen & Definitionen

In der Literatur werden zwei Arten von Stationarität unterschieden: *starke* und *schwache* Stationarität. In dieser Arbeit benötigt man aber nur das Konzept der schwachen Stationarität, die hier für mehrdimensionale Prozesse definiert wird.<sup>1</sup>

**Definition 2.1 (Schwache Stationarität)** Ein  $n$ -dimensionaler stochastischer Prozess  $\mathbf{y}_t$  ist schwach stationär, wenn gilt:

$$\mu = \mathbf{E}[\mathbf{y}_t] = \mathbf{E} \begin{pmatrix} y_{1,t} \\ \vdots \\ y_{n,t} \end{pmatrix} \quad \forall t \in \mathbb{Z} \quad \text{und} \quad (2.1)$$

$$h \in \mathbb{Z} : \mathbf{\Gamma}(h) = \begin{pmatrix} \mathbf{Cov}[y_{1,t+h}, y_{1,t}] & \dots & \mathbf{Cov}[y_{1,t+h}, y_{n,t}] \\ \vdots & \ddots & \vdots \\ \mathbf{Cov}[y_{n,t+h}, y_{1,t}] & \dots & \mathbf{Cov}[y_{n,t+h}, y_{n,t}] \end{pmatrix} \quad \forall t \in \mathbb{Z}. \quad (2.2)$$

---

<sup>1</sup>Die Bezeichnung schwache Stationarität ist nicht immer treffend, da es Fälle gibt, wo die starke Stationarität schwächer ist als die schwache Stationarität. Darum wird oft auch die Bezeichnung Kovarianz Stationarität verwendet.

Im Kapitel “Stationäre ARMA Prozesse“ (siehe Anhang A) werden beide Definitionen für den eindimensionalen Fall eingeführt. Wird im Verlauf dieser Arbeit von stationären Prozessen gesprochen, so ist damit immer die Schwache Stationarität gemeint.

Wie auch im eindimensionalen Fall wird sehr oft ein einfacher stochastischer Prozess mit folgenden Eigenschaften benötigt:

**Definition 2.2** („white noise“) *Ein  $n$ -dimensionaler stochastischer Prozess  $\varepsilon_t$  mit*

$$\mathbf{E}[\varepsilon_t] = 0 \quad \forall t \in \mathbb{Z} \quad \text{und} \quad (2.3)$$

$$\Gamma(h) = \begin{cases} \Sigma, & \text{wenn } h = 0 \\ 0, & \text{sonst} \end{cases} \quad (2.4)$$

heißt „white noise“ Prozess, i. Z.  $\varepsilon_t \sim WN(0, \Sigma)$ , mit  $\Sigma \geq 0$  einer  $n \times n$ -Matrix.

Ist der Prozess  $\varepsilon_t$  unabhängig und identisch verteilt, schreibt man  $\varepsilon_t \sim IID(0, \Sigma)$  (IID = independent identically distributed); ist dieser sogar normalverteilt bedeutet das in Zeichen  $\varepsilon_t \sim NID(0, \Sigma)$  (NID = normally identically distributed).

Wie die beiden vorherigen Definitionen zeigen, spielt die Kovarianz in der Betrachtung von stochastischen Prozessen eine wichtige Rolle. Sie wird daher als eigene Funktion definiert.

**Definition 2.3 (Autokovarianz- und Autokorrelationsfunktion)** *Die Autokovarianzfunktion eines stationären Prozesses  $\mathbf{y}_t$  ist definiert als*

$$h \in \mathbb{Z} : \quad \Gamma(h) = \begin{pmatrix} \gamma_{11}(h) & \dots & \gamma_{1n}(h) \\ \vdots & \ddots & \vdots \\ \gamma_{n1}(h) & \dots & \gamma_{nn}(h) \end{pmatrix} \quad \forall t \in \mathbb{Z} \quad (2.5)$$

mit  $\gamma_{ij}(h) = \mathbf{Cov}[y_{i,t+h}, y_{j,t}]$ .

Die Autokorrelationsfunktion (ACF) eines stationären Prozesses  $\mathbf{y}_t$  ist definiert als

$$h \in \mathbb{Z} : \quad \mathcal{R}(h) = \begin{pmatrix} \rho_{11}(h) & \dots & \rho_{1n}(h) \\ \vdots & \ddots & \vdots \\ \rho_{n1}(h) & \dots & \rho_{nn}(h) \end{pmatrix} \quad \forall t \in \mathbb{Z} \quad (2.6)$$

mit

$$\rho_{ij}(h) = \frac{\gamma_{ij}(h)}{\sqrt{\gamma_{ii}(h)\gamma_{jj}(h)}}.$$

Die beiden Funktion  $\gamma_{ij}(h)$  und  $\rho_{ij}(h)$  werden „cross-covariance“- und „cross-correlation“-Funktion genannt. Im Allgemeinen ist  $\gamma_{ij}(h)$  nicht gleich  $\gamma_{ji}(h)$  und damit auch

$\rho_{ij}(h) \neq \rho_{ji}(h)$ . Analog zum eindimensionalen Fall, wo  $\gamma_y(h) = \gamma_y(-h)$  gilt (siehe Anhang A), gibt es auch für die Kovarianzmatrix eine ähnliche Beziehung:

$$\Gamma(h) = \Gamma'(-h).$$

Schätzer für den Erwartungswert und die Kovarianzmatrix von  $\mathbf{y}_t$  sind das Stichprobenmittel

$$\bar{\mathbf{y}} = \begin{pmatrix} \frac{1}{T} \sum_{t=1}^T y_{1,t} \\ \vdots \\ \frac{1}{T} \sum_{t=1}^T y_{n,t} \end{pmatrix} \quad (2.7)$$

und die Stichprobenkovarianz

$$\hat{\Gamma}(h) = \frac{1}{T} \sum_{t=1}^{T-h} (\mathbf{y}_{t+h} - \bar{\mathbf{y}})(\mathbf{y}_t - \bar{\mathbf{y}})' \quad 0 \leq h < T. \quad (2.8)$$

## 2.2. Vektor-Autoregressive Prozesse

Das Basismodell für die spätere Betrachtung von Kointegrationsbeziehungen zwischen verschiedenen Zeitreihen ist das Vektor-Autoregressive Modell.

**Definition 2.4 (Vektor-Autoregressiver Prozess)** *Ein Vektor-Autoregressiver Prozess der Ordnung  $p$  (VAR( $p$ )) ist die Lösung der Differenzgleichung*

$$\mathbf{y}_t = \Phi_1 \mathbf{y}_{t-1} + \dots + \Phi_p \mathbf{y}_{t-p} + \kappa D_t + \varepsilon_t \quad (2.9)$$

wobei  $\varepsilon_t \sim WN(0, \Sigma)$  ist,  $\Phi_i$   $n \times n$ -Matrizen sind und  $\kappa$  eine  $n \times k$ -Matrix ist.

In  $D_t$  steckt der deterministische Teil des VAR Modells. Dieser kann zum Beispiel eine Konstante, ein linearer Trend, Saisondummies, usw. sein. Enthält  $D_t$  eine Konstante und einen linearen Trend ( $D_t = [1, t]'$ ) so ist  $\mathbf{y}_t$  sicher nicht stationär und die unten angegebene Lösung (Satz 2.1) teilt sich in einen stationären und nichtstationären Teil.

Mit Hilfe des „Backshift“-Operators:

$$B^l \mathbf{z}_t \equiv \mathbf{z}_{t-l}, \quad t \in \mathbb{Z} \quad (2.10)$$

lässt sich obiger Prozess kurz schreiben als

$$\Phi(B) \mathbf{y}_t = \kappa D_t + \varepsilon_t$$

wobei  $\Phi(B) = (1 - \Phi_1 B - \dots - \Phi_p B^p)$ .

Für die Darstellung einer Lösung ist folgende Definition hilfreich:

**Definition 2.5 (Kausaler Prozess)** Ein  $n$ -dimensionaler stochastischer Prozess  $\mathbf{y}_t$  wird auch kausal genannt, falls der Prozess folgende Darstellung besitzt:

$$\mathbf{y}_t = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j} \quad \forall t \in \mathbb{Z} \quad (2.11)$$

wobei  $\varepsilon_t \sim WN(0, \Sigma)$  und  $\Psi_j$   $n \times n$  Matrizen mit  $\sum_{j=0}^{\infty} \sqrt{sp\{\Psi_j \Psi_j'\}} < \infty$  sind.

Aufgrund der letzten Bedingung in der Definition eines kausalen Prozesses, sieht man sofort, dass aus der Kausalität die Stationarität folgt.

Die Frage nach der Stationarität und Gestalt einer Lösung von  $\mathbf{y}_t - \mathbf{E}[\mathbf{y}_t]$  beantwortet folgender Satz:

**Satz 2.1** Ist der deterministische Teil  $D_t$  durch ein Polynom in  $t$  beschränkt und gilt

$$\det \Phi(\mathbf{z}) = \det (I_n - \Phi_1 \mathbf{z} - \dots - \Phi_p \mathbf{z}^p) \neq 0 \quad \forall |\mathbf{z}| \leq 1$$

dann existiert eine kausale (und daher stationäre) Lösung von  $\mathbf{y}_t - \mathbf{E}[\mathbf{y}_t]$ . In diesem Fall besitzt der Prozess  $\mathbf{y}_t$  folgende Darstellung:

$$\mathbf{y}_t = \Phi^{-1}(B)(\varepsilon_t + \kappa D_t) + \mathbf{y}_t^0 = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j} + \Phi^{-1}(B)\kappa D_t + \mathbf{y}_t^0$$

wobei  $\mathbf{y}_t^0$  eine deterministische Lösung der homogenen Gleichung

$$\mathbf{y}_t = \Phi_1 \mathbf{y}_{t-1} + \dots + \Phi_p \mathbf{y}_{t-p}$$

ist; im Weiteren wird  $\mathbf{y}_t^0 = 0$  gesetzt.

**Beweis 2.1** siehe Johansen (1995)

Johansen (1995) beweist obigen Satz zuerst für  $p = 1$ . Für den allgemeinen Fall transformiert er den VAR( $p$ ) Prozess in einen VAR(1) und verwendet eine ähnliche Beweisführung wie im Fall  $p = 1$ .

### 2.2.1. Schätzung von VAR-Prozessen

Für die konkrete Schätzung eines VAR Prozesses mit deterministischen Teil  $D_t$

$$\Phi(B)\mathbf{y}_t = \kappa D_t + \varepsilon_t \quad (2.12)$$

werden die Fehler als multivariant normalverteilt  $\varepsilon_t \sim NID(0, \Sigma)$  angenommen. Die Startwerte  $\mathbf{y}_0, \dots, \mathbf{y}_{-p+1}$  sind fix gegeben und die Parameter

$$(\Phi_1, \dots, \Phi_p, \kappa, \Sigma)$$

haben keine Einschränkung. Um eine kompaktere Form zu erhalten wird folgende Notation eingeführt:  $Z_t' = (y_{t-1}', \dots, y_{t-p}', D_t')$  und  $\tilde{\Phi}' = (\Phi_1, \dots, \Phi_p, \kappa)$ . Damit hat der Vektor  $Z_t$  die Dimension  $np+k$  und  $\tilde{\Phi}'$  ist eine Matrix der Dimension  $n \times np+k$ . Die Gleichung (2.12) verändert sich somit zu

$$y_t = \tilde{\Phi}' Z_t + \varepsilon_t \quad t = 1, \dots, T. \quad (2.13)$$

Durch die normalverteilten Innovationen ist auch der Vektor-Autoregressive Prozess normalverteilt und somit eine Schätzung der Parameter mit Hilfe der Likelihood Funktion möglich. Die logarithmierte Likelihood Funktion

$$\log L(\tilde{\Phi}, \Sigma) = -\frac{1}{2}T \log(2\pi) - \frac{1}{2}T \log(\det \Sigma) - \frac{1}{2} \sum_{t=1}^T (\mathbf{y}_t - \tilde{\Phi}' Z_t)' \Sigma^{-1} (\mathbf{y}_t - \tilde{\Phi}' Z_t)$$

führt zur Gleichung für den Parameter  $B$

$$\sum_{t=1}^T \mathbf{y}_t Z_t' = \hat{\tilde{\Phi}}' \sum_{t=1}^T Z_t Z_t'$$

und zu den „gewöhnlichen“ Regressionsschätzern:

$$\hat{\tilde{\Phi}} = \left( \sum_{t=1}^T Z_t Z_t' \right)^{-1} \left( \sum_{t=1}^T Z_t \mathbf{y}_t' \right) \quad (2.14)$$

$$\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_t \hat{\varepsilon}_t' \quad (2.15)$$

mit  $\hat{\varepsilon}_t = \mathbf{y}_t - \hat{\tilde{\Phi}}' Z_t$ . Diese Schätzer erhält man auch mit dem OLS-Verfahren.

Die asymptotische Verteilung des Parameters  $\tilde{\Phi}'$  zeigt folgender Satz:

**Satz 2.2 (asymptotische Verteilung von  $\tilde{\Phi}'$ )** *Sind die  $\varepsilon_t$  unabhängig und identisch verteilt mit Erwartungswert Null und Varianz  $\Sigma$ ,  $\mathbf{y}_t$  stationär und  $\kappa D_t = \mu$ , dann ist die asymptotische Verteilung des Maximum Likelihood Schätzers für  $\tilde{\Phi}' = (\Phi_1, \dots, \Phi_p, \kappa)$  normalverteilt und gegeben durch*

$$T^{\frac{1}{2}}(\hat{\tilde{\Phi}} - \tilde{\Phi}) \xrightarrow{d} N(0, \Omega^{-1} \otimes \Sigma)$$

wobei

$$\hat{\Omega} = T^{-1} \sum_{t=1}^T Z_t Z_t' \xrightarrow{p} \Omega, \quad \text{and} \quad \hat{\Sigma} \xrightarrow{p} \Sigma.$$

**Beweis 2.2** *siehe Johansen (1995)*

Oben angeführte Aussagen beziehen sich auf das uneingeschränkte VAR Modell. Gelten für die Parameter bestimmte Restriktionen, können natürlich analoge Aussagen gemacht werden. Werden einige Lags ausgelassen, d.h. es werden zum Beispiel nur die Verzögerungen 1, 2 und 4 betrachtet, so ist die Maximum Likelihood Schätzung für so ein Modell wieder gleich der OLS-Schätzung. Andere, jedoch asymptotisch äquivalente Schätzer sind zum Beispiel:

- Yule-Walker Schätzer
- Burg Schätzer.

### 2.2.2. Spezifikationstests

Die Schätzung eines VAR Modells mit bekannter Ordnung  $p$  kann mit jedem herkömmlichen Regressionsprogramm durchgeführt werden. Damit die Schätzer ihre Gültigkeit haben, müssen jedoch einige Annahmen erfüllt werden. Im Besonderen sollten die Residuen unkorreliert und keine bedingte Heteroskedastizität aufweisen. Gilt außerdem die Annahme  $\varepsilon_t \sim NID(0, \Sigma)$ , so sollten die geschätzten Fehler annähernd normalverteilt sein. In der Praxis ist die Ordnung des Modells unbekannt und muss zuerst ermittelt werden. Eine solche Schätzung steht natürlich in enger Verbindung mit den Eigenschaften der Residuen.

#### Bestimmung der Ordnung

Die Ordnung  $p$  eines VAR Modells kann durch eine der zahlreichen „Informations“-Kriterien berechnet werden. Als Beispiele werden hier das Akaike Informations Kriterium und das Hannan-Quinn Kriterium angegeben (für eine genauere Diskussion siehe Lütkepohl (1991))

$$AIC(p) = \log |\hat{\Sigma}(p)| + \frac{2pn^2}{T}$$

$$HQ(p) = \log |\hat{\Sigma}(p)| + \frac{2 \log \log T}{T} pn^2.$$

Das jeweils minimalste AIC oder HQ bestimmen die Ordnung des Modells.

Eine andere Möglichkeit besteht darin mit einem überparameterisierten Modell zu starten und sukzessive die Ordnungszahl zu verringern, bis der letzte Koeffizient  $\Phi_p$  signifikant von Null verschieden ist. Als Test bietet sich zum Beispiel der Likelihood Verhältnis Test an. Unter der Nullhypothese  $\Phi_p = 0$  ist die Teststatistik

$$\lambda_{LR}(p) = T(\log |\hat{\Sigma}(p-1)| - \log |\hat{\Sigma}(p)|)$$

asymptotisch  $\chi^2$  verteilt mit  $n^2$  Freiheitsgraden.

## Korrelation der Residuen

Ist die Laglänge und die Datenzusammenstellung richtig gewählt, sollte ein Autokorrelations- und „cross“-Korrelationsplot der Residuen keine wesentlichen Abhängigkeiten mehr zeigen. Dieser ersten graphischen Analyse kann nun ein statistischer Test folgen, wie zum Beispiel der Portmanteau Test. Die Teststatistik

$$P(s) = T^2 \sum_{j=1}^s \frac{1}{T-j} sp\{\hat{C}_i \hat{C}_0^{-1} \hat{C}_i' \hat{C}_0^{-1}\}$$

wobei

$$\hat{C}_i = T^{-1} \sum_{t=i+1}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-i}'$$

ist für große  $T$  und  $s$  asymptotisch  $\chi^2$  verteilt mit  $n^2(s-p)$  Freiheitsgraden.

## Autoregressive conditional heteroscedastic (ARCH)

R. Engle publizierte 1982 eine einfache Prozedur, unter der Annahme, dass  $\varepsilon_{i,t} \sim NID(0, \sigma^2)$ , um ARCH Effekte festzustellen. Für jede Komponente der Residuen wird folgende Regression durchgeführt:

$$\varepsilon_{i,t}^2 = \phi_0^2 + \phi_1^2 \varepsilon_{i,t-1}^2 + \dots + \phi_j^2 \varepsilon_{i,t-j}^2 + \xi_{i,t}.$$

Die Teststatistik  $TR^2$  mit

$$R^2 = 1 - \frac{\sum_{t=1}^T \xi_{i,t}^2}{\sum_{t=1}^T \left( \varepsilon_{i,t}^2 - \frac{1}{T} \sum_{t=1}^T \varepsilon_{i,t}^2 \right)^2}$$

ist asymptotisch  $\chi^2$  verteilt mit  $j$  Freiheitsgraden.

## Test auf Normalverteilung

Als erster schneller Check kann bereits ein Histogramm grundlegende Auskunft über die Verteilung der Residuen geben. Der mit Abstand jedoch wichtigste Test ist der Jarque-Bera Test, der die aus den Daten generierte Schiefe und Kurtosis mit jener der Normalverteilung vergleicht. Die Teststatistik lautet

$$JB = T \left( \frac{b_1' b_1}{6} + \frac{(b_2 - 3I_n)' (b_2 - 3I_n)}{24} \right)$$

wobei

$$b_1 = (b_{11}, \dots, b_{n1})' \quad \text{mit} \quad b_{k1} := \frac{1}{T} \sum_{t=1}^T w_{kt}^3 \quad k = 1, \dots, n$$

$$b_2 = (b_{12}, \dots, b_{n2})' \quad \text{mit} \quad b_{k2} := \frac{1}{T} \sum_{t=1}^T w_{kt}^4 \quad k = 1, \dots, n$$

$$w_t = (w_{1t}, \dots, w_{nt})' := P^{-1} \hat{\varepsilon}_t$$

und  $P$  eine untere Dreiecksmatrix mit positiven Diagonalelementen ist und  $PP' = \hat{\Sigma}$  gilt. JB ist asymptotisch  $\chi^2$  verteilt mit  $2n$  Freiheitsgraden.

### 2.2.3. Beispiel

Als Beispiel dient ein dreidimensionales System, das den „Dow Jones Industrial“ Index, den DAX und den ATX enthält. Diese drei Indizes wurden wöchentlich vom 23. Dezember 2002 bis 13. September 2004 aufgezeichnet (91 Beobachtungen). In Abbildung 2.1 sieht man die Originalzeitreihen und deren erste Differenzen. Die Differenzen weisen ein stationäres Verhalten auf, während die Indizes sich wie „random walks“ verhalten.

Die Kriterien für die Ordnungsschätzung des VAR Modells ergeben ein sehr unterschiedliches Bild (Tabelle 2.1). Während man nach dem AIC und LR-Test ein großes Modell wählen sollte, ist nach dem HQ- und SC-Kriterium ein Modell mit der Ordnung Eins am besten. Im Zweifelsfall erhält immer das Kriterium mit der kleinsten Modellgröße den Vorzug (hier HQ oder SC).<sup>2</sup> Es wird nun mit den Differenzen ein Modell der Gestalt

$$\Delta \mathbf{y}_t = \Phi_1 \Delta \mathbf{y}_{t-1} + \kappa D_t + \varepsilon_t$$

geschätzt, wobei  $D_t = (1)$ . Für die OLS Schätzung wird von den  $3 \times 90$  Beobachtungen (ein Datenvektor ist durch Differenzenbildung verloren gegangen) der erste als Startwert verwendet. Damit ist  $T = 89$ ,  $n = 3$ ,  $p = 1$  und  $k = 1$ .

Tabelle 2.1: Informationskriterien für die Ordnungsbestimmung des Modells

	AIC	HQ	SC	FPE	LR
Modellordnung	20	1	1	7	20

AIC: Akaike Information Criterion

HQ: Hannan Quinn Criterion

SC: Schwarz Criterion

FPE: Final Prediction Error

LR: Likelihood Ratio

<sup>2</sup>Ein weiterer Grund für deren Bevorzugung ist die Eigenschaft, dass das SC und HQ Kriterium konsistente Schätzer für die Ordnungsgröße sind.

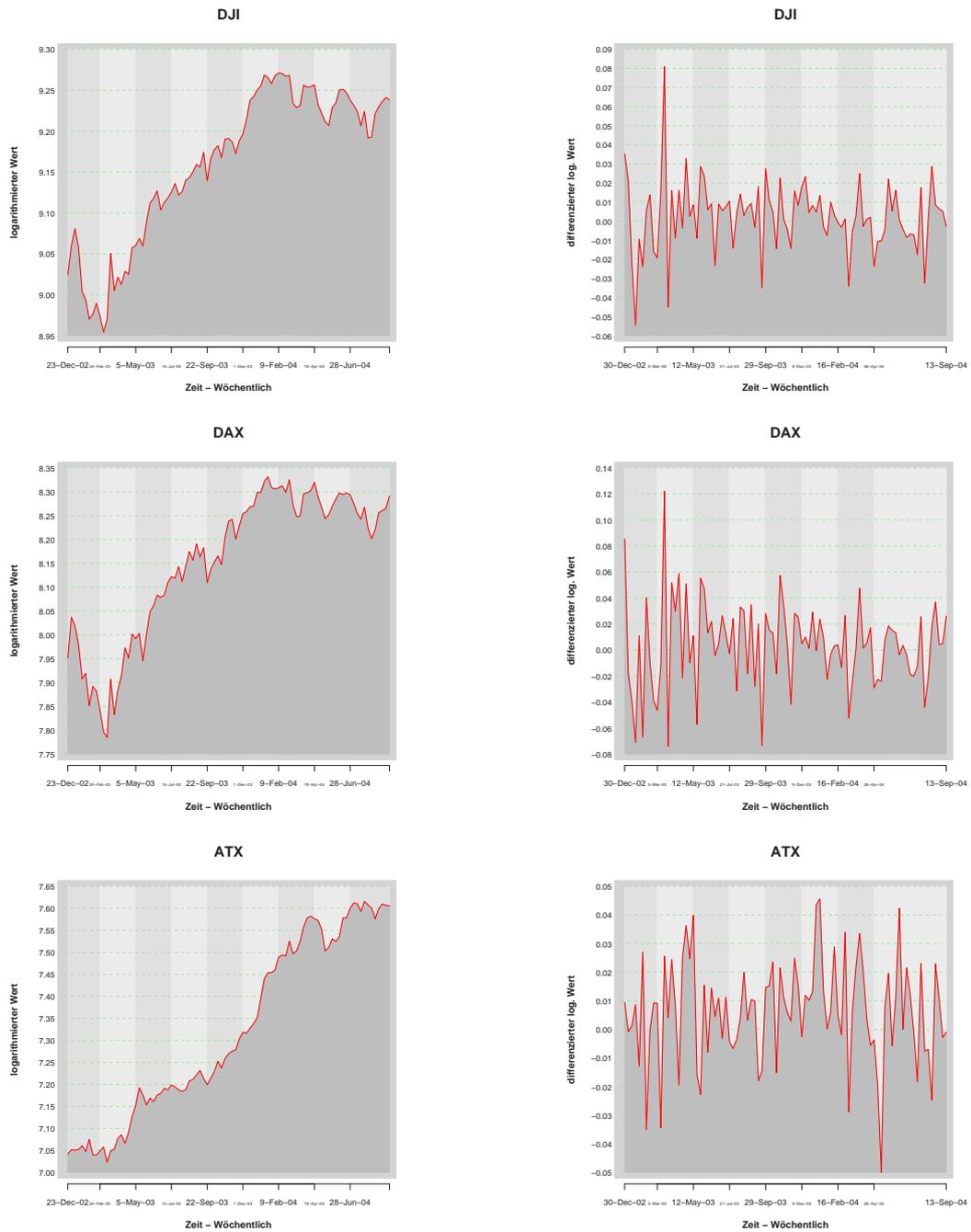


Abbildung 2.1: Plot der Zeitreihen und deren erste Differenzen

Die OLS Schätzung ergibt

$$\Phi_1 = \begin{pmatrix} 1.834 & -1.4747 & -1.72568 \\ 4.537 & -3.7236 & -1.53635 \\ 1.314 & 0.2419 & -0.02094 \end{pmatrix} \times 10^{-1}$$

$$\Sigma = \begin{pmatrix} 3.152 & 4.540 & 0.867 \\ 4.540 & 10.108 & 1.655 \\ 0.867 & 1.655 & 3.171 \end{pmatrix} \times 10^{-4}$$

und

$$\kappa = \begin{pmatrix} 3.175 \\ 4.038 \\ 5.836 \end{pmatrix}.$$

Die Residuen und deren Histogramm sind in Abbildung 2.2 dargestellt. Die schwarzen Balken im Histogramm resultieren aus einer simulierten Normalverteilung, die als Parameter das Stichprobenmittel und die Stichprobenvarianz der Residuen besitzt.

In Abbildung 2.3 sind die Auto- und „Cross“-Korrelationen der Residuen geplottet. Die Darstellung beginnt mit dem Lag Null, wo die geschätzte Varianz-Kovarianz Matrix eingeht. Damit die Fehler die Bedingungen eines „White Noise“ Prozesses erfüllen, sollten keine signifikanten Abhängigkeiten für Verzögerungen größer gleich Eins gefunden werden. Die strichlierten Linien stellen den Toleranzbereich (5%) für die Korrelationswerte dar. Nur zwei Werte überschreiten diese Intervallgrenzen geringfügig. Die visuelle Betrachtung lässt somit auf einen „white noise“ Prozess schließen. Der mehrdimensionale Portmanteautest bestätigt diesen Eindruck: Die Teststatistik, die Korrelationen bis zum Lag 15 berücksichtigt  $P(15) = 123.370$  mit  $n^2(s-p) = 9 \times (15-1) = 126$  Freiheitsgraden, ist kleiner als der dazugehörige  $\chi_{0.95}^2(126)$  Wert von 153.197.

Im nächsten Schritt werden die einzelnen Komponenten der Residuen auf bedingte Heteroskedastizität untersucht. In Tabelle 2.2 sieht man, dass es keine Abhängigkeiten gibt. Zu beachten ist jedoch, dass die Voraussetzung für den ARCH Test von Engle (Residuen sind normalverteilt) nicht gegeben und damit auch das Ergebnis nicht sehr aussagekräftig ist. Ein anschließend durchgeführter F-Test, der hier aber nicht tabelliert ist, kommt jedoch zum gleichen Ergebnis.

Tabelle 2.2: Test auf ARCH Effekte

	ARCH(2)	ARCH(5)	$\chi_{0.95}^2(2)$	$\chi_{0.95}^2(5)$
DJI	3.229	3.396	5.991	11.070
DAX	5.202	8.608	5.991	11.070
ATX	0.934	2.120	5.991	11.070

Das gefundene Modell repräsentiert die drei Indizes somit sehr gut. Durch Differen-

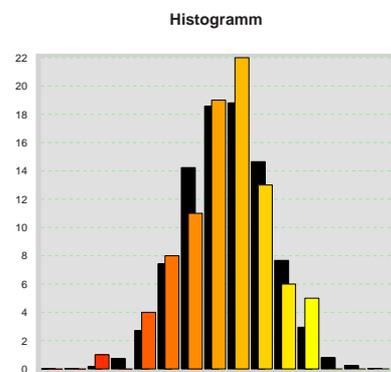
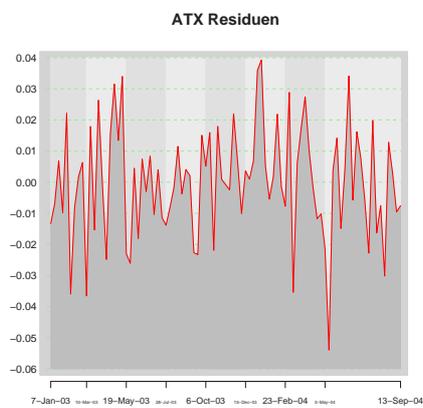
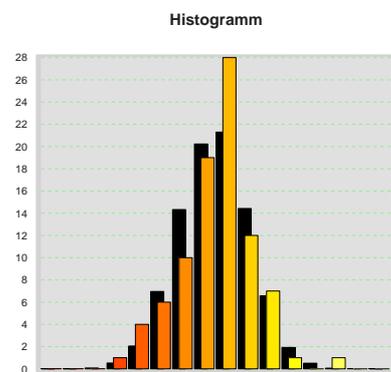
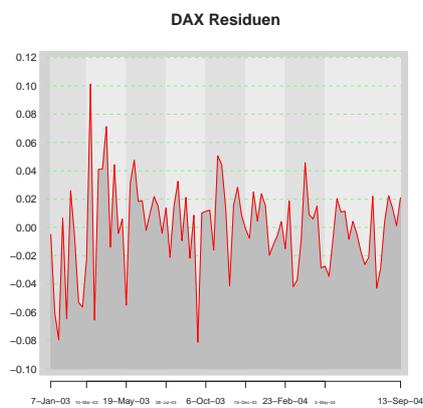
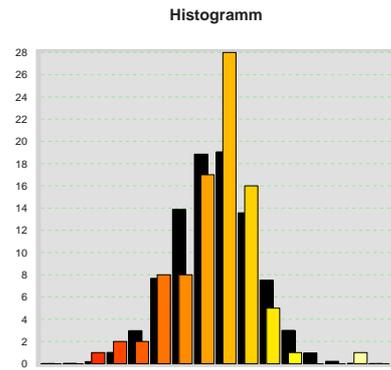
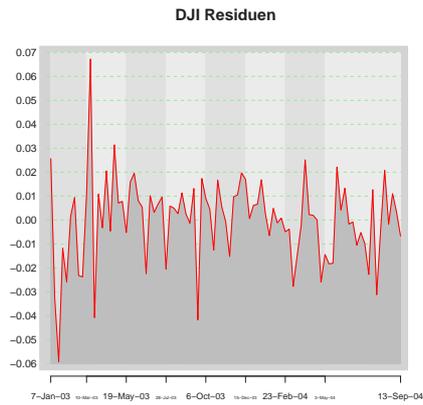


Abbildung 2.2: Plot der Residuen und deren Histogramm (die schwarzen Balken im Hintergrund entsprechen einer geschätzten Normalverteilung)

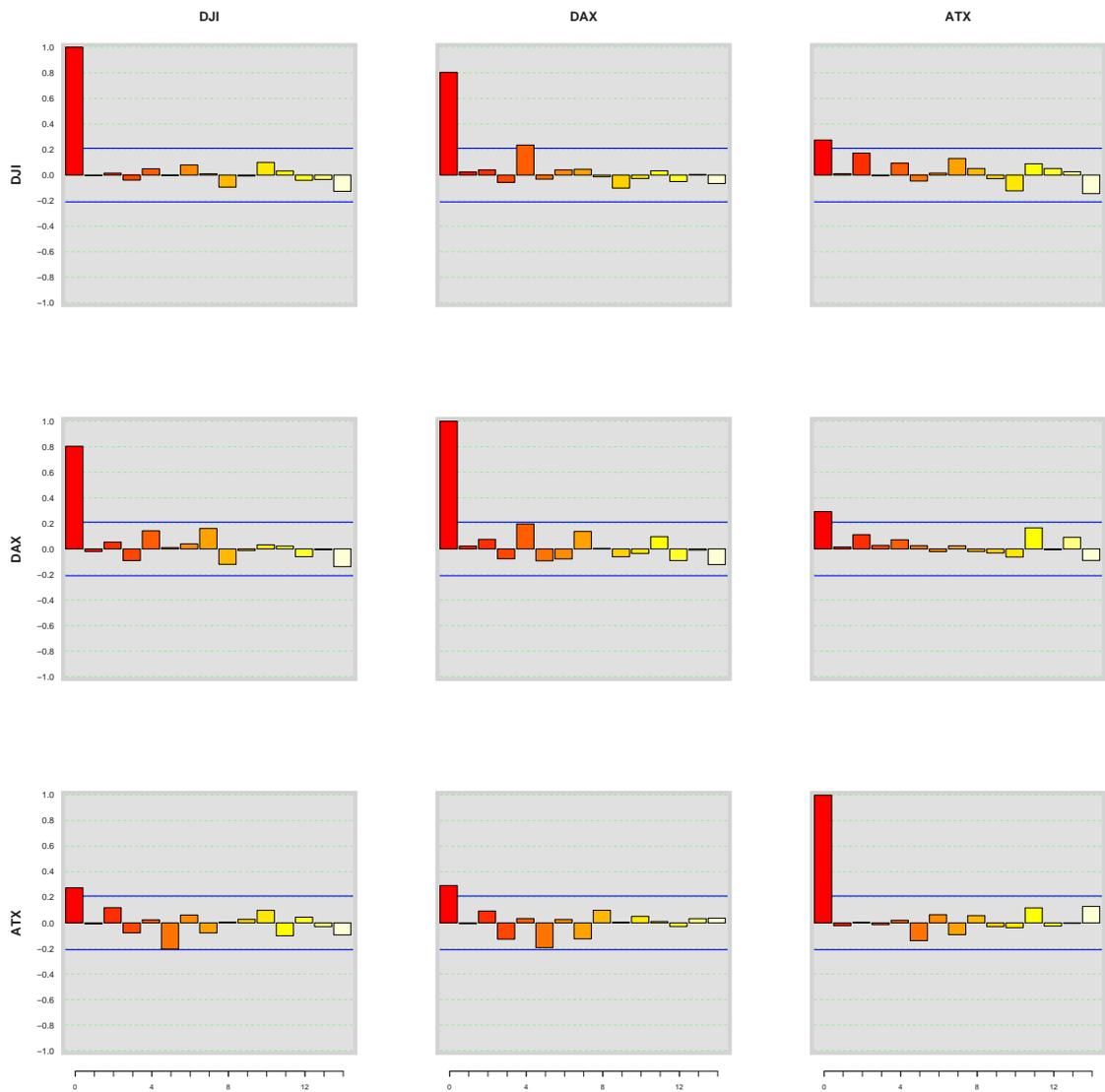


Abbildung 2.3: ACF und „Cross“ Korrelationsfunktion der Residuen

zenbildung der Daten ist der Prozess auch stationär im Sinne einer kausalen Lösung (vgl. Tabelle 2.3).

Der Prozess ist jedoch nicht normalverteilt, da die Residuen sich nicht wie normalverteilte Zufallsvariablen verhalten. Diese Tatsache spiegelt sich sowohl im Histogramm (Abb. 2.2), als auch im Jarque-Bera Test ( $JB = 5443$ ,  $\chi_{0.95}^2(10) = 12.591$ ) wieder.

Tabelle 2.3: Eigenwerte von  $\Phi_1$ 

Wurzeln	Re-Teil	Im-Teil	Betrag
$w_1$	-0.216	0.000	0.216
$w_2$	0.012	0.175	0.175
$w_3$	0.012	-0.175	0.175

## 2.3. Nichtstationäre Prozesse

Zur Erinnerung: ein stationärer VAR( $p$ ) Prozess

$$\mathbf{y}_t = \Phi_1 \mathbf{y}_{t-1} + \dots + \Phi_p \mathbf{y}_{t-p} + \kappa D_t + \varepsilon_t$$

besitzt eine kausale Lösung, wenn die Nullstellen von

$$\det(I_n - \Phi_1 \mathbf{z} - \dots - \Phi_p \mathbf{z}^p)$$

alle außerhalb des Einheitskreises liegen. Für einen eindimensionalen AR(1) Prozess  $y_t = \nu + \phi y_{t-1} + \xi_t$  lautet diese Bedingung

$$1 - \phi z \neq 0 \quad \forall |z| \leq 1 \quad (2.16)$$

oder equivalent dazu  $|\phi| < 1$  (vgl. Anhang A).

Was passiert nun, wenn  $\phi = 1$  ist? Für die Antwort betrachtet man zwei Fällen:  $\nu = 0$  und  $\nu \neq 0$ . Im Fall  $\nu = 0$ , erhält man einen sogenannten „random walk“ Prozess

$$y_t = y_{t-1} + \xi_t. \quad (2.17)$$

Setzt man  $y_0 = 0$  und lässt den Prozess bei  $t = 0$  starten, so ergibt sich folgende Darstellung:

$$y_t = y_{t-1} + \xi_t = y_{t-2} + \xi_{t-1} + \xi_t = \dots = \sum_{j=0}^t \xi_j.$$

Ein „random walk“ ist somit die Summe von vergangenen Innovationen.

Abbildung 2.4 zeigt vier „random walks“, die durch Simulation künstlich erzeugt wurden und jeweils 1000 Datenpunkte umfassen. Der Startwert  $y_0$  wurde dabei auf Null gesetzt. Die Residuen sind normalverteilt mit Mittel Null und Varianz Eins.

Ist  $\xi_t$  ein „white noise“ Prozess mit Mittel Null und Varianz  $\sigma^2$ , dann berechnet sich der Erwartungswert von  $y_t$  durch

$$\mathbf{E}[y_t] = \sum_{s=1}^t \mathbf{E}[\xi_s] = 0$$

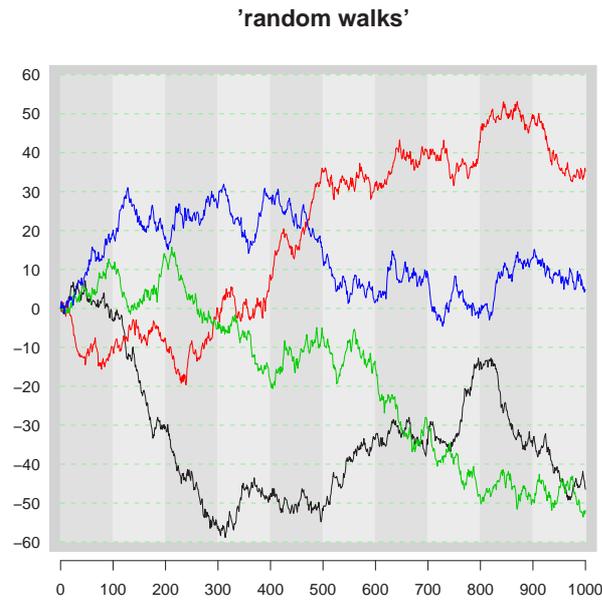


Abbildung 2.4: Simulationen von „random walk“ Prozessen mit normalverteilten Innovationen ( $\xi_t \sim NID(0, 1)$ )

und die Varianz beträgt

$$\mathbf{Var}[y_t] = \sum_{s=1}^t \mathbf{Var}[\xi_s] = t\sigma^2.$$

Die Varianz eines „random walks“ ist also nicht konstant und auch nicht unabhängig vom Zeitindex  $t$  und wächst somit proportional mit der Zeit gegen Unendlich. Durch diese Eigenschaft ist der Prozess  $y_t$  nicht stationär. Die ACF

$$\mathbf{Corr}[y_t, y_{t+h}] = \frac{\mathit{Cov}[y_t, y_{t+h}]}{\sqrt{\mathit{Var}[y_t]}\sqrt{\mathit{Var}[y_{t+h}]}} = \frac{t\sigma^2}{\sqrt{t\sigma^2}\sqrt{(t+h)\sigma^2}} = \frac{t}{\sqrt{t^2 + th}}$$

konvergiert für  $t \rightarrow \infty$  für alle  $h$  gegen 1. Das bedeutet, dass  $y_t$  und  $y_s$  selbst für  $s \gg t$  noch sehr stark korreliert sind.

Im zweiten Fall  $\nu \neq 0$ , wird der Prozess  $y_t = c + y_{t-1} + \xi$  „random walk“ mit Drift genannt. Beginnt der Prozess wieder bei  $t = 0$  und setzt man  $y_0 = 0$ , so enthält diesmal  $y_t$  einen deterministischen Trend:

$$y_t = \nu t + \sum_{j=0}^t \xi_j.$$

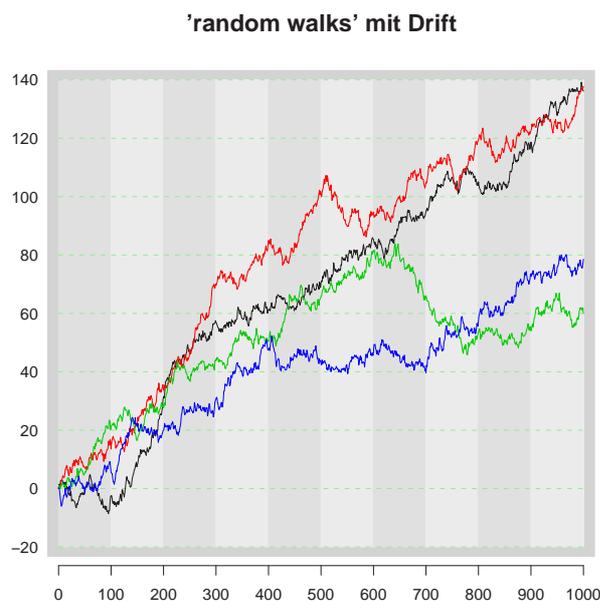


Abbildung 2.5: Simulationen von „random walk“ Prozessen mit Drift ( $c = 0.1$ ) und normalverteilten Innovationen ( $\xi_t \sim NID(0, 1)$ )

Der Erwartungswert beträgt  $\mathbf{E}[y_t] = \nu = t$  und die Varianz ist wieder  $\mathbf{Var}[y_t] = t\sigma^2$ . Abbildung 2.5 zeigt vier „random walks“ mit Drift mit standardnormalverteilten Innovationen.

Ein autoregressiver Prozess höherer Ordnung

$$y_t = \nu = +\phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \xi_t$$

zeigt ein ähnliches Verhalten wie ein „random walk“, wenn eine Wurzel von  $(1 + \phi_1 z + \dots + \phi_p z^p)$  auf dem Einheitskreis („unit root“) und alle anderen außerhalb liegen. Befindet sich eine Wurzel innerhalb, so wird der Prozess explosiv,<sup>3</sup> das bedeutet, dass die Varianz mit exponentieller Geschwindigkeit gegen Unendlich konvergiert für  $t \rightarrow \infty$ .

Nicht stationäre Prozesse, die nur eine „unit root“ enthalten, können jedoch sehr einfach zu stationären Prozessen transformiert werden. Durch Differenzbildung wird der stochastische Trend beseitigt und durch die Annahme  $\xi_t \sim WN(0, \sigma^2)$  ist der Prozess

$$\Delta y_t := (1 - B)y_t = y_t - y_{t-1} = \xi_t$$

<sup>3</sup>Betrachtet wird hier nur ein Prozess mit kausaler Lösung.

stationär. Diese Eigenschaft motiviert zu folgender Definition, die auch im mehrdimensionalen Fall durch Transformation des VAR( $p$ ) Prozesses und komponentenweise Betrachtung (siehe Lütkepohl (1991)) Sinn macht:

**Definition 2.6 (Integrierter Prozess)** *Ein Prozess  $\mathbf{y}_t$  ist integriert der Ordnung  $d$ , wenn  $\Delta^k \mathbf{y}_t$  nicht stationär ist für alle  $0 \leq k < d$  und  $\Delta^d \mathbf{y}_t$  stationär ist.  $\Delta$  ist der „Differenzen“ Operator:  $\Delta = (1 - B)$ , d.h.  $\Delta \mathbf{y}_t = \mathbf{y}_t - \mathbf{y}_{t-1}$ . Erfüllt ein Prozess diese Bedingung, so wird die Notation  $\mathbf{y}_t \sim I(d)$  verwendet.*

# Kapitel 3.

## Kointegration

### 3.1. „Spurious“ Regression

Abhängigkeiten zwischen zwei Beobachtungsreihen  $y_{1,t}$  und  $y_{2,t}$  werden meistens durch ein Regressionsmodell getestet. Die Gültigkeit und Richtigkeit der Ergebnisse und der erhaltenen Schätzer verlangt jedoch grundlegende Voraussetzungen. Die wichtigste Annahme ist die Stationarität der Daten. Was passiert jedoch wenn die Daten nicht stationär sind ?

Diese Frage stellten sich auch Granger and Newbold (1974) und untersuchten dazu folgendes Regressionsmodell:

$$y_{1,t} = \nu + \eta y_{2,t} + \xi_t \quad t = 1, 2, \dots, T. \quad (3.1)$$

$y_{1,t}$  und  $y_{2,t}$  sind zwei unkorrelierte „random walks“:

$$y_{1,t} = y_{1,t-1} + \xi_{1,t}, \quad \xi_{1,t} \sim IID(0, \sigma_{\xi_1}^2) \quad (3.2)$$

$$y_{2,t} = y_{2,t-1} + \xi_{2,t}, \quad \xi_{2,t} \sim IID(0, \sigma_{\xi_2}^2) \quad (3.3)$$

wobei

$$\mathbf{E}[\xi_{1,t}\xi_{2,s}] = 0 \quad \forall t, s \quad \text{und} \quad \mathbf{E}[\xi_{1,t}\xi_{1,t-h}] = \mathbf{E}[\xi_{2,t}\xi_{2,t-h}] = 0 \quad \forall h \neq 0, \quad \forall t$$

gilt. Da  $y_{1,t}$  keinen Einfluss auf  $y_{2,t}$  hat und umgekehrt, könnte man davon ausgehen, dass der Koeffizient  $\eta$  in der Gleichung (3.1) in Wahrscheinlichkeit gegen Null konvergiert. Doch das Gegenteil ist der Fall. Granger and Newbold (1974) testeten obiges Regressionsmodell mit den Annahmen  $y_0 = x_0 = 100$ ,  $\xi_{1,t}, \xi_{2,t} \sim NID(0, 1)$  und  $T = 50$  und stellten fest, dass die Nullhypothese

$$\hat{\eta} = 0$$

in 76 von 100 Fällen abgelehnt wird. Die beiden Autoren verwendeten dabei den gewöhnlichen  $t$ -Test um die Signifikanz von  $\hat{\eta}$  festzustellen.

In einer weiteren Untersuchung wurden vier unabhängige „random walks“ als Regressoren in die Gleichung (3.1) aufgenommen. Ein gewöhnlicher F-Test lehnte die Nullhypothese, die besagt, dass der Koeffizientenvektor gegen Null konvergiert, mit

einer Wahrscheinlichkeit von sogar 96 Prozent ab. Dieses Phänomen wird von Granger and Newbold (1974) als „Spurious“ Regression bezeichnet.

Weitere empirische Untersuchungen mit unabhängigen, integrierten Prozessen als Regressoren wurden von Granger und Newbold (1977) und Plosser und Schwert (1978) durchgeführt. Die empirischen Schlüsse aus diesen und allen vorangegangenen Arbeiten bekamen erst 8 Jahre später von Phillips (1986) ein theoretisches Fundament, auf das hier jedoch nicht näher eingegangen wird.

Um das Problem von Scheinzusammenhängen bei der Regressionsanalyse mit nichtstationären Zeitreihen zu vermeiden, wurde in den 80er Jahren meist Differenzbildung zur Trendbereinigung von Daten vor einer Analyse des Zusammenhangs der untersuchten Variablen empfohlen. Diese Vorgehensweise hat aber einen entscheidenden Nachteil: Durch die Differenzbildung gehen die Informationen über die Trendkomponenten verloren. Beispielsweise sind die Paare Dividende und Preis pro Aktie, kurz- und langfristiger Zinssatz oder Konsum und Bruttosozialprodukt durch einen gemeinsamen Trend gekennzeichnet. Dementsprechend können sie sich langfristig nicht beliebig weit voneinander entfernen. Mit anderen Worten, es besteht langfristig eine ökonomische Gleichgewichtsbeziehung. Die Zeitreihen können individuell zwar als nichtstationär betrachtet werden, Linearkombinationen wie die logarithmierte Dividendenrendite, die Zinsdifferenz oder die logarithmierte Konsumquote erscheinen aber stationär (vgl. Drobetz (2003)). Dementsprechend verliert man bei Differenzbildung wertvolle Information. Diese wichtige Beobachtung hat zur Entwicklung des Konzepts der Kointegration geführt, das erstmal von Granger (1981) vorgestellt und von Engle and Granger (1987) u.a. weiterentwickelt wurde. Im nächsten Kapitel wird der Begriff Kointegration definiert und Folgerungen und Zusammenhänge vorgestellt und analysiert.

### 3.2. Definition

Betrachtet man zwei Prozesse  $y_{1,t}$  und  $y_{2,t}$  die integriert mit der Ordnung  $d$  sind, dann liefert die Linearkombination

$$z_t = y_{1,t} - \eta y_{2,t}$$

im Allgemeinen einen  $I(d)$  Prozess. Es ist aber auch möglich, dass  $z_t \sim I(d-b)$ ,  $b > 0$  ist. In diesem Fall hat man eine Linearkombination gefunden, die die Ordnung von  $z_t$  reduziert. Dieser Sachverhalt wird Kointegration genannt und führt zu der von Engle and Granger (1987) vorgestellten Definition:

**Definition 3.1 (Kointegration (Originalversion))** *Die Komponenten eines  $n$ -dimensionalen Vektors  $\mathbf{y}_t$  sind mit der Ordnung  $d$ ,  $b$  kointegriert, i. Z.  $\mathbf{y}_t \sim CI(d, b)$ , wenn*

1. alle Komponenten von  $\mathbf{y}_t$  mit der Ordnung  $d$  integriert sind ( $\mathbf{y}_t \sim I(d)$ ) und

2. ein Vektor  $\beta_1 \in \mathbb{R}^n$ ,  $\beta_1 \neq 0$  existiert, mit  $\varepsilon_t = \beta_1' \mathbf{y}_t \sim I(d - b)$ .

Der Vektor  $\beta_1$  wird Kointegrationsvektor genannt.

In dieser Arbeit wird aber eine leicht abgewandelte Definition von Kointegration verwendet.

**Definition 3.2 (Kointegration (Version 2))** Ist ein  $n$ -dimensionalen Vektors  $\mathbf{y}_t$  integriert mit der Ordnung  $d$ , i. Z.  $\mathbf{y}_t \sim I(d)$  und existiert ein Vektor  $\beta_1 \in \mathbb{R}^n$ , so dass  $\beta_1' \mathbf{y}_t$  integriert ist mit einer Ordnung kleiner  $d$ , dann wird der Prozess  $\mathbf{y}_t$  kointegriert genannt.

Diese Version unterscheidet sich von der Originalversion in der Hinsicht, dass nicht alle Komponenten von  $\mathbf{y}_t$  integriert mit Ordnung  $d$  sein müssen, sondern auch eine kleinere Ordnung besitzen können. Ist nur eine Komponente  $I(d)$  und alle anderen  $I(0)$ , dann ist nach Definition 2.6  $\mathbf{y}_t$  ein  $I(d)$  Prozess. Die Relation  $\beta_1' \mathbf{y}_t$ , die nur die stationären Komponenten enthält, liefert nach Version 2 der Kointegrationsdefinition eine gültige Kointegrationsbeziehung. Solche Fälle entsprechen eigentlich nicht der Idee von Engle and Granger (1987). Trotzdem wird hier nur die Version 2 zum Einsatz kommen, da die Bezeichnung einfacher ist, weil nicht mehr zwischen den Ordnungen der Komponenten des Vektors unterschieden werden muss. Der Leser sollte aber die Originalidee von Kointegration im Hinterkopf behalten, vor allem, wenn es zur Interpretation von Kointegrationsbeziehungen kommt.

Hat man einen Kointegrationsvektor gefunden, erhält man durch Multiplizieren mit einer von Null verschiedenen Konstanten sofort einen zweiten Vektor. Gibt es  $r$  verschiedene Kointegrationsbeziehungen, im Sinne von linear unabhängig, so ist auch jede Linearkombination der Vektoren eine weitere Kointegrationsrelation. Die  $r$  Vektoren werden zu einer Matrix zusammengefasst, die durch die lineare Unabhängigkeit der Vektoren den Rang  $r$  hat. Der Kointegrationsrang von  $\mathbf{y}_t$  ist die maximale Zahl linear unabhängiger Vektoren.

Abbildung 3.1 und 3.2 zeigen zwei kointegrierte Prozesse, einen ohne Drift und einen anderen mit Drift. Der Kointegrationsvektor ist in im ersten Fall  $\beta_1 = [1, -0.7]'$  und im zweiten Fall  $\beta_1 = [1, 0.7]'$ . Man sieht in Abbildungen 3.1 sehr gut, wie die beiden Prozesse „einen gemeinsamen Weg gehen“. Ähnliche Bilder, und das ist hier das Interessante, können unter anderem auch auf den Aktienmärkten beobachtet werden. Untersuchungen in diese Richtung werden in Kapitel 4 vorgenommen.

Für die Analyse von Kointegrationsbeziehungen ist die Darstellung eines Prozesses als Fehlerkorrekturmodell von Vorteil. Im Folgenden wird nun diese Darstellung von der VAR Darstellung eines Prozesses abgeleitet.

Besitzt ein Prozess  $y_t$  eine VAR Darstellung

$$\begin{aligned} \mathbf{y}_t &= \Phi_1 \mathbf{y}_{t-1} + \dots + \Phi_p \mathbf{y}_{t-p} + \kappa D_t + \varepsilon_t \\ \Phi(B) \mathbf{y}_t &= \kappa D_t + \varepsilon_t \end{aligned}$$

### Kointegrierter Prozess

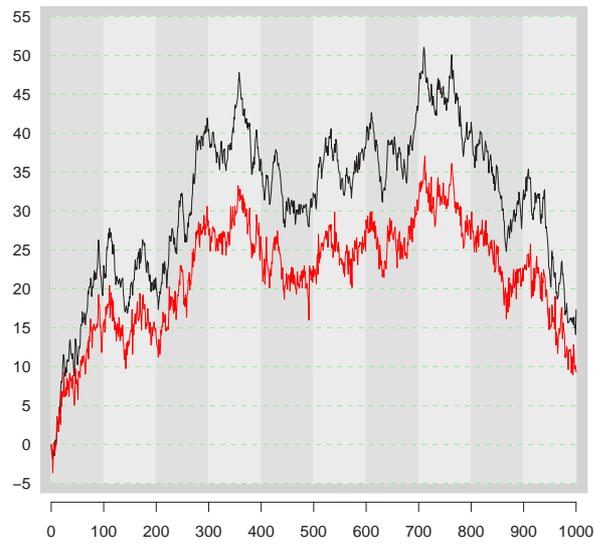


Abbildung 3.1: Simulation von zwei kointegrierten Prozessen

$$y_{1,t} = y_{1,t-1} + \varepsilon_{1,t}$$

$$y_{2,t} = 0.7y_{1,t} + \varepsilon_{2,t}$$

### Kointegrierter Prozess mit Drift

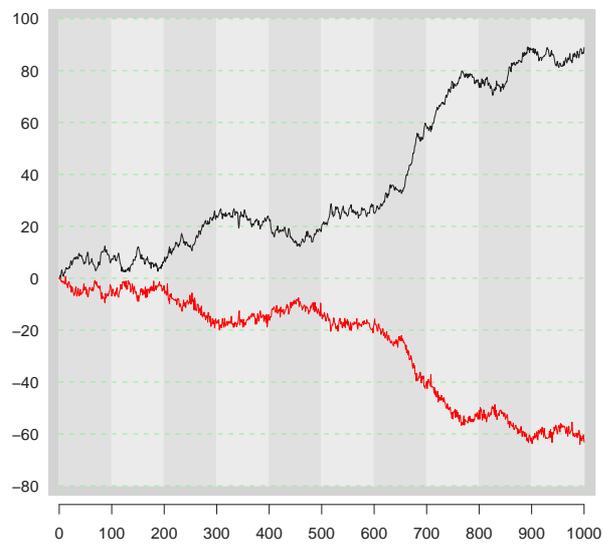


Abbildung 3.2: Simulation von zwei kointegrierten Prozessen mit Drift

$$y_{1,t} = 0.1 + y_{1,t-1} + \varepsilon_{1,t}$$

$$y_{2,t} = -0.7y_{1,t} + \varepsilon_{2,t}$$

wobei  $\Phi(\mathbf{z}) = (I_n - \Phi_1\mathbf{z} - \dots - \Phi_p\mathbf{z}^p)$  und  $\varepsilon \sim NID(0, \Sigma)$ ,  $\det \Sigma > 0$  gilt, dann kann mit Hilfe der Beziehungen

$$\begin{aligned}(\Phi(\mathbf{z}) - \Phi(1)\mathbf{z})|_{\mathbf{z}=1} &= 0 \\(\Phi(\mathbf{z}) - \Phi(1)\mathbf{z})|_{\mathbf{z}=0} &= I_n\end{aligned}$$

$\Phi(\mathbf{z})$  geschrieben werden als

$$\begin{aligned}\Phi(z) &= \Phi(1)\mathbf{z} + (\Phi(\mathbf{z}) - \Phi(1)\mathbf{z}) \\&= -\Pi\mathbf{z} + \Gamma(\mathbf{z})(1 - \mathbf{z}) \\&= -\Pi\mathbf{z} + I(1 - \mathbf{z}) - \Gamma_1\mathbf{z}(1 - \mathbf{z}) - \dots - \Gamma_{p-1}\mathbf{z}^{p-1}(1 - \mathbf{z})\end{aligned}$$

wobei  $\Gamma(\mathbf{z}) = I_n - \Gamma_1\mathbf{z} - \dots - \Gamma_{p-1}\mathbf{z}^{p-1}$  und  $\Pi = -\Phi(1)$ . Die Koeffizienten  $\Gamma_i$  erhält man durch Koeffizientenvergleich:

$$\begin{aligned}\Phi(z) &= I_n - \Phi_1\mathbf{z} - \dots - \Phi_p\mathbf{z}^p \\&= -\Pi\mathbf{z} + I - \Gamma_1\mathbf{z} - \dots - \Gamma_{p-1}\mathbf{z}^{p-1} - I_n\mathbf{z} + \Gamma_1\mathbf{z}^2 + \dots + \Gamma_{p-1}\mathbf{z}^p \\&= I_n - (I_n + \Pi + \Gamma_1)\mathbf{z} - (\Gamma_2 - \Gamma_1)\mathbf{z}^2 - \dots - (\Gamma_{p-1} - \Gamma_{p-2})\mathbf{z}^{p-1} + \Gamma_{p-1}\mathbf{z}^p \\&\quad \downarrow \\ \Gamma_{p-1} &= -\Phi_p \\ \Gamma_{p-2} &= -(\Phi_{p-1} + \Phi_p) \\ &\vdots \\ \Gamma_1 &= -(\Phi_2 + \dots + \Phi_p) \\ \Pi &= \Phi_1 + \dots + \Phi_p - I_n = -\Phi(1).\end{aligned}$$

Damit kann das VAR Modell

$$\Phi(B)\mathbf{y}_t = (-\Pi B + I_n(1 - B) - \Gamma_1 B(1 - B) - \dots - \Gamma_{p-1}B^{p-1}(1 - B))\mathbf{y}_t = \kappa D_t + \varepsilon_t$$

als Fehlerkorrekturmodell geschrieben werden:

**Definition 3.3 (Fehlerkorrekturmodell)** *Das Fehlerkorrekturmodell (ECM = Error Correction Model) eines Vektorprozesses  $\mathbf{y}_t$  ist die Darstellung:*

$$\Delta\mathbf{y}_t = \Pi\mathbf{y}_{t-1} + \Gamma_1\Delta\mathbf{y}_{t-1} + \dots + \Gamma_{p-1}\Delta\mathbf{y}_{t-p+1} + \kappa D_t + \varepsilon_t \quad (3.4)$$

wobei  $(\varepsilon_t) \sim WN(0, \Sigma)$  ein multivarianter „white noise“ Prozess mit Varianz-Kovarianz Matrix  $\Sigma$  ist und  $\Pi$  und  $\Gamma_i$   $i = 1, \dots, p - 1$   $n \times n$ -Matrizen sind;  $\Delta\mathbf{y}_t = \mathbf{y}_t - \mathbf{y}_{t-1}$ .

Setzt man die Koeffizienten  $\Gamma_i$  und  $\kappa$  gleich Null, dann erhält man ein einfaches ECM der Form

$$\Delta\mathbf{y}_t = \Pi\mathbf{y}_{t-1} + \varepsilon_t$$

mit der Interpretation: Die Veränderung des Prozesses  $\mathbf{y}_t$  vom Zeitpunkt  $t - 1$  zum Zeitpunkt  $t$  wird durch den Term  $\Pi\mathbf{y}_{t-1}$  korrigiert (Fehlerkorrekturterm).

Für die empirischen Untersuchungen sind vor allem die Fälle  $I(0)$  und  $I(1)$  von Bedeutung, weshalb von nun an, wenn nicht anders angegeben, nur mehr integrierte Prozesse mit der Ordnung Eins betrachtet werden.

Der Begriff Kointegrationsrang wurde vorher schon für den allgemeinen Fall eingeführt. Jetzt wird dieser im Licht des Fehlerkorrekturmodells nochmals analysiert. Ist der Prozess  $\mathbf{y}_t$  integriert mit der Ordnung Eins und die Matrix  $\Pi$ , die in zwei Matrizen  $\alpha$  und  $\beta$  mit der Dimension  $n \times r$  aufgespalten wird

$$\Pi = \alpha\beta', \quad (3.5)$$

besitzt den Rang  $r$

$$\text{rang } \Pi = r,$$

dann folgt aus

$$\beta'\mathbf{y}_t = \bar{\alpha}'\Pi\mathbf{y}_t = \bar{\alpha}'\Delta\mathbf{y}_{t+1} - \bar{\alpha}'\Gamma_1\Delta\mathbf{y}_t - \dots - \bar{\alpha}'\Gamma_{p-1}\Delta\mathbf{y}_{t+1-p} - \bar{\alpha}'\varepsilon_t$$

mit  $\bar{\alpha}' = I_r$ , dass die Spalten von  $\beta$  Kointegrationsvektoren sind, da der rechte Teil nur stationäre Komponenten enthält ( $\mathbf{y}_t \sim I(1)$ ). Die Matrix  $\beta$  ist somit die Kointegrationsmatrix und  $\alpha$  wird oft „loading“ Matrix genannt.

### 3.3. “Granger Representation Theorem“

Im Kapitel „Vektor Autoregressive Prozesse“ wurde unter gewissen Bedingungen eine Vektor „Moving Average“ Darstellung eines VAR Prozesses gefunden (vgl. Satz 2.1).

Für einen kointegrierten Prozess  $\mathbf{y}_t$  zeigt das „Granger Representations Theorem“, dass man unter gewissen Bedingungen ebenfalls eine Vektor „Moving Average“ Darstellung von  $\mathbf{y}_t$  finden kann.

Für die Formulierung des Satzes benötigt man aber vorher noch einige Definitionen: Für die Matrix  $\alpha \in \mathbb{R}^{n \times r}$  mit vollem Rang  $r$  ist die „links-inverse“ Matrix  $\bar{\alpha} \in \mathbb{R}^{n \times r}$  mit  $\bar{\alpha}'\alpha = I_r$  durch  $\bar{\alpha} = \alpha(\alpha'\alpha)^{-1}$  gegeben. Die Matrix  $\alpha_\perp \in \mathbb{R}^{n \times n-r}$  mit vollem Rang  $n - r$  erfüllt die Bedingung  $\alpha'_\perp\alpha = 0$  und die Matrix  $\bar{\alpha}_\perp \in \mathbb{R}^{n \times n-r}$  ist durch  $\bar{\alpha}_\perp = \alpha_\perp(\alpha'_\perp\alpha_\perp)^{-1}$  gegeben. Analoge Definitionen gelten für die Matrizen  $\bar{\beta}$ ,  $\beta_\perp$  und  $\bar{\beta}_\perp$ .

Die Polynommatrix eines VAR Prozesses

$$\Phi(\mathbf{z}) = (I_n - \Phi_1\mathbf{z} - \dots - \Phi_p\mathbf{z}^p)$$

besitzt an der Stelle  $\mathbf{z} = 1$  den Wert  $\Pi$  ( $\Phi(1) = \Pi$ ) und die erste Ableitung beträgt an der Stelle  $\mathbf{z} = 1$

$$\begin{aligned} \dot{\Phi}(\mathbf{1}) &= \frac{d}{d\mathbf{z}}\Phi(\mathbf{z})|_{\mathbf{z}=1} = -\Phi_1 - 2\Phi_2\mathbf{z} - \dots - p\Phi_p\mathbf{z}^{p-1}|_{\mathbf{z}=1} \\ &= -\Pi - I_n + \Gamma_1 + \dots + \Gamma_{p-1} \\ &=: -\Pi - \Gamma. \end{aligned}$$

Entwickelt man die Polynommatrix um  $\mathbf{z} = 1$ , siehe Johansen (1995, Lemma 4.1), dann lässt sich  $\Phi(\mathbf{z})$  schreiben als

$$\begin{aligned}\Phi(\mathbf{z}) &= \Phi(1) + \dot{\Phi}(1)(1 - \mathbf{z}) + \Phi^*(\mathbf{z})(1 - \mathbf{z})^2 \\ &= -\Pi + (\Gamma + \Pi)(1 - \mathbf{z}) + \Phi^*(\mathbf{z})(1 - \mathbf{z})^2\end{aligned}\quad (3.6)$$

und der VAR Prozess transformiert sich zu

$$\Phi(B)\mathbf{y}_t = -\Pi\mathbf{y}_t + (\Gamma + \Pi)\Delta\mathbf{y}_t + \Phi^*(B)\Delta^2\mathbf{y}_t = \varepsilon_t + \kappa D_t.$$

Mit diesen Definitionen, Transformationen und Notationen lässt sich nun folgender Satz formulieren und beweisen:

**Satz 3.1** (“Granger Representation Theorem“) *Gelten die folgenden drei Annahmen*

1. *Die Wurzeln von  $\Phi(\mathbf{z})$  sind entweder gleich Eins oder echt größer als Eins, d.h. aus  $\det \Phi(\mathbf{z}) = 0$  folgt  $|\mathbf{z}| > 1$  oder  $\mathbf{z} = 1$*
2. *Der Rang von  $\Pi = \alpha\beta'$  ist  $r < n$*
3. *Die  $(n - r) \times (n - r)$  Matrix  $\bar{\alpha}'_{\perp}\Gamma\bar{\beta}_{\perp}$  hat vollen Rang, d.h.  $\det(\bar{\alpha}'_{\perp}\Gamma\bar{\beta}_{\perp}) \neq 0$*

dann besitzt  $\mathbf{y}_t$  die Darstellung

$$\mathbf{y}_t = C \sum_{i=1}^t (\varepsilon_i + \kappa D_i) + C(B)(\varepsilon_t + \kappa D_t) + P \quad (3.7)$$

wobei

$$C = \bar{\beta}_{\perp}(\bar{\alpha}'_{\perp}\Gamma\bar{\beta})^{-1}\bar{\alpha}'_{\perp}$$

und  $C(\mathbf{z}) = \Phi(\mathbf{z})^{-1} - C(1 - \mathbf{z})^{-1}$  hat eine konvergente Potenzreihenentwicklung für  $|\mathbf{z}| < 1 + c$ ,  $c > 0$ .

Die erste und die dritte Annahme garantieren, dass der Prozess  $\mathbf{y}_t$  höchstens integriert mit der Ordnung Eins ist. Die zweite Annahme schließt stationäre Prozesse aus und impliziert somit, dass der Prozess  $I(1)$  ist und dass es daher  $r$  unabhängige Kointegrationsvektoren gibt.

**Beweis 3.1** *Multipliziert man Gleichung (3.3) von links mit  $(\bar{\alpha}, \bar{\alpha}_{\perp})'$  und fügt man die Identität  $I_n = \bar{\beta}\beta' + \bar{\beta}_{\perp}\beta'_{\perp}$  ein, so erhält man einen transformierten VAR Prozess*

$$\begin{pmatrix} \bar{\alpha}' \\ \bar{\alpha}'_{\perp} \end{pmatrix} \Phi(B)(\bar{\beta}, \bar{\beta}_{\perp}\Delta^{-1}) \begin{pmatrix} \beta' \\ \beta'_{\perp}\Delta \end{pmatrix} \mathbf{y}_t = \begin{pmatrix} \bar{\alpha}' \\ \bar{\alpha}'_{\perp} \end{pmatrix} (\varepsilon_t + \kappa D_t) \quad (3.8)$$

mit der kompakten Gestalt

$$\tilde{\Phi}(B)\tilde{\mathbf{y}}_t = \tilde{\varepsilon}_t$$

wobei

$$\tilde{\mathbf{y}}_t = \begin{pmatrix} \beta' \mathbf{y}_t \\ \beta'_\perp \Delta \mathbf{y}_t \end{pmatrix} = \begin{pmatrix} Z_t \\ U_t \end{pmatrix}, \quad \tilde{\varepsilon}_t = \begin{pmatrix} \bar{\alpha}' \\ \bar{\alpha}'_\perp \end{pmatrix} (\varepsilon_t + \kappa D_t)$$

und  $\tilde{\Phi}(\mathbf{z})$  mit Hilfe der Beziehung

$$\begin{pmatrix} \bar{\alpha}' \\ \bar{\alpha}'_\perp \end{pmatrix} \alpha \beta' (\bar{\beta}, \bar{\beta}_\perp (1 - \mathbf{z})^{-1}) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

und der Darstellung 3.6 folgende Gestalt besitzt

$$\begin{aligned} \tilde{\Phi}(\mathbf{z}) &= \begin{pmatrix} \bar{\alpha}' \\ \bar{\alpha}'_\perp \end{pmatrix} \Phi(z) (\bar{\beta}, \bar{\beta}_\perp (1 - \mathbf{z})^{-1}) \\ &= \begin{pmatrix} -I_r & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} (I_r + \bar{\alpha}' \Gamma \bar{\beta})(1 - \mathbf{z}) & \bar{\alpha}'_\perp \Gamma \bar{\beta} \\ (\bar{\alpha}'_\perp \Gamma \bar{\beta})(1 - \mathbf{z}) & \bar{\alpha}'_\perp \Gamma \bar{\beta} \end{pmatrix} + \\ &\quad + \begin{pmatrix} \bar{\alpha}' \\ \bar{\alpha}'_\perp \end{pmatrix} \Phi^*(B) (\bar{\beta}(1 - \mathbf{z}), \bar{\beta}_\perp(1 - \mathbf{z})). \end{aligned}$$

Aufgrund dieser Darstellung sieht man, dass  $\tilde{\Phi}(\mathbf{z})$  dieselben Wurzeln besitzt wie  $\Phi(\mathbf{z})$ , außer an der Stelle  $\mathbf{z} = 1$ , i.Z.

$$\det \Phi(\mathbf{z}) = 0 \iff \det \tilde{\Phi}(\mathbf{z}) = 0 \quad \text{für } z \neq 1.$$

Weiters besitzt die Matrix

$$\tilde{\Phi}(1) = \begin{pmatrix} -I_r & \bar{\alpha}'_\perp \Gamma \bar{\beta} \\ 0 & \bar{\alpha}'_\perp \Gamma \bar{\beta} \end{pmatrix}$$

den vollen Rang, da laut Voraussetzung 3

$$\det \tilde{\Phi}(1) = (-1)^r \det(\bar{\alpha}'_\perp \Gamma \bar{\beta}_\perp) \neq 0$$

gilt. Damit besitzt  $\tilde{\Phi}(\mathbf{z})$  keine Einheitswurzel und alle anderen liegen laut Voraussetzung 1 außerhalb des Einheitskreises. Dadurch ist die Matrix  $\tilde{\Phi}(z)$  invertierbar:

$$\begin{aligned} \tilde{\mathbf{y}}_t &= \tilde{\Phi}^{-1}(B) \tilde{\varepsilon}_t \\ &= \tilde{\Phi}^{-1}(1) \tilde{\varepsilon}_t + \tilde{\Phi}^*(B) \Delta \tilde{\varepsilon}_t \end{aligned}$$

wobei  $\tilde{\Phi}^*(\mathbf{z}) = (\tilde{\Phi}^{-1}(\mathbf{z}) - \tilde{\Phi}^{-1}(1))(1 - \mathbf{z})^{-1}$  nach Johansen (1995, Lemma 4.1) für  $|\mathbf{z}| < 1 + c$ ,  $c > 0$  eine konvergente Potenzreihenentwicklung hat. Im letzten Schritt

wird mit Hilfe der Beziehung

$$\begin{aligned}
\mathbf{y}_t &= \sum_{i=1}^t \Delta \mathbf{y}_i + \mathbf{y}_0 \\
&= (\bar{\beta} \beta' + \bar{\beta}_\perp \beta'_\perp) \left( \sum_{i=1}^t \Delta \mathbf{y}_i + \mathbf{y}_0 \right) \\
&= \bar{\beta} \sum_{i=1}^t \Delta Z_i + \bar{\beta}_\perp \sum_{i=1}^t U_i + \bar{\beta} Z_0 + \bar{\beta}_\perp \beta'_\perp \mathbf{y}_0 \\
&= \bar{\beta} Z_t + \bar{\beta}_\perp \sum_{i=1}^t U_i + P_{\beta_\perp} \mathbf{y}_0
\end{aligned} \tag{3.9}$$

der transformierte VAR Prozess  $\tilde{\mathbf{y}}_t$  wieder rücktransformiert, indem die „Moving Average“ Darstellung in 3.9 eingesetzt wird

$$\begin{aligned}
\mathbf{y}_t &= \bar{\beta} Z_t + \bar{\beta}_\perp \sum_{i=1}^t U_i + P_{\beta_\perp} \mathbf{y}_0 \\
&= \bar{\beta}(I_r, 0) \tilde{\Phi}^{-1}(B) \tilde{\varepsilon}_t + \bar{\beta}_\perp(0, I_{n-r}) \tilde{\Phi}^{-1}(1) \sum_{i=1}^t \tilde{\varepsilon}_i + \bar{\beta}_\perp(0, I_{n-r}) \tilde{\Phi}^*(B) \tilde{\varepsilon}_t \\
&\quad + P_{\beta_\perp} \mathbf{y}_0 - \bar{\beta}_\perp(0, I_{n-r}) \tilde{\Phi}^*(B) \tilde{\varepsilon}_0 \\
&= \bar{\beta}_\perp(0, I_{n-r}) \begin{pmatrix} -I_r & \bar{\alpha}'_\perp \Gamma \bar{\beta} \\ 0 & \bar{\alpha}'_\perp \Gamma \bar{\beta} \end{pmatrix}^{-1} \begin{pmatrix} \bar{\alpha}' \\ \bar{\alpha}'_\perp \end{pmatrix} \sum_{i=1}^t (\varepsilon_i + \kappa D_i) + \\
&\quad + \left( \bar{\beta}(I_r, 0) \tilde{\Phi}^{-1}(B) + \bar{\beta}_\perp(0, I_{n-r}) (\tilde{\Phi}^{-1}(B) - \tilde{\Phi}^{-1}(1)) \Delta^{-1} \right) \begin{pmatrix} \bar{\alpha}' \\ \bar{\alpha}'_\perp \end{pmatrix} (\varepsilon_t + \kappa D_t) \\
&\quad + P_{\beta_\perp} \mathbf{y}_0 - \bar{\beta}_\perp(0, I_{n-r}) \tilde{\Phi}^*(B) \tilde{\varepsilon}_0 \\
&= \bar{\beta}_\perp (\bar{\alpha}'_\perp \Gamma \bar{\beta})^{-1} \bar{\alpha}'_\perp \sum_{i=1}^t (\varepsilon_i + \kappa D_i) + \\
&\quad + \left( (\bar{\beta}, \bar{\beta}_\perp) \begin{pmatrix} I_r & 0 \\ 0 & I_{n-r} \Delta^{-1} \end{pmatrix} \tilde{\Phi}^{-1}(B) \begin{pmatrix} \bar{\alpha}' \\ \bar{\alpha}'_\perp \end{pmatrix} - \bar{\beta}_\perp (\bar{\alpha}'_\perp \Gamma \bar{\beta})^{-1} \bar{\alpha}'_\perp \Delta^{-1} \right) (\varepsilon_t + \kappa D_t) + \\
&\quad + P_{\beta_\perp} \mathbf{y}_0 - \bar{\beta}_\perp(0, I_{n-r}) \tilde{\Phi}^*(B) \tilde{\varepsilon}_0 \\
&= C \sum_{i=1}^t (\varepsilon_i + \kappa D_i) + (\Phi(B)^{-1} - C \Delta^{-1}) (\varepsilon_t + \kappa D_t) + P_{\beta_\perp} \mathbf{y}_0 - \bar{\beta}_\perp(0, I_{n-r}) \tilde{\Phi}^*(B) \tilde{\varepsilon}_0 \\
&= C \sum_{i=1}^t (\varepsilon_i + \kappa D_i) + C(B) (\varepsilon_t + \kappa D_t) + P.
\end{aligned}$$

■

Mit dieser Formulierung und Beweisführung liefert das „Granger Representation Theorem“ nur für den Koeffizienten  $C$  eine konkrete Darstellung, die beiden anderen Variablen  $P$  und  $C(\mathbf{z})$  hängen entweder von den Startwerten ab, oder sind implizit definiert. In der asymptotischen Analysis von kointegrierten Prozessen reichen diese Informationen aus, in anderen Bereichen aber (vgl. Engle und Kozicki (1993), Lütkepohl und Reimers(1992), Warne (1993), ...) wird auch die konkrete Darstellung von  $C(\mathbf{z})$  benötigt. Eine explizite Angabe aller berechneten Variablen und Koeffizienten liefert Hansen (2000), indem er einen anderen Weg der Beweisführung wählt.

Die Darstellung, die man durch das „Granger Representation Theorem“ erhält teilt  $\mathbf{y}_t$  in drei Teile: Der erste Teil stellt die sogenannten „common trends“ dar, der zweite Teil ist abgesehen von der deterministischen Komponente  $\kappa D_t$ , stationär und der letzte Teil ist abhängig von den Startwerten des Prozesses.

### 3.4. Schätzung von kointegrierten Systemen

Das Hauptaugenmerk in diesem Kapitel liegt auf dem Johansen-Verfahren, dass auch bei den empirischen Untersuchungen verwendet wird, da es sehr viele unterschiedliche Modellspezifikationen (deterministische Komponenten) ermöglicht. Die statistische Analyse von Restriktionen der Parameter  $\alpha$  und  $\beta$  wird hier aber nicht durchgeführt, sondern auf Johansen (1995, Kapitel 7 und 8) verwiesen.

#### 3.4.1. Zweistufiges Verfahren von Engle-Granger (1987)

Für den Fall, dass es nur eine Kointegrationsbeziehung ( $r = 1$ ) gibt, haben Engle und Granger ein sehr populäres zweistufiges Verfahren entwickelt. Im ersten Schritt wird der Kointegrationsvektor mit einer OLS Schätzung bestimmt und dann werden, bedingt auf diesen Schätzer, die restlichen Parameter ermittelt. Für eine einfachere Analyse werden die Koeffizienten der deterministischen Komponenten auf Null gesetzt. Vor Anwendung der Prozedur wird von den einzelnen Zeitreihen der Integrationsrang mit Einheitswurzeltests bestimmt. Jede Zeitreihe sollte den gleichen Integrationsrang, hier ( $d = 1$ ), besitzen.

Ausgangspunkt der Analyse ist die VEC Darstellung eines Prozesses:

$$\Delta \mathbf{y}_t = \alpha \beta' \mathbf{y}_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta \mathbf{y}_{t-i} + \varepsilon_t, \quad t = 1, \dots, T. \quad (3.10)$$

Da ( $r = 1$ ) ist, ist  $\beta = (\beta_1, \dots, \beta_n)$  ein  $(1 \times n)$  Vektor und  $\beta \mathbf{y}_t = \beta_1 y_{1,t} + \dots + \beta_n y_{n,t}$ . Wird angenommen, dass das erste Element des Kointegrationsvektors ungleich Null ist, dann kann der Kointegrationsvektor so normiert werden, dass

$$(1, \tilde{\beta}) = \left( 1, -\frac{\beta_2}{\beta_1}, \dots, -\frac{\beta_n}{\beta_1} \right) = (1, -\tilde{\beta}_2, \dots, -\tilde{\beta}_n)$$

gilt. Die Kointegrationsbeziehung lautet dann

$$y_{1,t} = \tilde{\beta}_2 y_{2,t} + \cdots + \tilde{\beta}_n y_{n,t} + \xi_t$$

mit  $\xi_t$  einem stationären Prozess. Die Koeffizienten werden nun durch die Methode der Kleinsten-Quadrate geschätzt:

$$\hat{\beta} = y_{(1)} \mathbf{y}'_{(2)} (\mathbf{y}_{(2)} \mathbf{y}'_{(2)})^{-1} \quad (3.11)$$

wobei

$$y_{(1)} = \begin{pmatrix} y_{1,1} \\ \vdots \\ y_{1,T} \end{pmatrix} \quad \mathbf{y}_{(2)} = \begin{pmatrix} y_{2,1} & \cdots & y_{n,1} \\ \vdots & \ddots & \vdots \\ y_{2,T} & \cdots & y_{n,T} \end{pmatrix}.$$

In der zweiten Stufe werden die restlichen Parameter von (3.10), bedingt auf  $\hat{\beta} = (1, \hat{\beta})$ , bestimmt:

$$(\hat{\alpha}, \hat{\Gamma}) = Z_0 \left( \begin{pmatrix} \hat{\beta}' Z_1 \\ Z_2 \end{pmatrix} \right)' \left( \left( \begin{pmatrix} \hat{\beta}' Z_1 \\ Z_2 \end{pmatrix} \right) \left( \begin{pmatrix} \hat{\beta}' Z_1 \\ Z_2 \end{pmatrix} \right)' \right)^{-1} \quad (3.12)$$

wobei

$$\begin{aligned} Z_0 &= (\Delta \mathbf{y}_1, \dots, \Delta \mathbf{y}_T) \\ Z_1 &= (\mathbf{y}_0, \dots, \mathbf{y}_{T-1}) \\ Z_2 &= \begin{pmatrix} \Delta \mathbf{y}_0 & \cdots & \Delta \mathbf{y}_{T-1} \\ \vdots & \ddots & \vdots \\ \Delta \mathbf{y}_{-p+2} & \cdots & \Delta \mathbf{y}_{T-p+1} \end{pmatrix} \quad \text{mit Startwerten } \mathbf{y}_{-p+1}, \dots, \mathbf{y}_0 \\ \tilde{\Gamma}_1 &= (\Gamma_1, \dots, \Gamma_{p-1}). \end{aligned}$$

Mit den Eigenschaften dieser Schätzer beschäftigten sich Stock (1987), Phillips & Durlauf (1986), Park & Phillips (1989) und Johansen (1989a), um nur einige zu nennen.

Im Gegensatz zur ML-Schätzung (vgl. Abschnitt 3.4.2) muss beim mehrstufigen Verfahren bekannt sein, dass es eine Variable in der Kointegrationsbeziehung gibt, die einen Koeffizienten  $\beta_1 \neq 0$  besitzt. Diese Variable wird als abhängige Variable in der Kointegrationsregression (3.11) benutzt.

### 3.4.2. Johansen Verfahren (Maximum Likelihood)

Eine effiziente Schätzung des Systems von kointegrierten Variablen ist mittels einer Maximum-Likelihood-Prozedur möglich, die von Johansen abgeleitet wurde. Ausgangspunkt seiner Schätzung ist die VECM Darstellung eines Prozesses:

$$\Delta \mathbf{y}_t = \alpha \beta' \mathbf{y}_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta \mathbf{y}_{t-i} + \kappa D_t + \varepsilon_t, \quad t = 1, \dots, T \quad (3.13)$$

mit unabhängigen, normalverteilten  $\varepsilon_t$  ( $\varepsilon_t \sim NID(0, \Sigma)$ ) und uneingeschränkten Parametern  $(\alpha, \beta, \Gamma_1, \dots, \Gamma_{p-1}, \Phi, \Sigma)$ .

Manchmal wird auch die Darstellung

$$\Delta \mathbf{y}_t = \sum_{i=1}^{p-1} \tilde{\Gamma}_i \Delta \mathbf{y}_{t-i} + \alpha \beta' \mathbf{y}_{t-p} + \kappa D_t + \varepsilon_t$$

mit  $\tilde{\Gamma}_i = \Phi_1 + \dots + \Phi_i - I_n$ , benutzt. Diese Reparameterisierung führt aber zur gleichen statistischen Analyse.

Nach dem Granger Representation Satz besitzt die Lösung  $\mathbf{y}_t$  der Gleichung (3.13) folgende Darstellung

$$\mathbf{y}_t = C \sum_{i=1}^t (\varepsilon_i + \kappa D_i) + C(B)(\varepsilon_t + \kappa D_t) + A \quad (3.14)$$

mit  $C = \bar{\beta}_\perp (\bar{\alpha}'_\perp \Gamma \bar{\beta}_\perp)^{-1} \bar{\alpha}'_\perp$ . Man sieht, dass der Prozess ganz allgemein einen deterministischen Teil der Form  $C \kappa \sum_{i=1}^t (D_i) + C(B) \kappa D_t$  enthält. Ist  $\kappa D_t$  konstant, so besitzt der Prozess  $\mathbf{y}_t$  einen linearen Trend und ein linearer Trend in  $D_t$  erzeugt einen quadratischen Trend im Prozess. Durch diesen Umstand kann es sehr leicht zu sprachlichen Missverständnissen kommen, wenn man nicht genau spezifiziert, ob man von  $D_t$  oder vom deterministischen Teil des Prozesses  $\mathbf{y}_t$  spricht. Für die weitere Analyse wird  $\kappa D_t = \nu + \delta t$  angenommen und die Parameter  $\nu$  und  $\delta$  in die Teile  $\alpha$  und  $\alpha_\perp$  aufgespalten:

$$\nu = \alpha \tau_0 + \alpha_\perp \omega_0 \quad (3.15)$$

$$\delta = \alpha \tau_1 + \alpha_\perp \omega_1 \quad (3.16)$$

wobei  $\tau_0 = (\alpha' \alpha)^{-1} \alpha' \nu$ ,  $\tau_1 = (\alpha' \alpha)^{-1} \alpha' \delta$ ,  $\omega_0 = (\alpha'_\perp \alpha_\perp)^{-1} \alpha'_\perp \nu$  und  $\omega_1 = (\alpha'_\perp \alpha_\perp)^{-1} \alpha'_\perp \delta$ . Durch diese Betrachtungsweise lassen sich fünf Modellklassen definieren:

$$\begin{aligned} M(r) : \quad \kappa D_t &= \alpha \tau_0 + \alpha_\perp \omega_0 + (\alpha \tau_1 + \alpha_\perp \omega_1) t \\ M^*(r) : \quad \kappa D_t &= \alpha \tau_0 + \alpha_\perp \omega_0 + \alpha \tau_1 t \\ M_1(r) : \quad \kappa D_t &= \alpha \tau_0 + \alpha_\perp \omega_0 \\ M_1^*(r) : \quad \kappa D_t &= \alpha \tau_0 \\ M_2(r) : \quad \kappa D_t &= 0. \end{aligned}$$

Für ein konkretes Modell in einer bestimmten Klasse wird die Bedingung formuliert, dass der Rang von  $\Pi$  kleiner oder gleich  $r$  ist. Tabelle 3.1 zeigt wie die Modelle ineinander verschachtelt sind.

Für den Fall  $r = n$  erhält man einen VAR Prozess, in dem alle Variablen stationär sind. Ist  $r = 0$  dann ist  $\Pi = 0$  und man erhält einen VAR Prozess für die Differenzen von  $\mathbf{y}_t$ . Die Modelle, die zwischen diesen beiden Extremen liegen, stellen den Grad der Kointegrationsbeziehungen dar. In Abschnitt 3.5.2 wird für die einzelnen

Tabelle 3.1: Die Beziehung zwischen den einzelnen  $I(1)$  Modellklassen

$M(0)$	$\subset$	$\dots$	$\subset$	$M(r)$	$\subset$	$\dots$	$\subset$	$M(n)$
$\cup$				$\cup$				$\cup$
$M^*(0)$	$\subset$	$\dots$	$\subset$	$M^*(r)$	$\subset$	$\dots$	$\subset$	$M^*(n)$
$\parallel$				$\cup$				$M \cup$
$M_1(0)$	$\subset$	$\dots$	$\subset$	$M_1(r)$	$\subset$	$\dots$	$\subset$	$M_1(n)$
$\cup$				$\cup$				$\cup$
$M_1^*(0)$	$\subset$	$\dots$	$\subset$	$M_1^*(r)$	$\subset$	$\dots$	$\subset$	$M_1^*(n)$
$\parallel$				$\cup$				$\cup$
$M_2(0)$	$\subset$	$\dots$	$\subset$	$M_2(r)$	$\subset$	$\dots$	$\subset$	$M_2(n)$

Quelle: [JOHANSEN 1995]

Klassen ein Test beschrieben um entweder eine gegebene Kointegrationsbeziehung zu überprüfen oder um durch eine Reihe von Tests einen Schätzer für den Kointegrationsrang zu bestimmen.

Die Maximum Likelihood Schätzung wird nun in zwei Schritten durchgeführt. Zuerst wird für die Klassen  $M(r)$ ,  $M_1(r)$  und  $M_2(r)$  ein gemeinsames Verfahren entwickelt (es wird wieder die Notation  $\kappa D_t$  verwendet) und anschließend werden die beiden anderen Fälle auf die allgemeine Prozedur zurückgeführt.

Damit die Gleichung (3.13) kompakter geschrieben werden kann wird folgende Notation eingeführt:

$$\begin{aligned}
 Z_{0t} &= \Delta \mathbf{y}_t \\
 Z_{1t} &= \mathbf{y}_{t-1} \\
 Z_{2t} &= (\Delta \mathbf{y}'_{t-1}, \dots, \Delta \mathbf{y}'_{t-p+1}, D'_t)' \\
 \tilde{\Gamma} &= (\Gamma_1, \dots, \Gamma_{p-1}, \kappa).
 \end{aligned}$$

Mit diesen Variablen verändert sich das VECM zu

$$Z_{0t} = \alpha \beta' Z_{1t} + \tilde{\Gamma} Z_{2t} + \varepsilon_t. \quad (3.17)$$

Damit erhält man ein nichtlineares Regressionsmodell mit uneingeschränktem Parameter  $\tilde{\Gamma}$ . Die Koeffizientenmatrix von  $Z_{1t}$  besitzt nicht den vollen Rang ( $\text{rank } \Pi = r < n$ ), man spricht daher auch von einer „reduced rank“ Regression. Die logarithmierte Likelihoodfunktion lautet, abgesehen von der Konstanten  $-(T/2) \log(2\pi)$

$$\begin{aligned}
 \log L(\tilde{\Gamma}, \alpha, \beta, \Sigma) &= -\frac{1}{2} T \log(\det \Sigma) - \\
 &\quad - \frac{1}{2} \sum_{t=1}^T (Z_{0t} - \alpha \beta' Z_{1t} - \tilde{\Gamma} Z_{2t})' \Sigma^{-1} (Z_{0t} - \alpha \beta' Z_{1t} - \tilde{\Gamma} Z_{2t}).
 \end{aligned}$$

Ableiten nach  $\tilde{\Gamma}$  und Null setzen liefert folgende Bedingung

$$\sum_{t=1}^T (Z_{0t} - \alpha \beta' Z_{1t} - \hat{\Gamma} Z_{2t}) Z'_{2t} = 0. \quad (3.18)$$

Definiert man die empirischen zweiten Momente mit

$$M_{ij} = \frac{1}{T} \sum_{t=1}^T Z_{it} Z'_{jt} \quad i, j = 0, 1, 2$$

wobei gilt

$$M_{ij} = M'_{ji} \quad i, j = 0, 1, 2$$

dann lässt sich Gleichung (3.18) schreiben als

$$M_{02} = \alpha \beta' M_{12} + \hat{\Gamma} M_{22}$$

und der optimale Wert für  $\hat{\Gamma}$  (für gegebene  $\alpha$  und  $\beta$ ) ist gleich

$$\hat{\Gamma}(\alpha, \beta) = M_{02} M_{22}^{-1} - \alpha \beta' M_{12} M_{22}^{-1}.$$

Dieses Ergebnis führt zur Definition der Residuen

$$\begin{aligned} R_{0t} &= Z_{0t} - M_{02} M_{22}^{-1} Z_{2t} \\ R_{1t} &= Z_{1t} - M_{12} M_{22}^{-1} Z_{2t} \end{aligned}$$

der Regression von  $Z_{0t}$  bzw.  $Z_{1t}$  auf  $Z_{2t}$ . Damit ist die konzentrierte logarithmierte Likelihoodfunktion gleich

$$\log L(\alpha, \beta, \Sigma) = -\frac{T}{2} \log \det \Sigma - \frac{1}{2} \sum_{t=1}^T (R_{0t} - \alpha \beta' R_{1t})' \Sigma^{-1} (R_{0t} - \alpha \beta' R_{1t}). \quad (3.19)$$

Als letzte Notation werden die empirischen Momente der Residuen  $R_{0t}$  und  $R_{1t}$  eingeführt:

$$S_{ij} = \frac{1}{T} \sum_{t=1}^T R_{it} R'_{jt} = M_{ij} - M_{i2} M_{22}^{-1} M_{2j} \quad i, j = 0, 1.$$

Für gegebenes  $\beta$  ist der optimale Wert für  $\alpha$  und  $\Sigma$  durch die Regression von  $R_{0t}$  auf  $\beta R_{1t}$  gegeben:

$$\begin{aligned} \hat{\alpha}(\beta) &= S_{01} \beta (\beta' S_{11} \beta)^{-1} \\ \hat{\Sigma}(\beta) &= S_{00} - S_{01} \beta (\beta' S_{11} \beta)^{-1} \beta' S_{10} = S_{00} - \hat{\alpha}(\beta) (\beta' S_{11} \beta)^{-1} \hat{\alpha}(\beta)' \end{aligned}$$

und die konzentrierte Likelihood ist gleich:

$$L_{max}(\hat{\alpha}(\beta), \beta, \hat{\Sigma}(\beta)) = (\det \hat{\Sigma})^{-\frac{T}{2}} = (\det [S_{00} - S_{01} \beta (\beta' S_{11} \beta)^{-1} \beta' S_{10}])^{-\frac{T}{2}}.$$

Dieser Ausdruck soll nun maximiert werden. Dazu wird folgende Identität benutzt:

$$\det \begin{bmatrix} \Upsilon_{11} & \Upsilon_{12} \\ \Upsilon_{21} & \Upsilon_{22} \end{bmatrix} = \det \Upsilon_{11} \det [\Upsilon_{22} - \Upsilon_{21} \Upsilon_{11}^{-1} \Upsilon_{12}] = \det \Upsilon_{22} \det [\Upsilon_{11} - \Upsilon_{12} \Upsilon_{22}^{-1} \Upsilon_{21}].$$

Angewendet auf die Matrix

$$\begin{aligned} \det \begin{bmatrix} S_{00} & S_{01}\beta \\ \beta' S_{10} & \beta' S_{11}\beta \end{bmatrix} &= \det S_{00} \det[\beta'(S_{11} - S_{10}S_{00}^{-1}S_{01})\beta] \\ &= \det[\beta' S_{11}\beta] \det[(S_{00} - S_{01}\beta(\beta' S_{11}\beta)^{-1}\beta' S_{10})] \end{aligned}$$

folgt daraus nun

$$\begin{aligned} &\det[S_{00} - S_{01}\beta(\beta' S_{11}\beta)^{-1}\beta' S_{10}] \\ &= \det S_{00} \det[\beta' S_{11}\beta - \beta' S_{01}S_{00}^{-1}S_{01}\beta] / \det[\beta' S_{11}\beta] \\ &= \det S_{00} \det[\beta'(S_{11} - S_{10}S_{00}^{-1}S_{01})\beta] / \det[\beta' S_{11}\beta]. \end{aligned} \quad (3.20)$$

Die Maximierung der Likelihoodfunktion ist damit equivalent zur Minimierung des letzten Faktors von Gleichung (3.20). Dafür benötigt man folgendes Hilfsresultat:

**Satz 3.2** Sind die Matrizen  $M \in \mathbb{R}^{n \times n}$  und  $N \in \mathbb{R}^{n \times n}$  symmetrisch und positiv semi-definit und  $X \in \mathbb{R}^{n \times r}$ , dann wird die Funktion

$$f(X) = \det[X' M X] / \det[X' N X]$$

maximal, wenn  $X = (v_1, \dots, v_r)$ , wobei  $\lambda_i$  und  $v_i$  die Eigenwerte und die dazugehörigen Eigenvektoren des Eigenwertproblems

$$\det[\lambda N - M] = 0$$

sind. Das Maximum der Funktion ist durch

$$\max f(X) = \prod_{i=1}^r \lambda_i$$

gegeben.

**Beweis 3.2** siehe Johansen (1995)

Den Schätzer für  $\beta$  erhält man daher durch Lösen des Eigenwertproblems

$$\det(\tilde{\lambda} S_{11} - (S_{11} - S_{10}S_{00}^{-1}S_{01})) = 0$$

bzw. für  $\lambda = 1 - \tilde{\lambda}$  des Eigenwertproblems

$$\det(\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}) = 0. \quad (3.21)$$

Der ML Schätzer für  $\beta$  ist dann gegeben durch

$$\hat{\beta} = (\hat{v}_1, \dots, \hat{v}_r) \hat{T}$$

wobei  $\hat{v}_1, \hat{v}_2, \dots, \hat{v}_n$  die Eigenvektoren des Eigenwertproblems 3.21 zu den Eigenwerten  $0 \leq \lambda_1 \leq \dots \leq \lambda_n \leq 1$  sind und die Matrix  $\hat{T} \in \mathbb{R}^{r \times r}$  so gewählt wird, dass die „Normierungsbedingungen“

$$\begin{aligned}\hat{\beta} S_{11} \hat{\beta} &= I_r \\ \hat{\beta}' (S_{10} S_{00}^{-1} S_{01}) \hat{\beta} &= \text{diag}(\lambda_1, \dots, \lambda_n)\end{aligned}$$

erfüllt sind. Das Maximum der Likelihoodfunktion ist gegeben durch:

$$L_{max}(\hat{\alpha}, \hat{\beta}, \hat{\Gamma}, \hat{\Sigma}) = \left( \det S_{00} \prod_{i=1}^r (1 - \lambda_i) \right)^{-\frac{T}{2}}. \quad (3.22)$$

Damit hat man nun ein Schätzverfahren, das allgemein für einen deterministischen Term  $D_t$  gilt. Nimmt man  $\kappa D_t = \nu + \delta t$ , so erhält man das Modell  $M(r)$ , für  $\kappa D_t = \nu$  erhält man das Modell  $M_1(r)$  und für  $\kappa D_t = 0$  erhält man das Modell  $M_2(r)$ . Was noch verbleibt ist die Analysis der Modelle, in welchen der deterministische Teil eingeschränkt ist:  $M^*(r)$  wo  $\alpha'_{\perp} \delta = 0$  und  $M_1^*(r)$  wo  $\delta = 0$  und  $\alpha'_{\perp} \nu = 0$  gilt. Die Analysis dieser Modelle ist den uneingeschränkten Modellen sehr ähnlich.

Für den Fall  $M^*(r)$  ist  $\Pi = \alpha \beta'$ ,  $\kappa D_t = \nu + \delta t$  und  $\alpha'_{\perp} \delta = 0$ . Aufgrund der Relation

$$\alpha \beta' \mathbf{y}_t + \delta t = \alpha \beta' \mathbf{y}_t + \alpha \tau_1 t = \alpha (\beta', \tau_1) (\mathbf{y}'_t, t)' = \alpha \beta^{*'} Z_{1t}^*$$

definiert man folgende Variablen

$$\begin{aligned}Z_{0t}^* &= \Delta \mathbf{y}_t \\ Z_{1t}^* &= (\mathbf{y}'_{t-1}, t)' \\ Z_{2t}^* &= (\Delta \mathbf{y}'_{t-1}, \dots, \Delta \mathbf{y}'_{t-p+1}, 1)' \\ \beta^* &= (\beta', \tau_1)' \\ \tilde{\Gamma}^* &= (\Gamma_1, \dots, \Gamma_{p-1}, \nu)\end{aligned}$$

und erhält das Regressionsmodell

$$Z_{0t}^* = \alpha \beta^{*'} Z_{1t}^* + \tilde{\Gamma}^* Z_{2t}^* + \varepsilon_t. \quad (3.23)$$

Die Schätzung des Modells erfolgt nun mit derselben Prozedur wie oben.

Analog dazu behandelt man das Modell  $M_1^*(r)$ . In diesem Fall werden durch die Tatsache

$$\alpha \beta' \mathbf{y}_t + \alpha \tau_0 = \alpha (\beta', \tau_0) (\mathbf{y}'_t, 1)' = \alpha \beta^{*'} Z_{1t}^*$$

die Variablen wie folgt definiert:

$$\begin{aligned}Z_{0t}^* &= \Delta \mathbf{y}_t \\ Z_{1t}^* &= (\mathbf{y}'_{t-1}, 1)' \\ Z_{2t}^* &= (\Delta \mathbf{y}'_{t-1}, \dots, \Delta \mathbf{y}'_{t-p+1})' \\ \beta^* &= (\beta', \tau_0)' \\ \tilde{\Gamma}^* &= (\Gamma_1, \dots, \Gamma_{p-1}).\end{aligned}$$

Die Schätzung des Regressionsmodells ist nun wieder analog zum uneingeschränkten Fall durchzuführen.

### 3.4.3. Andere Verfahren

#### Nichtlinearer Ansatz von Stock

Beim zweistufigen Verfahren von Engle und Granger werden superkonsistente<sup>1</sup> Schätzer für den Kointegrationsvektor erzielt. Stock (1987) weist darauf hin, dass die Engle-Granger Schätzer für die Kointegrationsmatrix nicht effizient sind, da die Dynamik des Systems nicht beachtet wird. Stock schlägt für  $r = 1$  deshalb die Schätzung folgender Regressionsgleichung vor

$$\Delta y_{1,t} = f_1 y_{1,t-1} + f_2 y_{2,t-1} + \dots + f_n y_{n,t-1} + \sum_{i=1}^{p-1} \Gamma_{1,i} \Delta y_{t-i} + \xi_t. \quad (3.24)$$

Die Schätzung erfolgt mit der Methode der Kleinsten-Quadrate. Dieser Ansatz wird nicht linear genannt, da die Kointegrationsparameter eine nichtlineare Kombination der  $f_i$  sind. Die Schätzung der Kointegrationsparameter lautet

$$\begin{aligned} \hat{\beta}_1 &= 1 \\ \hat{\beta}_i &= \frac{\hat{f}_i}{\hat{f}_1} \quad \text{für } i = 2, \dots, n. \end{aligned}$$

Mit diesem Ergebnis steigt man in den zweiten Teil des Engle-Granger Verfahrens ein.

## 3.5. Kointegrationstests

In den vorangegangenen Kapiteln wurde die Charakterisierung von kointegrierten Systemen dargestellt und verschiedene Schätzmethoden für kointegrierte Systeme oder Kointegrationsbeziehungen aufgeführt. Bei der Analyse wurde ein bekannter Kointegrationsrang unterstellt. In empirischen Arbeiten ist der Kointegrationsrang nicht bekannt und muss ermittelt werden. Zur Bestimmung des Kointegrationsranges wurden unterschiedliche Teststrategien entwickelt. Ein kleiner Teil der bereits zahlreichen Tests wird nun vorgestellt.

### 3.5.1. Erweiterte „Unit Root“ Tests

In diesen Testansätzen wird die Hypothese überprüft, ob es eine Kointegrationsbeziehung zwischen zwei oder allgemeiner:  $n$  Variablen gibt. In diesen Tests wird unterstellt, dass alle Komponenten von  $\mathbf{y}_t$  nicht stationär, genauer  $y_t \sim I(1)$  sind

<sup>1</sup>Der Schätzer konvergiert in Verteilung mit  $T$  (siehe Phillips(1989))

(vgl. Engle & Granger (1987)). Anschließend wird mit den integrierten Daten die von Engle und Granger vorgeschlagene Kointegrationsregression (vgl. 3.11)

$$y_{1,t} = \gamma_2 y_{2,t} + \cdots + \gamma_n y_{n,t} + \xi_t$$

oder die von Stock (vgl. 3.24)

$$\Delta y_{1,t} = f_1 y_{1,t-1} + f_2 y_{2,t-1} + \cdots + f_n y_{n,t-1} + \sum_{i=1}^{p-1} \Pi_{1,i} \Delta y_{t-i} + \xi_t$$

durchgeführt. Beide Regressionen liegt die Annahme zugrunde, dass der erste Parameter  $\beta_1$  des allgemeinen Kointegrationsvektors  $\beta$  ungleich Null ist. Der Kointegrationsvektor wird für die erste Variable auf Eins normiert. Die aus der Regression erhaltenen Residuen  $\hat{\xi}_t$  werden nun einem Einheitswurzeltest unterzogen. Unter der Nullhypothese

$$H_0 : r = 0 \text{ (keine Kointegration)}$$

mit der alternativen Hypothese

$$H_1 : r = 1$$

ist der Residuenprozess integriert mit der Ordnung Eins.

### 3.5.2. Johansen

Der in Abschnitt 3.5.2 entwickelte Maximum-Likelihood-Schätzansatz erlaubt die Formulierung eines Likelihoodverhältnistests (Johansen (1995)) zur Bestimmung des Kointegrationsranges.

Das Maximum der logarithmierten Likelihoodfunktion ist gleich

$$L_{max} = \left( \det S_{00}^{\times} \prod_{i=1}^r (1 - \hat{\lambda}_i^{\times}) \right)^{-\frac{T}{2}} \quad (3.25)$$

wobei die Matrix  $S_{00}^{\times}$  und die Eigenwerte  $\hat{\lambda}_i^{\times}$  für jedes gewählte Modell (z. B.:  $M(r)$ ,  $M^*(r)$ ,  $M_1(r), \dots$ ) unterschiedlich sind. Dieser Umstand soll durch das Zeichen  $\times$  symbolisiert werden.

Die Likelihoodverhältnisteststatistik der Nullhypothese

$$H_0 : \text{rank } \Pi \leq r$$

mit der alternativen Hypothese

$$H_1 : \text{rank } \Pi = n \quad ((\mathbf{y}_t) \text{ ist stationär})$$

ist gegeben durch die sogenannte „trace“ Statistik

$$-2 \log Q(H^\times(r)|H^\times(p)) = Q_r^\times = -T \sum_{i=r+1}^p \log(1 - \hat{\lambda}_i^\times) \quad (3.26)$$

Diese Teststatistik konvergiert in Verteilung gegen ein Funktional einer  $(n - r)$ -dimensionalen Brownschen Bewegung, das von der Spezifikation des deterministischen Teils des Modells abhängt. Die Quantile der asymptotischen Verteilungen sind in den Tabellen 3.2 bis 3.6 angegeben.

Tabelle 3.2: Kritische Werte des Likelihoodverhältnistests im Modell  $M_2(r)$

n-r	10.0%	5.0%	2.5%	1.0%
1	2.98	4.14	5.30	7.02
2	10.35	12.21	13.94	16.16
3	21.58	24.08	26.42	29.19
4	36.58	39.71	42.59	46.00
5	55.54	59.24	62.68	66.71
6	75.48	82.61	86.36	91.12
7	104.93	109.93	114.24	119.58
8	135.16	140.74	145.80	151.70
9	169.30	175.47	181.20	187.82
10	207.21	214.07	220.00	226.95
11	248.77	256.23	262.69	270.47
12	293.83	301.95	309.08	318.14

Hinweis:  $n - r$  stellt die Anzahl der sogenannten „common trends“ dar; wird  $n - r = 1$  abgelehnt so sind alle Variablen stationär (vgl. Absatz).

Quelle: [JOHANSEN 1995]

Johansen schlägt folgende Prozedur zur Bestimmung des Kointegrationsranges vor: Zuerst vergleicht man  $Q_0^\times$  mit  $c_0^\times$ ; gilt  $Q_0^\times < c_0^\times$  dann ist der Kointegrationsrang gleich Null; gilt  $Q_0^\times \geq c_0^\times$  dann berechnet man  $Q_1^\times$  und vergleicht es mit  $c_1^\times$ ; ist nun  $Q_1^\times < c_1^\times$  dann ist der Kointegrationsrang gleich Eins, andernfalls berechnet man  $Q_2^\times$  und vergleicht es mit  $c_2^\times$  usw. Für eine genauere Analyse, vor allem in Bezug auf die verschiedenen Varianten des deterministischen Teils, siehe Johansen (1995) Kapitel 6.3.

Beim Likelihoodverhältnistest ist anzumerken, dass eine gesonderte Überprüfung der I(1)-Eigenschaft der einzelnen Komponenten in  $y_t$  nicht notwendig ist. Beim Test  $H_0 : r = n - 1$  gegen  $H_1 : r = n$  wird ein entsprechender Test durchgeführt. Wird die Nullhypothese verworfen, dann sind alle Variablen stationär. Wenn die Nullhypothese nicht verworfen wird, gibt es nichtstationäre Komponenten in  $y_t$ .

Tabelle 3.3: Kritische Werte des Likelihoodverhältnistests im Modell  $M_1^*(r)$

n-r	10.0%	5.0%	2.5%	1.0%
1	7.50	9.13	10.73	12.73
2	17.79	19.99	22.07	24.74
3	31.88	34.80	37.56	40.84
4	49.92	53.42	56.57	60.42
5	71.66	75.74	79.60	83.93
6	97.17	101.84	106.07	111.38
7	126.71	132.00	136.69	142.34
8	159.74	165.73	171.18	177.42
9	196.66	203.34	209.28	216.08
10	237.35	244.56	250.86	258.31
11	281.63	289.71	296.70	304.89
12	333.26	338.10	354.77	354.32

Hinweis:  $n-r$  stellt die Anzahl der sogenannten „common trends“ dar; wird  $n-r = 1$  abgelehnt so sind alle Variablen stationär (vgl. Absatz).

Quelle: [JOHANSEN 1995]

Tabelle 3.4: Kritische Werte des Likelihoodverhältnistests im Modell  $M_1(r)$

n-r	10.0%	5.0%	2.5%	1.0%
1	2.71	3.84	5.02	6.64
2	13.31	15.34	17.24	19.69
3	26.70	29.38	31.76	34.87
4	43.84	47.21	50.19	53.91
5	64.74	68.68	72.21	76.37
6	89.37	93.92	97.97	102.95
7	117.73	123.04	127.59	133.04
8	149.99	155.75	160.87	166.95
9	185.83	192.30	198.19	204.64
10	228.55	232.60	238.69	246.17
11	272.03	276.37	283.24	291.58
12	319.21	323.93	331.22	339.64

Hinweis:  $n-r$  stellt die Anzahl der sogenannten „common trends“ dar; wird  $n-r = 1$  abgelehnt so sind alle Variablen stationär (vgl. Absatz).

Quelle: [JOHANSEN 1995]

Für kleine Stichproben schlagen Reinsel und Ahn (1992) eine Korrektur der Teststatistik vor. Anstatt die Stichprobengröße  $T$  zu verwenden, sollte man den Faktor

Tabelle 3.5: Kritische Werte des Likelihoodverhältnistests im Modell  $M^*(r)$

n-r	10.0%	5.0%	2.5%	1.0%
1	10.56	12.39	14.13	16.39
2	22.95	25.47	27.87	30.65
3	39.08	42.20	45.04	48.59
4	58.96	62.61	66.04	70.22
5	82.68	86.96	90.87	95.38
6	110.00	114.96	119.43	124.61
7	141.31	146.75	151.70	157.53
8	176.13	182.45	187.80	194.12
9	214.72	221.56	227.61	234.65
10	257.08	264.23	270.90	278.80
11	306.47	311.13	318.03	326.73
12	356.39	361.07	368.75	377.54

Hinweis:  $n-r$  stellt die Anzahl der sogenannten „common trends“ dar; wird  $n-r = 1$  abgelehnt so sind alle Variablen stationär (vgl. Absatz).

Quelle: [JOHANSEN 1995]

Tabelle 3.6: Kritische Werte des Likelihoodverhältnistests im Modell  $M(r)$

n-r	10.0%	5.0%	2.5%	1.0%
1	2.71	3.84	5.02	6.64
2	15.94	18.15	20.26	22.78
3	31.57	34.56	37.21	40.61
4	50.67	54.11	57.41	61.28
5	73.62	77.79	81.57	86.11
6	99.97	104.76	109.04	114.23
7	130.39	135.66	140.49	146.05
8	160.32	170.15	175.38	181.62
9	201.83	208.53	214.51	221.47
10	243.29	250.53	256.96	264.52
11	291.51	296.02	302.86	311.41
12	340.54	345.27	352.98	361.69

Hinweis:  $n-r$  stellt die Anzahl der sogenannten „common trends“ dar; wird  $n-r = 1$  abgelehnt so sind alle Variablen stationär (vgl. Absatz).

Quelle: [JOHANSEN 1995]

$(T - pn)$  in die Kalkulation einbauen. Diese Modifikation passt die Teststatistik für kleine Stichproben besser an die asymptotische Verteilung an.

Ausgehend vom ML-Ansatz von Johansen können die Kointegrationsbeziehungen auch mit Hilfe von Ordnungskriterien untersucht werden. Darauf wird hier jedoch nicht näher eingegangen, sondern auf Reimers (1991) verwiesen.

### 3.5.3. Phillips und Ouliaris (1990)

Die beiden Phillips und Ouliaris Tests bauen auf einem Spektralschätzer der Residuen einer Kointegrationsregression auf. Im Gegensatz zu sehr vielen anderen Tests ist bei ihnen die Nullhypothese: keine Kointegration

$$H_0 : r = 0$$

und die alternative Hypothese: es gibt mindestens eine Kointegrationsbeziehung

$$H_1 : r > 0.$$

Das Problem bei dieser Testart ist klar: im mehrdimensionalen Fall ist keine exakte Aussage über die Anzahl der Kointegrationsvektoren möglich. Einer der beiden Tests besitzt jedoch einen entscheidenden Vorteil im Vergleich zu anderen Residuen bezogenen Tests. Dieser Test ist unabhängig von der Normalisierung der Kointegrationsregression, die in der empirischen Analyse ein nicht zu unterschätzendes Problem darstellt.

Für die Berechnung der Teststatistik startet man bei einer einfachen Vektorregression

$$\mathbf{y}_t = \Phi_1 \mathbf{y}_{t-1} + \xi_t. \quad (3.27)$$

Mit den Residuen  $\hat{\xi}_t$  berechnet einen Schätzer für das Spektrum des Prozesses  $\xi_t$  an der Frequenz Null:

$$\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T \hat{\xi}_t \hat{\xi}_t' + \frac{1}{T} \sum_{s=1}^l w_{sl} \frac{1}{T} \sum_{t=s+1}^T \hat{\xi}_t \hat{\xi}_{t-s}' + \hat{\xi}_{t-s} \hat{\xi}_t'$$

wobei  $w_{sl} = 1 - s/(l+1)$  (siehe Phillips und Durlauf (1986)) die Gewichtsfunktion und  $l$  ein Abschneidungsparameter ist (z.B.  $l = T^{1/2}$ ). Partitioniert man die Matrix

$$\hat{\Sigma} = \begin{pmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{21}' \\ \hat{\sigma}_{21} & \hat{\Sigma}_{22} \end{pmatrix}$$

so kann man  $\hat{\sigma}_{11.2} = \hat{\sigma}_{11} - \hat{\sigma}_{21}' \hat{\Sigma}_{22}^{-1} \hat{\sigma}_{21}$  berechnen.

Die erste Teststatistik berechnet sich nun durch

$$\hat{P}_u = \frac{T \hat{\sigma}_{11.2}}{T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t^2} \quad (3.28)$$

wobei  $\hat{\varepsilon}_t$  die Residuen aus

$$\mathbf{y}_{1,t} = \hat{\beta}' \mathbf{y}_{2-n,t} + \varepsilon_t \quad (3.29)$$

sind. Die zweite Teststatistik ergibt sich durch

$$\hat{P}_z = T \text{sp}\{\hat{\Sigma} M_{zz}^{-1}\} \quad (3.30)$$

wobei  $M_{zz} = T^{-1} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t'$ . Diese Teststatistik benötigt die Residuen der Regression (3.29) nicht und ist damit nicht abhängig von deren Normalisierung.

Tabelle 3.7: Kritische Werte für die  $\hat{P}_u$  Teststatistik

no constant					
$n - 1$	10.0 %	7.5%	5.0%	2.5%	1.0%
1	20.3933	22.7588	25.9711	31.8337	38.3413
2	26.7022	29.4114	32.9392	39.2236	46.4097
3	33.5359	36.5407	40.1220	46.3395	55.7341
4	39.2826	41.8969	46.2691	53.3683	63.2149
5	44.3725	47.6970	51.8614	59.6040	69.4939
constant					
$n - 1$	10.0 %	7.5%	5.0%	2.5%	1.0%
1	27.8536	30.3123	33.7130	39.9288	48.0021
2	33.6955	36.4757	40.5252	46.6707	53.8731
3	39.6949	42.8111	46.7281	53.9710	63.4128
4	45.3308	48.6675	53.2502	61.2555	71.5214
5	50.3537	53.5654	57.7855	65.8230	76.7705
trend					
$n - 1$	10.0 %	7.5%	5.0%	2.5%	1.0%
1	41.2488	44.2614	48.8439	56.0886	65.1714
2	46.1061	49.3671	53.8300	60.8745	69.2629
3	52.0015	55.4625	60.2384	68.4051	78.3470
4	57.3667	60.8175	65.8706	74.4712	84.5480
5	61.6155	65.3514	70.7416	79.0043	91.0392

Quelle: [PHILLIPS AND OULIARIS 1990]

Werden die Regressionen (3.27) und (3.29) mit einer Konstanten durchgeführt und wird  $\bar{M}_{zz} = T^{-1} \sum_{t=1}^T T(\mathbf{y}_t - \bar{\mathbf{y}}_t)(\mathbf{y}_t - \bar{\mathbf{y}}_t)'$  verwendet, verändern sich die asymptotischen, kritischen Werte. In allen drei Gleichungen kann auch ein Trend miteinbezogen werden, jedoch gelten wieder andere Werte. In der Tabelle 3.7 und in der Tabelle 3.8 sind die kritischen Werte für unterschiedliche Ablehnungsniveaus, unterschiedliche Dimensionen des Prozesses und unterschiedliche deterministische Komponenten dargestellt.

Tabelle 3.8: Kritische Werte für die  $\hat{P}_z$  Teststatistik

no constant					
$n - 1$	10.0 %	7.5%	5.0%	2.5%	1.0%
1	33.9267	36.6646	40.8217	47.2452	55.1911
2	62.1436	65.6162	71.2751	79.5177	89.6679
3	99.2664	103.8454	109.7426	119.3793	131.5716
4	143.0775	148.4109	155.8019	166.3516	180.4845
5	195.6202	201.9621	210.2910	224.0976	237.7723
constant					
$n - 1$	10.0 %	7.5%	5.0%	2.5%	1.0%
1	47.5877	50.7511	55.2202	61.4556	71.9273
2	80.2034	84.4027	89.7619	97.8734	109.4525
3	120.3035	125.4579	132.2207	142.5992	153.4504
4	168.8572	174.2575	182.0749	194.7555	209.8054
5	225.2303	232.4652	241.3316	255.5091	270.5018
trend					
$n - 1$	10.0 %	7.5%	5.0%	2.5%	1.0%
1	71.9586	75.7349	81.3812	90.2944	102.0167
2	113.4929	118.3710	124.3933	133.6963	145.8644
3	163.1050	168.7736	175.9902	188.1265	201.0905
4	219.5098	225.6645	234.2865	247.3640	264.4988
5	284.0100	291.2705	301.0949	315.2799	335.9054

Quelle: [PHILLIPS AND OULIARIS 1990]

### 3.6. Prognose in mehrdimensionalen Systemen

Für die Berechnung der Prognose wird ein VAR mit der Ordnung  $p$  betrachtet

$$\mathbf{y}_t = \Phi_1 \mathbf{y}_{t-1} + \dots + \Phi_p \mathbf{y}_{t-p} + \kappa D_t + \varepsilon_t \quad (3.31)$$

Die optimale  $h$ -Schritt Prognose ab dem Zeitpunkt  $t$  mit minimalen MSE ist durch den bedingten Erwartungswert  $E_t[\mathbf{y}_{t+h}] := E[\mathbf{y}_{t+h} | \{y_s | s < t\}]$  gegeben und unter der Annahme, dass  $\varepsilon_t \sim NID(0, \Sigma)$  ist, gleich

$$\mathbf{y}_t(h) = \Phi_1 \mathbf{y}_t(h-1) + \dots + \Phi_p \mathbf{y}_t(h-1) + \kappa D_t \quad (3.32)$$

wobei  $\mathbf{y}_t(j) := \mathbf{y}_{t+j}$  für  $j \leq 0$  ist. Durch die Einführung der Variablen

$$Y_t = \begin{pmatrix} \mathbf{y}_t \\ \vdots \\ \mathbf{y}_{t-p-1} \end{pmatrix}, \quad \nu_t = \begin{pmatrix} \kappa D_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_{p-1} & \Phi_p \\ I_n & 0 & \dots & 0 & 0 \\ 0 & I_n & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I_n & 0 \end{pmatrix}, \quad \Xi_t = \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

lässt sich ein VAR( $p$ ) Prozess als ein VAR(1) Prozess

$$Y_t = \nu_t + AY_{t-1} + \Xi_t$$

schreiben. Mit Hilfe der zwei Beziehungen

$$\begin{aligned} Y_t(h) &= \nu_{t+h} + AY_t(h-1) = \nu_{t+h} + A^h Y_t \\ Y_{t+h} &= \nu_{t+h} + AY_{t+h-1} + \Xi_t \\ &= \nu_{t+h} + A^h Y_t + \Xi_{t+h} + A\Xi_{t+h-1} + A^{h-1}\Xi_{t+1} \end{aligned}$$

ist der Prognosefehler gleich

$$Y_{t+h} - Y_t(h) = \Xi_{t+h} + A\Xi_{t+h-1} + \dots + A^{h-1}\Xi_{t+1}.$$

Multipliziert man das Ergebnis von links mit der Matrix  $J := [I_n, 0, \dots, 0]$  und verwendet man die Beziehung  $J'J\Xi_t = \Xi_t$ , dann ist

$$\begin{aligned} y_{t+h} - y_t(h) &= J\Xi_{t+h} + JAJ'J\Xi_{t+h-1} + \dots + JA^{h-1}J'J\Xi_{t+1} \\ &= \varepsilon_{t+h} + H_1\varepsilon_{t+h-1} + \dots + H_{h-1}\varepsilon_{t+1} \end{aligned}$$

wobei  $H_i = JA^iJ'$ . Die  $H_i$  lassen sich sehr einfach mit folgender Rekursion berechnen:

$$H_i = \sum_{j=1}^i H_{j-i}\Phi_j \quad i = 1, 2, \dots \quad (3.33)$$

mit  $H_0 = I_n$ . Daraus ergibt sich eine Matrix der mittleren quadratischen Prognosefehler (MSE-Matrix)

$$\Sigma_y(h) = \sum_{i=0}^{h-1} H_i \Sigma_\varepsilon H_i'$$

Im Konvergenzverhalten von  $\Sigma_y(h)$  muss man nun erstmals zwischen einem VAR Prozess, der die Bedingung  $\det \Phi(\mathbf{z}) \neq 0 \quad \forall |\mathbf{z}| \leq 1$  erfüllt, und einem integrierten Prozess unterscheiden. Im ersten Fall konvergieren die Matrizen  $H_i$  gegen Null für  $i \rightarrow \infty$  und  $\Sigma_y(h)$  konvergiert gegen die Kovarianzmatrix von  $y_t$  für  $h \rightarrow \infty$ . Im kointegrierten Fall konvergieren die  $H_i$  nicht gegen Null und daher werden einige Elemente von  $\Sigma_y(h)$  beliebig groß werden.

In der Praxis müssen natürlich die Koeffizienten von (3.31) geschätzt werden. Dadurch erhält man den Schätzer für die  $h$ -Schritt Prognose  $\hat{y}_t(h)$  indem man in (3.32) die Koeffizienten durch deren Schätzer ersetzt. Der Prognosefehler ist gleich

$$\begin{aligned} \mathbf{y}_{t+h} - \hat{\mathbf{y}}_t(h) &= [\mathbf{y}_{t+h} - \mathbf{y}_t(h)] + [\mathbf{y}_t(h) - \hat{\mathbf{y}}_t(h)] \\ &= \sum_{i=0}^{h-1} H_i \varepsilon_{t+h-i} + [\mathbf{y}_t(h) - \hat{\mathbf{y}}_t(h)] \end{aligned}$$

wobei die letzten beiden Terme unkorreliert sind. Sind die Schätzer für die Parameter des VAR Prozesses konsistent, dann konvergiert der Term  $[\mathbf{y}_t(h) - \hat{\mathbf{y}}_t(h)]$  für  $h \rightarrow \infty$  gegen Null. Ein möglicher Schätzer der MSE-Matrix ist

$$\hat{\Sigma}_y(h) = \sum_{i=0}^{h-1} \hat{H}_i \hat{\Sigma}_\varepsilon \hat{H}_i' \quad (3.34)$$

wobei man die  $H_i$  aus den geschätzten  $\Phi_i$  mit der Rekursion (3.33) erhält.

# Kapitel 4.

## Eine empirische Untersuchung

### 4.1. Daten

Das Problem bei der Datenauswahl für die Kointegrationsanalyse zwischen Rohstoff- und Aktienmärkten besteht darin, Unternehmen zu finden, die nur von wenigen Rohstoffen abhängig sind, das bedeutet, dass die Unternehmensidee bzw. das Unternehmensgeschäft im direkten Zusammenhang mit den Rohstoffen stehen sollte. Indirekt ist zwar jedes produzierende Gewerbe von irgendwelchen Rohstoffentwicklungen abhängig aber dennoch nicht für die Untersuchungen geeignet. Nach reifen Überlegungen wurden schließlich folgende zwei Datensets ausgewählt:

1. Die erste Datenauswahl umfasst Unternehmen aus der Ölbranche und natürlich den Rohstoff Öl. Im Speziellen werden die großen erdölfördernden Unternehmen aus diesem Bereich ausgewählt, von welchen angenommen wird, dass Öl die einzig wichtige Rohstoffkomponente im Unternehmen darstellt.

#### Aktien Datenset 1:

- Exxon Mobil Corporation (XOM), BK<sup>1</sup>: 325 Mrd. \$
- BP Plc (BP), BK: 221.1 Mrd. \$
- Total SA (TOT), BK: 134.3 Mrd. \$
- Chevron Texaco Corporation (CVX), BK: 114.2 Mrd. \$
- Royal Dutch Corporation (RD), BK: 112.9 Mrd. \$
- Shell (SC), BK: 77.6 Mrd. \$

Die Auswahl umfasst die fünf größten Erdölfirmer der Welt, gereiht nach der Börsenkapitalisierung (Stand: 22.11.2004) <sup>2</sup> und außerdem das Unternehmen Shell, die Nummer acht nach der Börsenkapitalisierung. Die Abkürzungen in den Klammern, die aufgrund der einfacheren Handhabung in den meisten Tabellen verwendet werden, bezeichnen das Börsenkurzzeichen des jeweiligen Unternehmens.

---

<sup>1</sup>BK steht für Börsenkapitalisierung

<sup>2</sup>Quelle: <http://biz.yahoo.com/ic/oilint.html>

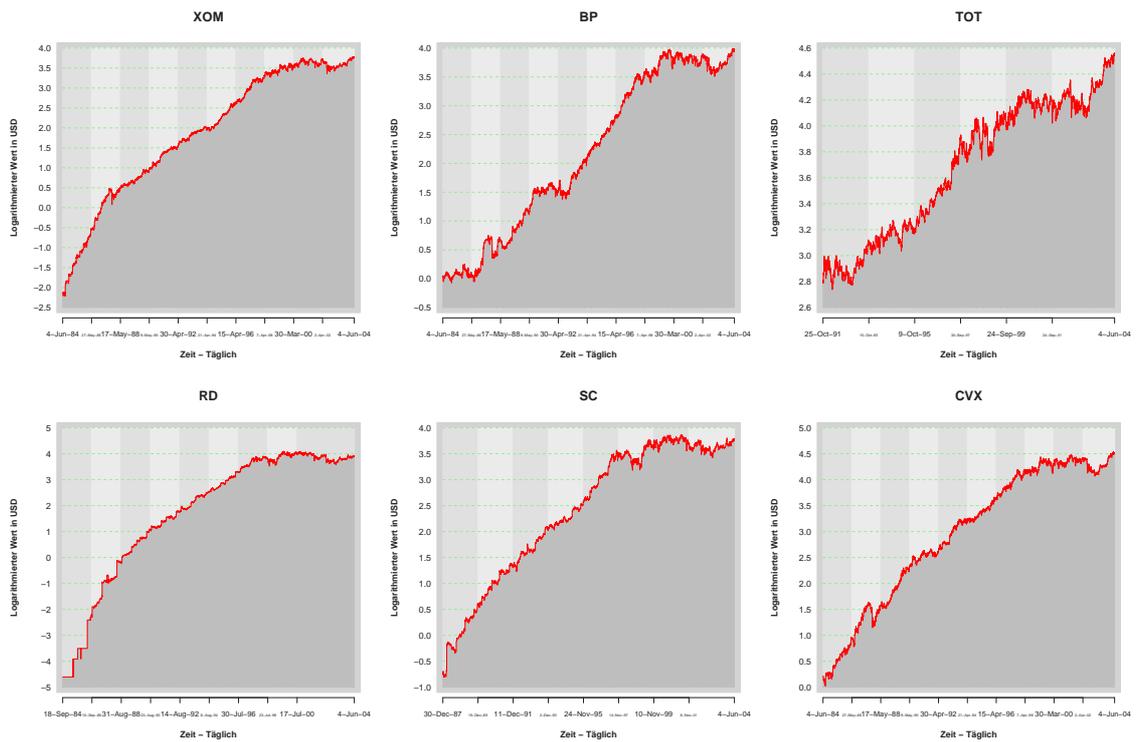


Abbildung 4.1: Aktien Datenset 1

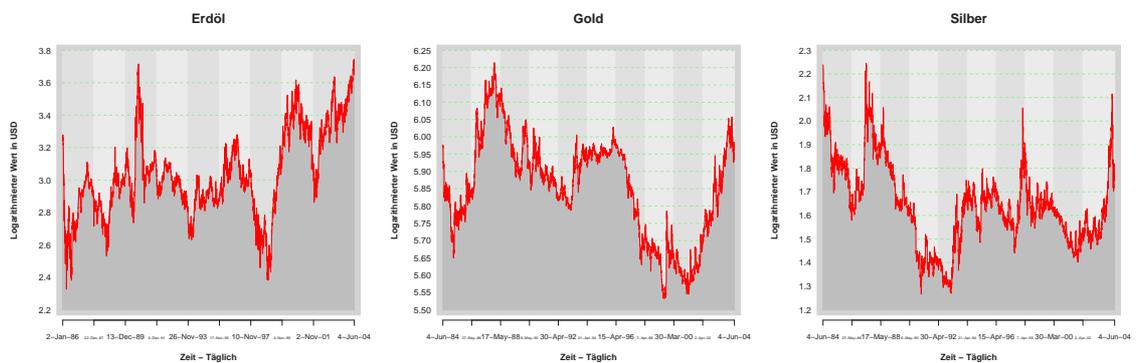


Abbildung 4.2: Rohstoffe

- Die zweite Datenzusammenstellung ist leider nicht mehr so einfach und klar. Ausgewählt wurden Bergbauunternehmen, die hauptsächlich Edelmetalle fördern, wobei auch hier wieder angenommen wird, dass diese Unternehmen im Wesentlichen nur von Gold und Silber abhängig sind.

### Aktien Datenset 2:

- Anglo American PLC (AAUK), BK: 35.6 Mrd. \$
- Newmont Mining Corporation (NEM), BK: 21.8 Mrd. \$
- Barrick Gold Corp (ABX), BK: 13.12 Mrd. \$
- AngloGold Ashanti Ltd (AU), BK:10.6 Mrd. \$
- Placer Dome Inc (PDG), BK: 9.2 Mrd. \$

Auch in dieser Auswahl wurden die fünf größten Gesellschaften ihrer Branche, erneut gemessen an der Börsenkaptalisierung (Stand: 22.11.2004)<sup>3</sup>, aufgenommen.

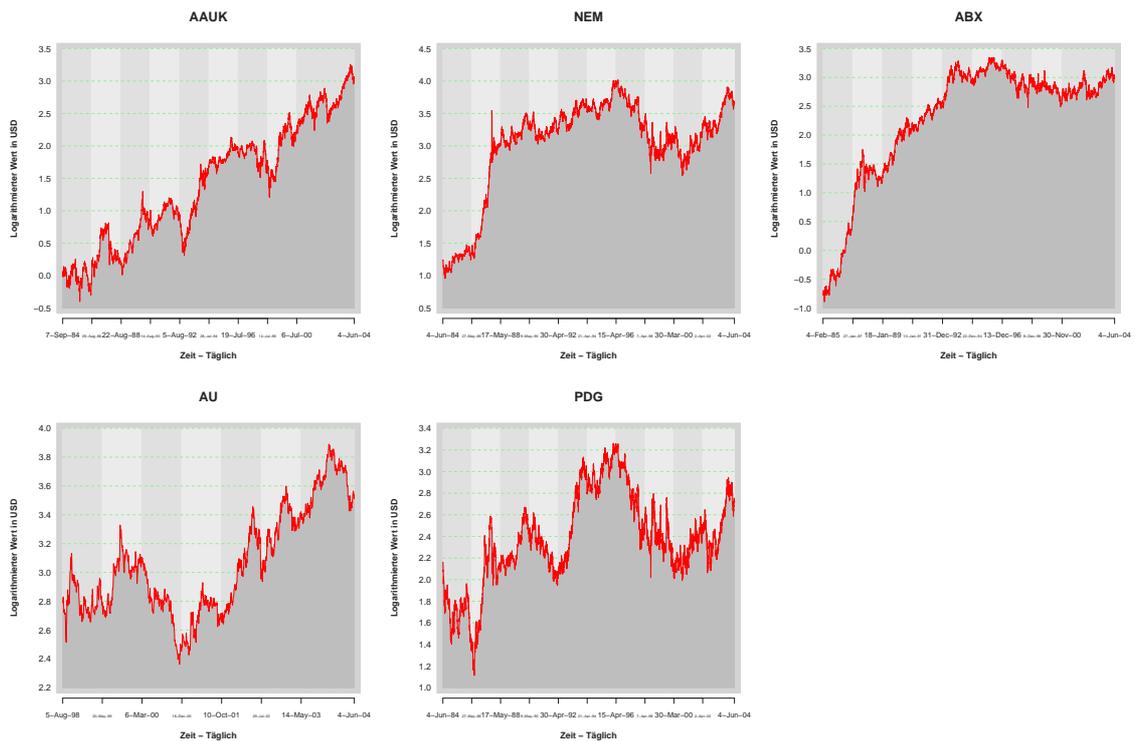


Abbildung 4.3: Aktien Datenset 2

Der verfügbare Datenzeitraum erstreckt sich bei den meisten Aktien und Rohstoffen vom 04.06.1984 bis zum 04.06.2004, also 20 Jahre zurück. Einen genauen Überblick bietet die Tabelle 4.1.

Verwendet werden die täglichen Schlusskurse, obwohl auch Eröffnungskurs und Tagesmittelwert zur Verfügung stehen. Aktiensplits und Aktienzusammenführung sind genauso berücksichtigt wie Dividendenabschläge.

<sup>3</sup>Quelle: <http://biz.yahoo.com/ic/gldslv.html>

Tabelle 4.1: Übersicht

Daten	# Beob.	Zeitraum	Daten	# Beob.	Zeitraum
XOM	5049	04.06.1984 - 04.06.2004	AAUK	4982	07.09.1984 - 04.06.2004
BP	5049	04.06.1984 - 04.06.2004	NEM	5049	04.06.1984 - 04.06.2004
TOT	3179	25.10.1991 - 04.06.2004	ABX	4879	04.02.1985 - 04.06.2004
RD	4975	18.09.1984 - 04.06.2004	AU	1467	05.08.1998 - 04.06.2004
CVX	5049	04.06.1984 - 04.06.2004	PDG	5049	04.06.1984 - 04.06.2004
SC	4146	30.12.1987 - 04.06.2004	Gold	5049	04.06.1984 - 04.06.2004
Öl	4650	02.01.1986 - 04.06.2004	Silber	5049	04.06.1984 - 04.06.2004

Für die Untersuchungen wurden die Kurse logarithmiert, um in erster Linie die Varianz zu stabilisieren.

Durch das Logarithmieren der Daten und durch die geringe Anzahl der Nachkommastellen der Daten kam es bei einigen Zeitreihenanfängen zu sehr großen Sprüngen, wie man in den Abbildungen von RD, SC und XOM sehen kann. Damit außerdem für alle Aktien dieselbe Beobachtungsanzahl verwendet wird, werden die ersten tausend Werte weggelassen und die Aktien TOT und AU aus der Liste gestrichen. Der neue Beobachtungszeitraum erstreckt sich nun von 18.08.1988 bis 04.06.2004 mit 4049 Datenpunkten. Die beiden Datensets enthalten jeweils 6 Zeitreihen.

Die verwendeten Daten stammen aus folgenden Quellen:

**Aktien:** <http://finance.yahoo.com>

**Öl:** [http://www.eia.doe.gov/oil\\_gas/petroleum/info\\_glance/crudeoil.html](http://www.eia.doe.gov/oil_gas/petroleum/info_glance/crudeoil.html)

**Gold & Silber:** <http://www.markt-daten.de/daten/daten.htm>

Sowohl die Aktien als auch die Rohstoffe werden in Dollar gehandelt und müssen nicht extra umgerechnet werden. Ein Problem, das jedoch bei der Datenaufbereitung sichtbar wurde, sind die unterschiedlichen Feiertage in den einzelnen Ländern. Für die Aktien und dem Rohstoff Öl gelten die Bestimmungen der USA, während Gold und Silber in London mit den Feiertagsbestimmungen Großbritanniens notieren. Gelöst wurde dieses Problem, indem der größte gemeinsame Nenner gewählt wurde und die fehlenden Daten durch die letzten verfügbaren Notierungen ergänzt wurden, d.h. falls in der Datenmatrix auch nur für eine Zeitreihe ein aktueller Kurs für ein bestimmtes Datum vorhanden ist, werden alle anderen Zeitreihen ergänzt. Beispielsweise muss der 4.Juli bei allen Aktien und beim Rohstoff Öl mit den Daten vom Vortag gefüllt werden, da am Nationalfeiertag der USA die Börsen geschlossen haben, während in London ein normaler Börsentag ist.

## 4.2. Programm

Für die Analyse der Daten wurde das frei verfügbare Statistikprogramm **R** verwendet.<sup>4</sup>



Im Basispaket sind die wichtigsten Funktionen enthalten, doch reichen diese oft nicht aus um spezielle Problemstellungen programmtechnisch zu lösen. Es besteht daher die Möglichkeit als Programmierer eigene Pakete für ganz bestimmte Aufgaben zu erstellen.

In dieser Arbeit wurden die folgenden Zusatzpakete für die Auswertung der Daten und die Erstellung der Beispiele verwendet:

**tseries** allgemeines Paket für die Zeitreihenanalyse

**urca** ergänzt die UR Tests des vorherigen Pakets und enthält außerdem die Kointegrationstests: Johansen und Phillips-Ouliaris

**oekfinm** von diesem Paket wird die sehr umfangreiche Johansen Prozedur verwendet, die einige Vorteile gegenüber der Johansenfunktion aus dem Paket *urca* besitzt

Für alle benötigten Funktionen, die noch nicht in **R** oder in einem der Zusatzpakete vorhanden sind, ist der „source code“ im Anhang angegeben.

## 4.3. Modellbildung

Bevor man zur Kointegrationsanalyse der beiden Datensets schreiten kann, sind noch einige Voruntersuchungen notwendig.

Im Kapitel „Kointegration“ wurden zwei, leicht unterschiedliche, Definitionen von Kointegration gegeben. Die Originalversion setzt voraus, dass alle Zeitreihen mit derselben Ordnung integriert sind (in unserem Fall  $I(1)$ ), während die zweite Definition auch stationäre Prozesse zulässt. Obwohl für die Kointegrationsuntersuchungen nur der Johansen Test verwendet wird (Definition 2 ist für diesen Test ausreichend), ist es trotzdem von Interesse die Integrationsordnung der einzelnen Prozesse zu wissen. Es werden daher die folgenden „unit root“-Tests durchgeführt:<sup>5</sup>

**Augmented Dickey Fuller (ADF):** Bei diesem Test wird die Nullhypothese: „random walk“ mit Drift getestet, wobei das geschätzte Modell die Gestalt  $y_t = \nu + \phi y_{t-1} + \delta t + \sum_{j=1}^{p+1} \zeta_j \Delta y_{t-j} + \xi_t$  besitzt. Die maximale Laglänge  $p$  ist abhängig

<sup>4</sup>Das Programm kann unter der im Literaturverzeichnis angegebenen Webadresse gratis und ohne Anmeldung bezogen werden.

<sup>5</sup>Eine ausführliche Zusammenfassung der einzelnen Tests gibt es im Anhangkapitel „UR - Integrationstests“

von der Anzahl der Beobachtungen und wird mit der in der Literatur üblichen Formel  $12\sqrt[4]{T/100}$  berechnet.

Tabelle 4.2: Wahrscheinlichkeitswerte für die Annahme der Null-Hypothese des ADF-Tests

lag	Erdöl	XOM	BP	CVX	RD	SC
1	0.0529	0.9035	0.8499	0.6438	0.9495	0.9028
2	0.0879	0.9289	0.8785	0.6726	0.9502	0.8990
3	0.1905	0.9435	0.9023	0.6901	0.9481	0.9050
4	0.1880	0.9554	0.9151	0.7080	0.9510	0.9083
5	0.2349	0.9632	0.9221	0.7193	0.9516	0.9087
6	0.2863	0.9691	0.9330	0.7526	0.9513	0.9192
7	0.2706	0.9709	0.9315	0.7533	0.9506	0.9189
8	0.3339	0.9748	0.9314	0.7768	0.9534	0.9177
9	0.3373	0.9742	0.9339	0.7895	0.9567	0.9271
10	0.3459	0.9723	0.9437	0.7874	0.9568	0.9268
11	0.3268	0.9751	0.9505	0.8035	0.9573	0.9278
12	0.3632	0.9751	0.9556	0.8206	0.9604	0.9327
13	0.3670	0.9790	0.9490	0.8301	0.9568	0.9285
14	0.3642	0.9817	0.9485	0.8509	0.9602	0.9333
15	0.3278	0.9797	0.9487	0.8492	0.9619	0.9391
16	0.3189	0.9782	0.9494	0.8610	0.9648	0.9460
17	0.2805	0.9806	0.9527	0.8467	0.9636	0.9435
18	0.2398	0.9841	0.9515	0.8537	0.9669	0.9490
19	0.2406	0.9829	0.9524	0.8405	0.9657	0.9467
20	0.2575	0.9831	0.9545	0.8270	0.9641	0.9405
21	0.2172	0.9832	0.9520	0.8273	0.9642	0.9414
22	0.2117	0.9827	0.9589	0.8255	0.9651	0.9423
23	0.1996	0.9868	0.9624	0.8314	0.9670	0.9441
24	0.1421	0.9877	0.9553	0.8342	0.9634	0.9368
25	0.0999	0.9883	0.9570	0.8347	0.9621	0.9276
26	0.1014	0.9900	0.9589	0.8217	0.9632	0.9269
27	0.1401	0.9892	0.9566	0.8113	0.9606	0.9221
28	0.1896	0.9900	0.9588	0.8265	0.9633	0.9249
29	0.1647	0.9900	0.9568	0.8117	0.9627	0.9268
30	0.1990	0.9900	0.9518	0.7742	0.9595	0.9211

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot). Die Wahrscheinlichkeitswerte werden im Intervall (0.01, 0.99) genau tabelliert, Werte die außerhalb dieses Bereiches liegen werden mit der Intervallgrenze angegeben.

**Kwiatkowski, Phillips, Schmidt und Shin (KPSS):** Beim KPSS wird im Gegensatz zu den anderen Test die Nullhypothese: Level-Stationarität und Trend-Stationarität getestet, die alternative Hypothese ist ein integrierter Prozess mit der Ordnung Eins. Die Anzahl der Lags geht beim KPSS und PP-Test in die Gewichtsfunktion  $w_{s,l}$  ein und beträgt wieder  $12\sqrt[4]{T/100}$ .

**Phillips Perron (PP):** Der PP-Test wird sowohl für die Nullhypothese: „random walk“ als auch für die Nullhypothese: „random walk“ mit Drift durchgeführt; die jeweils geschätzten Regressionen lauten  $y_t = \nu + \phi y_{t-1} + \xi_t$  und  $y_t =$

Tabelle 4.3: Wahrscheinlichkeitswerte für die Annahme der Null-Hypothese des ADF-Tests

lag	Gold	Silber	AAUK	NEM	ABX	PDG
1	0.6766	0.0484	0.0840	0.3145	0.4515	0.2888
2	0.6688	0.0511	0.0686	0.3650	0.4749	0.3681
3	0.6468	0.0552	0.0696	0.3993	0.5239	0.3764
4	0.6105	0.0592	0.0793	0.3635	0.5279	0.3663
5	0.6090	0.0551	0.0856	0.3436	0.5150	0.3644
6	0.6272	0.0488	0.0997	0.3342	0.5211	0.3875
7	0.6410	0.0458	0.1290	0.3542	0.5218	0.3681
8	0.6229	0.0567	0.1501	0.3879	0.5440	0.4066
9	0.6259	0.0613	0.1341	0.3775	0.5471	0.4111
10	0.6184	0.0515	0.1432	0.3683	0.5621	0.3968
11	0.6085	0.0277	0.1386	0.4126	0.5915	0.4260
12	0.6436	0.0316	0.1281	0.4061	0.5953	0.4291
13	0.6793	0.0400	0.1538	0.4130	0.6029	0.4444
14	0.6844	0.0466	0.1290	0.4238	0.6047	0.4668
15	0.6980	0.0665	0.1340	0.4311	0.6176	0.4823
16	0.7025	0.0482	0.1067	0.4435	0.6051	0.4769
17	0.7178	0.0633	0.1085	0.4430	0.6067	0.4659
18	0.7237	0.0745	0.1194	0.4333	0.6062	0.4522
19	0.6898	0.0672	0.1281	0.4513	0.6157	0.4886
20	0.6623	0.0444	0.0992	0.4447	0.6119	0.4784
21	0.6716	0.0460	0.0972	0.4311	0.5991	0.4657
22	0.6442	0.0455	0.1155	0.4645	0.6160	0.4924
23	0.6808	0.0481	0.1159	0.4552	0.6127	0.4968
24	0.6818	0.0551	0.0860	0.4405	0.6093	0.5027
25	0.7283	0.0617	0.1200	0.4633	0.6075	0.4970
26	0.7610	0.0738	0.1181	0.4959	0.6095	0.5205
27	0.7793	0.1143	0.1033	0.5144	0.6185	0.5232
28	0.8103	0.1126	0.1206	0.5245	0.6128	0.5179
29	0.8402	0.1632	0.1323	0.5057	0.5980	0.5037
30	0.8272	0.1681	0.1273	0.5180	0.5885	0.5190

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot). Die Wahrscheinlichkeitswerte werden im Intervall (0.01, 0.99) genau tabelliert, Werte die außerhalb dieses Bereiches liegen werden mit der Intervallgrenze angegeben.

$$\nu + \delta t + \phi y_{t-1} + \xi_t.$$

Wie die Tabellen 4.2 und 4.3 zeigen wird, bis auf Silber, die Nullhypothese des ADF-Tests von allen anderen Aktien oder Rohstoffen angenommen, das bedeutet, dass Silber ein trendstationärer Prozess ist. Für alle anderen Zeitreihen wird angenommen, dass sie integriert mit der Ordnung Eins sind. Es besteht natürlich die Möglichkeit, dass der eine oder andere Prozess mit einer höheren Ordnung integriert ist. Zahlreiche empirische Untersuchungen zeigen jedoch, dass Finanzdaten im Allgemeinen einem  $I(1)$  Prozess folgen, darum werden hier keine Untersuchungen in diese Richtung angestellt.

Die Ergebnisse der anderen beiden Tests sind im Anhang dargestellt. Der KPSS-Test lehnt für alle Zeitreihen die Nullhypothese ab, das bedeutet, dass alle Zeitreihen nicht stationär sind. Dieses Resultat steht im Widerspruch zum ADF-Test und auch zum PP-Test, der bezüglich Silber, ein analoges Ergebnis liefert. Silber wird daher zumindest als trendstationär angesehen. Bei der Aktie AAUK zeigt der PP-Test gerade noch Trendstationarität an, jedoch wird hier der ADF-Test stärker bewertet und AAUK somit als „random walk“ mit Drift angesehen. Interessant sind die beiden Aktien RD und SC: Der ADF-Test zeigt hier ganz klar integrierte Prozesse an, während der PP-Test einerseits dieses Ergebnis bestätigt, andererseits aber auch ganz eindeutig ablehnt (dies hängt davon ab, ob ein Trend in die Regression aufgenommen wird). Auch in diesem Fall erhält das Ergebnis des ADF-Tests (nichtstationäre Prozesse) den Vorzug, trotzdem sollte man die Unstimmigkeiten der einzelnen Tests für die Kointegrationsanalyse im Hinterkopf behalten.

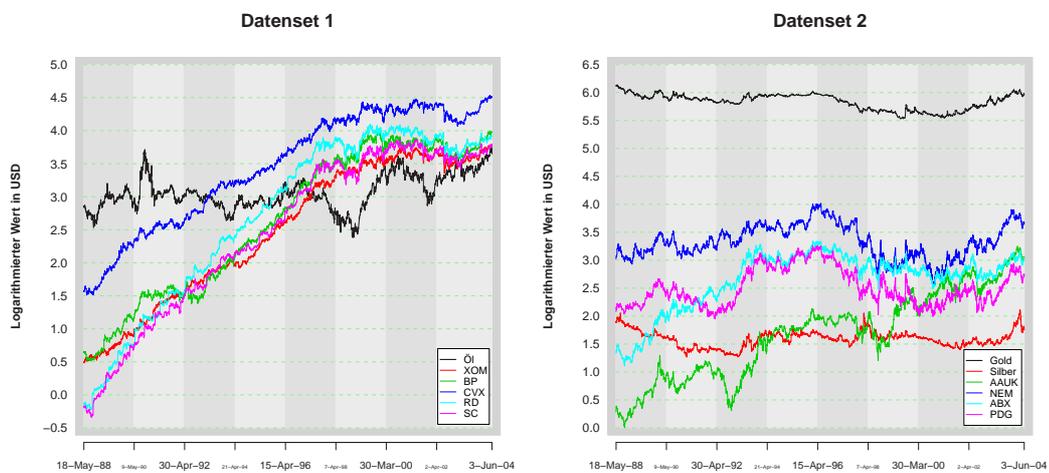


Abbildung 4.4: Gesamtansicht

In Abbildung 4.4 sind die beiden Datenzusammenstellungen gemeinsam geplottet. Man kann dabei erkennen, dass die Öllaktien zuerst gemeinsam fast linear ansteigen und über den gesamten Zeitraum ein sehr ähnliches Verhalten aufweisen. Die Ursache dafür könnte möglicherweise ein gemeinsamer stochastischer Prozess sein. Bei den Minengesellschaften ist ein solches Verhalten nicht zu beobachten. Auffällig ist in beiden Graphen jedoch der völlig unterschiedliche Verlauf der Rohstoffe verglichen mit den Aktien.

Zusammenfassend kann man festhalten, dass bei Betrachtung des gesamten Datenbereiches (4049 Beobachtungen) der Rohstoff Silber nicht integriert ist, während alle anderen Zeitreihen mehr oder weniger eindeutig nichtstationär sind.

Die zweite wichtige Voruntersuchung ist die Bestimmung der Ordnung der VAR Prozesse der beiden Datensätze. Für die Ermittlung werden die Informationskriterien AIC, SC und HQ für alle vier Modellklassen berechnet. Die maximal mögliche Modellgröße wird mit 30 festgesetzt.

Tabelle 4.4: Informationskriterien

max. Ordnung des VAR Prozesses: 30									
	Set 1	AIC	SC	HQ	Set 2	AIC	SC	HQ	
$r = 0$	$\mathbf{M}^*$ :	4	1	4	$\mathbf{M}^*$ :	3	2	3	
$r = 1$		4	1	4		3	2	3	
$r = 2$		4	1	4		3	2	3	
$r = 3$		4	1	4		3	2	3	
$r = 4$		4	1	4		3	2	3	
$r = 5$		4	1	4		3	2	3	
$r = 6$		4	1	4		3	2	3	
$r = 0$	$\mathbf{M}_1$ :	4	1	4	$\mathbf{M}_1$ :	3	2	3	
$r = 1$		4	1	4		3	2	3	
$r = 2$		4	1	4		3	2	3	
$r = 3$		4	1	4		3	2	3	
$r = 4$		4	1	4		3	2	3	
$r = 5$		4	1	4		3	2	3	
$r = 6$		4	1	4		3	2	3	
$r = 0$	$\mathbf{M}_1^*$ :	4	1	4	$\mathbf{M}_1^*$ :	3	2	3	
$r = 1$		4	1	4		3	2	3	
$r = 2$		4	1	4		3	2	3	
$r = 3$		4	1	4		3	2	3	
$r = 4$		4	1	4		3	2	3	
$r = 5$		4	1	4		3	2	3	
$r = 6$		4	1	4		3	2	3	
$r = 0$	$\mathbf{M}_2$ :	4	1	4	$\mathbf{M}_2$ :	3	2	3	
$r = 1$		4	1	4		3	2	3	
$r = 2$		4	1	4		3	2	3	
$r = 3$		4	1	4		3	2	3	
$r = 4$		4	1	4		3	2	3	
$r = 5$		4	1	4		3	2	3	
$r = 6$		4	1	4		3	2	3	

Wie man in Tabelle 4.4 sehen kann, haben der Kointegrationsrang und die Wahl des Modells keinen Einfluss auf die Größe des VAR-Prozesses. Obwohl hier die Informationskriterien besser harmonieren als im Beispiel des Kapitels „mehrdimensionale Prozesse“, werden die Kointegrationsanalysen trotzdem für fünf verschiedene Modellgrößen durchgeführt, um nicht im Vorhinein durch eine falsche Festlegung die Ergebnisse zu verfälschen. Für die Interpretation der Kointegrationstabellen sollten jedoch die Ordnungsgrößen 1,2 und 4 am meisten Gewicht bekommen.

## 4.4. Auswertung

Für die Untersuchungen der Kointegrationsbeziehungen wurden der Johantest und der  $\hat{P}_z$ -Test von Phillips & Ouliaris verwendet. Der Vorteil dieser beiden Tests ist die Unabhängigkeit gegenüber der Normalisierung einer Kointegrationsregres-

Tabelle 4.5: Teststatistik und kritische Werte des Johansen - Kointegrationstests

<b>Daten:</b> Erdöl, XOM, BP, CVX, RD, SC									
<b>Beobachtungszeitraum:</b> 18.05.1988 - 04.06.2004									
<b>Anzahl der Beobachtungen:</b> 4049									
<b>verwendetes Modell:</b> $M^*$									
$p$	Teststatistik (trace)					Kritische Werte			
	1	2	5	9	16	10%	5%	2.5%	1%
$r \leq 5$	7.7329	8.4300	9.3058	10.4530	9.3436	10.56	12.39	14.13	16.39
$r \leq 4$	18.2452	19.1174	19.9290	22.3695	21.1120	22.95	25.47	27.87	30.65
$r \leq 3$	34.1212	36.9086	35.9183	<b>39.3202</b>	37.5390	39.08	42.20	45.04	48.59
$r \leq 2$	57.4383	58.8958	55.2604	57.1935	56.5325	58.96	62.61	66.04	70.22
$r \leq 1$	<b>91.3219</b>	<b>90.4215</b>	<b>83.0133</b>	80.6946	75.8860	82.68	86.96	90.87	95.38
$r = 0$	<b>168.2741</b>	<b>163.4816</b>	<b>147.9197</b>	<b>134.5195</b>	<b>122.0525</b>	110.00	114.96	119.43	124.61
<b>Daten:</b> Erdöl, XOM, BP, CVX, RD, SC									
<b>Beobachtungszeitraum:</b> 18.05.1988 - 04.06.2004									
<b>Anzahl der Beobachtungen:</b> 4049									
<b>verwendetes Modell:</b> $M_1$									
$p$	Teststatistik (trace)					Kritische Werte			
	1	2	5	9	16	10%	5%	2.5%	1%
$r \leq 5$	<b>5.9354</b>	<b>6.7219</b>	<b>6.9201</b>	<b>5.9814</b>	<b>5.4794</b>	2.71	3.84	5.02	6.64
$r \leq 4$	<b>15.0620</b>	<b>16.1264</b>	<b>16.7741</b>	<b>16.4346</b>	<b>15.2623</b>	13.31	15.34	17.24	19.69
$r \leq 3$	<b>27.1633</b>	<b>28.1797</b>	<b>27.4717</b>	<b>28.4393</b>	<b>27.9836</b>	26.70	29.38	31.76	34.87
$r \leq 2$	<b>44.0766</b>	<b>46.4420</b>	43.4968	<b>45.4073</b>	<b>44.6495</b>	43.84	47.21	50.19	53.91
$r \leq 1$	<b>73.5664</b>	<b>72.9368</b>	<b>67.8960</b>	<b>68.1160</b>	<b>64.0023</b>	64.74	68.68	72.21	76.37
$r = 0$	<b>146.8199</b>	<b>142.5785</b>	<b>131.3594</b>	<b>121.3960</b>	<b>109.8164</b>	89.37	93.92	97.97	102.95
<b>Daten:</b> Erdöl, XOM, BP, CVX, RD, SC									
<b>Beobachtungszeitraum:</b> 18.05.1988 - 04.06.2004									
<b>Anzahl der Beobachtungen:</b> 4049									
<b>verwendetes Modell:</b> $M_1^*$									
$p$	Teststatistik (trace)					Kritische Werte			
	1	2	5	9	16	10%	5%	2.5%	1%
$r \leq 5$	<b>7.6580</b>	<b>7.8885</b>	<b>7.8681</b>	6.9739	7.4006	7.50	9.13	10.73	12.73
$r \leq 4$	<b>19.1963</b>	<b>19.2186</b>	<b>18.0149</b>	17.4300	17.2666	17.79	19.99	22.07	24.74
$r \leq 3$	<b>35.4534</b>	<b>37.1429</b>	<b>33.8776</b>	<b>34.3460</b>	<b>32.7634</b>	31.88	34.80	37.56	40.84
$r \leq 2$	<b>63.2893</b>	<b>63.3192</b>	<b>58.2069</b>	<b>57.0128</b>	<b>51.5822</b>	49.92	53.42	56.57	60.42
$r \leq 1$	<b>93.6147</b>	<b>93.8800</b>	<b>93.2512</b>	<b>98.1175</b>	<b>91.6614</b>	71.66	75.74	79.60	83.93
$r = 0$	<b>166.8690</b>	<b>163.6754</b>	<b>156.7458</b>	<b>151.6465</b>	<b>140.4105</b>	97.17	101.84	106.07	111.38
<b>Daten:</b> Erdöl, XOM, BP, CVX, RD, SC									
<b>Beobachtungszeitraum:</b> 18.05.1988 - 04.06.2004									
<b>Anzahl der Beobachtungen:</b> 4049									
<b>verwendetes Modell:</b> $M_2$									
$p$	Teststatistik (trace)					Kritische Werte			
	1	2	5	9	16	10%	5%	2.5%	1%
$r \leq 5$	0.2996	0.2519	0.3047	0.5733	0.9168	2.98	4.14	5.30	7.02
$r \leq 4$	8.8601	8.8729	9.6832	<b>10.6161</b>	<b>10.5963</b>	10.35	12.21	13.94	16.16
$r \leq 3$	<b>21.6580</b>	<b>21.9806</b>	21.3468	<b>22.6339</b>	<b>22.7964</b>	21.58	24.08	26.42	29.19
$r \leq 2$	<b>42.3379</b>	<b>43.6571</b>	<b>40.6294</b>	<b>42.3740</b>	<b>41.5014</b>	36.58	39.71	42.59	46.00
$r \leq 1$	<b>70.8102</b>	<b>73.9368</b>	<b>75.2062</b>	<b>74.2017</b>	<b>66.0155</b>	55.54	59.24	62.68	66.71
$r = 0$	<b>112.6724</b>	<b>111.8265</b>	<b>112.2009</b>	<b>116.0640</b>	<b>109.6628</b>	75.48	82.61	86.36	91.12

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot).

sion, die bei einigen anderen Tests benötigt wird. Für mehrdimensionale Systeme gibt es beim Phillips Ouliaris Test aber keine konkrete Aussage über die Anzahl der

Tabelle 4.6: Teststatistik und kritische Werte des Johansen - Kointegrationstests

<b>Daten:</b> Gold, Silber, AAUK, NEM, ABX, PDG									
<b>Beobachtungszeitraum:</b> 18.05.1988 - 04.06.2004									
<b>Anzahl der Beobachtungen:</b> 4049									
<b>verwendetes Modell:</b> $M^*$									
$p$	Teststatistik (trace)					Kritische Werte			
	1	2	5	9	16	10%	5%	2.5%	1%
$r \leq 5$	3.8817	3.4257	3.8978	4.1498	4.0786	10.56	12.39	14.13	16.39
$r \leq 4$	12.0549	11.8504	11.2332	11.8070	12.7917	22.95	25.47	27.87	30.65
$r \leq 3$	25.2824	26.8913	25.7760	25.2602	25.1885	39.08	42.20	45.04	48.59
$r \leq 2$	45.4600	49.1564	48.8046	47.6326	48.4963	58.96	62.61	66.04	70.22
$r \leq 1$	73.7923	75.4158	74.2790	73.6502	74.9453	82.68	86.96	90.87	95.38
$r = 0$	149.9640	115.2770	110.6582	109.1974	110.2839	110.00	114.96	119.43	124.61
<b>Daten:</b> Gold, Silber, AAUK, NEM, ABX, PDG									
<b>Beobachtungszeitraum:</b> 18.05.1988 - 04.06.2004									
<b>Anzahl der Beobachtungen:</b> 4049									
<b>verwendetes Modell:</b> $M_1$									
$p$	Teststatistik (trace)					Kritische Werte			
	1	2	5	9	16	10%	5%	2.5%	1%
$r \leq 5$	3.5719	3.3261	3.8494	4.1056	3.6506	2.71	3.84	5.02	6.64
$r \leq 4$	8.0502	7.9329	8.8658	8.6017	8.0965	13.31	15.34	17.24	19.69
$r \leq 3$	17.2473	17.3989	17.3408	17.7388	17.5732	26.70	29.38	31.76	34.87
$r \leq 2$	30.5331	32.4728	31.9895	31.2579	31.4246	43.84	47.21	50.19	53.91
$r \leq 1$	58.2252	57.7073	55.4605	54.2305	54.8326	64.74	68.68	72.21	76.37
$r = 0$	130.0208	95.0101	88.0593	86.7969	88.5718	89.37	93.92	97.97	102.95
<b>Daten:</b> Gold, Silber, AAUK, NEM, ABX, PDG									
<b>Beobachtungszeitraum:</b> 18.05.1988 - 04.06.2004									
<b>Anzahl der Beobachtungen:</b> 4049									
<b>verwendetes Modell:</b> $M_1^*$									
$p$	Teststatistik (trace)					Kritische Werte			
	1	2	5	9	16	10%	5%	2.5%	1%
$r \leq 5$	3.6871	3.3523	3.9029	4.1060	3.8205	7.50	9.13	10.73	12.73
$r \leq 4$	8.8989	8.4402	9.1420	8.7772	8.3662	17.79	19.99	22.07	24.74
$r \leq 3$	20.4778	19.5614	19.1177	19.3455	18.4483	31.88	34.80	37.56	40.84
$r \leq 2$	34.9158	37.1828	36.4210	36.7629	37.7286	49.92	53.42	56.57	60.42
$r \leq 1$	63.8623	63.5269	61.5082	60.4317	61.1992	71.66	75.74	79.60	83.93
$r = 0$	135.6584	101.1322	94.2648	93.7224	96.0699	97.17	101.84	106.07	111.38
<b>Daten:</b> Gold, Silber, AAUK, NEM, ABX, PDG									
<b>Beobachtungszeitraum:</b> 18.05.1988 - 04.06.2004									
<b>Anzahl der Beobachtungen:</b> 4049									
<b>verwendetes Modell:</b> $M_2$									
$p$	Teststatistik (trace)					Kritische Werte			
	1	2	5	9	16	10%	5%	2.5%	1%
$r \leq 5$	0.2096	0.2118	0.2029	0.2392	0.3011	2.98	4.14	5.30	7.02
$r \leq 4$	4.9604	5.2877	5.4237	4.8568	4.5470	10.35	12.21	13.94	16.16
$r \leq 3$	13.9647	13.0654	13.1626	12.0018	11.6028	21.58	24.08	26.42	29.19
$r \leq 2$	26.8943	25.5357	25.0449	24.6423	24.7542	36.58	39.71	42.59	46.00
$r \leq 1$	49.7294	45.4115	43.2860	44.1991	47.4206	55.54	59.24	62.68	66.71
$r = 0$	81.3891	72.0600	68.4968	68.6602	72.5280	75.48	82.61	86.36	91.12

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot).

Kointegrationsbeziehungen, darum liegt das Hauptaugenmerk beim Johansentest.

Tabelle 4.7: Teststatistik und kritische Werte des Phillips&Ouliaris - Kointegrationstests

<b>Dimension des VAR Prozesses: 2</b>			
<b>Nullhypothese: keine Kointegration</b>			
<b>Anzahl der Beobachtungen: 4049</b>			
kritische Werte	linear	constant	none
10.0%	284.0100	225.2303	195.6202
7.5%	291.2705	232.4652	201.9621
5.0%	301.0949	241.3316	210.2910
2.5%	315.2799	255.5091	224.0976
1.0%	335.9054	270.5018	237.7723
Daten	Teststatistik		
<b>Set 1</b>	234.8000	<b>232.3050</b>	<b>227.2823</b>
<b>Set 2</b>	134.2083	134.7419	<b>227.2823</b>
kritische Werte	linear	constant	none
10.0%	219.5098	168.8572	143.0775
7.5%	225.6645	174.2575	148.4109
5.0%	234.2865	182.0749	155.8019
2.5%	247.3640	194.7555	166.3516
1.0%	264.4988	209.8054	180.4845
Daten	Teststatistik		
<b>Set 3</b>	76.6009	77.7713	75.5638

**Set 1:** Erdöl, XOM, BP, CVX, RD, SC; **Set 2:** Gold, Silber, AAUK, NEM, ABX, PDG; **Set 3:** Gold, AAUK, NEM, ABX, PDG.  
Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot).

Zusätzlich zur Kointegrationsanalyse der beiden Datensets wurden auch mehrere Zweierkombinationen zwischen Rohstoffen und Aktien aus derselben Branche untersucht. Die Tabellen mit den Ergebnissen sind im Anhang dargestellt.

Alle Tests werden für verschiedene Modellklassen (unterschiedliche deterministische Komponenten) durchgeführt. Beim Phillips Ouliaris Test wird zuerst kein deterministischer Teil in das VAR Modell aufgenommen, danach eine Konstante mitgeschätzt und beim dritten Mal noch zusätzlich ein Trend in die Gleichung aufgenommen. Beim Johansen Test werden vier Modellklassen untersucht:

- $M^*$  :
- $\mathbf{y}_t$  ist ein I(1) Prozess plus linearer deterministischer Trend
  - $\beta' \mathbf{y}_t$  ist ein I(0) Prozess plus linearer deterministischer Trend
- $M_1$  :
- $\mathbf{y}_t$  ist ein I(1) Prozess plus linearer deterministischer Trend
  - $\beta' \mathbf{y}_t$  ist ein I(0) Prozess plus Konstante
- $M_1^*$  :
- $\mathbf{y}_t$  ist ein I(1) Prozess plus Konstante
  - $\beta' \mathbf{y}_t$  ist ein I(0) Prozess plus Konstante

- $M_2$  :
- $\mathbf{y}_t$  ist ein I(1) Prozess plus Konstante
  - $\beta' \mathbf{y}_t$  ist ein I(0) Prozess mit Erwartungswert Null

In der fünften Modellklasse wäre  $\mathbf{y}_t$  ein nichtstationärer Prozess, kombiniert mit einem quadratischen deterministischen Trend. Diese Möglichkeit wird aufgrund der visuellen Betrachtung der Daten ausgeschlossen.

Für die erste Datenauswahl (Tabelle 4.5) findet der Johantest eindeutig Kointegrationsbeziehungen, jedoch ist deren Anzahl in den einzelnen Modellklassen sehr unterschiedlich. Das Hauptaugenmerk sollte laut Informationskriterien auf den VAR-Modellgrößen 1, 2 und 4 liegen. In der ersten Klasse wird ein Kointegrationsrang von zwei nicht mehr abgelehnt, in der dritten Klassen gibt es sogar fünf Kointegrationsbeziehungen und in der vierten Klasse beträgt der Kointegrationsrang drei. Die Klasse  $M_1$  weist ein sehr eigenartiges Verhalten auf, einerseits wird ein Kointegrationsrang von zwei fast nicht mehr abgelehnt, andererseits wird für die VAR-Größe von zwei die Teststatistik eindeutig abgelehnt. Dieses Fehlverhalten wird aber nicht weiter gewertet und am wahrscheinlichsten ist daher ein Kointegrationsrang von 2 oder 3.

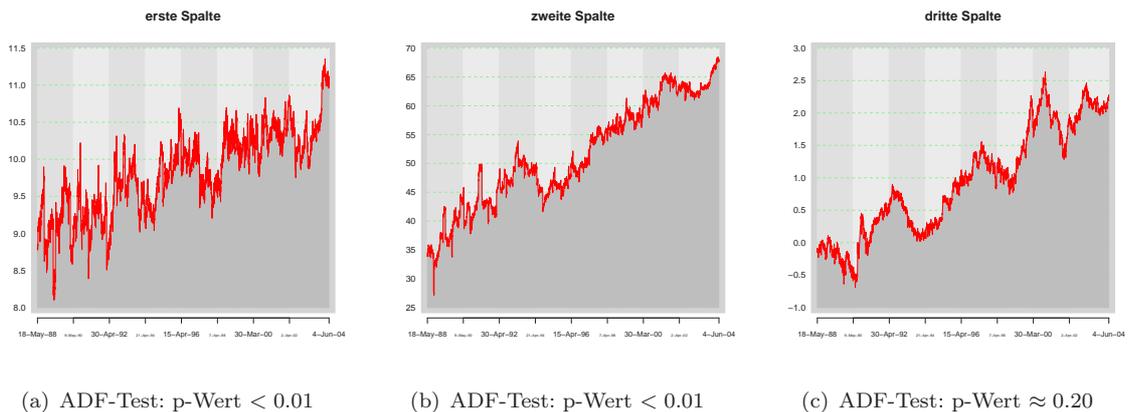


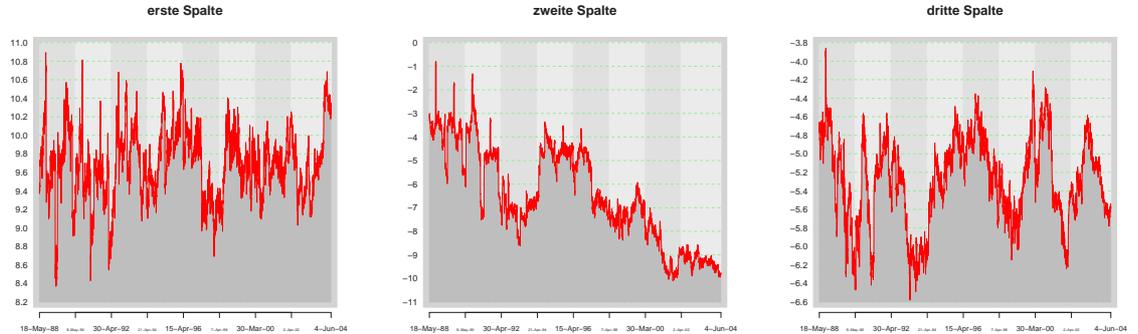
Abbildung 4.5:  $\beta' \mathbf{y}_t$  für die Klasse  $M^*$

$$M^* : \beta^6 = \begin{pmatrix} 1.0000 & 1.0000 & 1.0000 \\ 2.4047 & 7.000 & 2.3031 \\ -1.2367 & -9.8259 & -0.3949 \\ 3.1427 & 22.0826 & -2.3944 \\ 2.1420 & -28.0361 & 0.2226 \\ -5.1152 & 23.0193 & 0.3529 \end{pmatrix}$$

In der Abbildung 4.5 sind die drei Spalten von  $\beta' \mathbf{y}_t$  für die Modellklasse  $M^*$  geplottet. Die ersten beiden Zeitreihen scheinen trendstationär zu sein. Ein ADF-Test

<sup>6</sup> $\beta$  wurde auf die erste Komponente normalisiert.

bestätigt dieses Ergebnis. Für die anderen Modellklassen erhält man ähnliche Resultate, nur die Klasse  $M_1^*$  liefert eine dritte Kointegrationsbeziehung (vgl. Abbildung 4.6).



(a) ADF-Test: p-Wert < 0.01

(b) ADF-Test: p-Wert  $\approx 0.02$

(c) ADF-Test: p-Wert  $\approx 0.04$

Abbildung 4.6:  $\beta' \mathbf{y}_t$  für die Klasse  $M_1^*$

$$M_1^* : \beta = \begin{pmatrix} 1.0000 & 1.0000 & 1.0000 \\ 2.2621 & -4.4240 & 1.6825 \\ -1.3022 & 6.3418 & 1.1376 \\ 3.5146 & -5.3619 & -5.7740 \\ 3.3031 & 11.3336 & 3.5310 \\ -6.7531 & -11.1661 & -2.0960 \end{pmatrix}$$

Folglich kann man festhalten, dass im ersten Datenset entweder zwei oder drei Kointegrationsbeziehungen existieren, abhängig von der Wahl der Modellklasse. Im Kapitel „Prognose“ wird diese Information genutzt um zukünftige Entwicklungen der Daten vorherzusagen. Möglicherweise erhält man dadurch auch eine Antwort auf die Frage, welche Klasse die Daten am besten beschreibt.

Für die Paare Erdöl - Aktie (vgl. Tabelle C.11) wird in den meisten Modellklassen eine Kointegrationsbeziehung gefunden. Eine Analyse der Beta - Matrix zeigt, dass der zu Erdöl zugehörige Koeffizient ungefähr neun Mal größer ist als der Koeffizient der jeweiligen Aktie. Im Gegensatz dazu sind die Koeffizienten der Beta - Matrix des Datensets 1 ungefähr alle von der selben Größenordnung. Ein Kointegrationstest der ersten Datenzusammenstellung ohne Erdöl liefert die gleichen Ergebnisse wie zuvor beschrieben.

Der Phillips Ouliaris Test findet zumindest eine Kointegrationsbeziehung sowohl für das Datenset 1 als auch für die zweite Datenzusammenstellung (vgl. Tabelle 4.7). Bei der Betrachtung Rohstoff - Aktie wird die Nullhypothese (keine Kointegration) der einzelnen Paare jedoch nicht verworfen (vgl. Tabelle C.9).

Die Auswertung des Johansenstest der zweiten Zusammenstellung (vgl. Tabelle 4.6) ist wesentlich einfacher. Drei der vier Klassen finden genau eine Kointegrationsbeziehung in den laut Informationskriterien relevanten VAR-Modellgrößen (vgl. Abbildung 4.7). Dieses Ergebnis korrespondiert mit dem Resultat des Phillips Oulliaris

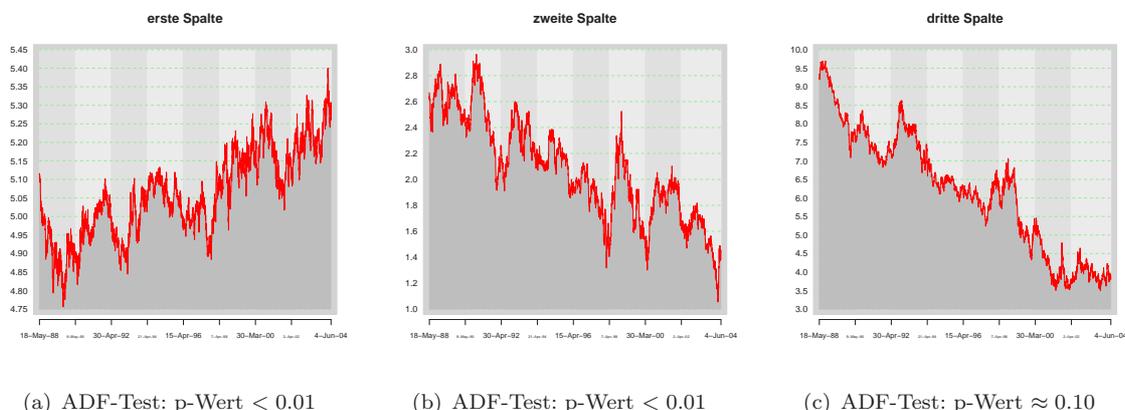


Abbildung 4.7:  $\beta' \mathbf{y}_t$  für die Klasse  $M^*$

Tests (vgl. vorheriger Absatz). Das Resultat der beiden Tests war aufgrund der Analyse der einzelnen Zeitreihen (UR-Tests) eigentlich zu erwarten, da Silber zumindest trendstationär ist und der Vektor  $[0, 1, 0, 0, 0, 0]'$  eine Kointegrationsbeziehung darstellt. Die Analyse der Beta-Matrix des Modells  $M^*$  liefert jedoch ein anderes Bild

$$M^* : \quad \beta = \begin{pmatrix} 1.0000 & 1.0000 & 1.0000 \\ 0.2321 & -1.1062 & 1.6400 \\ 0.1021 & -0.2632 & -2.1740 \\ -0.4313 & -1.0006 & 0.0474 \\ 0.3088 & -0.2972 & 0.4858 \\ -0.2789 & 1.0231 & -0.0145 \end{pmatrix}$$

Daraus kann man erkennen, dass Silber keine wesentliche Rolle spielt und die Kointegrationsbeziehung ein Mix von allen Komponenten ist, wobei Gold am meisten Gewicht besitzt. Ein neuerlicher Johansenstest des Sets 2 ohne Silber liefert dasselbe Resultat (vgl. Tabelle C.10).

Paarweise Untersuchungen von Gold - Aktie (vgl. Tabelle C.12) bzw. Silber - Aktie (vgl. Tabelle C.13) liefern keine eindeutigen Ergebnisse. Zwei Beobachtungen sind jedoch auffällig: Bei fast allen Kombinationen wird in der Klasse  $M_1$  das System als stationär angesehen und die Kombination Silber - AAUK besitzt nicht einmal eine Kointegrationsbeziehung obwohl Silber zumindest trendstationär ist.

## 4.5. Prognose

Mit den Ergebnissen des vorherigen Kapitels werden nun Prognosedaten der beiden Datensätze geschätzt und mit Originaldaten verglichen. Dazu werden fünfzig neue Datenpunkte eingelesen (07.06.2004 - 16.08.2004). Analog dazu werden Prognosedaten der eindimensionalen Zeitreihen (geschätzte ARIMA-Modelle) ermittelt und ebenfalls mit den Originaldaten verglichen. Die Funktion ist im Anhang ersichtlich und die Ergebnisse sind in der Tabelle 4.8 dargestellt.

Tabelle 4.8: Prognosevergleich

Prognosehorizont: 1 Anzahl der Prognosen: 50 Ordnung des VAR Prozesses: 2			
Set 1		Set 2	
$M^*$	$\hat{e} = 5.9789$	$M^*$	$\hat{e} = 37.5482$
$r = 2$	$\tilde{e} = 1.4426$	$r = 1$	$\tilde{e} = 2.1008$
$M_1$	$\hat{e} = 4.4175$	$M_1$	$\hat{e} = 36.2212$
$r = 2$	$\tilde{e} = 1.4426$	$r = 1$	$\tilde{e} = 2.1008$
$M_1^*$	$\hat{e} = 4.4515$	$M_1^*$	$\hat{e} = 35.9878$
$r = 3$	$\tilde{e} = 1.4426$	$r = 1$	$\tilde{e} = 2.1008$
$M_2$	$\hat{e} = 0.0000$	$M_2$	$\hat{e} = 35.3016$
$r = 3$	$\tilde{e} = 1.4426$	$r = 1$	$\tilde{e} = 2.1008$

$\tilde{e}$  ist der Betrag des Fehlervektors des mehrdimensionalen kointegrierten Systems.  
 $\hat{e}$  ist der Betrag des Fehlervektors der eindimensionalen Modellen.

Wie man in Tabelle 4.8 erkennen kann, sind die Prognosen der univarianten Zeitreihen mit Abstand besser als die Vorhersagen der kointegrierten Systeme. Die Wahl der deterministischen Komponenten spielen dabei kaum eine Rolle. Weitere Untersuchungen, die hier jedoch nicht abgebildet sind, zeigen aber, dass die Qualität der Prognoserichtung in beiden Fällen in etwa gleich gut ist. Mit Prognoserichtung ist hier der folgende Vergleich gemeint: sagt die Prognose steigende Kurse voraus, dann stimmt die Prognoserichtung, wenn auch die tatsächlichen Kurse steigen.

## 4.6. Zusammenfassung der Untersuchungen

Die eindimensionale Betrachtung der Rohstoff- und Aktienzeitreihen zeigte, dass eine „ad-hoc“ Annahme, dass alle zu untersuchenden Datenreihen nichtstationär wären, eindeutig falsch läge. Silber wird sowohl vom ADF-Test als auch vom PP-Test als zumindest trendstationär angesehen. Nur der KPSS-Test mit seinem Resultat passt nicht ganz in dieses Bild. Bei Erdöl und der Aktie AAUK wird die Trendstationarität gerade noch abgelehnt (ADF-Test), und alle anderen Zeitreihen sind eindeutig nichtstationär.

Die anschließend durchgeführten Kointegrationsuntersuchungen brachten für das

erste Datenset ebenfalls kein eindeutiges Bild: Abhängig vom gewählten Modell liefert der Johantestentest entweder zwei oder drei Kointegrationsbeziehungen. Für die zweite Datenzusammenstellung wird in allen Klassen eine Kointegrationsbeziehung gefunden. Silber spielt dabei aber keine herausragende Rolle, wie man eigentlich nach den „unit root“ Tests vermuten hätte können.

Sehr oft ist die Qualität der Prognose Motivation für die langwierige Suche nach einem passenden Modell für gegebene Daten. In diesem Fall waren die Prognosetests nur Zusatzuntersuchungen und nicht das eigentliche Ziel. Diese zusätzliche Analyse zeigte in einem sehr einfachen Vergleichsverfahren, dass die kointegrierten Systeme um ein Vielfaches schlechtere Resultate bei der Vorhersage liefern als die geschätzten ARIMA Modelle. Mögliche Rückschlüsse auf die Wahl der Modellklasse sind daher wahrscheinlich nicht sehr sinnvoll.

Betrachtet man die Daten in einer gemeinsamen graphischen Darstellung (vgl. Abbildung 4.4), so kann man daraus auf mögliche Strukturbrüche, vorallem bei der ersten Datenauswahl, schließen. Die Theorie der Kointegration unter Berücksichtigung von Strukturbrüchen wurde erst in neuerer Zeit entwickelt und wäre eine Fortsetzung der Untersuchungen.

# Anhang A.

## Stationäre ARMA Prozesse

Dieses Kapitel gibt eine kurze Einführung in die große Klasse der eindimensionalen ARMA Prozesse. Diese bieten eine sehr gute Möglichkeit verschiedenste Zeitreihen zu beschreiben.

### A.1. Stochastische Prozesse, Stationarität, ...

Für die Analyse von Finanzzeitreihen ist es sehr hilfreich, wenn man die beobachteten Daten  $(y_1, y_2, \dots, y_t)$  als eine spezielle Realisation eines stochastischen Prozesses sieht. Diese Realisation wird oft mit  $\{y_t\}_1^T$  bezeichnet, während der stochastische Prozess im Allgemeinen als eine Familie von Zufallsvariablen  $\{Y_t | t \in \mathcal{T}\}$  in einem passenden Wahrscheinlichkeitsraum definiert ist. Für diese Arbeit genügt es vollkommen, wenn die Indexmenge des Prozesses diskret ist ( $\mathcal{T} = \mathbb{Z}$ ). Ist außerdem eindeutig klar, ob es sich um die beobachtete Zeitreihe oder den zugrunde liegenden stochastischen Prozess handelt, wird für beides die Bezeichnung  $y_t$  verwendet.

Mit diesen Konventionen lässt sich der stochastische Prozess durch eine  $\mathcal{T}$ -dimensionale Wahrscheinlichkeitsverteilung beschreiben. Doch diese Betrachtungsweise würde zu viele unbekannte Parameter enthalten, die alle mit Hilfe der vorhandenen Daten geschätzt werden müssten. Stattdessen konzentriert man sich nur auf die ersten und zweiten Momente der Familie von Zufallsvariablen: die Erwartungswerte

$$\mathbf{E}[y_t] \quad \forall t$$

die Varianzen

$$\mathbf{Var}[y_t] \quad \forall t$$

und die Kovarianzen

$$\mathbf{Cov}[y_t, y_s], \quad \forall t \neq s.$$

In dem speziellen Fall, wo die gemeinsame Verteilung von  $y_t$  eine Normalverteilung ist, reichen diese Bedingungen aus um die Eigenschaften des stochastischen Prozesses zu charakterisieren. Im Allgemeinen ist jedoch ohne vereinfachende Annahmen keine Aussage möglich. Weiters ist auch eine sinnvolle Schätzung der ersten und zweiten Momente ohne weitere Annahmen nicht möglich, da nur  $T$  Beobachtungen zur Verfügung stehen.

Eine wichtige Annahme ist das Konzept der Stationarität. Dazu folgende Definition:

**Definition A.1 (Starke Stationarität)** Ein stochastischer Prozess  $(y_t | t \in \mathbb{Z})$  ist stark stationär, wenn seine Eigenschaften nicht von der Zeit abhängig sind, d.h. die Menge  $\{y_{t_1}, y_{t_2}, \dots, y_{t_m}\}$ , besitzt für beliebige Zeitpunkte  $t_i$ , dieselbe Wahrscheinlichkeitsverteilung wie die Menge  $\{y_{t_1+k}, y_{t_2+k}, \dots, y_{t_m+k}\}$  für alle Zeitverschiebungen  $k$ , i. Z.:

$$(y_{t_1}, y_{t_2}, \dots, y_{t_m}) \stackrel{d}{=} (y_{t_1+k}, y_{t_2+k}, \dots, y_{t_m+k}) \quad \forall k \in \mathbb{Z} \text{ und } m \geq 1 \quad (\text{A.1})$$

Für  $m = 1$  folgt daraus, dass (unter der Annahme:  $\mathbf{E}[(y_t)^2] < \infty$ ) der Erwartungswert und die Varianz von  $y_t$  konstant sind, und für  $m = 2$ , dass die Kovarianz nur von der betrachteten Zeitverschiebung abhängt. Das führt zu einer weiteren und in den meisten Fällen schwächeren (vor allem hier in diesem Kapitel) Definition von Stationarität:

**Definition A.2 (Schwache Stationarität)** Ein stochastischer Prozess  $(y_t | t \in \mathbb{Z})$  ist schwach stationär, wenn gilt:

$$\mathbf{E}[y_t] = \mu \quad \forall t \in \mathbb{Z} \quad (\text{A.2})$$

und

$$h \in \mathbb{Z}: \quad \mathbf{Cov}[y_{t+h}, y_t] = E[(y_t - \mu)(y_{t+h} - \mu)] = \gamma_y(h) \quad \forall t \in \mathbb{Z}. \quad (\text{A.3})$$

Im Allgemeinen folgt aus der *schwachen Stationarität* nicht die *starke Stationarität*. Für den Fall von normalverteilten  $y_t$ 's, wo, wie oben schon erwähnt, die Mittelwerte und Varianzen komplett ausreichen um den Prozess zu beschreiben, jedoch schon.

Die Autokovarianzen und -korrelationen spielen in der Datenanalyse eine so wichtige Rolle, dass sie als eigene Funktion nochmals eingeführt werden.

**Definition A.3 (Autokovarianz- und Autokorrelationsfunktion)** Die Autokovarianzfunktion eines stationären Prozesses  $y_t$  ist definiert als

$$h \in \mathbb{Z}: \quad \gamma_y(h) = \mathbf{Cov}[y_{t+h}, y_t] \quad \forall t \in \mathbb{Z}. \quad (\text{A.4})$$

Die Autokorrelationsfunktion (ACF) eines stationären Prozesses  $y_t$  ist definiert als

$$h \in \mathbb{Z}: \quad \rho_y(h) = \frac{\gamma_y(h)}{\gamma_y(0)} = \mathbf{Cor}[y_{t+h}, y_t] \quad \forall t \in \mathbb{Z}. \quad (\text{A.5})$$

Analog dazu wird für Zeitreihen die Stichprobenautokovarianzfunktion und -korrelationsfunktion eingeführt.

**Definition A.4 (Stichprobenautokovarianz- und Stichprobenautokorrelationsfunktion)** Die Stichprobenautokovarianzfunktion einer Zeitreihe  $y_t$   $t \in [1, T]$  ist definiert als

$$\hat{\gamma}(h) = T^{-1} \sum_{t=1}^{T-|h|} (y_{t+|h|} - \bar{y})(y_t - \bar{y}) \quad \text{mit } \bar{y} = \frac{1}{T} \sum_{t=1}^T y_t \quad -T < h < T. \quad (\text{A.6})$$

Die Stichprobenautokorrelationsfunktion (SACF) einer Zeitreihe  $y_t$   $t \in [1, T]$  ist definiert als

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \quad -T < h < T. \quad (\text{A.7})$$

Aus der Definition eines schwach stationären Prozesses folgt, dass für die Autokovarianzfunktion und für die ACF folgende Beziehung gilt:

$$\gamma_y(h) = \gamma_y(-h)$$

und

$$\rho_y(h) = \rho_y(-h).$$

Dieser Zusammenhang ergibt sich durch die Betrachtung der einfachen Umformungen:

$$\gamma_y(h) = \mathbf{Cov}[y_{t+h}, y_t] = \mathbf{Cov}[y_t, y_{t+h}] = \mathbf{Cov}[y_{t-h}, y_t] = \gamma_y(-h).$$

Im ersten Schritt können aufgrund der Definition der Kovarianz durch die Erwartungswerte die Argumente vertauscht werden. Im zweiten Schritt wird einfach der Index  $t$  durch  $t - h$  substituiert.

Die ACF und SACF spielen eine wichtige Rolle in der Modellierung von Datenreihen. Durch den Vergleich der theoretischen mit der aus den Daten generierten tatsächlichen Autokorrelationsfunktion können Zeitreihen möglichen Modellen zugeordnet werden.

## A.2. „White Noise“ - Weißes Rauschen

Der einfachste stochastische aber auch zugleich wichtigste Prozesse ist der „white noise“ Prozess

**Definition A.5 („white noise“ Prozess)** Ein stochastischer Prozess  $(\xi_t | t \in \mathbb{Z})$  mit

$$\mathbf{E}[\xi_t] = 0 \quad \forall t \in \mathbb{Z}, \quad (\text{A.8})$$

$$\mathbf{V}[\xi_t] = \sigma^2 \quad \forall t \in \mathbb{Z} \quad (\text{A.9})$$

und

$$\mathbf{Cov}[\xi_{t+h}, \xi_t] = 0 \quad \forall t, h \in \mathbb{Z} \setminus \{0\} \quad (\text{A.10})$$

heißt „white noise“ Prozess, i. Z.  $\xi_t \sim WN(0, \sigma^2)$ .

Wird die Bedingung (A.10) durch die stärkere Bedingung (A.11)

$$\xi_{t+h}, \xi_t \text{ sind unabhängig} \quad \forall h \in \mathbb{Z} \setminus \{0\} \quad (\text{A.11})$$

ersetzt, so spricht man von einem „strict white noise“ Prozess oder „independent white noise“ Prozess ( $\xi_t \sim SWN(0, \sigma^2)$ ). Es ist offensichtlich, dass aus (A.11) die

Bedingung (A.10) folgt, aber nicht umgekehrt.

Ist der Prozess  $\xi_t$  unabhängig und identisch verteilt, schreibt man  $\xi_t \sim IID(0, \sigma^2)$  (IID = independent identically distributed), ist dieser sogar normalverteilt bedeutet das in Zeichen  $\xi_t \sim NID(0, \sigma^2)$  (NID = normally identically distributed).

In Abbildung A.1(a) sieht man eine Realisation eines stochastischen Prozesses  $\xi_t \sim NID(0, 1)$ . Für normalverteilte Daten sollte die SACF für  $h \neq 0$  im Intervall

$$\left( \frac{1.96}{\sqrt{T}}, \frac{1.96}{\sqrt{T}} \right)$$

liegen (5% Niveau), damit sie als nicht signifikant angesehen wird (vgl. die ACF in A.1(c)). In Abbildung A.1(b) sieht man jedoch, dass diese Bedingung nicht für alle Lags erfüllt wird (die blauen Linien stellen die Intervallgrenzen dar). Neben der natürlichen Erklärung, dass die Testkonstruktion sowieso Fehler im Ausmaß von 5% zulässt, könnte aber auch die Stichprobengröße  $T = 150$  dieses Phänomen entstehen lassen.

### A.3. „Moving Average“ Prozesse

Mit  $\xi_t \sim WN(0, \sigma^2)$  kann man den folgenden einfachen stochastischen Prozess betrachten:

$$y_t = \mu + \xi_t + \theta \xi_{t-1} \quad (\text{A.12})$$

wobei  $\mu$  und  $\theta$  beliebige Konstanten sind. Dieser wird „Moving Average“ Prozess der Ordnung Eins genannt.  $\mu$  wird hier deshalb gewählt, weil die Konstante den Mittelwert des Prozess bestimmt, wie in den nächsten Zeilen gezeigt wird.

Der Erwartungswert von  $y_t$  ist gegeben durch

$$\mathbf{E}[y_t] = \mathbf{E}[\mu + \xi_t + \theta \xi_{t-1}] = \mu + \mathbf{E}[\xi_t] + \theta \mathbf{E}[\xi_{t-1}] = \mu. \quad (\text{A.13})$$

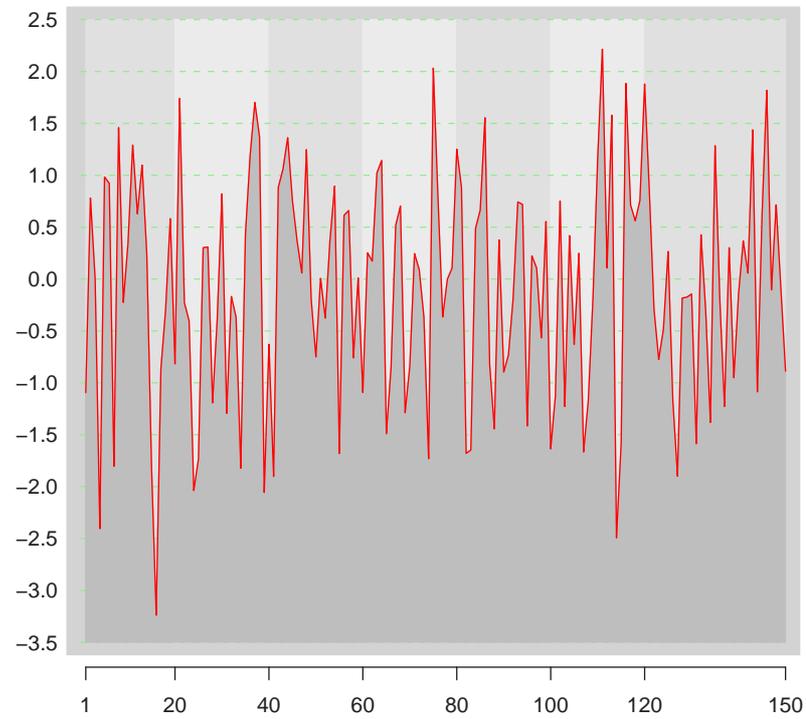
Die Varianz von  $y_t$  ist

$$\begin{aligned} \mathbf{Var}[y_t] = \mathbf{E}[(y_t - \mu)^2] &= \mathbf{E}[(\xi_t + \theta \xi_{t-1})^2] & (\text{A.14}) \\ &= \mathbf{E}[\xi_t^2 + 2\xi_t \theta \xi_{t-1} + \theta^2 \xi_{t-1}^2] \\ &= \sigma^2 + 0 + \theta^2 \sigma^2 \\ &= (1 + \theta^2) \sigma^2. \end{aligned}$$

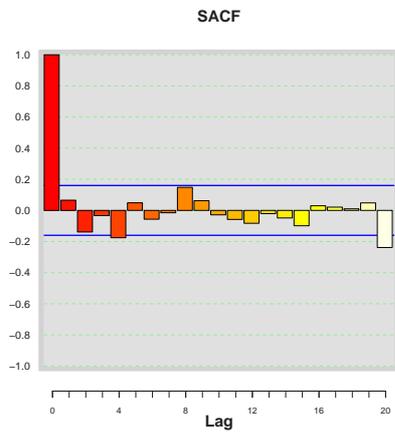
Die erste Autokovarianz beträgt

$$\begin{aligned} \mathbf{Cov}[y_t, y_{t-1}] = \mathbf{E}[(y_t - \mu)(y_{t-1} - \mu)] &= \mathbf{E}[(\xi_t + \theta \xi_{t-1})(\xi_{t-1} + \theta \xi_{t-2})] & (\text{A.15}) \\ &= \mathbf{E}[\xi_t \xi_{t-1} + \theta \xi_{t-1}^2 + \xi_t \theta \xi_{t-2} + \theta^2 \xi_{t-1} \xi_{t-2}] \\ &= 0 + \theta \sigma^2 + 0 + 0 \end{aligned}$$

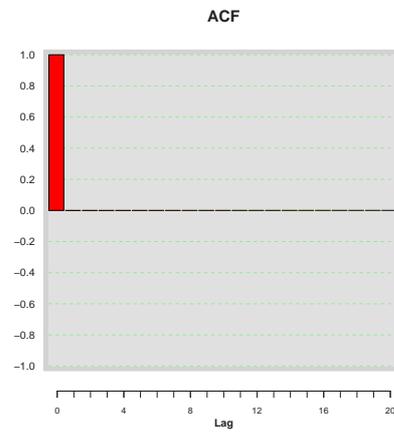
### White Noise



(a)  $\xi_t \sim NID(0, 1)$



(b) Stichprobenautokorrelationsfunktion



(c) Autokorrelationsfunktion

Abbildung A.1: Simuliertes „Weißes Rauschen“

und alle anderen sind gleich Null:

$$\mathbf{Cov}[y_t, y_{t-h}] = \mathbf{E}[(\xi_t + \theta\xi_{t-1})(\xi_{t-h} + \theta\xi_{t-h-1})] = 0 \quad \forall |h| > 1. \quad (\text{A.16})$$

Mit diesen Ergebnissen ist klar ersichtlich, dass der MA(1) Prozess  $y_t$  ein stationärer Prozess ist.

Mit Hilfe von (A.14), (A.15) und (A.16) ist die ACF eines MA(1) Prozesses gegeben durch:

$$\rho_y(h) = \begin{cases} \frac{\theta}{(1+\theta^2)} & \text{für } |h| = 1 \\ 0 & \text{für } |h| > 1 \end{cases} \quad (\text{A.17})$$

Die Verallgemeinerung des MA(1) Prozesses ist der MA( $q$ ) Prozess mit folgender Definition:

**Definition A.6 (MA( $q$ ) Prozess)** *Ein stochastischer Prozess  $y_t$  ist ein ‘‘Moving Average’’ Prozess der Ordnung  $q$  (MA( $q$ )) wenn*

$$y_t = \mu + \xi_t + \theta_1\xi_{t-1} + \dots + \theta_q\xi_{t-q} \quad (\text{A.18})$$

wobei  $\theta_q \neq 0$  und  $\xi_t \sim WN(0, \sigma^2)$ .

Mit Hilfe des ‘‘Backshift’’-Operators:

$$B^m z_t \equiv z_{t-m}, \quad t \in \mathbb{Z}$$

lässt sich obiger Prozess kurz schreiben als

$$y_t = \mu + \theta(B)\xi_t$$

wobei  $\theta(B) = (1 + \theta_1 B + \dots + \theta_q B^q)$ .

Der Erwartungswert von (A.18) ist wieder gegeben durch  $\mu$ :

$$\mathbf{E}[y_t] = \mu + \mathbf{E}[\xi_t] + \theta_1\mathbf{E}[\xi_{t-1}] + \dots + \theta_q\mathbf{E}[\xi_{t-q}] = \mu. \quad (\text{A.19})$$

Die Varianz des MA( $q$ ) Prozesses ist

$$\begin{aligned} \mathbf{V}[y_t] = \mathbf{E}[y_t - \mu]^2 &= \mathbf{E}[(\xi_t + \theta_1\xi_{t-1} + \dots + \theta_q\xi_{t-q})^2] & (\text{A.20}) \\ &= \mathbf{E}[\xi_t^2] + \mathbf{E}[\theta_1\xi_{t-1}\xi_t] + \dots + \mathbf{E}[\theta_q\xi_{t-q}\xi_t] \\ &\quad + \mathbf{E}[\xi_t\theta_1\xi_{t-1}] + \mathbf{E}[\theta_1^2\xi_{t-1}^2] + \dots + \mathbf{E}[\theta_q\theta_1\xi_{t-q}\xi_{t-1}] \\ &\quad \vdots \\ &\quad + \mathbf{E}[\theta_q\xi_t\xi_{t-q}] + \mathbf{E}[\theta_1\theta_q\xi_{t-1}\xi_{t-q}] + \dots + \mathbf{E}[\theta_q^2\xi_{t-q}^2] \\ &= \sigma^2 + \theta_1^2\sigma^2 + \dots + \theta_q^2\sigma^2 \\ &= (1 + \theta_1^2 + \dots + \theta_q^2)\sigma^2. \end{aligned}$$

und die Autokovarianzen  $\gamma_y(h)$  lassen sich durch

$$\gamma_y(h) = \begin{cases} (\theta_h + \theta_{h+1}\theta_1 + \theta_{h+2}\theta_2 + \cdots + \theta_q\theta_{q-h})\sigma^2 & \forall h = 1, 2, \dots, q \\ 0 & h > q \end{cases} \quad (\text{A.21})$$

anschreiben.

**Beispiel A.1** In Abbildung A.2 wurden zwei „Moving Average“ Prozesse der Ordnung 1 und 4

$$\begin{aligned} y_t &= \xi_t + 0.8\xi_t \\ y_t &= \xi_t - 0.6\xi_t + 0.3\xi_t - 0.5\xi_t + 0.5\xi_t \end{aligned}$$

simuliert ( $\xi \sim WN(0, 1)$ ). Sowohl beim MA(1) Prozess als auch beim MA(4) Prozess liefert die SACF schlechte Ergebnisse (siehe A.2(c) und A.2(d)).

In Abbildung A.3 wurden die beiden Prozesse nochmals simuliert, jedoch wurden 1500 Daten erzeugt um Umstimmigkeiten zwischen der SACF und ACF aufgrund der Stichprobengröße auszuschließen.

Betrachtet man den Fall  $q \rightarrow \infty$ , führt das zu folgender allgemeinen Definition:

**Definition A.7 („Moving Average“ Prozess)** Ein stochastischer Prozess  $y_t$  wird auch „Moving Average“ Prozess MA( $\infty$ ) genannt, falls er sich darstellen lässt als

$$y_t = \sum_{j=0}^{\infty} \psi_j \xi_{t-j} \quad \forall t \quad (\text{A.22})$$

wobei  $\xi_t \sim WN(0, \sigma^2)$  und  $\psi_j$  eine Folge von Konstanten mit  $\sum_{j=0}^{\infty} |\psi_j| < \infty$  ist.<sup>1</sup>

Kompakter anschreiben lässt sich (A.22) wieder mit Hilfe des „Backshift“ Operators:

$$y_t = \psi(B)\xi_t$$

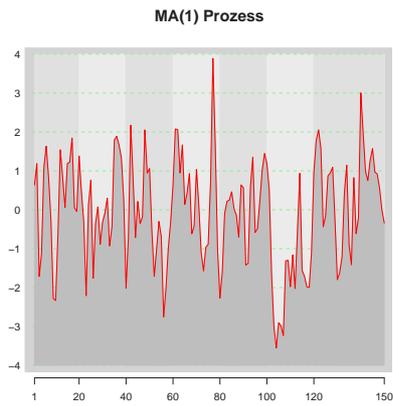
wobei  $\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j$ .

Durch eine einfache Extrapolation der Ergebnisse für einen MA( $q$ ) Prozess können der Erwartungswert und die Autokovarianzen eines MA( $\infty$ ) Prozesses mit absolut summierbaren Koeffizienten

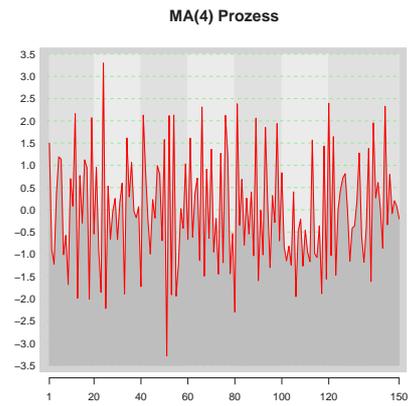
$$\sum_{j=0}^{\infty} |\psi_j| < \infty \quad (\text{A.23})$$

<sup>1</sup>Ausreichend wäre auch bereits die quadratische Summierbarkeit der Koeffizienten:

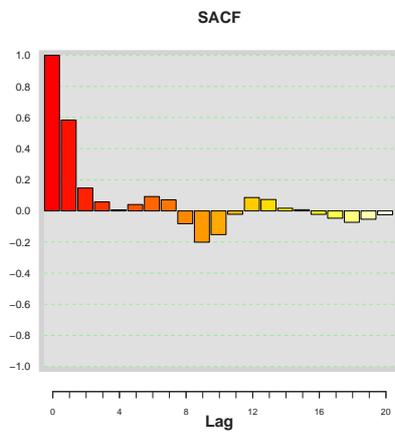
$$\sum_{j=0}^{\infty} \psi_j^2 < \infty.$$



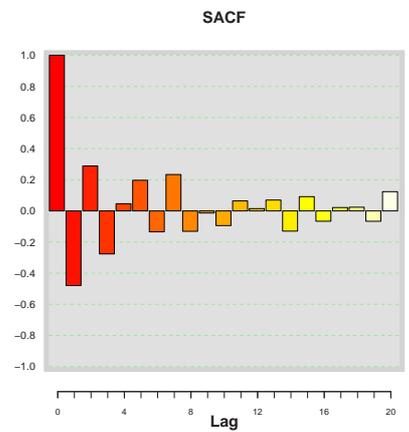
(a) Simulierter MA(1) Prozess mit normalverteilten Innovationen  $\xi_t \sim NID(0, 1)$



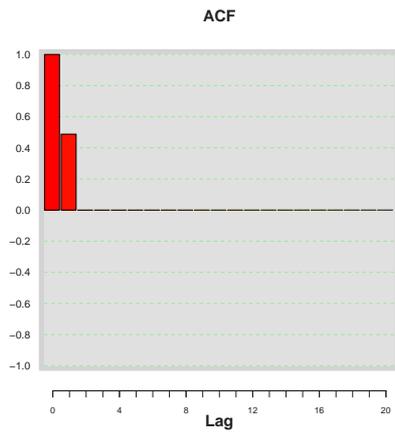
(b) Simulierter MA(4) Prozess mit normalverteilten Innovationen  $\xi_t \sim NID(0, 1)$



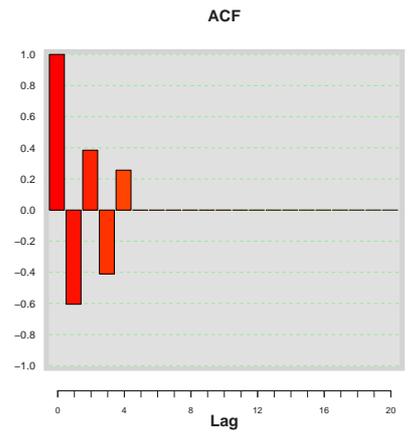
(c) Stichprobenautokorrelationsfunktion eines MA(1) Prozesses



(d) Stichprobenautokorrelationsfunktion eines MA(4) Prozesses

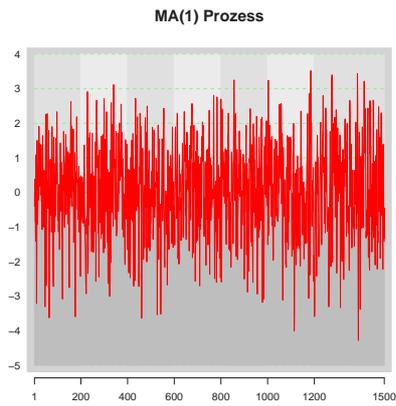


(e) Autokorrelationsfunktion eines MA(1) Prozesses

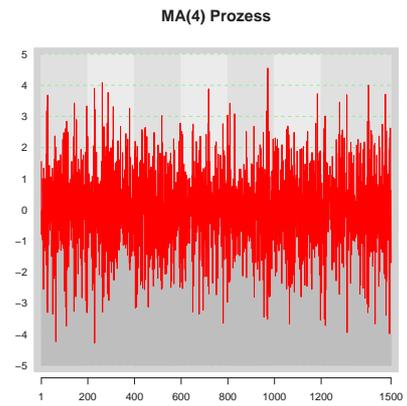


(f) Autokorrelationsfunktion eines MA(4) Prozesses

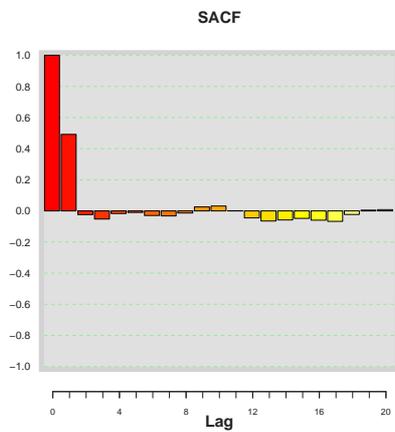
Abbildung A.2: „Moving Average“ Prozesse mit unterschiedlichen Ordnungen



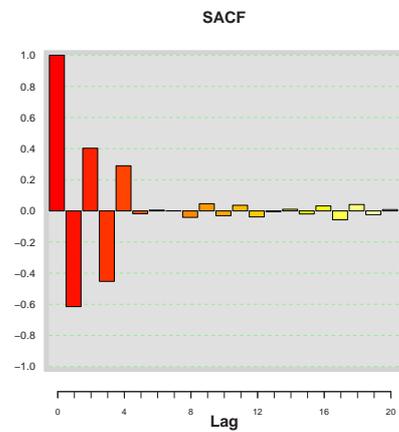
(a) Simulierter MA(1) Prozess mit normalverteilten Innovationen  $\xi_t \sim NID(0, 1)$



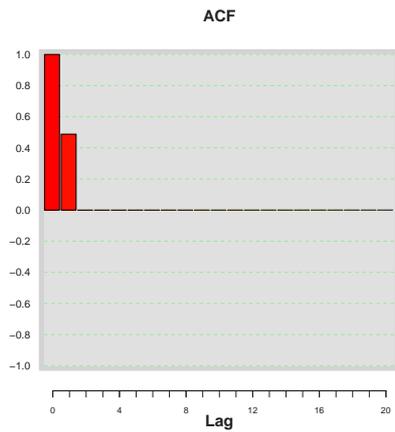
(b) Simulierter MA(4) Prozess mit normalverteilten Innovationen  $\xi_t \sim NID(0, 1)$



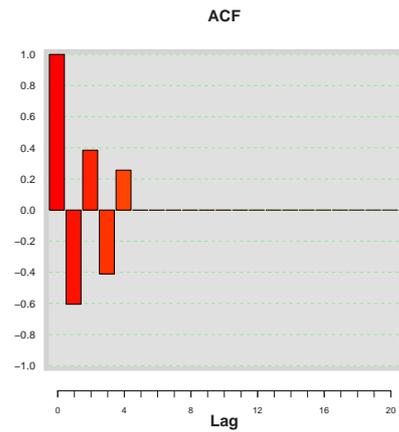
(c) Stichprobenautokorrelationsfunktion eines MA(1) Prozesses



(d) Stichprobenautokorrelationsfunktion eines MA(4) Prozesses



(e) Autokorrelationsfunktion eines MA(1) Prozesses



(f) Autokorrelationsfunktion eines MA(4) Prozesses

Abbildung A.3: „Moving Average“ Prozesse mit unterschiedlichen Ordnungen

berechnet werden <sup>2</sup>:

$$\begin{aligned} \mathbf{E}[y_t] &= \lim_{T \rightarrow \infty} \mathbf{E}[\mu + \psi_0 \xi_t + \psi_1 \xi_{t-1} + \dots + \psi_q \xi_{t-T}] \\ &= \mu \end{aligned} \quad (\text{A.24})$$

$$\begin{aligned} \gamma_y(0) &= \mathbf{E}[(y_t - \mu)^2] \\ &= \lim_{T \rightarrow \infty} \mathbf{E}[(\psi_0 \xi_t + \psi_1 \xi_{t-1} + \dots + \psi_q \xi_{t-T})^2] \\ &= \lim_{T \rightarrow \infty} (\psi_0^2 + \psi_1^2 + \dots + \psi_T^2) \sigma^2 \end{aligned} \quad (\text{A.25})$$

$$\begin{aligned} &= \sigma^2 \sum_{j=0}^{\infty} \psi_j^2 \\ \gamma_y(h) &= \mathbf{E}[(y_t - \mu)(y_{t-h} - \mu)] \\ &= \sigma^2 (\psi_h \psi_0 + \psi_{h+1} \psi_1 + \psi_{h+2} \psi_2 + \psi_{h+3} \psi_3 + \dots) \\ &= \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}. \end{aligned} \quad (\text{A.26})$$

## A.4. Autoregressive Prozesse

Ein autoregressiver Prozess der Ordnung Eins ist die stationäre Lösung der Differenzgleichung

$$y_t = \nu + \phi y_{t-1} + \xi_t \quad (\text{A.27})$$

wobei  $\xi_t \sim WN(0, \sigma^2)$ ,  $c$  und  $\phi$  Konstanten sind. Ist der Betrag des Koeffizienten  $|\phi|$  ungleich Eins so existiert eine stationäre Lösung, für  $\phi = 1$  ist (A.27) ein „random walk“, der nichtstationär ist. Für  $|\phi| < 1$  besitzt die stationäre Lösung von (A.27) folgende Gestalt:

$$\begin{aligned} y_t &= (\nu + \xi_t) + \phi(\nu + \xi_{t-1}) + \phi^2(\nu + \xi_{t-2}) + \dots \\ &= [\nu/(1 - \phi)] + \xi_t + \phi \xi_{t-1} + \phi^2 \xi_{t-2} + \dots \end{aligned} \quad (\text{A.28})$$

$y_t$  hat damit eine Darstellung eines MA( $\infty$ ) Prozesses (vgl. (A.22)) mit  $\psi_j = \phi^j$ . Gilt  $|\phi| < 1$  so wird auch die Bedingung A.23 erfüllt:

$$\sum_{j=0}^{\infty} |\psi_j| = \sum_{j=0}^{\infty} |\phi^j| = \frac{1}{1 - |\phi|}.$$

Motiviert durch die MA( $\infty$ ) Darstellung eines AR(1) Prozesses wird folgende Bezeichnung eingeführt:

<sup>2</sup>Absolut summierbare Koeffizienten und die Existenz des zweiten Momentes sind ausreichende Bedingungen für das Vertauschen von Integration und Summation, i. Z.

$$\mathbf{E} \left[ \sum_{j=1}^{\infty} y_t \right] = \sum_{j=1}^{\infty} \mathbf{E}[y_t]$$

mit  $y_t$  ein Folge von Zufallsvariablen.

**Definition A.8 (Kausalität)** Ein stochastischer Prozess  $y_t$  wird kausal genannt, wenn eine Folge von Konstanten  $\psi_j$  mit  $\sum_{j=0}^{\infty} |\psi_j| < \infty$  existiert und

$$y_t = \sum_{j=0}^{\infty} \psi_j \xi_{t-j} \quad \forall t. \quad (\text{A.29})$$

Wie man aus der Definition sofort erkennt, folgt aus der Kausalität eines Prozess die Stationarität.

Mit Hilfe der kausalen Lösung des AR(1) Prozesses (A.28) lassen sich nun der Erwartungswert, die Varianz und die Autokovarianzen berechnen:

$$\begin{aligned} \mu = \mathbf{E}[y_t] &= [\nu/(1 - \phi)] + \mathbf{E}[\xi_t] + \mathbf{E}[\phi\xi_{t-1}] + \mathbf{E}[\phi^2\xi_{t-2}] + \dots \quad (\text{A.30}) \\ &= [\nu/(1 - \phi)] + 0 + 0 + \dots \end{aligned}$$

$$\begin{aligned} \gamma_y(0) = \mathbf{V}[y_t] &= \mathbf{E}[(y_t - \mu)^2] \quad (\text{A.31}) \\ &= \mathbf{E}[(\xi_t + \phi\xi_{t-1} + \phi^2\xi_{t-2} + \dots)^2] \\ &= (1 + \phi^2 + \phi^4 + \dots)\sigma^2 \\ &= \sigma^2/(1 - \phi^2) \end{aligned}$$

$$\begin{aligned} \gamma_y(h) = \mathbf{Cov}[y_t, y_{t-h}] &= \mathbf{E}[(y_t - \mu)(y_{t-h} - \mu)] \quad (\text{A.32}) \\ &= \mathbf{E}[\xi_t + \phi\xi_{t-1} + \phi^2\xi_{t-2} + \dots + \phi^h\xi_{t-h} + \phi^{h+1}\xi_{t-h-1} \\ &\quad + \phi^{h+2}\xi_{t-h-2}] \times [\xi_{t-h} + \phi\xi_{t-h-1} + \phi^2\xi_{t-h-2} + \dots] \\ &= [\phi^h + \phi^{h+2} + \phi^{h+4} + \dots]\sigma^2 \\ &= \phi^h[1 + \phi^2 + \phi^4 + \dots]\sigma^2 \\ &= [\phi^h/(1 - \phi^2)]\sigma^2 \text{ für } h > 0. \end{aligned}$$

Aus (A.31) und (A.32) folgt, dass die Autokorrelationsfunktion:

$$\rho_y(h) = \frac{\gamma_y(h)}{\gamma_y(0)} = \phi^h \quad (\text{A.33})$$

eine exponentiell fallende Funktion ist (siehe A.4(e) und A.4(f)).

Die Momente des AR(1) Prozesses wurden mit Hilfe der MA( $\infty$ ) Darstellung berechnet. Ein anderer Weg wäre, die Momente direkt aus der Gleichung (A.27), mit der Annahme, dass der Prozess stationär ist, zu berechnen. Dies wird jedoch hier nicht ausgeführt, sondern auf Hamilton (1994) verwiesen.

Die Verallgemeinerung des AR(1) Prozesses ist der AR( $p$ ) Prozess mit folgender Definition:

**Definition A.9 (AR( $p$ ) Prozess)** Ein stochastischer Prozess  $y_t$  ist ein stationärer autoregressiver Prozess der Ordnung  $p$  (AR( $p$ )) wenn

$$y_t = \nu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \xi_t \quad (\text{A.34})$$

wobei  $\phi_p \neq 0$  und  $\xi_t \sim WN(0, \sigma^2)$ .

Unter Verwendung des „Backshift“-Operators lässt sich obiger Prozess kurz schreiben als

$$\phi(B)y_t = \nu + \xi_t$$

wobei  $\phi(B) = (1 + \phi_1 B + \dots + \phi_p B^p)$ .

Liegen die Wurzeln von  $\phi(z)$  alle außerhalb des Einheitskreises, so existiert eine eindeutig kausale Lösung von (A.34) in der Form

$$y_t - \mathbf{E}[y_t] = \psi(B)\xi_t \quad (\text{A.35})$$

mit  $\psi(B) = \frac{1}{\phi(B)}$  und  $\sum_{j=0}^{\infty} |\psi_j| < \infty$ . Angenommen, diese Bedingung wird erfüllt, dann kann das Mittel des Prozesses dadurch berechnet werden, dass auf beiden Seiten von (A.34) die Erwartungswerte gebildet werden:

$$\begin{aligned} \mu &= \nu + \phi_1 \mu + \phi_2 \mu + \dots + \phi_p \mu \\ &= \nu / (1 - \phi_1 - \phi_2 - \dots - \phi_p) \end{aligned} \quad (\text{A.36})$$

Mit Hilfe von (A.36) kann (A.34) geschrieben werden als

$$\begin{aligned} y_t - \mu &= \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + \dots \\ &\quad + \phi_p(y_{t-p} - \mu) + \xi_t. \end{aligned} \quad (\text{A.37})$$

Werden beide Seiten von Gleichung (A.37) mit  $(y_{t-h} - \mu)$  multipliziert und anschließend der Erwartungswert davon berechnet, erhält man die Autokovarianzen:

$$\gamma_y(h) = \begin{cases} \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + \dots + \phi_p \gamma_{h-p} & \text{für } h = 1, 2, \dots \\ \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma^2 & \text{für } h = 0. \end{cases} \quad (\text{A.38})$$

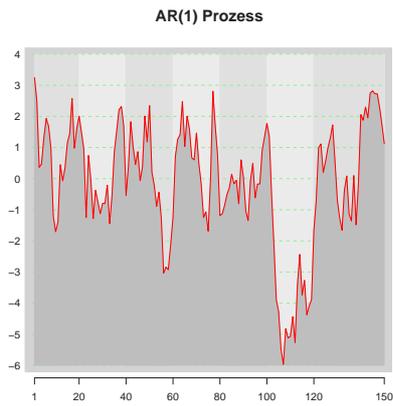
Mit der Tatsache  $\gamma_y(h) = \gamma_y(-h)$  kann das Gleichungssystem (A.38) für  $h = 0, 1, \dots, p$  gelöst werden.  $\gamma_y(0), \gamma_y(1), \dots, \gamma_y(p)$  sind dabei eine Funktion von  $\sigma^2, \phi_1, \phi_2, \dots, \phi_p$ .

**Beispiel A.2** In Abbildung A.4 wurden zwei autoregressive Prozesse der Ordnung Eins und Zwei

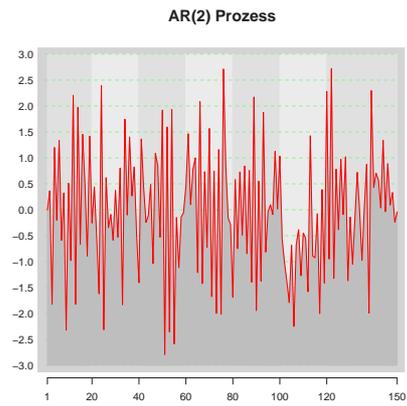
$$\begin{aligned} y_t &= 0.8y_{t-1} + \xi_t \\ y_t &= -0.5y_{t-1} + 0.3y_{t-2} + \xi_t \end{aligned}$$

simuliert ( $\xi \sim WN(0, 1)$ ). Die visuelle Betrachtung vermittelt einem auch in diesem Beispiel den Eindruck, dass sich das Verhalten der SACF nicht mit der ACF deckt. Im Gegensatz zum MA Beispiel ist aber hier der Wert der ACF nie gleich Null und ein Vergleich (ACF ungleich/gleich Null vs. SACF ungleich/gleich Null) nicht möglich.

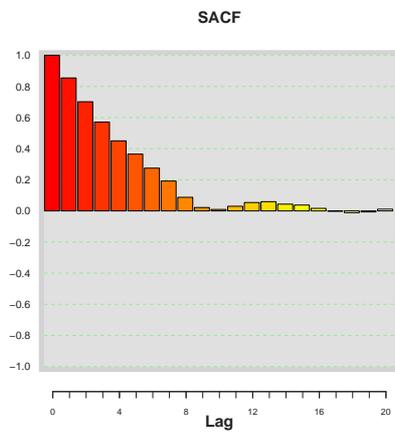
Wie im Kapitel MA-Prozesse ist es auch hier möglich den Fall  $p \rightarrow \infty$  zu betrachten und folgenden Prozess zu definieren:



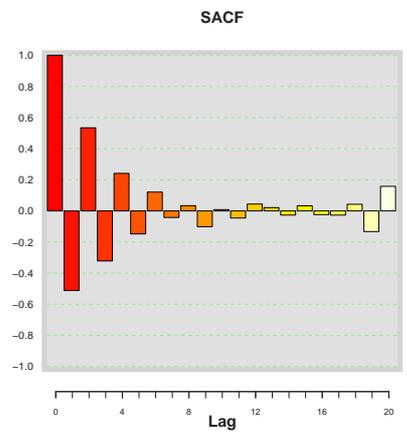
(a) Simulierter AR(1) Prozess mit normalverteilten Innovationen  $\xi_t \sim NID(0, 1)$



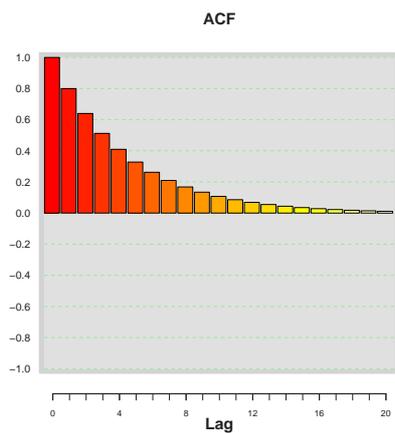
(b) Simulierter AR(2) Prozess mit normalverteilten Innovationen  $\xi_t \sim NID(0, 1)$



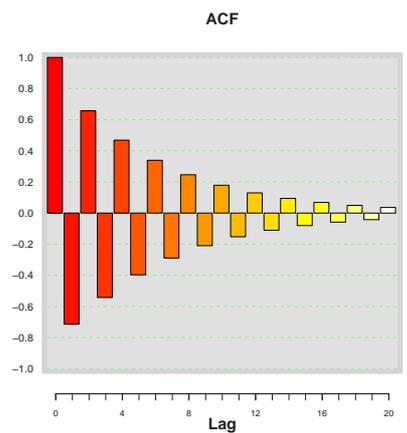
(c) Stichprobenautokorrelationsfunktion eines AR(1) Prozesses



(d) Stichprobenautokorrelationsfunktion eines AR(2) Prozesses



(e) Autokorrelationsfunktion eines AR(1) Prozesses



(f) Autokorrelationsfunktion eines AR(2) Prozesses

Abbildung A.4: Autoregressive Prozesse mit unterschiedlichen Ordnungen

**Definition A.10 (AR( $\infty$ ) Prozess)** Ein stochastischer Prozess  $y_t$  ist ein AR( $\infty$ ) Prozess, falls

$$\xi_t = \sum_{j=0}^{\infty} \pi_j y_{t-j} \quad \forall t \quad (\text{A.39})$$

wobei  $\xi_t \sim WN(0, \sigma^2)$  und  $\pi_j$  eine Folge von Konstanten mit  $\sum_{j=0}^{\infty} |\pi_j| < \infty$  ist.

Betrachtet man nochmals den MA( $q$ ) Prozess aus dem vorherigen Kapitel

$$y_t = \mu + \theta(B)\xi_t \quad (\text{A.40})$$

und liegen alle Wurzeln von

$$(1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q) = 0$$

außerhalb des Einheitskreises, so kann (A.40) als ein AR( $\infty$ ) Prozess geschrieben werden:

$$\left( \sum_{j=0}^{\infty} \pi_j B^j \right) y_t = \theta^{-1}(B)y_t = \xi_t. \quad (\text{A.41})$$

Diese Eigenschaft nennt man Invertierbarkeit des MA( $q$ ) Prozesses (vgl. Definition A.12).

## A.5. Autoregressive „Moving Average“ Prozesse

Die Verknüpfung eines MA( $q$ ) und AR( $p$ ) Prozess führt zu folgendem Prozess:

**Definition A.11 (ARMA( $p, q$ ) Prozess)** Ein stochastischer Prozess  $y_t$  ist ein stationärer autoregressiver „Moving Average“ (ARMA) Prozess der Ordnung ( $p, q$ ) wenn für alle  $t$  gilt:

$$y_t - \phi_1 y_{t-1} - \dots - \phi_p y_{t-p} = \nu + \xi_t + \theta_1 \xi_{t-1} + \dots + \theta_p \xi_{t-p} \quad (\text{A.42})$$

wobei  $\phi_q, \theta_p \neq 0$ ,  $a_t \sim WN(0, \sigma^2)$ .

Bezeichnet man mit  $\phi(\cdot)$  und  $\theta(\cdot)$  die beiden  $p$ - und  $q$ -gradigen Polynome

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p$$

und

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_p z^p$$

und ist  $B$  der oben definierte „Backshift“-Operator, so lässt sich die Gleichung (A.42) kurz schreiben als

$$\phi(B)y_t = \nu + \theta(B)\xi_t.$$

Die Existenz und Eindeutigkeit ist durch folgenden Satz gegeben:

**Satz A.1 (Existenz und Eindeutigkeit)** Eine kausale Lösung der Gleichung A.42 existiert (und ist auch die einzige kausale Lösung) dann und nur dann, wenn

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \neq 0 \quad \forall |z| \leq 1. \quad (\text{A.43})$$

**Beweis A.1** siehe Brockwell and Davis (1996)

Die „Moving Average“ Parameter des Prozesses  $(\theta_1, \theta_2, \dots, \theta_q)$  haben somit keinen Einfluss auf die Existenz einer kausalen Lösung.

Betrachtet man

$$\psi(B) = \frac{\theta(B)}{\phi(B)}$$

und

$$\mu = \frac{\nu}{\phi(B)}$$

so erhält man eine  $\text{MA}(\infty)$  Darstellung des ARMA Prozesses:

$$y_t = \mu + \psi(B)\xi_t. \quad (\text{A.44})$$

Durch Koeffizientenvergleich von  $z^j$  der Gleichung  $\sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}$  oder ident dazu von

$$(1 - \phi_1 z - \dots - \phi_p z^p)(\psi_0 + \psi_1 z + \dots) = 1 + \theta_1 z + \dots + \theta_p z^p$$

erhält man

$$\begin{aligned} \psi_0 &= 1 \\ \psi_1 &= \theta_1 + \psi_0 \phi_1 \\ \psi_2 &= \theta_2 + \psi_1 \phi_1 - \psi_0 \phi_2 \\ &\vdots \\ &\vdots \\ &\vdots \end{aligned}$$

oder ganz allgemein

$$\psi_j = \theta_j + \sum_{k=1}^p \phi_k \psi_{j-k} \quad j = 0, 1, \dots \quad (\text{A.45})$$

mit  $\theta_0 := 1$ ,  $\theta_j := 0$  wenn  $j > q$  und  $\psi_j := 0$  wenn  $j < 0$ .

Die Autokovarianzen für einen ARMA(p,q) Prozess sind nicht mehr so einfach zu berechnen wie für die beiden Komponenten AR(p) und MA(q) und werden aus diesem Grund hier nicht explizit dargestellt.

Abschließend noch eine weitere Charakterisierung eines ARMA Prozesses:

**Definition A.12 (Invertierbarkeit)** Ein ARMA( $p, q$ ) Prozess  $y_t$  ist invertierbar, wenn eine Folge von Konstanten  $\pi_j$  mit  $\sum_{j=0}^{\infty} |\pi_j| < \infty$  existiert und

$$\xi_t = \sum_{j=0}^{\infty} \pi_j y_{t-j} \quad \forall t. \quad (\text{A.46})$$

Invertierbarkeit ist gleichwertig mit der Bedingung:

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_p z^p \neq 0 \quad \forall |z| \leq 1. \quad (\text{A.47})$$

Vertauscht man die Rollen der AR und MA-Polynome in der Gleichung (A.45) so erhält man die Koeffizienten  $\pi_j$  durch folgende Gleichungen:

$$\pi_j = -\phi_j - \sum_{k=1}^q \theta_k \pi_{j-k} \quad j = 0, 1, \dots \quad (\text{A.48})$$

mit  $\phi_0 := 1$ ,  $\phi_j := 0$  wenn  $j > q$  und  $\pi_j := 0$  wenn  $j < 0$ .

Ein ARMA( $p, q$ ) Prozess hat somit unter gegebenen Voraussetzungen eine AR( $\infty$ ) und/oder eine MA( $\infty$ ) Darstellung.

**Beispiel A.3** In Abbildung A.5 wurde ein ARMA Prozess der Ordnung 1,1

$$y_t - 0.5y_{t-1} = \xi_t + 0.4\xi_{t-1}$$

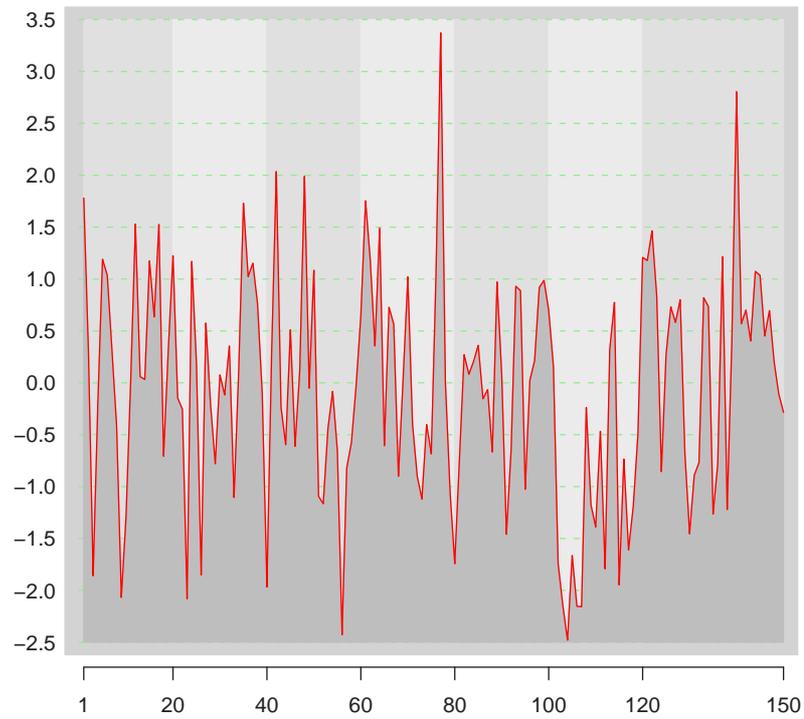
simuliert ( $\xi \sim WN(0, 1)$ ). Das AR Polynom des Prozesses  $\phi(z) = 1 - 0.5z$  besitzt eine Nullstelle bei  $z = 2$ . Aufgrund (A.43) besitzt der betrachtete Prozess eine kausale Darstellung. Die Koeffizienten  $\psi_j$  der MA( $\infty$ ) Darstellung von  $y_t$  können mit Hilfe von (A.45) berechnet werden:

$$\begin{aligned} \psi_0 &= 1 \\ \psi_1 &= 0.4 + 0.5 \\ \psi_2 &= 0.5(0.4 + 0.5) \\ \psi_j &= 0.5^{j-1}(0.4 + 0.5) \quad j = 3, 4, \dots \end{aligned}$$

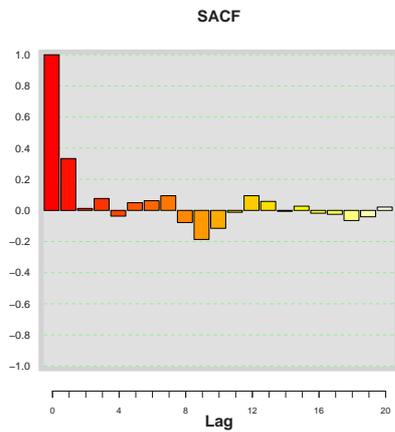
Auch das MA Polynom  $\theta(z) = 1 + 0.4z$  besitzt eine Nullstelle, die außerhalb des Einheitskreises liegt ( $z = -10/4$ ). Daraus folgt, dass  $y_t$  auch invertierbar ist mit den Koeffizienten  $\pi_j$ :

$$\begin{aligned} \pi_0 &= 1 \\ \pi_1 &= -(0.4 + 0.5) \\ \pi_2 &= -(-0.4)(0.4 + 0.5) \\ \pi_j &= -(-0.4)^{j-1}(0.4 + 0.5) \quad j = 3, 4, \dots \end{aligned}$$

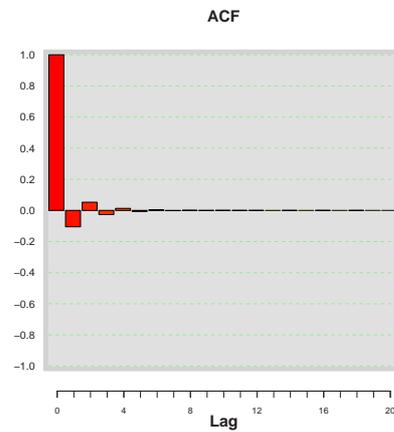
### ARMA(1,1) Prozess



(a) Simulierter ARMA(1,1) Prozess mit normalverteilten Innovationen  $\xi_t \sim NID(0, 1)$



(b) Stichprobenautokorrelationsfunktion



(c) Autokorrelationsfunktion

Abbildung A.5: Autoregressiver „Moving Average“ Prozess der Ordnung 1,1

## Anhang B.

### „Unit Root“- Integrationstests

Der Hauptteil dieser Arbeit handelt von Kointegration und kointegrierten Daten. Wie das Wort „kointegriert“ selber schon vermuten lässt, benötigt man dazu integrierte Zeitreihen (zumindest eine).

In diesem Kapitel werden vier Tests vorgestellt, die überprüfen ob eine Zeitreihe integriert der Ordnung Eins ist oder nicht ( $y_t \sim I(1)$ ). Den allgemeinen Fall, dass Daten von höherem Grade integriert sind, wird hier gänzlich ausser Acht gelassen, da in dieser Arbeit nur stationäre oder mit der Ordnung Eins integrierte Daten betrachtet werden.

Aufgrund der ausführlichen Datenanalyse im Hauptteil dieser Arbeit, in welchem alle hier erwähnten Tests nocheinmal vorkommen, sind in diesem Kapitel nur spärlich Beispiele vorhanden.

Die kritischen Werte, die für die Tests benötigt werden, wurden durch computer-gestützte Simulationen erzeugt (vor allem Monte Carlo). Dass diese numerische Auswertung der berechneten Verteilungen kein unerheblicher Aufwand ist sei hier nur kurz anhand von MacKinnon (1991) erwähnt: In Summe wurden 27 Millionen Simulationen in 1080 Experimenten durchgeführt. Der Rechenzeitaufwand betrug ungefähr neunhundert Stunden!

Bei der Beschreibung der Tests wurde nach folgendem Muster vorgegangen:

- Nullhypothese/Modelle
- Bedingungen/Voraussetzungen
- Teststatistiken
- asymptotische Konvergenz der Statistiken
- kritische Werte der Verteilungen

Nicht immer konnte diese Reihenfolge aufgrund des Leseflusses einhalten werden und auch die asymptotische Konvergenz der Statistiken ist meistens nur in Worten beschrieben, da die Formeln zu lange wären.

## B.1. Dickey - Fuller Tests

Im Jahre 1976 beschäftigte sich Fuller mit der Integration von Zeitreihen und schon 1979 veröffentlichte er gemeinsam mit Dickey den ersten „unit root“ Test. Sehr viele Tests folgten dem Dickey-Fuller (DF) Test, doch durch seine Einfachheit und durch die Qualität der Ergebnisse, vor allem des später publizierten „Augmented“ DF-Testes, ist er bis in die heutige Zeit weit verbreitet.

### B.1.1. Dickey - Fuller t-Test

Beim einfachsten aller „unit root“-Tests werden, ganz allgemein formuliert, folgende zwei Hypothesen gegeneinander getestet:

$H_0$ : Nichtstationarität gegen  $H_1$ : Stationarität

Man unterscheidet drei Fälle:

**Fall 1** Man betrachtet zunächst eine gewöhnliche AR(1)-Regression

$$y_t = \phi y_{t-1} + \xi_t \quad (\text{B.1})$$

mit  $\xi_t \sim IID(0, \sigma^2)$ . Nach einer OLS Schätzung der Parameter lässt sich die t-Statistik von  $\phi$  berechnen

$$t_T = \frac{\hat{\phi}_T - 1}{\hat{\sigma}_{\hat{\phi}_T}} \quad (\text{B.2})$$

wobei  $\hat{\sigma}_{\hat{\phi}_T}$  die OLS Standardabweichung des geschätzten Koeffizienten bezeichnet,

$$\hat{\sigma}_{\hat{\phi}_T} = \frac{\frac{1}{T-1} \sum_{t=1}^T (y_t - \hat{\phi} y_{t-1})^2}{\sum_{t=1}^T y_{t-1}^2}$$

und

$$\hat{\phi}_T = \frac{\sum_{t=1}^T y_{t-1} y_t}{\sum_{t=1}^T y_{t-1}^2}.$$

Von Interesse ist das Verhalten von  $\phi$  unter der Nullhypothese

$$H_0 : \phi = 1$$

mit der alternativen Hypothese

$$H_1 : \phi < 1.$$

Unter  $H_0$  ist das wahre Modell ein „random walk“:

$$H_0 : y_t = y_0 + \xi_1 + \xi_2 + \dots + \xi_t, \quad (\text{B.3})$$

die Teststatistik von  $\phi$  folgt jedoch nicht wie sonst üblich einer „Student“-Verteilung, sondern konvergiert gegen ein Funktional der „Brownschen“-Bewegung

$$t_T \xrightarrow{d} \frac{\frac{1}{2}(W^2(1) - 1)}{\left(\int_0^1 W^2(r) dr\right)^{1/2}}.$$

**Fall 2** Wie in Fall 1 wird der Zeitreihe  $y_t$  unterstellt, dass sie von einem „random walk“ Prozess erzeugt wird:

$$H_0 : y_t = y_{t-1} + \xi_t \quad (\text{B.4})$$

mit  $\xi_t \sim IID(0, \sigma^2)$ . Der zu schätzenden Regression wird jedoch noch ein konstanter Term hinzugefügt

$$y_t = \nu + \phi y_{t-1} + \xi_t, \quad (\text{B.5})$$

bevor wieder mittels OLS-Verfahren die Parameter geschätzt werden. Von Interesse ist nun das Verhalten von  $\nu$  und  $\phi$  unter der Nullhypothese

$$H_0 : \nu = 0, \phi = 1$$

mit der alternativen Hypothese

$$H_1 : \phi < 1.$$

Fuller zeigt in seinem Artikel, dass die t-Statistik von  $\phi$  wieder einer „Nicht-standard“-Verteilung folgt

$$t_T \xrightarrow{d} \frac{\frac{1}{2}(W^2(1) - 1) - W(1) \int_0^1 W(r) dr}{\left(\int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr\right)^2\right)^{1/2}}.$$

**Fall 3** Im letzten Fall wird angenommen, dass das wahre Modell der Zeitreihe ein „random walk“ mit Drift ist:

$$H_0 : y_t = \nu + y_{t-1} + \xi_t \quad (\text{B.6})$$

wobei  $\xi_t \sim IID(0, \sigma^2)$ . In die Regression wird ein Trend eingebunden und mittels OLS-Verfahren werden die Parameter geschätzt

$$y_t = \nu + \phi y_{t-1} + \delta t + \xi_t. \quad (\text{B.7})$$

Die Nullhypothese ist in diesem Fall

$$H_0 : \delta = 0, \phi = 1$$

mit der alternativen Hypothese

$$H_1 : \phi < 1.$$

Man beachte, dass der Wert von  $\nu$  beliebig ist. Wiederrum wird die Verteilung der Teststatistik  $t_T$  von  $\phi$  berechnet, die aber wegen der nicht mehr ganz so einfachen Gestalt hier nicht dargestellt wird (siehe zum Beispiel Hamilton (1994)).

Fuller (1976) berechnete für alle drei Modelle mit Hilfe von Monte Carlo Simulationen kritische Werte der Verteilung von  $t_T$ . Für kleine Stichprobengrößen  $T$  sind diese Werte nur dann exakt, wenn die Innovationen normalverteilt sind  $\xi_t \sim NID(0, \sigma^2)$ . Konvergiert  $T$  gegen Unendlich charakterisieren die angegebenen Werte auch die asymptotische Verteilung für nicht normalverteilte Innovationen. Heute werden meistens die kritischen Werte aufgrund der neueren, umfangreicheren und genaueren Berechnungen von MacKinnon (1991) verwendet (siehe Tabelle B.1).

Es drängt sich natürlich die Frage auf, welches der drei Modelle gewählt werden soll um die Nullhypothese einer „unit root“ zu testen. Steht eine ökonomische Theorie hinter den Daten, so wird diese für die Entscheidung ausschlaggebend sein. Fehlt eine solche Theorie, so wird wahrscheinlich eine visuelle Betrachtung der Daten von Vorteil sein. Erkennt man dabei einen Trend in der Zeitreihe wird man Modell 3 wählen, fehlt ein solcher ist Modell 2 zu bevorzugen.

**Beispiel B.1 (Beispiel zu Fall 3)** *Abbildung B.1.1 zeigt den wichtigsten amerikanischen Index, den „Dow Jones Industrial“ Index der USA, im Zeitraum vom 1.10.1928 bis 1.9.2004 mit monatlicher Beobachtungsfrequenz. Die OLS Schätzung eines AR(1) Prozesses mit einer Konstanten und einem Trend liefert folgendes Ergebnis:*

$$y_t = \underset{(0.020)}{0.0592} + \underset{(0.004556)}{0.986000}y_{t-1} + \underset{(0.00002354)}{0.00008098}t$$

*In den Klammern ist der Standardfehler des Parameters angegeben. Die OLS t-Statistik von  $H_0 : \phi = 1$  ist:*

$$t = (0.986 - 1)/0.004556 = -3.07287.$$

*Die Stichprobengröße beträgt  $T = 912$ . Vergleicht man den kritischen Wert aus der Tabelle B.1*

$$\beta_\infty + \beta_1/T + \beta_2/T^2 = -3.9638 - \frac{8.353}{912} - \frac{47.44}{912^2} = -3.972$$

*mit der Teststatistik, so kann die Nullhypothese zum 1% Niveau nicht abgelehnt werden, d.h. die Daten verhalten sich wie ein „random walk“ mit Drift. Auch beim 5% und 10% Niveau wird die Nullhypothese nicht abgelehnt ( $-3.417 < -3.072$ ,  $-3.13 < -3.072$ ). Der F-Test der gemeinsamen Nullhypothese ( $\delta = 0$  und  $\phi = 1$ ) beträgt 4313. Dieser Wert übersteigt den kritischen Wert 8.27 aus Tabelle B.2 um ein Vielfaches, die gemeinsame Nullhypothese wird daher ganz klar abgelehnt.*

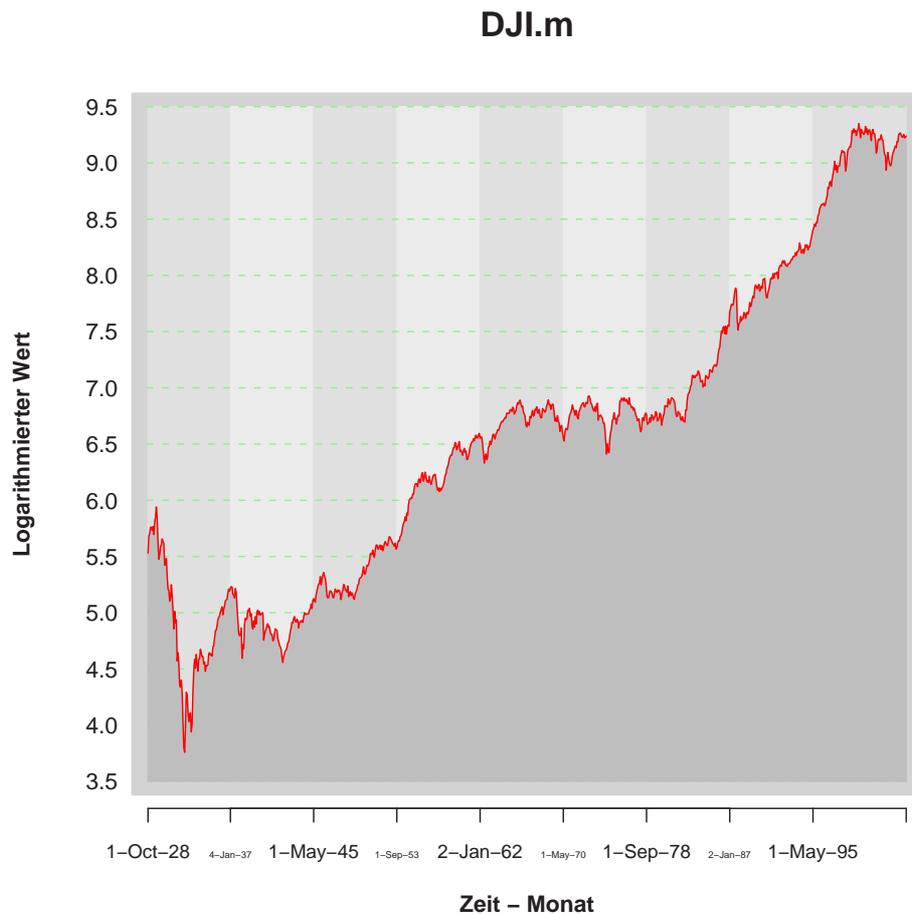


Abbildung B.1: „Dow Jones Industrial“ Index vom 1. Oktober 1928 bis 1. September 2004, monatliche Beobachtungen

### B.1.2. „Augmented“ Dickey - Fuller Test

Der DF t-Test basiert auf der Auswertung der t-Statistik einer normalen OLS-Regression, die aber nur dann gültig ist, wenn die geschätzten Residuen die „white noise“ Eigenschaften (Erwartungswert von Null, fehlende Autokorrelation, Homoskedastizität) aufweisen.

In den Gleichungen (B.5) und (B.7) werden die Residuen um Null zentriert. Ein Verstoß gegen die Annahme der Homoskedastizität ist beim Dickey-Fuller Test unproblematisch (wie Phillips (1987) zeigte). Um schließlich das Problem autokorrelierter Residuen zu berücksichtigen, kann der „Augmented“ Dickey-Fuller Test (ADF-Test) durchgeführt werden. Dabei werden verzögerte endogene Variablen als weitere Regressoren aufgenommen. Die Gleichungen (B.1), (B.5) und (B.7) verändern sich zu:

Tabelle B.1: Kritische Werte für den DF-Test

	Quantil	$\beta_\infty$	$\beta_0$	$\beta_1$
<b>Modell (1):</b>	1.0%	-2.5658	-1.960	-10.04
	5.0%	-1.9393	-0.398	-0.0
	10.0%	-1.6156	-0.181	-0.0
<b>Modell (2):</b>	1.0%	-3.4336	-5.999	-29.25
	5.0%	-2.8621	-2.738	-8.36
	10.0%	-2.5671	-1.438	-4.48
<b>Modell (3):</b>	1.0%	-3.9638	-8.353	-47.44
	5.0%	-3.4126	-4.039	-17.83
	10.0%	-3.1279	-2.418	-7.58

Für eine beliebige Stichprobengröße  $T$  berechnet sich der kritische Wert durch  $\beta_\infty + \beta_1/T + \beta_2/T^2$ .

Quelle: [MACKINNON 1990]

$$(1) : \quad y_t = \phi y_{t-1} + \sum_{j=1}^{p+1} \zeta_j \Delta y_{t-j} + \xi_t \quad (B.8)$$

$$(2) : \quad y_t = \nu + \phi y_{t-1} + \sum_{j=1}^{p+1} \zeta_j \Delta y_{t-j} + \xi_t \quad (B.9)$$

$$(3) : \quad y_t = \nu + \phi y_{t-1} + \delta t + \sum_{j=1}^{p+1} \zeta_j \Delta y_{t-j} + \xi_t. \quad (B.10)$$

Die Anzahl der verzögerten Terme, die in die Regressionsgleichungen aufgenommen werden, kann empirisch ermittelt werden. Eine Möglichkeit ist zum Beispiel, in den Gleichungen B.8, B.9 und B.10 die Anzahl der verzögerten Terme sukzessive zu erhöhen, bis die geschätzten OLS-Residuen frei von Autokorrelation sind. Diese kann durch Anwendung des „Durbin-Watson“ Tests überprüft werden. Eine alternative Vorgangsweise ist die sequenzielle Testprozedur von NG und Perron (1995). Beginnend mit einem relativ hohen Wert  $\bar{m}$  wird die Lagzahl jeweils um eine Einheit reduziert, solange bis der letzte Koeffizient auf dem 10% Niveau signifikant von Null verschieden ist.

Zur Durchführung des ADF-Tests wird wieder die OLS  $t$ -Statistik bestimmt. Diese Teststatistik besitzt dieselbe asymptotische Verteilung wie im vorherigen Test, man kann daher die kritischen Werte aus der Tabelle B.1 verwenden.

### B.1.3. Dickey - Fuller F-Test

Sowohl beim DF- als auch beim ADF-Test konzentriert sich die Aufmerksamkeit auf die Teststatistik des Parameters  $\phi$ . Im sogenannten Dickey-Fuller F-Test (Wald-Test) wird die F-Statistik der gemeinsamen Nullhypothese  $\phi = 0$  **und**  $\nu = 0$  der Gleichungen (B.5) und (B.9) bzw.  $\phi = 0$  **und**  $\delta = 0$  der Gleichungen (B.7) und (B.10) berechnet und getestet. Diese Statistik entspricht wieder einer „Nichtstandard“-Verteilung. Kritische Werte sind in Tabelle (B.2) dargestellt.

Tabelle B.2: Kritische Werte für den Dickey-Fuller F-Test

<b>Modell (2)</b>	$T$	10.0%	5.0%	2.5%	1.0%
	25	4.12	5.18	6.30	7.88
	50	3.94	4.86	5.80	7.06
	100	3.86	4.71	5.57	6.70
	250	3.81	4.63	5.45	6.52
	500	3.79	4.61	5.41	6.47
	$\infty$	3.78	4.59	5.38	6.43
<b>Modell (3)</b>	$T$	10.0%	5.0%	2.5%	1.0%
	25	5.91	7.24	8.65	10.61
	50	5.61	6.73	7.81	9.31
	100	5.47	6.49	7.44	8.73
	250	5.39	6.34	7.25	8.43
	500	5.36	6.30	7.20	8.34
	$\infty$	5.27	7.16	6.25	5.34

Hinweis: Für Stichproben die keiner angegebenen Größe entsprechen werden einfach die kritischen Werte interpoliert.

Quelle: [DICKEY AND FULLER 1981]

## B.2. Kwiatkowski, Phillips, Schmidt und Shin Test

Im Gegensatz zu den meisten anderen Tests ist beim Kwiatkowski, Phillips, Schmidt und Shin (KPSS) Test die Null-Hypothese durch die Stationarität und damit die Alternativhypothese durch den „unit root“ Fall definiert.

Die vier Autoren betrachten zwei Modelle:

### Modell 1: Level Stationarität

Die beobachtete Zeitreihe setzt sich aus einem „random walk“ und einem stationären Prozess zusammen.

### Modell 2: Trend Stationarität

Im zweiten Fall kommt zu Modell 1 noch ein deterministischer Trend hinzu.

Tabelle B.3: Zusammenfassung der Dickey-Fuller Tests

**Fall 1:**

Geschätzte Regression:  $y_t = \phi y_{t-1} + \xi_t \quad (+ \sum_{j=1}^{p+1} \zeta_j \Delta y_{t-j})$

Wahrer Prozess:  $y_t = y_{t-1} + \xi_t \quad u_t \sim NID(0, \sigma^2)$

$(\hat{\phi}_T - 1)/(\hat{\sigma}_{\hat{\phi}_T})$  hat die in Tabelle B.1, Fall 1 beschriebene Verteilung.

**Fall 2:**

Geschätzte Regression:  $y_t = \nu + \phi y_{t-1} + \xi_t \quad (+ \sum_{j=1}^{p+1} \zeta_j \Delta y_{t-j})$

Wahrer Prozess:  $y_t = y_{t-1} + \xi_t \quad u_t \sim NID(0, \sigma^2)$

$(\hat{\phi}_T - 1)/(\hat{\sigma}_{\hat{\phi}_T})$  hat die in Tabelle B.1, Fall 2 beschriebene Verteilung.

Der OLS F-Test der gemeinsamen Hypothese von  $\nu = 0$  und  $\phi = 1$  hat die in Tabelle B.2, Fall 2 beschriebene Verteilung.

**Fall 3:**

Geschätzte Regression:  $y_t = \nu + \phi y_{t-1} + \delta t + \xi_t \quad (+ \sum_{j=1}^{p+1} \zeta_j \Delta y_{t-j})$

Wahrer Prozess:  $y_t = \nu + y_{t-1} + \xi_t \quad \nu$  beliebig,  $u_t \sim NID(0, \sigma^2)$

$(\hat{\phi}_T - 1)/(\hat{\sigma}_{\hat{\phi}_T})$  hat die in Tabelle B.1, Fall 3 beschriebene Verteilung.

Der OLS F-Test der gemeinsamen Hypothese von  $\delta = 0$  und  $\phi = 1$  hat die in Tabelle B.2, Fall 3 beschriebene Verteilung.

**Hinweise:**

- ☞ Wird zur "Geschätzten" Regressionen der Term in Klammer hinzugenommen (ADF-Test), so gelten asymptotisch die selben kritischen Werte
- ☞  $\hat{\phi}_T$  ist die OLS Schätzung von  $\phi$
- ☞  $(\hat{\phi}_T - 1)/(\hat{\sigma}_{\hat{\phi}_T})$  ist die OLS t-Statistik für  $\phi = 1$
- ☞ Sind die Innovationen  $u_t$  nicht normalverteilt gelten die Ergebnisse der Tabellen B.1 und B.2 nur für große Stichproben  $T$
- ☞ "Wahrer Prozess" beschreibt die Nullhypothese unter der die Verteilung von  $t_T$  berechnet wurde.

Das ergibt:

$$y_t = r_t + \xi_t, \quad t = 1, 2, \dots, T \quad (\text{B.11})$$

$$y_t = \delta t + r_t + \xi_t, \quad t = 1, 2, \dots, T \quad (\text{B.12})$$

mit  $y_t$  die beobachtete Zeitreihe,  $r_t = r_{t-1} + u_t$  ein „random walk“ (mit  $r_0$  als Startwert), wobei  $u_t \sim IID(0, \sigma_u^2)$  ist, und  $\xi_t$  ein stationärer Prozess.

Unter der Null-Hypothese,  $\sigma_u^2 = 0$ , gilt:

$$H_0 : y_t = r_0 + \xi_t \quad \text{level stationär, oder} \quad (\text{B.13})$$

$$H_0 : y_t = \delta t + r_0 + \xi_t \quad \text{trend stationär,} \quad (\text{B.14})$$

während  $y_t$  unter der Alternative,  $\sigma_u^2 > 0$ , infolge der „random walk“-Komponente integriert der Ordnung Eins ist.

Nimmt man zunächst an, dass die  $u_t \sim NID(0, \sigma_u^2)$  und die  $\xi_t \sim IID(0, \sigma_\xi^2)$  sind, dann ergibt sich ein sehr einfaches Resultat: Sei  $\hat{\xi}_t$ ,  $t = 1, 2, \dots, T$  die Menge der Residuen von  $y_t = r_0 + \xi_t$  oder analog dazu von  $y_t = \delta t + r_0 + \xi_t$  und sei  $\hat{\sigma}_\xi^2$  der Standardschätzer für die Fehlervarianz der betrachteten Regression (d.h.  $\sigma_\xi^2 = 1/T \sum_{i=1}^T \hat{\xi}_i^2$ ), dann ergibt sich die Teststatistik durch

$$LM = \sum_{t=1}^T \frac{S_t^2}{\hat{\sigma}_\xi^2}$$

mit

$$S_t = \sum_{i=1}^t \hat{\xi}_i, \quad t = 1, 2, \dots, T$$

die Partialsumme der Residuen.

Die Annahmen aus dem vorhergehenden Absatz sind jedoch in der Realität eher unwahrscheinlich, da die Fehlervariable  $\xi_t$  in den meisten Fällen zeitlich von sich selbst abhängig ist. Kwiatkowski & Co verweisen für den Fall von zeitlich abhängigen  $\xi_t$  auf die Annahmen von Phillips und Perron (1988, S. 336), die im nächsten Kapitel auf Seite 88 beschrieben sind, oder Phillips und Solo (1989, Theorem 3.3, 3.14). Mit diesen Annahmen erhält man für den level-stationären Fall die KPSS Teststatistik

$$\hat{\eta}_\mu = \frac{1}{T^2} \sum_{t=1}^T \frac{\hat{S}_t^2}{s^2(l)} \quad (\text{B.15})$$

mit

$$\begin{aligned} \hat{\xi}_t &= y_t - \hat{r}_0 \\ \hat{S}_t &= \sum_{i=1}^t \hat{\xi}_i \\ s^2(l) &= \frac{1}{T} \sum_{t=1}^T \hat{\xi}_t^2 + \frac{2}{T} \sum_{s=1}^l w(s, l) \sum_{t=s+1}^T \hat{\xi}_t \hat{\xi}_{t-s} \\ w(s, l) &= 1 - \frac{s}{l+1} \end{aligned}$$

und für den trend-stationären Fall

$$\hat{\eta}_\tau = \frac{1}{T^2} \sum_{t=1}^T \frac{\hat{S}_t^2}{s^2(l)} \quad (\text{B.16})$$

mit

$$\begin{aligned}\hat{\xi}_t &= y_t - \hat{\xi}t - \hat{r}_0 \\ \hat{S}_t &= \sum_{i=1}^t \hat{\xi}_i \\ s^2(l) &= \frac{1}{T} \sum_{t=1}^T \hat{\xi}_t^2 + \frac{2}{T} \sum_{s=1}^l w(s,l) \sum_{t=s+1}^T \hat{\xi}_t \hat{\xi}_{t-s} \\ w(s,l) &= 1 - \frac{s}{l+1}.\end{aligned}$$

Um konsistente Schätzer zu erhalten muss die Zahl der Lags  $l$  langsam mit der Stichprobengröße wachsen (z.B.:  $l = 12\sqrt[4]{T/100}$ ).

Die Teststatistiken konvergieren asymptotisch gegen das Integral der „Brownschen“ Brücke (level-stationärer Fall) bzw. gegen das in Integral der sogenannten „second level Brownian Bridge“ (trend-stationärer Fall), i.Z.

$$\begin{aligned}\hat{\eta}_\mu &\longrightarrow \int_0^1 V^2(r)dr \\ \hat{\eta}_\tau &\longrightarrow \int_0^1 V_2^2(r)dr.\end{aligned}$$

Für genauere Information siehe Kwiatkowski, Phillips, Schmidt und Shin (1991, S. 164-167). Tabelle B.4 gibt die asymptotisch kritischen Werte für die beiden Teststatistiken wieder.

Tabelle B.4: Kritische Werte für die KPSS-Teststatistiken  $\hat{\eta}_\mu$  und  $\hat{\eta}_\tau$

Quantile der Verteilung von $\int_0^1 V^2(r)dr$				
Kritischer Level:	0.100	0.050	0.025	0.010
Kritischer Wert:	0.347	0.463	0.574	0.739
Quantile der Verteilung von $\int_0^1 V_2^2(r)dr$				
Kritischer Level:	0.100	0.050	0.025	0.010
Kritischer Wert:	0.119	0.146	0.176	0.216

Quelle: [D. KWIATKOWSKI AND SHIN 1992]

### B.3. Phillips - Perron Test

Der Phillips-Perron Test stellt neben dem KPSS-Test und dem ADF-Test eine weitere Möglichkeit dar, zeitlich von einander abhängige „Residuen“  $\xi_t$  zu berücksichtigen.

tigen. Dazu werden folgende allgemeine Annahmen über das Verhalten dieser  $\xi_t$ 's getroffen:

1.  $E(\xi_t) = 0 \quad \forall t$ ;
2.  $\sup_t E|\xi_t|^{\beta+\kappa} < \infty$  für  $\beta > 2$  und  $\kappa > 0$ ;
3. für  $T \rightarrow \infty$  existiert  $\sigma^2 = \lim E(T^{-1}S_T^2)$  und  $\sigma^2 > 0$ , wobei  $S_t = \xi_1 + \dots + \xi_t$ ;
4.  $\{\xi_t\}$  ist stark gemischt mit gemischten Koeffizienten  $\nu_m$  die  $\sum_{m=1}^{\infty} \nu_m^{1-2/\beta} < \infty$  erfüllen.

Weitere Informationen zu diesen Bedingungen siehe Phillips (1987a), für die genaue Definition von „stark gemischt mit gemischten Koeffizienten“ siehe Hall & Heyde (1980, S. 132).

Der Zeitreihe wird unterstellt, dass sie durch

$$H_0: y_t = \phi y_{t-1} + \xi_t \quad (t = 1, 2, \dots) \quad (\text{B.17})$$

mit  $\phi = 1$

erzeugt wird.

Mit vorherigen Bedingungen werden zwei Kleinst-Quadrat Regressionen betrachtet

$$y_t = \hat{\nu} + \hat{\phi} y_{t-1} + \hat{\xi}_t \quad \text{und} \quad (\text{B.18})$$

$$y_t = \tilde{\nu} + \tilde{\delta}(t - \frac{1}{2}T) + \tilde{\phi} y_{t-1} + \tilde{\xi}_t. \quad (\text{B.19})$$

Um den Einfluss von seriell korrelierten und heteroskedastisch verteilten Innovationen  $\xi_t$  zu berücksichtigen, wurden die t-Statistiken der OLS geschätzten Parameter der beiden Regressionen transformiert. Das Ergebnis sind die von Phillips & Perron genannten Z-Statistiken:

$$Z(t_{\hat{\nu}}) = \left(\frac{\hat{s}}{\hat{\sigma}_{Tl}}\right)t_{\hat{\phi}} - \frac{\hat{\lambda}'\hat{\sigma}_{Tl}}{\hat{m}_{yy}^{\frac{1}{2}}}, \quad (\text{B.20})$$

$$Z(t_{\tilde{\nu}}) = \left(\frac{\tilde{s}}{\tilde{\sigma}_{Tl}}\right)t_{\tilde{\phi}} - \frac{\tilde{\lambda}'\tilde{\sigma}_{Tl}}{M^{\frac{1}{2}}} \quad (\text{B.21})$$

mit

$$m_{yy} = T^{-2} \sum_{t=1}^T y_t^2, \quad \bar{m}_{yy} = T^{-2} \sum_{t=1}^T (y_t - \bar{y})^2, \quad m_y = T^{-3/2} \sum_{t=1}^T y_t, \quad m_{ty} = T^{-5/2} \sum_{t=1}^T ty_t,$$

$$M = (1 - T^{-2})m_{yy} - 12m_{ty}^2 + 12(1 + T^{-1})m_{ty}m_y - (4 + 6T^{-1} + 2T^{-2})m_y^2,$$

$$\hat{\lambda} = \frac{1}{2}(\hat{\sigma}_{Tl}^2 - s^2), \quad \hat{\lambda}' = \hat{\lambda}/\hat{\sigma}_{Tl}^2, \quad \tilde{\lambda} = \frac{1}{2}(\tilde{\sigma}_{Tl}^2 - s^2), \quad \tilde{\lambda}' = \tilde{\lambda}/\tilde{\sigma}_{Tl}^2$$

$$\hat{\sigma}_{Tl}^2 = T^{-1} \sum_{t=1}^T \hat{\xi}_2^t + 2T^{-1} \sum_{s=1}^l w_{sl} \sum_{t=s+1}^T \hat{\xi}_t \hat{\xi}_{t-s}$$

$$\tilde{\sigma}_{Tl}^2 = T^{-1} \sum_{t=1}^T \tilde{\xi}_2^t + 2T^{-1} \sum_{s=1}^l w_{sl} \sum_{t=s+1}^T \tilde{\xi}_t \tilde{\xi}_{t-s}$$

$$\hat{s} = \left( (T-k)^{-1} \sum_{t=1}^T \hat{\xi}_t \right)^{1/2}, \quad \tilde{s} = \left( (T-k)^{-1} \sum_{t=1}^T \tilde{\xi}_t \right)^{1/2},$$

wobei

$$w_{sl} = 1 - \frac{s}{l+1}.$$

Wie beim KPSS-Test muss die Zahl der Lags  $l$  langsam mit der Stichprobengröße wachsen (z.B.:  $l = 12\sqrt[4]{T/100}$ ).  $t_{T\hat{\phi}}$  und  $t_{T\tilde{\phi}}$  sind die t-Statistiken des Parameters  $\nu$  in den Gleichungen B.18 und B.19.

Die beiden Standardfehler  $\hat{s}$  und  $\tilde{s}$ , die eine große Rolle bei den üblichen t-Statistiken spielen, wurden um die allgemeinen Schätzer  $\hat{\sigma}_{Tl}$  und  $\tilde{\sigma}_{Tl}$  korrigiert. Die Differenzterme  $\hat{\sigma}_{Tl}^2 - s^2$  und  $\tilde{\sigma}_{Tl}^2 - \tilde{s}^2$  neutralisieren asymptotisch den Effekt der seriellen Korrelation in den Z-Statistiken.

Die von Phillips & Perron entwickelten Teststatistiken konvergieren gegen Funktionale der "Brownschen Bewegung" und sind dieselben wie von Fuller (1976) und Dicky & Fuller (1981) tabelliert. Dadurch kann die Tabelle B.1 (Modell (2) & Modell (3)) auf Seite 83 zur Berechnung der kritischen Werte für diesen Test verwendet werden.

**Beispiel B.2** Die OLS-Regression im PP-Test liefert mit den Daten von Beispiel B.1 folgendes Ergebnis:

$$y_t = \underset{(0.02998)}{0.09612} + \underset{(0.00002354)}{0.00008098} \left(t - \frac{T}{2}\right) + \underset{(0.004556)}{0.986000} y_{t-1}$$

In den Klammern ist der Standardfehler des Parameters angegeben. Die Stichprobengröße beträgt  $T = 912$ . Mit Hilfe der Regression lassen sich die Zwischenergebnisse anschreiben:

$$M = (1 - T^{-2})0.002034238 - 120.1214922^2 + 12(1 + T^{-1})(0.1214922)(0.2178311) - (4 + 6T^{-1} + 2T^{-2})0.2178311^2 = 0.0001720863$$

$$\tilde{\lambda}' = \frac{\frac{1}{2}(\tilde{\sigma}_{Tl}^2 - s^2)}{\tilde{\sigma}_{Tl}^2} = \frac{\frac{1}{2}(0.003198656 - 0.002981115)}{0.003198656} = 0.03400505$$

$$\begin{aligned}\tilde{\sigma}_{Tl}^2 &= T^{-1} \sum_{t=1}^T \tilde{u}_2^t + 2T^{-1} \sum_{s=1}^l w_{sl} \sum_{t=s+1}^T \tilde{u}_t \tilde{u}_{t-s} = 0.003198656 \\ \tilde{s}^2 &= (T-k)^{-1} \sum_{t=1}^T \tilde{u}_t^2 = 0.002981115 \\ t_{\tilde{\nu}} &= -3.076470.\end{aligned}$$

Die Phillips Perron Z-Statistik betragt:

$$Z(t_{\tilde{\nu}}) = \left( \frac{0.002981^{1/2}}{0.00319} \right) (-3.07647) - \frac{(0.03400)(0.00319)}{0.00017^{1/2}} = -3.1166.$$

Da  $-3.13 < -3.1166$  wird auch beim PP-Test die Nullhypothese angenommen (10% Niveau).

## B.4. Zivot - Andrews Test

Im Jahre 1987 entwickelte Perron eine Prozedur um die Nullhypothese, dass gegebene Daten eine Einheitswurzel mit Drift besitzen und ein Strukturbruch zum Zeitpunkt  $1 < T_B < T$  eingetreten ist, gegen die Alternativhypothese, dass die Zeitreihen stationar um einen deterministischen Trend mit einer exogenen anderung dieser Trendfunktion zum Zeitpunkt  $T_B$  sind, zu testen.

Zivot und Andrews griffen diese Untersuchungen auf. Im Gegensatz zu Perron, der den Strukturbruchpunkt selbst festlegte, nahmen die beiden Autoren an, dass sie nicht genau wissen wann ein solcher eintritt. Unter ihrer Nullhypothese folgen die Daten einem „unit root“ Prozess mit Drift, der aber keine strukturellen anderungen besitzt,

$$H_0 : y_t = \mu + y_{t-1} + e_t, \quad (\text{B.22})$$

mit  $e_t$  einem stationaren Prozess. Die Alternative ist je nach Regressionsmethode ein stationarer, levelstationarer oder trend-stationarer Prozess der aber nun einen Bruch in der Trendfunktion besitzen darf.

Perrons erweiterter Dickey-Fuller Teststrategie folgend, werden drei Regressionsgleichungen (A,B und C) auf die mogliche Prasenz einer „unit root“ untersucht:

$$\begin{aligned}(\mathbf{A}) : \quad y_t &= \hat{\nu}^A + \hat{\theta}^A DU_t(\lambda) + \hat{\delta}^A t + \hat{\phi}^A y_{t-1} + \\ &+ \sum_{j=1}^k \hat{\zeta}_j^A \Delta y_{t-j} + \hat{\xi}_t, \quad (\text{B.23})\end{aligned}$$

$$\begin{aligned}(\mathbf{B}) : \quad y_t &= \hat{\nu}^B + \hat{\delta}^B t + \hat{\gamma}^B DT_t^*(t\lambda) + \hat{\phi}^B y_{t-1} + \\ &+ \sum_{j=1}^k \hat{\zeta}_j^B \Delta y_{t-j} + \hat{\xi}_t \quad (\text{B.24})\end{aligned}$$

und

$$\begin{aligned}
 \text{(C)} : \quad y_t &= \hat{\nu}^C + \hat{\theta}^C DU_t(\lambda) + \hat{\delta}^C t + \hat{\gamma}^C DT_t^*(\lambda) + \hat{\phi}^C y_{t-1} + \\
 &\quad + \sum_{j=1}^k \hat{\zeta}_j^C \Delta y_{t-j} + \hat{\xi}_t \quad (\text{B.25})
 \end{aligned}$$

wobei  $DU_t = 1$  wenn  $t > T\lambda$ , 0 sonst und  $DT_t^*(\lambda) = t - T\lambda$  wenn  $t > T\lambda$ , 0 sonst.

Die erste Gleichung berücksichtigt durch den Term  $\hat{\theta}^A DU_t(\lambda)$  einen Levelbruch der Zeitreihe, die zweite Zeile berücksichtigt durch  $\hat{\gamma}^B DT_t^*(\lambda)$  eine Änderung in der Wachstumsrate und die dritte Gleichung vereint die ersten beiden Veränderungen.  $\lambda = T_B/T$  stellt den Zeitpunkt des Strukturbruches dar. Die  $k$  zusätzlichen Regressoren sollen störende Abhängigkeiten der Zeitreihe zu sich selbst eliminieren.

Getestet wird nun  $\phi^i = 1$  ( $i = A, B$  und  $C$ ) durch Ermittlung der t-Statistik, wobei für jedes einzelne  $\lambda$  dieser Wert berechnet und von dieser Menge das Minimum (i.Z.  $\inf_{\lambda \in \Lambda} t_{\hat{\phi}^i}(\lambda)$ ) gewählt wird.

Die Vorgangsweise von Zivot und Andrews ist folgende: Man bestimme für jeden Strukturbruchpunkt  $\lambda$  ( $\lambda$  reicht dabei von  $2/T$  bis  $(T-1)/T$ ) die Anzahl der Extraregressoren  $k$  mittels Rückwärtsrechnung; beginnend mit  $k = \bar{k}$  wähle man jenes  $k$  wo die t-Statistik von  $\hat{\zeta}_k$  absolut größer als 1.6 und für  $l > k$  die t-Statistik von  $\hat{\zeta}_l$  kleiner als 1.6 ist. Für jedes einzelne  $\lambda$  und dazugehörige  $k$  schätze man nun die Parameter der drei Gleichungen und bestimme die jeweilige t-Statistik von  $\phi^i$ . Aus den erhaltenen Mengen wähle man nun das minimalste Element und vergleiche es mit den kritischen Werten in Tabelle B.5. Gilt

$$\inf_{\lambda \in \Lambda} t_{\hat{\phi}^i}(\lambda) < \kappa_{\text{inf},\phi}^i, \quad i = A, B \text{ und } C \quad (\text{B.26})$$

wobei  $\kappa_{\text{inf},\phi}^i$  den kritischen Wert der asymptotischen Verteilung von  $\inf_{\lambda \in \Lambda} t_{\hat{\phi}^i}(\lambda)$  bezeichne, dann wird die Nullhypothese verworfen.

Tabelle B.5: Kritische Werte der ZA-Teststatistik  $\inf_{\lambda \in \Lambda} t_{\hat{\phi}^i}(\lambda)$   $i = A, B$  und  $C$

	Quantile			
	1.0%	2.5%	5.0%	10.0%
<b>Modell (A):</b>	-5.34	-5.02	-4.80	-4.58
<b>Modell (B):</b>	-4.93	-4.67	-4.42	-4.11
<b>Modell (C):</b>	-5.57	-5.30	-5.08	-4.82

Quelle: [ZIVOT AND ANDREWS 1992]

# Anhang C.

## Ergänzende Tabellen

Tabelle C.1: Wahrscheinlichkeitswerte für die Annahme der Null-Hypothese des KPSS-Tests

lag	Erdöl	XOM	BP	RD	CVX	SC
1	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
2	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
3	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
4	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
5	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
6	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
7	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
8	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
9	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
10	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
11	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
12	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
13	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
14	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
15	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
16	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
17	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
18	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
19	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
20	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
21	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
22	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
23	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
24	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
25	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
26	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
27	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
28	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
29	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
30	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot). Die Wahrscheinlichkeitswerte werden im Intervall (0.0100, 0.1000) genau tabelliert, Werte die außerhalb dieses Bereiches liegen werden mit der Intervallgrenze angegeben.

Tabelle C.2: Wahrscheinlichkeitswerte für die Annahme der Null-Hypothese des KPSS-Tests

lag	Erdöl	XOM	BP	RD	CVX	SC
1	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
2	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
3	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
4	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
5	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
6	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
7	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
8	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
9	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
10	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
11	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
12	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
13	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
14	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
15	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
16	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
17	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
18	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
19	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
20	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
21	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
22	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
23	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
24	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
25	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
26	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
27	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
28	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
29	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
30	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot). Die Wahrscheinlichkeitswerte werden im Intervall (0.0100, 0.1000) genau tabelliert, Werte die außerhalb dieses Bereiches liegen werden mit der Intervallgrenze angegeben.

Tabelle C.3: Wahrscheinlichkeitswerte  
für die Annahme der Null-  
Hypothese des PP-Tests

---

Daten: Öl.Auswahl  
 Test-Methode: Phillips-Perron Test  
 Geschätzte Regression:  $y_t = \nu + \phi y_{t-1} + \xi_t$   
 Null-Hypothese:  $y_t = \nu + y_{t-1} + \xi_t$   
 Anzahl der Beobachtungen: 4049

---

Name	p-Wert	lag
Erdöl	0.1000	30
Exxon Mobil	0.1000	30
British Petrol	0.1000	30
Chevron	0.1000	30
Royal Dutch	0.0205	30
SC	0.0704	30

---

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot). Die Wahrscheinlichkeitswerte werden im Intervall (0.0100, 0.1000) genau tabelliert, Werte die außerhalb dieses Bereiches liegen werden mit der Intervallgrenze angegeben.

Tabelle C.4: Wahrscheinlichkeitswerte  
für die Annahme der Null-  
Hypothese des PP-Tests

---

Daten: Öl.Auswahl  
 Test-Methode: Phillips-Perron Test  
 Geschätzte Regression:  $y_t = \nu + \delta t + \phi y_{t-1} + \xi_t$   
 Null-Hypothese:  $y_t = \nu + y_{t-1} + \xi_t$   
 Anzahl der Beobachtungen: 4049

---

Name	p-Wert	lag
Erdöl	0.1000	30
Exxon Mobil	0.1000	30
British Petrol	0.1000	30
Chevron	0.1000	30
Royal Dutch	0.1000	30
SC	0.1000	30

---

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot). Die Wahrscheinlichkeitswerte werden im Intervall (0.0100, 0.1000) genau tabelliert, Werte die außerhalb dieses Bereiches liegen werden mit der Intervallgrenze angegeben.

Tabelle C.5: Wahrscheinlichkeitswerte für die Annahme der Null-Hypothese des KPSS-Tests

lag	Gold	Silber	AAUK	NEM	ABX	PDG
1	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
2	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
3	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
4	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
5	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
6	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
7	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
8	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
9	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
10	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
11	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
12	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
13	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
14	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
15	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
16	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
17	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
18	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
19	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
20	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
21	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
22	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
23	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
24	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
25	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
26	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
27	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
28	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
29	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
30	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
31	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot). Die Wahrscheinlichkeitswerte werden im Intervall (0.0100, 0.1000) genau tabelliert, Werte die außerhalb dieses Bereiches liegen werden mit der Intervallgrenze angegeben.

Tabelle C.6: Wahrscheinlichkeitswerte für die Annahme der Null-Hypothese des KPSS-Tests

lag	Gold	Silber	AAUK	NEM	ABX	PDG
1	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
2	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
3	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
4	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
5	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
6	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
7	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
8	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
9	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
10	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
11	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
12	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
13	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
14	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
15	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
16	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
17	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
18	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
19	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
20	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
21	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
22	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
23	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
24	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
25	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
26	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
27	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
28	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
29	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
30	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
31	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot). Die Wahrscheinlichkeitswerte werden im Intervall (0.0100, 0.1000) genau tabelliert, Werte die außerhalb dieses Bereiches liegen werden mit der Intervallgrenze angegeben.

Tabelle C.7: Wahrscheinlichkeitswerte  
für die Annahme der Null-  
Hypothese des PP-Tests

---

Daten: GoldSilber.Auswahl  
 Test-Methode: Phillips-Perron Test  
 Geschätzte Regression:  $y_t = \nu + \phi y_{t-1} + \xi_t$   
 Null-Hypothese:  $y_t = \nu + y_{t-1} + \xi_t$   
 Anzahl der Beobachtungen: 4049

---

Name	p-Wert	lag
Gold	0.1000	30
Silber	0.0230	30
AAUK	0.1000	30
NEM	0.1000	30
ABX	0.1000	30
PDG	0.1000	30

---

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot). Die Wahrscheinlichkeitswerte werden im Intervall (0.0100, 0.1000) genau tabelliert, Werte die außerhalb dieses Bereiches liegen werden mit der Intervallgrenze angegeben.

Tabelle C.8: Wahrscheinlichkeitswerte  
für die Annahme der Null-  
Hypothese des PP-Tests

---

Daten: GoldSilber.Auswahl  
 Test-Methode: Phillips-Perron Test  
 Geschätzte Regression:  $y_t = \nu + \delta t + \phi y_{t-1} + \xi_t$   
 Null-Hypothese:  $y_t = \nu + y_{t-1} + \xi_t$   
 Anzahl der Beobachtungen: 4049

---

Name	p-Wert	lag
Gold	0.1000	30
Silber	0.0761	30
AAUK	0.0949	30
NEM	0.1000	30
ABX	0.1000	30
PDG	0.1000	30

---

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot). Die Wahrscheinlichkeitswerte werden im Intervall (0.0100, 0.1000) genau tabelliert, Werte die außerhalb dieses Bereiches liegen werden mit der Intervallgrenze angegeben.

Tabelle C.9: Teststatistik und kritische Werte des Phillips&Oulliaris - Kointegrationstests

<b>Dimension des VAR Prozesses: 2</b>			
<b>Nullhypothese: keine Kointegration</b>			
<b>Anzahl der Beobachtungen: 4049</b>			
kritische Werte	linear	constant	none
10.0%	71.958	47.587	33.926
7.5%	75.734	50.751	36.664
5.0%	81.381	55.220	40.821
2.5%	90.294	61.4556	47.245
1.0%	102.0167	71.927	55.191
Daten	Teststatistik		
Erdöl, KOM	1.8163	1.8145	1.6751
Erdöl, BP	1.6394	1.6260	1.4919
Erdöl, CVX	1.5456	1.5441	1.4393
Erdöl, RD	4.1697	4.1171	3.8179
Erdöl, SC	1.4363	1.4487	1.3417
Gold, AAUK	3.4963	3.3034	3.2826
Gold, NEM	30.3219	29.9938	29.9320
Gold, ABX	4.8300	4.8403	4.6987
Gold, PDG	20.7716	20.7358	20.7244
Silber, AUUK	4.1851	3.9053	3.8541
Silber, NEM	24.5377	24.6937	24.8663
Silber, ABX	6.4905	6.3995	6.5283
Silber, PDG	24.3490	24.4829	24.5876

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% **blau**, 5% **grün** und 1% **rot**).

Tabelle C.10: Teststatistik und kritische Werte des Johansen - Kointegrationstests

Daten: Gold, AAUK, NEM, ABX, PDG									
Beobachtungszeitraum: 18.08.1988 - 04.06.2004									
Anzahl der Beobachtungen: 4049									
verwendetes Modell: $M^*$									
$p$	Teststatistik (trace)					Kritische Werte			
	1	2	5	9	16	10%	5%	2.5%	1%
$r \leq 3$	4.0465	4.5438	4.8746	4.6302	3.8918	10.56	12.39	14.13	16.39
$r \leq 2$	14.1650	15.1063	14.1638	13.6307	12.5761	22.95	25.47	27.87	30.65
$r \leq 1$	29.8595	30.2504	27.9211	27.5616	26.3694	39.08	42.20	45.04	48.59
$r = 0$	51.1689	53.4606	53.0772	53.0121	52.3133	58.96	62.61	66.04	70.22
$r \leq 1$	<b>112.0186</b>	<b>88.8302</b>	<b>83.6859</b>	<b>82.9100</b>	81.6712	82.68	86.96	90.87	95.38

Daten: Gold, AAUK, NEM, ABX, PDG									
Beobachtungszeitraum: 18.08.1988 - 04.06.2004									
Anzahl der Beobachtungen: 4049									
verwendetes Modell: $M_1$									
$p$	Teststatistik (trace)					Kritische Werte			
	1	2	5	9	16	10%	5%	2.5%	1%
$r \leq 3$	<b>3.6981</b>	<b>4.0417</b>	<b>4.5105</b>	<b>4.4942</b>	<b>3.5669</b>	2.71	3.84	5.02	6.64
$r \leq 2$	8.1258	8.8850	9.8063	9.2226	7.4588	13.31	15.34	17.24	19.69
$r \leq 1$	18.9031	20.6055	20.1090	19.6578	18.4964	26.70	29.38	31.76	34.87
$r = 0$	35.5485	35.7511	34.0198	33.5959	33.2025	43.84	47.21	50.19	53.91
$r \leq 2$	<b>95.0647</b>	<b>70.2123</b>	63.0550	62.9567	62.4151	64.74	68.68	72.21	76.37

Daten: Gold, AAUK, NEM, ABX, PDG									
Beobachtungszeitraum: 18.08.1988 - 04.06.2004									
Anzahl der Beobachtungen: 4049									
verwendetes Modell: $M_1^*$									
$p$	Teststatistik (trace)					Kritische Werte			
	1	2	5	9	16	10%	5%	2.5%	1%
$r \leq 3$	3.9115	4.2835	4.6845	4.6079	3.7011	7.50	9.13	10.73	12.73
$r \leq 2$	8.9449	9.3655	10.0876	9.3896	7.6549	17.79	19.99	22.07	24.74
$r \leq 1$	23.080	23.9433	22.3284	22.7355	20.9314	31.88	34.80	37.56	40.84
$r = 0$	41.1474	41.3567	39.7828	39.5235	39.7652	49.92	53.42	56.57	60.42
$r \leq 2$	<b>100.7007</b>	<b>76.2236</b>	69.0168	69.4682	69.4319	71.66	75.74	79.60	83.93

Daten: Gold, AAUK, NEM, ABX, PDG									
Beobachtungszeitraum: 18.08.1988 - 04.06.2004									
Anzahl der Beobachtungen: 4049									
verwendetes Modell: $M_2$									
$p$	Teststatistik (trace)					Kritische Werte			
	1	2	5	9	16	10%	5%	2.5%	1%
$r \leq 3$	0.3746	0.3034	0.2732	0.2894	0.3648	2.98	4.14	5.30	7.02
$r \leq 2$	5.0658	5.1384	5.4242	5.0458	4.0676	10.35	12.21	13.94	16.16
$r \leq 1$	13.6210	12.6826	12.5028	11.4567	10.0957	21.58	24.08	26.42	29.19
$r = 0$	29.3183	28.1920	26.6837	26.7147	27.2495	36.58	29.71	42.59	46.00
$r \leq 2$	50.6846	48.7731	45.1495	46.3211	48.0519	55.54	59.24	62.68	66.71

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% **blau**, 5% **grün** und 1% **rot**).

Tabelle C.11: Teststatistik und kritische Werte des Johansen - Kointegrationstests

Klasse:	M*		M <sub>1</sub>		M <sub>1</sub> *		M <sub>2</sub>	
Niveau	krit. Werte		krit. Werte		krit. Werte		krit. Werte	
10.0%	10.56	22.95	2.71	13.31	7.50	17.79	2.98	10.35
5.0%	12.39	25.47	3.84	15.34	9.13	19.99	4.14	12.21
2.5%	14.13	27.87	5.02	17.24	10.73	22.07	5.30	13.94
1.0%	16.39	30.65	6.64	19.69	12.73	24.74	7.02	16.16
<b>Daten: Erdöl, XOM</b>								
<b>Anzahl der Beobachtungen: 4049</b>								
p	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$
1	3.4144	22.0744	2.7395	12.4778	9.5928	25.3664	0.7394	15.7279
2	3.3503	21.8609	3.0030	12.6066	9.2269	27.0804	0.7537	17.2342
5	3.8398	18.4044	3.7871	10.6747	6.8083	31.1093	0.7079	23.6901
9	4.2977	16.5305	4.2425	9.5471	5.2998	33.9542	0.7623	28.0542
16	4.2013	15.9145	4.1634	9.6488	5.4853	34.9928	0.7779	28.8673
<b>Daten: Erdöl, BP</b>								
<b>Anzahl der Beobachtungen: 4049</b>								
p	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$
1	2.8240	18.7194	2.4344	13.0385	8.0996	24.1215	0.9112	12.9498
2	2.8083	19.8584	2.4638	13.4149	7.1837	23.4680	0.8918	11.7377
5	2.6763	16.6365	2.6443	11.3899	5.5638	23.7129	0.8738	13.8571
9	3.0268	16.0284	2.9450	11.1885	4.3875	25.4763	0.9126	16.0156
16	3.1992	15.6182	3.0169	11.1411	4.3947	26.8852	0.8723	17.5468
<b>Daten: Erdöl, CVX</b>								
<b>Anzahl der Beobachtungen: 4049</b>								
p	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$
1	5.3020	24.7706	4.5818	14.2291	8.9556	24.0230	1.4269	14.1474
2	5.1791	25.4447	4.8383	14.6907	8.4331	25.1532	1.4573	14.5805
5	5.2367	20.7357	5.2366	12.5852	6.3582	25.1405	1.3057	16.7457
9	5.5315	18.6365	4.8956	11.8630	4.9923	26.7939	1.3578	19.2860
16	5.4863	17.9062	4.8965	11.6437	5.0382	27.8860	1.3424	20.4148
<b>Daten: Erdöl, RD</b>								
<b>Anzahl der Beobachtungen: 4049</b>								
p	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$
1	8.2268	24.7706	7.9950	17.4577	8.9019	31.4531	0.5168	21.2550
2	8.8992	25.4447	7.8256	18.2307	8.5669	33.5496	0.5270	22.7878
5	9.4350	22.8189	6.1200	16.0749	6.1619	31.6582	0.5171	23.3626
9	11.0348	22.3733	4.7092	15.8625	4.7221	32.7942	0.5645	25.8453
16	10.5633	22.78632	4.6943	16.0006	4.7483	34.0976	0.54628	27.6386
<b>Daten: Erdöl, SC</b>								
<b>Anzahl der Beobachtungen: 4049</b>								
p	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$
1	6.2704	22.4338	5.9213	15.5349	9.5846	29.3098	0.4975	19.7694
2	6.4063	22.6394	6.2817	15.5866	9.2829	29.3936	0.4898	19.7309
5	6.7765	20.1229	6.3444	13.2935	6.6573	27.7273	0.4727	20.6831
9	8.2346	20.3680	4.8860	13.2433	5.0985	30.0215	0.4846	24.4633
16	8.0546	20.7008	4.6113	13.5285	5.0741	32.1263	0.4822	26.9691

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot).

Tabelle C.12: Teststatistik und kritische Werte des Johansen - Kointegrationstests

Klasse:	M*		M <sub>1</sub>		M <sub>1</sub> *		M <sub>2</sub>	
Niveau	krit. Werte		krit. Werte		krit. Werte		krit. Werte	
10.0%	10.56	22.95	2.71	13.31	7.50	17.79	2.98	10.35
5.0%	12.39	25.47	3.84	15.34	9.13	19.99	4.14	12.21
2.5%	14.13	27.87	5.02	17.24	10.73	22.07	5.30	13.94
1.0%	16.39	30.65	6.64	19.69	12.73	24.74	7.02	16.16
<b>Daten: Gold, AAUK</b>								
<b>Anzahl der Beobachtungen: 4049</b>								
p	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$
1	4.7843	18.1454	2.4654	7.9221	5.2049	11.7660	0.9854	7.5459
2	5.1290	17.5636	2.5794	7.9125	4.8280	11.6975	0.6258	7.4767
5	5.4516	17.8889	2.0911	7.6402	4.9257	11.0942	0.5557	6.4993
9	5.3440	16.1003	2.3001	7.7514	5.2180	12.0205	0.6410	7.4388
16	4.9129	15.7038	2.1665	7.3193	4.3890	11.7995	0.6412	8.0078
<b>Daten: Gold, NEM</b>								
<b>Anzahl der Beobachtungen: 4049</b>								
p	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$
1	3.8972	34.4105	3.8581	29.0575	3.8789	29.3894	0.0105	15.9427
2	4.1092	23.6966	4.0980	21.7694	4.0980	22.3610	0.1338	11.9151
5	4.5449	20.8345	4.5392	19.3790	4.5392	20.0432	0.2088	10.1027
9	4.3695	19.8689	4.3162	18.2133	4.3162	18.8533	0.2362	9.0613
16	4.2755	18.1144	4.1009	16.3081	4.1035	16.7839	0.2275	8.0318
<b>Daten: Gold, ABX</b>								
<b>Anzahl der Beobachtungen: 4049</b>								
p	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$
1	5.2589	16.5540	4.9142	16.1704	4.9615	17.8766	0.0841	12.7934
2	4.6651	17.9233	4.6608	17.7197	4.6637	20.5556	0.0022	15.8663
5	5.1719	17.4536	5.1719	17.0611	5.1822	20.2927	0.0000	15.1047
9	4.8916	16.4634	4.8879	16.0993	4.9295	19.3095	0.0000	14.2460
16	4.0512	15.2728	4.0436	14.6840	4.1439	17.9519	0.0020	13.5147
<b>Daten: Gold, PDG</b>								
<b>Anzahl der Beobachtungen: 4049</b>								
p	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$
1	4.7290	23.5561	4.4907	22.2163	4.5043	22.5331	0.0114	14.3488
2	4.5536	17.6846	4.5499	17.6626	4.5531	18.2687	0.1303	11.4996
5	4.9413	15.5259	4.9318	15.5129	4.9333	16.1939	0.1940	9.4230
9	4.7038	14.6473	4.7033	14.6433	4.7034	15.3593	0.2223	8.4359
16	4.1657	13.0804	4.1448	13.0335	4.1462	13.7279	0.2426	7.1446

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot).

Tabelle C.13: Teststatistik und kritische Werte des Johansen - Kointegrationstests

Klasse:	M*		M <sub>1</sub>		M <sub>1</sub> *		M <sub>2</sub>	
Niveau	krit. Werte		krit. Werte		krit. Werte		krit. Werte	
10.0%	10.56	22.95	2.71	13.31	7.50	17.79	2.98	10.35
5.0%	12.39	25.47	3.84	15.34	9.13	19.99	4.14	12.21
2.5%	14.13	27.87	5.02	17.24	10.73	22.07	5.30	13.94
1.0%	16.39	30.65	6.64	19.69	12.73	24.74	7.02	16.16
<b>Daten:</b> Silber, AAUK								
<b>Anzahl der Beobachtungen:</b> 4049								
p	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$
1	9.3963	24.4888	1.3533	15.5316	4.9439	19.1231	0.9951	6.9193
2	10.4840	23.7533	1.3553	14.2008	4.6184	17.5872	0.6414	6.7557
5	10.4974	22.8054	1.2500	13.5041	4.3735	16.8011	0.5609	6.4750
9	8.9739	21.1614	1.1088	13.2904	4.8236	17.2133	0.6598	6.9264
16	9.1848	21.6480	1.1154	13.5733	4.5779	17.8375	0.6171	8.1548
<b>Daten:</b> Silber, NEM								
<b>Anzahl der Beobachtungen:</b> 4049								
p	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$
1	11.0977	30.8224	11.0685	29.7636	11.2403	29.9492	0.0194	18.0250
2	10.0332	23.7533	9.3829	22.3392	11.3326	27.6852	0.0192	10.7097
5	9.1816	22.3058	8.4934	20.8444	8.7998	21.1540	0.0231	9.8236
9	7.9908	20.8707	7.6489	19.5648	7.9230	19.8399	0.0243	9.7337
16	7.8413	21.6615	7.7259	19.9047	7.9226	20.1386	0.0068	10.7568
<b>Daten:</b> Silber, ABX								
<b>Anzahl der Beobachtungen:</b> 4049								
p	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$
1	10.8023	29.4452	10.2942	28.5667	11.1599	29.9198	0.3992	18.0250
2	10.6906	26.6816	10.6429	25.7361	11.3326	27.6852	0.2370	12.2710
5	10.2928	25.7860	10.1633	24.9787	10.8879	27.1773	0.3052	12.0059
9	9.8476	23.3423	9.7556	22.5114	10.8074	24.7214	0.3089	11.5491
16	10.3872	23.3681	10.3866	22.6517	11.9820	24.8801	0.3900	12.5520
<b>Daten:</b> Silber, PDG								
<b>Anzahl der Beobachtungen:</b> 4049								
p	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$	$r \leq 1$	$r = 0$
1	10.2637	32.3546	10.2635	31.8822	10.4122	32.0550	0.0014	19.5492
2	9.7470	24.2066	9.6911	23.1566	9.9116	23.4447	0.0000	12.2918
5	8.4203	22.0080	8.3379	20.9019	8.5736	21.1861	0.0000	10.9202
9	7.3497	21.2020	7.2937	20.1910	7.5210	20.4642	0.0011	10.6536
16	6.5049	21.1523	6.4770	19.6983	6.6843	20.0111	0.0000	11.0732

Hinweis: Färbige Werte charakterisieren das Niveau der Ablehnung (10% blau, 5% grün und 1% rot).

# Anhang D.

## „Source Code“

```
# Informationscriteria
# Dimension of incoming data: T rows (number of observations), n columns (number of different time series)

Order.VAR <- function(data,method=c("AIC","SC","HQ","FPE","LR")) {

  method <- match.arg(method)
  T <- nrow(data)
  n <- ncol(data)
  order.selection <- matrix(data=0,nrow=round(10*log10(T)),ncol=1)

  if (method == "LR") {
    for (i in round(10*log10(T)):1) {
      Sigma1 <- matrix(data=0,nrow=n,ncol=n)
      Sigma2 <- matrix(data=0,nrow=n,ncol=n)
      VM.ols1 <- ar.ols(data,aic=FALSE,order.max=i)
      VM.ols2 <- ar.ols(data,aic=FALSE,order.max=i-1)

      for (t in (i+1):T) {
        Sigma1 <- Sigma1 + (1/(T-round(10*log10(T)))) * (VM.ols1$resid[t,] %*% t(VM.ols1$resid[t,]))
      }

      for (t in (i-1+1):T) {
        Sigma2 <- Sigma2 + (1/(T-round(10*log10(T))-1)) * (VM.ols2$resid[t,] %*% t(VM.ols2$resid[t,]))
      }
      order.selection[i] <- (T-round(10*log10(T)))*(log(det(Sigma2)) - log(det(Sigma1)))
    }
    return(order.selection)
  } else if (method == "AIC") {
    for (i in round(10*log10(T)):1) {
      Sigma <- matrix(data=0,nrow=n,ncol=n)
      VM.ols <- ar.ols(VarModel,aic=FALSE,order.max=i)

      for (t in (i+1):T) {
        Sigma <- Sigma + (1/(T-round(10*log10(T)))) * (VM.ols$resid[t,] %*% t(VM.ols$resid[t,]))
      }
      order.selection[i] <- log(det(Sigma)) + (2*i*n*n)/(T-round(10*log10(T)))
    }
    return(which.min(order.selection))
  } else if (method == "SC") {
    for (i in round(10*log10(T)):1) {
      Sigma <- matrix(data=0,nrow=n,ncol=n)
      VM.ols <- ar.ols(VarModel,aic=FALSE,order.max=i)

      for (t in (i+1):T) {
        Sigma <- Sigma + (1/(T-round(10*log10(T)))) * (VM.ols$resid[t,] %*% t(VM.ols$resid[t,]))
      }
      order.selection[i] <- log(det(Sigma)) + (i*n*n*log(T-round(10*log10(T)))/(T-round(10*log10(T)))
    }
    return(which.min(order.selection))
  }
}
```

```

} else if (method == "HQ") {
  for (i in round(10*log10(T)):1) {
    Sigma <- matrix(data=0,nrow=n,ncol=n)
    VM.ols <- ar.ols(VarModel,aic=FALSE,order.max=i)

    for (t in (i+1):T) {
      Sigma <- Sigma + (1/(T-round(10*log10(T)))) * (VM.ols$resid[t,] %*% t(VM.ols$resid[t,]))
    }
    order.selection[i] <- log(det(Sigma))+(2*i*n*n*log(log((T-round(10*log10(T))))))/(T-round(10*log10(T)))
  }
  return(which.min(order.selection))
} else if (method == "FPE") {
  for (i in round(10*log10(T)):1) {
    Sigma <- matrix(data=0,nrow=n,ncol=n)
    VM.ols <- ar.ols(VarModel,aic=FALSE,order.max=i)

    for (t in (i+1):T) {
      Sigma <- Sigma + (1/(T-round(10*log10(T)))) * (VM.ols$resid[t,] %*% t(VM.ols$resid[t,]))
    }
    order.selection[i] <- det(Sigma) * (((T-round(10*log10(T)))+n*i+1)/((T-round(10*log10(T)))-n*i-1))^3
  }
  return(which.min(order.selection))
}
}

```

```

# multiple Jarque-Bera Test
# Dimension of incoming data: n columns (number of different time series), T rows (number of observations)

jarque.bera.ck <- function(data) {

  data <- as.matrix(data)
  n <- ncol(data)
  T <- nrow(data)
  Sigma <- matrix(data=0,nrow=n,ncol=n)

  for (i in 1:T) {
    Sigma <- Sigma + (1/T) * (t(data[i,,drop=FALSE]) %*% data[i,,drop=FALSE])
  }

  Pinv <- solve(chol(Sigma))
  w <- data %*% Pinv
  b1 <- matrix(data=1/T,nrow=1,ncol=T) %*% (w*w*w)
  b2 <- matrix(data=1/T,nrow=1,ncol=T) %*% (w*w*w*w)
  JB <- T*((b1 %*% t(b1))/6 + ((b2-matrix(data=3,nrow=1,ncol=n))%*%(t(b2-matrix(data=3,nrow=1,ncol=n)))))/24

  ARG <- c(T,n);
  names(ARG) <- c('T','n');
  STAT <- JB
  names(STAT) <- "Jarque-Bera Teststatistic";
  PVAL <- 1-pchisq(STAT,2*n)
  names(PVAL) <- "p value";
  PARAMETER <- 2*n
  names(PARAMETER) <- "df";
  METHOD <- "multivariate Jarque-Bera Test"
  alternative <- "no gaussian data"

  structure(list(statistic = STAT,

```

```

        p.value = PVAL,
        parameter = PARAMETER,
        alternative = alternative,
        data.name = "-----",
        arguments = ARG,
        method = METHOD),
    class = "htest")
}

# ARCH Test based on Engle (1982)
#

arch.test.ck <- function(data, nlags=10) {

  T <- length(data)
  nlags = max(nlags,1)
  if (nlags>=T)
    stop('too many lags!')

  x2 = as.numeric(data)
  x2 = x2*x2
  m = 'x2 ~ B(x2,1)'
  if (nlags>1) {
    for (lag in (2:nlags)) {
      m = paste(m,'+B(x2,',lag,')',sep='')
    }
  }
  m <- summary(lm(as.formula(m)))

  DNAME=deparse(substitute(data))
  ARG <- c(T);
  names(ARG) <- c('T');
  STAT <- m$r.squared * T
  names(STAT) <- "ARCH Teststatistic";
  PVAL <- 1-pchisq(STAT,nlags)
  names(PVAL) <- "p value";
  PARAMETER <- nlags
  names(PARAMETER) <- "df";
  METHOD <- "ARCH Test"
  alternative <- "no ARCH effects"

  structure(list(statistic = STAT,
                p.value = PVAL,
                parameter = PARAMETER,
                alternative = alternative,
                data.name = DNAME,
                arguments = ARG,
                method = METHOD),
            class = "htest")
}

# multiple Portmanteau Test
# Dimension of incoming data: n columns (number of different time series), T rows (number of observations)

port.test.ck <- function(data,s=10,lags) {

  n <- ncol(data)
  T <- nrow(data)
  CO <- matrix(data=0,nrow=n,ncol=n)
  P <- 0

  for (t in 1:n) {

```

```

    CO <- CO + (1/T) * (t(data[t,,drop=FALSE]) %>% data[t,])
  }

  for (i in 1:s) {
    Ci <- matrix(data=0,nrow=n,ncol=n)
    for (t in (i+1):T) {
      Ci <- Ci + (1/T) * (t(data[t,,drop=FALSE]) %>% data[t-i,])
    }
    P <- P + (1/(T-i))*diag( t(Ci) %>% solve(CO) %>% Ci %>% solve(CO) ) %>% matrix(data=1,nrow=n,ncol=1)
  }

  P <- T*T*P

  DNAME=deparse(substitute(data))
  ARG <- c(T,n,lags,s);
  names(ARG) <- c('T','n','lags','s');
  STAT <- P
  names(STAT) <- "Portmanteau Teststatistic";
  PVAL <- 1-pchisq(STAT,n*n*(s-lags))
  names(PVAL) <- "p value";
  PARAMETER <- n*n*(s-lags)
  names(PARAMETER) <- "df";
  METHOD <- "multivariate Portmanteau Teststatistic"
  alternative <- "no white noise data"

  structure(list(statistic = STAT,
                p.value = PVAL,
                parameter = PARAMETER,
                alternative = alternative,
                data.name = DNAME,
                arguments = ARG,
                method = METHOD),
            class = "htest")
}

# one step forecasting
# multivariate data: n columns, T rows

prog.ck <- function(data,obs=30,step=1,p=2,r=1,det="none",restricted=FALSE) {

  data <- as.matrix(data)
  T <- nrow(data)
  n <- ncol(data)
  Ordnung.AIC <- array(0,c(5,5,n))
  Ordnung <- matrix(0,nrow=n,ncol=2)

  for (k in 1:n) {
    for (i in 1:5) {
      for (j in 1:5) {
        z <- arima(data[1:(T-obs-1),k],order=c(i,1,j))
        Ordnung.AIC[i,j,k] <- z$aic
      }
    }
  }

  for (k in 1:n) {
    t <- which.min(Ordnung.AIC[, ,k])

    if (t %% 5 == 0) {
      Ordnung[k,1] <- 5
      Ordnung[k,2] <- t %% 5
    } else {

```

```

    Ordnung[k,1] <- t %% 5
    Ordnung[k,2] <- t %/% 5 + 1
  }
}

fehlerA <- 0
fehlerB <- 0
zp <- matrix(0,ncol=1,nrow=n)
fehlerA1 <- matrix(0,ncol=1,nrow=n)
fehlerB1 <- matrix(0,ncol=1,nrow=n)

for (i in (T-obs-1):T-1) {
  for (k in 1:n) {
    z <- arima(data[1:i,k],order=c(Ordnung[k,1],1,Ordnung[k,2]))
    zp[k] <- predict(z,se.fit=FALSE,n.ahead=step)
  }

  for (j in 1:n) {
    if ((zp[j] > data[i,j] && data[i+1,j]>data[i,j]) || (zp[j] < data[i,j] && data[i+1,j]<data[i,j])) {
      fehlerA1[j] = fehlerA1[j] + 1
    }
  }

  fehlerA <- fehlerA + sqrt(sum((zp-data[i+1,])^2))
  t1 <- johansen(data[1:i,], r = r, p = p, det = det, restricted = restricted)

  if (p == 1) {
    Z2 <- 0
  } else if (p > 1) {
    for (j in 1:(p-1)) {
      if (j == 1) {
        Z2 <- t(data[i-j,,drop=FALSE]-data[i-j-1,drop=FALSE])
      } else {
        Z2 <- rbind(Z2,t(data[i-j,,drop=FALSE]-data[i-j-1,drop=FALSE]))
      }
    }
  }

  if (det == "none") {
    tp <- t(data[i,,drop=FALSE]) + t1$pi[,1:n] %*% t(data[i,,drop=FALSE]) +
      + t1$gamma %*% Z2
  } else if (det == "const") {
    tp <- t(data[i,,drop=FALSE]) + t1$pi[,1:n] %*% t(data[i,,drop=FALSE]) +
      + t1$gamma %*% Z2
      + t1$mu0
  } else if (det == "linear") {
    tp <- t(data[i,,drop=FALSE]) + t1$pi[,1:n] %*% t(data[i,,drop=FALSE]) +
      + t1$gamma %*% Z2
      + t1$mu0 + t1$mu1 * i
  }

  for (j in 1:n) {
    if ((tp[j] > data[i,j] && data[i+1,j]>data[i,j]) || (tp[j]<data[i,j] && data[i+1,j]<data[i,j])) {
      fehlerB1[j] = fehlerB1[j] + 1
    }
  }

  fehlerB <- fehlerB + sqrt(sum((t(tp)-data[i+1,])^2))
}
return(cbind(fehlerA,fehlerB,fehlerA1,fehlerB1))
}

```

# Anhang E.

## Zeichenerklärung

### E.1. Griechische Buchstaben

alpha	$\alpha$	„loading“ Matrix, s. S. 23
beta	$\beta$	Kointegrationsmatrix, s. S. 23
	$\beta_i$	Kointegrationsvektoren, s. S. 19
gamma	$\gamma$	Autokovarianzfunktion, s. S. 62
	$\hat{\gamma}$	Koeffizient im Zivot-Andrews Test, Stichprobenautokovarianzfunktion, s. S. 90, 62
	$\gamma_{ij}$	Element von $\Gamma$ , s. S. 3
	$\Gamma$	Kovarianzmatrix, s. S. 2
	$\hat{\Gamma}$	Stichprobenkovarianzmatrix, s. S. 4
	$\Gamma_i$	Koeffizientenmatrix im VECM, s. S. 22
	$\tilde{\Gamma}$	Kompaktere Schreibweise eines VEC Prozesses, s. S. 28
delta	$\delta$	Koeffizient eines Trends, s. S. 29
	$\Delta$	Differenzenoperator, s. S. 17
epsilon	$\varepsilon$	mehrdimensionaler stationärer Prozess, s. S. 3
zeta	$\zeta$	Koeffizient der differenzierten Verzögerung, s. S. 83
eta	$\eta$	Koeffizient in der „Spurious“ Regression, s. S. 18
theta	$\theta$	Koeffizient der „Moving Average“ Prozesse, s. S. 64
	$\hat{\theta}$	Koeffizient im Zivot-Andrews Test, s. S. 90
kappa	$\kappa$	Koeffizientenmatrix des deterministischen Teils der VAR Prozesse, s. S. 4
lambda	$\lambda$	Zeitpunkt eines Strukturbruches, Eigenwert, s. S. 90 32
	$\Lambda$	Menge der möglichen Strukturbruchpunkte, s. S. 90
mu	$\mu$	ein- oder merdimensionaler Erwartungswert, s. S. 2
nu	$\nu$	Konstante , s. S. 14, 29
xi	$\xi$	eindimensionaler stationärer Prozess, s. S. 14
pi	$\pi$	Koeffizient der AR( $\infty$ )-Prozesse, s. S. 74
	$\Pi$	Koeffizientenmatrix im VECM, s. S. 22
rho	$\rho$	Autokorrelationsfunktion, s. S. 62
	$\rho_{ij}$	Element von $\mathcal{R}$ , s. S. 3
	$\hat{\rho}$	Stichprobenautokorrelationsfunktion, s. S. 63
sigma	$\sigma$	Standardabweichung, s. S. 63
	$\Sigma$	Varianz-Kovarianz Matrix, s. S. 3
tau	$\tau$	Aufspaltung von $D_t$ , s. S. 29
phi	$\phi$	Koeffizient der AR-Prozesse, s. S. 14
	$\Phi$	Koeffizientenmatrix der VAR Prozesse, s. S. 4
	$\tilde{\Phi}$	Kompaktere Schreibweise eines VAR Prozesses, s. S. 6
Psi	$\psi$	Koeffizient der MA( $\infty$ )-Prozesse, s. S. 67
	$\Psi$	Koeffizientenmatrix der kausalen Lösung von $\mathbf{y}$ , s. S. 5
Omega	$\omega$	Aufspaltung von $D_t$ , s. S. 29

Wird die Variable \* mit irgendeinem Verfahren geschätzt, so wird die Symbolik  $\hat{*}$  verwendet.

### E.2. Andere Zeichen

$y_t$	Zeitreihe bzw. stochastischer Prozess, s. S. 2
$\mathbf{y}_t$	mehrdimensionale Zeitreihe bzw. stochastischer Prozess, s. S. 2
$\bar{\mathbf{y}}$	mehrdimensionales Stichprobenmittel, s. S. 4
$t$	Zeitindex, s. S. 2
$T$	maximaler Zeitindex, Anzahl der Beobachtungen
$\mathcal{T}$	Zeitmenge
$u_t$	stationärer Prozess, s. S. 85
$r_t$	„random walk“, s. S. 85
$r_0$	Konstante, s. S. 85
$h$	Verzögerung, s. S. 2
$B$	„Backshift“-Operator, s. S. 4
$P$	Teil der Darstellung von $y_t$ im „GRT“ der abhängig ist von den Startwerten, s. S. 24
$q, p, d, b$	Ordnungen der Prozesse, s. S. 4, 19
$n$	Anzahl der verschiedenen Datenreihen, s. S. 2
$k$	Zeilen von $D_t$ , s. S. 6
$r$	Anzahl der Kointegrationsvektoren, Rang von $\Pi$ , s. S. 20
$\mathcal{R}$	Korrelationsmatrix, s. S. 3
$D_t$	Matrix die nichtstochastische Komponenten enthält (deterministischer Teil), s. S. 4
$I_n$	Einheitsmatrix der Dimension $n$ , s. S. 4
$w_i$	Wurzeln von $\Phi_1$ , s. S. 14
$f$	Koeffizient im Verfahren von Stock, s. S. 34
$c_i$	Kritische Werte, s. S. 36
$\mathbf{z}$	Hilfsvariable, s. S. 4
$Z_t$	Kompaktere Schreibweise eines VAR Prozesses, s. S. 6
$Z_i, Z_{i,t}$	Kompaktere Schreibweise eines VECM, s. S. 28, 30
$M_{ij}, R_{ij}, S_{ij}$	Hilfsmatrizen im ML-Schätzverfahren, s. S. 30
$H_i, A, \Xi, J$	Hilfsmatrizen bei der Prognose von kointegrierten Prozessen, s. S. 41

### E.3. Mathematische Symbole

$\xrightarrow{d}$	Konvergenz in Verteilung ( <u>d</u> istribution), $P( y_t - c  > \delta) < \varepsilon$
$\xrightarrow{p}$	Konvergenz in Wahrscheinlichkeit ( <u>p</u> robability), $\lim_{T \rightarrow \infty} F_{y_T}(y) = F_y(y)$
$\log(x)$	Logarithmus zur Basis $e$ von $x$
$ x $	Absolutbetrag von $x$
$sp\{X\}$	Spur der Matrix $X$
$\det X$	Determinante der Matrix $X$

# Literaturverzeichnis

- J.J. Dolado A. Banerjee and R. Mestre. Error-correction mechanism tests for cointegration in a single-equation framework. *Journal of Time Series Analysis*, 19(3):267–283, 1998.
- A. Geyer A. Trapletti and F. Leisch. Cointegration and exchange market efficiency: An analysis of high frequency data. *Working Paper Series*, 52, 1999.
- P.J. Brockwell and R.A. Davis. *Introduction to time series and forecasting*. Springer Texts in Statistics, New York, 1996.
- Compiled by Adrian Trapletti. *tseries: Time series analysis and computational finance*, 2004. R package version 0.9-24.
- Compiled by Wolfgang Scherrer. *oekfinm: konometrie der Finanzmrkte*, 2004. R package version 0.4-03.
- P. Schmidt D. Kwiatkowski, P.C.B. Phillips and Y. Shin. Testing the null hypothesis of stationarity against the slternative of a unit root: How sure are we that economic time series have a unit root. *Journal of Econometrics*, 54:159–178, 1992.
- D.A. Dickey and W.A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of American Statistical Association*, 74(366), 1979.
- D.A. Dickey and W.A. Fuller. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4):1057 – 1072, 1981.
- Wolfgang Drobetz. Nichtstationarität und kointegration: Eine einführung. Technical report, Universität Basel , Abteilung Finanzmarkttheorie, 2003.
- R.F. Engle and C.W.J. Granger. Cointegration and error correction: representation, estimation and testing. *Econometrica*, 35(2):251 – 276, 1987.
- T.J. Rothenberg G. Elliott and J.H. Stock. Efficient tests for an autoregressive unit root. *Econometrica*, 64 (4):813 – 836, 1996.
- C. Granger and P. Newbold. Spurious regressions in econometrics. *Journal of Econometrics*, 2:111 – 120, 1974.
- J.D. Hamilton. *Time series analysis*. Princeton University Press, Princeton, New Jersey, 1994.
- U. Hassler. Cointegration testing in single error-correction equations in the presence of linear time trends. *Oxford Bulletin of Economics and Statistics*, 62(5):621 – 632, 2000.
- S. Johansen. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59(6):1551 – 1580, 1991.
- S. Johansen. *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press, 1995.
- H. Lütkepohl K. Hubrich and P. Saikkonen. A review of systems cointegration tests. *Econometrics Reviews*, 20, 2001.
- Walter Krämer. Kointegration von aktienkursen. Technical report, Universität Dortmund, Fachbereich Statistik, 1997.

- H. Lütkepohl. *Introduction to multiple time series analysis*. Springer-Verlag, Kiel, 1991.
- J.G. MacKinnon. Critical values for cointegration tests. *Oxford University Press*, (4), 1990.
- Terence C. Mills. *The econometric modelling of financial time series*. Cambridge University Press, 1997.
- Evelyne Neurauner. Multi-kointegrierte zeitreihen. Master's thesis, Technische Universität Wien, Institut für Ökonometrie, keine Ahnung.
- Bernhard Pfaff. *urca: Unit root and cointegration tests for time series data*, 2004. URL <http://www.r-project.org>. R package version 0.6-1.
- P.C.B. Phillips and S. Ouliaris. Asymptotic properties of residual based tests for cointegration. *Econometrica*, 58(1):165 – 193, 1990.
- P.C.B. Phillips and P. Perron. Testing for a unit root in time series regression. *Biometrika*, 75(2):335 – 346, 1988.
- R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2004. URL <http://www.R-project.org>. ISBN 3-900051-07-0.
- H.-E. Reimers. *Analyse kointegrierter Variablen mittels vektorautoregressiver Modelle*. Physica-Verlag Heidelberg, Heidelberg, 1991.
- G. C. Reinsel and S. K. Ahn. Vector autoregressive models with unit roots and reduced rank structure: estimation, likelihood ratio test, and forecasting. *Journal of Time Series Analysis*, 33, 1992.
- S.E. Said and D.A. Dickey. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599 – 607, 1984.
- W. Scherrer. ökonometrie der finanzmärkte. Technical report, Technische Universität Wien, Institut für Wirtschaftsmathematik, 2004.
- P. Schmidt and P.C.B. Phillips. Lm test for a unit root in the presence of deterministic trends. *Oxford Bulletin of Economics and Statistics*, 54(3):257 – 287, 1992.
- T.C. Mills S.J. Leybourne and P. Newbold. Spurious rejections by dickey-fuller tests in the presence of a break under the null. *Journal of Econometrics*, 87:191 – 203, 1998.
- Leopold Steindl. Cointegration of futures markets. Master's thesis, Technische Universität Wien, Institut für Statistik und Wahrscheinlichkeitstheorie, 1994.
- H. J. Stock. Asymptotic properties of least square estimators of cointegrating vectors. *Econometrica*, 55, 1987.
- C. Trenkler. A new set of critical values for systems cointegration tests with a prior adjustment for deterministic terms. *Economics Bulletin*, 3(11):1 – 9, 2003.
- S. K. Ahn und G. C. Reinsel. Nested reduced-rank autoregressive models for multiple time series. *Journal of the American Statistical Association*, 83, 1988.
- E. Zivot and D.W.K. Andrews. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics*, 10(3):251 – 270, 1992.