

RUPRECHT-KARLS-UNIVERSITÄT HEIDELBERG
Fakultät für Verhaltens- und Empirische Kulturwissenschaften
Psychologisches Institut

*Multivariate Analyse instationärer Zeitreihen:
Integration und Kointegration
in Theorie und Simulation*

Diplomarbeit

Heidelberg, im Juli 2005

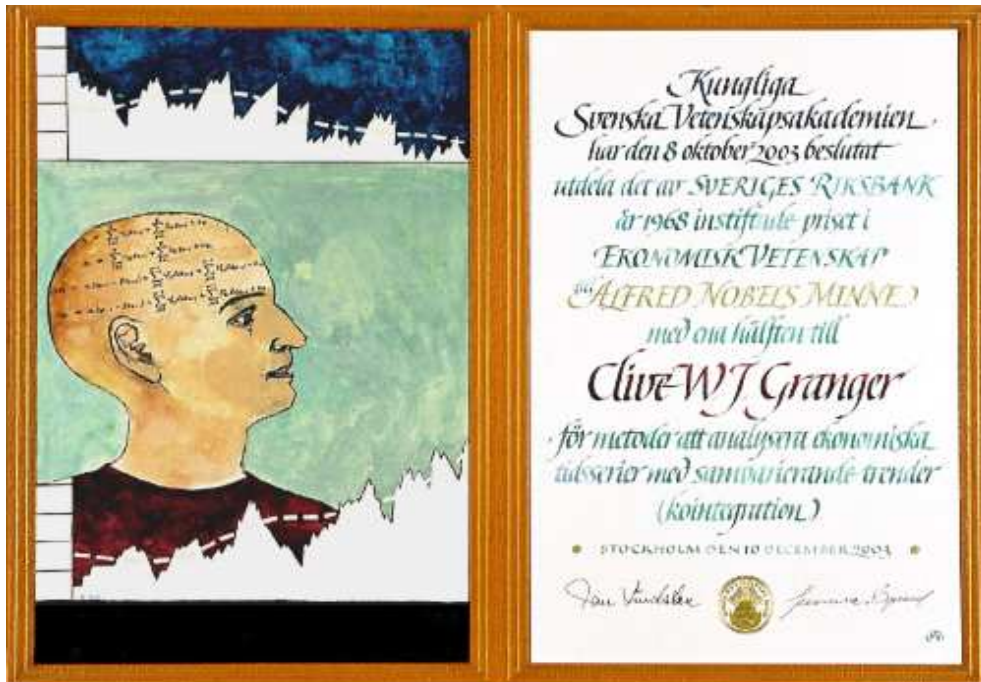
vorgelegt von
Esther Stroe-Kunold
Brunnengasse 10
69117 Heidelberg

Erstgutachter:
Prof. Dr. J. Werner

Zweitgutachter:
Prof. Dr. J. Funke

There obviously are pairs of (...) series, (...) which may or may not be co-integrated and a decision on this has to be determined by an appropriate theory or an empirical investigation. It might be interesting to undertake a wide-spread study to find out which pairs of (...) variables are co-integrated.

Granger (1981, S. 129)



Die im Jahr 2003 von der Nobelstiftung an Clive W.J. Granger verliehene Urkunde: *The Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel* „for methods of analyzing economic time series with common trends (cointegration)“. (Artist: N. G. Stenqvist, Calligrapher: A. Rucker; entnommen aus: <http://www.nobel.se/economics/laureates/2003/granger-diploma.jpg>). Allgemeine Informationen zur Verleihung des Nobelpreises bei Pöppe (2003).

Danksagung

Besonders danke ich meinem wissenschaftlichen Betreuer und Erstgutachter, Herrn Prof. Dr. Joachim Werner, der durch die Gestaltung der Forschungsgruppe zur Zeitreihenanalyse den Weg für diese Diplomarbeit bahnte und sich nicht nur im Rahmen der konstruktiven Betreuung der Diplomarbeit jederzeit offen für den Austausch über methodologische Fragestellungen zeigte. Zusätzlich danke ich Simone Braun und Tetiana Stadnytska für die anregenden Diskussionen in der genannten Gruppe.

Ebenfalls möchte ich Herrn Prof. Dr. Joachim Funke danken, der freundlicherweise die Zweitbegutachtung der Diplomarbeit übernommen hat und mir im Rahmen des Kolloquiums der Allgemeinen und Theoretischen Psychologie (ATP) ermöglichte, das methodologische Thema dieser Diplomarbeit vorzustellen und so den Austausch mit anderen psychologischen Teildisziplinen herzustellen.

Ausdrücklich möchte ich der Studienstiftung des deutschen Volkes danken, durch deren Stipendium mein Studium und damit auch diese Diplomarbeit in vielfältiger Form unterstützt wurden.

Der Dank gilt meinen Eltern, meiner Familie und meinen Freundinnen und Freunden, insbesondere aber meinem Mann Marian, der in unterschiedlichster Form durch die ihm eigene besondere Weise zum Gelingen dieser Diplomarbeit beigetragen hat.

Ihm widme ich diese Arbeit.

<i>Inhaltsverzeichnis</i>	4
Inhaltsverzeichnis	
Zusammenfassung	8
Abstract	9
Einleitung	10
Theoretische Grundlagen	15
1 Multivariate Zeitreihenanalyse	15
1.1 Vektorielle Autoregressive-(VAR-)Prozesse	16
1.2 Voraussetzungen multivariater stochastischer Prozesse	20
2 Theoretischer Exkurs	21
2.1 Stationarität	21
2.2 Integrierte Prozesse	23
2.2.1 $I(1)$ - und $I(0)$ -Prozesse	24
2.3 Differenzenbildung	25
2.4 Random-Walk-Prozesse	26
2.5 Überprüfung der Stationarität	28
2.5.1 Lag-Polynom und Charakteristische Gleichung	28
2.5.2 Unit-Root-Tests	29
Kointegration in Theorie und Simulation	31
3 Kointegration aus theoretischer Perspektive	32
3.1 Definitionen in der einschlägigen Fachliteratur	32
Weitgefasste Definition des Kointegrationskonzepts	33
Mehrheitlich geteilte Ausgangsdefinition und deren Weiterent-	
wicklungen	35
Auseinanderdriftende Zeitpfade der Komponentenprozesse	40

Gleichheit des Integrationsgrads der Komponentenprozesse . . .	41
3.2 Kointegration und Scheinregression	44
3.2.1 Abhängigkeitskonzept der Scheinregression	44
3.2.2 Kointegration vs. Scheinregression	47
3.3 Fehlerkorrekturmodelle	49
3.4 Engle-Granger-Repräsentationstheorem	51
3.4.1 Interpretation der Funktion des Anpassungskoeffizienten α und des Kointegrationsvektors β	54
Funktion des Anpassungskoeffizienten α	54
Funktion des Kointegrationsvektors β	57
3.4.2 Schätzung von Anpassungskoeffizient α und Kointegrations- vektor β mithilfe des für <i>eine</i> Kointegrationsbeziehung ver- einfachten Fehlerkorrekturmodells	59
3.5 Testverfahren	60
3.5.1 Dickey-Fuller-Kointegrationstest	61
3.5.2 Maximum-Likelihood-Analyse kointegrierter Systeme: Das Johansen-Testverfahren	62
3.6 Weiterentwicklungen des Kointegrationskonzepts: Ausblick	67
3.6.1 ARFIMA-Modelle und fraktionale Kointegration	67
Relevanz fraktional (ko-)integrierter Prozesse für die Psychologie	69
Fraktionale Kointegration	71
3.6.2 Saisonale Kointegration und Multi-Kointegration	73
3.7 Beispiele für die Anwendung des Kointegrationskonzepts	73
4 Kointegration als empirisches Konzept:	
Simulationen	77
4.1 Problemstellung	77
4.1.1 Fragestellung und Hypothesen	77
Hypothesen zur Klärung des Kointegrationskonzepts	77
Hypothesen zur Schätzung von α und β	79

<i>Inhaltsverzeichnis</i>	6
4.2 Methode	80
4.2.1 Monte-Carlo-Simulation	81
4.2.2 Datenerzeugung	83
4.2.3 Programmierung in SAS	85
4.2.4 Untersuchungsablauf	85
Testverfahren: Dickey-Fuller- und Johansen-Kointegrationstest	86
Güte der Parameterschätzungen von Kointegrationsvektor und Anpassungskoeffizient über das Fehlerkorrekturmodell	87
4.2.5 Übersicht über die im Rahmen der Simulationen verwendeten Matrizen	87
5 Interpretation der Ergebnisse und Überprüfung der Hypothesen	91
5.1 Testverfahren: Ergebnisse	93
5.1.1 Dickey-Fuller-Kointegrationstest	93
5.1.2 Johansen-Kointegrationstest	96
5.2 Überprüfung der Hypothesen zu den Testverfahren	100
Hypothese 0	100
Hypothese 1a	102
Hypothese 1b	104
Hypothese 1c	107
5.3 Güte der Parameterschätzungen von Kointegrationsvektor und An- passungskoeffizient: Ergebnisse	107
5.4 Überprüfung der Hypothesen zu den Parameterschätzungen	115
Hypothese 2a	115
Hypothese 2b	117
Hypothese 2c	121
Hypothese 2d	122

Diskussion und Einordnung der Ergebnisse	124
Literatur	128
Anhang	139
Anhang A: Backward-Shift-Operator	139
Anhang B: Granger-Kausalität	140
Anhang C: Wiener Prozess	141
Anhang D: SAS-Programme	141
Anhang D1: SAS-Programm zur Generierung von Φ_1 -Matrizen mit einer Einheitswurzel	146
Anhang D2: SAS-Programm zur Bestimmung der Populationspara- meter von Anpassungskoeffizient α und Kointegrationsvektor β über das Fehlerkorrekturmodell	147
Anhang D3: SAS-Simulationsprogramm zum Dickey-Fuller-Kointe- grationstest	147
Anhang D4: SAS-Simulationsprogramm zum Johansen-Kointegrati- onstest	153
Anhang D5: SAS-Simulationsprogramm zur Schätzung von Anpas- sungskoeffizient α und Kointegrationsvektor β	157
Anhang D6: SAS-Simulationsprogramm zur Replikation von Granger und Newbold (1974)	162
Anhang D7: SAS-Simulationsprogramm zur Überprüfung der Faust- regel von Granger und Newbold (1974)	165
Anhang D8: SAS-Programm zur Generierung von Diagrammen und Graphiken	168
Anhang E: Detail-Herleitung zum Engle-Granger-Repräsentationstheorem .	172
Tabellenverzeichnis	174
Abbildungsverzeichnis	176
Erklärungen	179

Zusammenfassung

Mit Hilfe der Kointegrationsmethodologie ist die multivariate Modellierung und Analyse instationärer bzw. integrierter Zeitreihen möglich. Für die Begründung des Kointegrationskonzepts, das im Zentrum der Diplomarbeit steht, erhielt der Ökonometriker C.W.J. Granger 2003 den Nobelpreis. Kointegrierte Modelle und deren Fehlerkorrekturrepräsentationen ermöglichen die Betrachtung der beiden Prozesseigenschaften des Gleichgewichts und der dynamischen Anpassungsprozesse. Dabei ist der multivariate Prozess, der sich aus den kointegrierten Komponentenprozessen zusammensetzt, als stabiles System - repräsentiert durch eine durch den Kointegrationsvektor gewichtete stationäre Linearkombination - aufzufassen, in dem Abweichungen vom langfristigen Gleichgewicht nur vorübergehender Natur sind. Das methodologische Thema wird im Rahmen der Arbeit für die psychologische Forschungspraxis erarbeitet, in deren Zusammenhang zeitbezogene und entsprechend zu modellierende Phänomene eine nicht unbedeutende Rolle spielen, die nicht selten miteinander in (Wechsel-)Beziehungen stehen. Übergeordnetes Ziel der Darstellung des Kointegrationsansatzes ist die theoretische und empirische Überprüfung notwendiger Bedingungen für das Vorliegen von Kointegration. Aus diesem Grund werden nach Darstellung der für das Verständnis des Ansatzes notwendigen Grundlagen der Zeitreihenanalyse die in der Fachliteratur in einigen Punkten uneinheitlichen Definitionen von Kointegration theoretisch kontrastiert und diskutiert. Der Zusammenhang zwischen Scheinregression und Kointegration wird beleuchtet und die bisher nur wenig interpretierten Funktionen von Kointegrationsvektor und Anpassungskoeffizient werden beschrieben und graphisch veranschaulicht. Außerdem werden die für den empirischen Teil relevanten Testverfahren dargestellt. Dem ausführlichen theoretischen Teil, der auch einen Literaturüberblick enthält und durch die Erläuterung fraktional integrierter und kointegrierter Modelle und aktuelle Anwendungsbeispiele abgerundet wird, folgt der empirische Teil, im Rahmen dessen Hypothesen zur Güte der Schätzung von Kointegrationsvektor und Anpassungskoeffizient durch Monte-Carlo-Simulationen überprüft werden. Die theoretisch vorgenommenen Annahmen zum Kointegrationskonzept werden durch systematische Simulationen zu den Testverfahren validiert. Eine Schlussbetrachtung vor psychologischem Kontext beschließt die Diplomarbeit.

Abstract

With the help of cointegration methodology it is possible to model and analyse instationary respectively integrated time series in a multivariate system. This is one of the reasons why in 2003 C.W.J. Granger was granted the Nobel Prize for Economics. Being in the focus of this diploma thesis, cointegrated models and their vector error correction representations incorporate the process characteristics of both equilibrium and dynamic adjustment processes. Consisting of its cointegrated component variables, the multivariate process can be regarded as a stable system, which is represented through a stationary linear combination weighted by the cointegration vector. As a result, these deviations from the long-run equilibrium are only temporary. As many psychological phenomena are time-dependent and connected with each other, they have to be modelled accordingly. That's why the methodological topic is analysed for psychological research practice with the superordinate goal of a theoretical and an empirical check of necessary conditions for the existence of cointegration. The diploma thesis is introduced with the basics of time series analysis necessary for understanding cointegration. This part is followed by the theoretical presentation and discussion of the definitions of cointegration, which are differing from each other in scientific literature. Furthermore the relation between spurious regression and cointegration is described and the seldomly mentioned functions of cointegration vector and adjustment coefficient are interpreted and the interpretations are supported by graphics. Moreover the testing procedures relevant for the empirical part are illustrated. The theoretical part which contains also a literature overview, an introduction into fractionally integrated and cointegrated series concludes with application examples of cointegration during the last years. In the following empirical part hypotheses about the efficiency of the estimated cointegration vector and adjustment coefficient are tested through monte-carlo simulations. The previously made theoretical assumptions about cointegration find validation through systematic simulations on the described testing procedures. To sum up, the diploma thesis concludes with a final statement from the psychological perspective.

Einleitung

Mithilfe zeitreihenanalytischer Verfahren können zeitliche Abhängigkeiten aufeinanderfolgender Beobachtungen analysiert werden.

Im Rahmen der multivariaten Zeitreihenanalyse sind Vektorielle Autoregressive Moving-Average-(VARMA)-Modelle als direkte vektorielle Verallgemeinerungen der herkömmlichen univariaten Modelle nach dem Box-Jenkins-Ansatz zu verstehen. Bei der zeitreihenanalytischen Modellierung stochastischer Prozesse stehen Zufallskomponenten im Mittelpunkt der Betrachtung. Eine Zeitreihe ist dabei als Realisation eines stochastischen Prozesses definiert, wobei eine aktuelle Beobachtung als gewichtete Summe zurückliegender Beobachtungen sowie zeitgleicher und zurückliegender Zufallseinflüsse beschrieben wird. Da in der Realität eine Variable auch von anderen Variablen abhängt, ist zu erwarten, die Prognosequalität jeder in einem System verbundenen Variablen zu verbessern, wenn man sie gemeinsam betrachtet und modelliert (Rinne & Specht, 2002).

Die Erforschung der Zeitreihenanalyse hat ihren Ursprung in Ökonometrie und Naturwissenschaften, was nicht zuletzt dazu führte, dass der *Nobelpreis für Wirtschaftswissenschaften im Jahr 2003* an Clive W.J. Granger für die Entdeckung und Erforschung der *Kointegrationsmethodologie* verliehen wurde. Im Rahmen der Diplomarbeit werden gewisse Teilaspekte hieraus vor psychologischem Hintergrund aufgegriffen.

Der Mensch kann als *dynamisches System* aufgefasst werden. In der Psychologie sind unter anderem Phänomene wie Lernen, Vergessen, Entwicklung, Reifung, Alterungsprozesse und zwischenmenschliche Interaktionsmuster Gegenstand der Forschung (Schmitz, 1989). „Allen diesen Phänomenen gemeinsam ist die Zeitdimension als konstituierendes Element“ (Werner, 2005, S. 1). So betonen Rinne und Specht (2002, S. 3ff.), die dem Thema „Zeit in der Psychologie“ ein Kapitel widmen:

In der Psychologie steht das **Zeiterlebnis** im Vordergrund: als Erinnerung (→ Vergangenheit), als momentanes Denken, Fühlen und Handeln

(→ Gegenwart) oder als Planung und Erwartung (→ Zukunft). (Hervorhebungen im Original)

Dennoch dominieren in der Psychologie in empirischen Untersuchungen eher querschnittliche als prozessbezogene Analyseverfahren und „die Prozessforschung hat in der sozialwissenschaftlichen Forschung demnach noch nicht den Stellenwert, der ihr aufgrund des Anwendungspotentials und der Aussagefähigkeit zukommen müsste“ (Schmitz & Wiese, 1999, S. 217). Gleichzeitig muss der Hinweis gegeben werden, dass beispielsweise in der allgemeinspsychologischen Anwendungsforschung das dynamische Lern- und Problemlöseverhalten des Menschen u.a. durch das Forschungsparadigma *finiter Automaten* (vgl. z.B. Buchner & Funke, 1993) und die *DYNAMIS-Szenarien* (vgl. u.a. Schulz, 2003) abgebildet werden kann. Aufgrund der zeitlichen Abhängigkeiten, die immer dann entstehen, wenn Daten über mehrere Messzeitpunkte hinweg erhoben werden, ist die Zeitreihenanalyse das methodologisch angemessene *Auswertungsverfahren für Prozess-Daten*. In diesem Sinne argumentieren Farrell, Wagenmakers und Ratcliff (2004, S.3) für die Wichtigkeit *dynamischer* Auswertungsverfahren:

Research in experimental psychology usually proceeds by collecting data from participants, averaging the data for each participant, and performing inferential statistics on the subject means. This standard practice necessarily hides any changes in behavior that occur over time, whether during an experiment or across a lifetime. Just as many psychophysiological measures such as electroencephalograms (EEG), event related potentials (ERP), or heart rate fluctuations critically depend on the ordering of a system's behavior in time, so too do psychological phenomena such as development, fatigue, practice, decision making, learning, and attention possess inherently temporal characteristics. This belief underlies the extension in focus, in some areas of psychology, to dynamical theories of human behavior.

In jüngster Zeit lässt sich jedoch in der psychologischen Forschungspraxis eine zunehmende Modellierung psychologischer Phänomene durch zeitbezogene Analyse-

verfahren verzeichnen. Die Voraussetzung für die Anwendung des zeitreihenanalytischen ARMA-Ansatzes von Box und Jenkins (1976), auf dem die multivariaten VARMA-Modelle basieren, ist **Stationarität**. Die Stationaritätsannahme bezieht sich auf die *zeitliche Stabilität* von Prozessen, weil für die Schätzung der Momente einer Zeitreihe die zeitliche Homogenität des zugrundeliegenden Prozesses notwendig ist. Die zu analysierenden Zeitreihen sollten in dem zu betrachtenden Zeitfenster keine starken Veränderungen im Niveau aufweisen. Dabei ist - wie im Rahmen der Arbeit deutlich wird - zwischen **differenzenstationären (DS-)** und **trendstationären (TS-)** Prozessen zu unterscheiden. Im empirischen Teil der Diplomarbeit spielen multivariate VAR(1)-Prozesse eine Rolle, deren Einzelreihen durch den einfachsten Fall eines integrierten bzw. instationären DS-Prozesses, nämlich durch sogenannte **Random Walks**, repräsentiert sind.

Werner (2005, S. 2) betont im Vorwort des zu Beginn dieses Jahres erschienenen Readers zum Thema *Zeitreihenanalysen mit Beispielen aus der Psychologie*, dass der Kointegrationsansatz intensiv verfolgt werden müsse, da durch ihn „eine Erfassung des Zusammenwirkens der beteiligten Prozesse trotz vorliegender Instationaritäten möglich“ werde. **Kointegration** als multivariates Abhängigkeitskonzept ermöglicht die Modellierung von Zusammenhängen instationärer Variablen und hat deshalb einen „Paradigmenwechsel“ (Hassler, 2003, S. 1) mit sich gebracht, worauf der „herausragende Stellenwert“ (S. 4) der Kointegrationsmethodologie zurück zu führen ist.

Das Ziel der Diplomarbeit besteht in der Untersuchung des Kointegrationskonzepts aus psychologischer Perspektive und umfasst sowohl die theoretische Klärung als auch die empirische Validierung der theoretischen Annahmen durch systematische Simulationen.

Die *multivariate Analyse instationärer Zeitreihen* mithilfe des Kointegrationsansatzes erhöht die Anwendungsrelevanz längsschnittlicher Verfahren für die psychologische Forschung *in doppelter Hinsicht*: 1. sind instationäre Prozesse modellier- und analysierbar und 2. ermöglicht die multivariate Analyse die Betrachtung mehrerer voneinander abhängiger Variablen im Zeitverlauf. Diese Charakteristika kointegrativer Verfahren implizieren im Vergleich zu herkömmlichen zeitreihenanalytischen

Verfahren die Abbildung eines breiteren Realitätsspektrums durch die multivariate zeitreihenanalytische psychologische Forschung mit kointegrierten Modellen, da psychologische Phänomene im Zeitverlauf instationär sein können und in Abhängigkeit von anderen Phänomenen zu betrachten sind.

Die Definitionen von Kointegration unterscheiden sich in der Fachliteratur v.a. in ihrer Restriktivität sehr deutlich voneinander. So herrscht in der *scientific community* z.B. Uneinigkeit darüber, ob ein multivariater Prozess mit auseinanderdriftenden Zeitpfaden der Komponentenprozesse auch als kointegriertes System aufzufassen ist. Aufgrund dieser uneinheitlichen theoretischen Ausgangssituation werden die unterschiedlichen definitorischen Ansätze dargestellt und kontrastiert, um gemeinsame Bedingungen für das Vorliegen von Kointegration herauszukristallisieren, damit Kointegration als ein Konzept definiert werden kann, das für den psychologischen Forscher *interpretier- und testbar* und damit anwendbar ist. Aus der Betrachtung der Unterschiede und der daraus teilweise resultierenden offenen Fragen lassen sich Hypothesen für den empirischen Teil der Arbeit ableiten.

Allgemein vereint das Kointegrationskonzept die beiden bis dahin nur getrennt voneinander betrachteten und betonten Prozesseigenschaften des Gleichgewichts einerseits und der dynamischen Anpassungsprozesse andererseits. Kointegrierte Variablen, die in einem multivariaten Prozess zusammengefasst sind, sind als ein stabiles System aufzufassen, in dem Abweichungen vom langfristigen Gleichgewicht nur vorübergehender Natur sind.

Da Kointegration und Scheinregression (*Spurious Regression*) als gegenläufige Konzepte auffassbar sind, wird ihr Verhältnis theoretisch erörtert.

Das Engle-Granger-Repräsentationstheorem postuliert, dass kointegrierte Reihen eine Fehlerkorrektur-Darstellung haben und umgekehrt. In Bezug darauf werden die Eigenschaften der Fehlerkorrektur-Form anhand der beiden Parameter Kointegrationsvektor β und Anpassungskoeffizient α erläutert. Für den bivariaten Fall mit *einer* Kointegrationsbeziehung wird die Rückrechnung aus dem Fehlerkorrekturmodell auf $\alpha\beta'$ theoretisch vereinfacht.

Kointegration ist durch unterschiedliche Testverfahren operationalisierbar, die im theoretischen Teil erläutert und im empirischen Teil miteinander verglichen werden. Gegenstand der Betrachtung sind dabei der aus der univariaten Zeitreihenanalyse bekannte Einheitswurzeltest (Unit-Root-Test) von Dickey und Fuller (1981) und der Johansen-Kointegrationstest (Johansen, 1988).

Da der Wertebereich des Differenzenoperators d nicht nur auf den ganzzahligen und positiven Wertebereich beschränkt sein muss, existieren neben den herkömmlichen integrierten Prozessen fraktional integrierte Prozesse, bzw. (V)ARFIMA-Modelle. Ein Literaturüberblick zeigt, dass diese Modelle sich durch eine erhöhte Anwendungsfrequenz in der aktuellen experimentellen, v.a. allgemeinspsychologischen, Forschung auszeichnen. Dies gibt Anlass, auf das ebenfalls sehr uneinheitlich und selten explizit definierte Konzept der fraktionalen Kointegration hinzuweisen.

Ziel des empirischen Teils der Arbeit ist es, einige in den Definitionen enthaltene Prämissen mithilfe von systematischen *Monte-Carlo-Simulationen* zu überprüfen. Zu diesem Zweck können die theoriegeleitet formulierten Detail-Hypothesen den folgenden Themenbereichen zugeordnet werden: *Klärung des Kointegrationskonzepts* und *Güte der Schätzung von α und β* .

Die Diplomarbeit ist eine psychologisch-methodologische Arbeit. Welche Wichtigkeit die Methodologie für die psychologische Forschungspraxis hat, wird von Erdfelder und Funke im Vorwort des 2004 erschienenen Buches folgendermaßen in Worte gefasst. Das Zitat soll das übergeordnete Ziel der Diplomarbeit verdeutlichen.

Gemeint ist damit, dass eine explizit formulierte Methodologie, deren Prämissen akzeptabel sind, ganz praktische Entscheidungshilfen bei allfälligen methodischen Entscheidungen in der Forschungspraxis bietet und somit unmittelbar anwendungsrelevant ist.

Theoretische Grundlagen

1 Multivariate Zeitreihenanalyse:

Vektorielle stochastische Prozesse

Wenn - wie für viele Fragestellungen angemessen - Abhängigkeiten zwischen mehreren Variablen modelliert und untersucht werden sollen, reicht der univariate Ansatz, d.h. die Betrachtung einzelner stochastischer Prozesse, nicht aus. So schreiben Rinne und Specht (2002, S. 485):

Die Erklärung und Prognose einer Variablen aus nur ihrer eigenen Entwicklung ist oft nicht sachgerecht und wenig treffsicher. In der Realität hängt eine Variable zeitgleich und/oder zeitverzögert auch von anderen Variablen ab, so dass man erwarten kann, die Prognosequalität jeder in einem System verbundenen Variablen zu verbessern, wenn man sie gemeinsam betrachtet und modelliert.

Wenn beispielsweise Zeitreihendaten für mehrere Variablen gegeben sind, gilt es zu klären, ob und welche Beziehungen zwischen den Variablen existieren. Als Möglichkeiten stehen einerseits einseitige bzw. kausale Beziehungen und andererseits Rückkopplungsbeziehungen bzw. Feedbacks zur Diskussion.

Vektorielle, d.h. multivariate AutoRegressive-Moving-Average- bzw. VARMA-Modelle sind als direkte Verallgemeinerungen der herkömmlichen univariaten Modelle nach dem Box-Jenkins-Ansatz (vgl. Box & Jenkins, 1976; Werner, 2005, S. 73ff.) zu verstehen. In der *stochastischen* Zeitreihenanalyse stehen Zufallskomponenten, d.h. Störeinflüsse unbekannter Herkunft, im Mittelpunkt der Betrachtung. Dabei ist eine Zeitreihe die Realisation eines stochastischen Prozesses. In der ARMA-Modellierung werden zeitliche Abhängigkeiten aufeinanderfolgender Beobachtungen analysiert. Der stochastische Ansatz zeichnet sich dabei dadurch aus, dass eine aktuelle Beobachtung als gewichtete Summe zurückliegender Beobachtungen sowie zeitgleicher und zurückliegender Zufallseinflüsse beschrieben wird. Im multivariaten Pendant

werden einzelne stochastische Prozesse zum Beispiel durch vektorielle Prozesse und skalare Parameter durch Parametermatrizen ersetzt.

Ein multivariater oder m -dimensionaler¹ Vektor-Prozess definiert sich durch den Vektor $\mathbf{x}_t = (x_{1t}, x_{2t}, \dots, x_{mt})$, dessen Elemente $x_{it}, i = 1, 2, \dots, m; t = 1, 2, \dots$ für stochastische Prozesse stehen.

Es sind sowohl Vektorielle Autoregressive-(VAR-)Prozesse als auch Vektorielle Moving-Average-(VMA-) oder Vektorielle Autoregressive-Moving-Average-(VARMA-)Prozesse im Rahmen der multivariaten Zeitreihenanalyse modellierbar. Da im Rahmen des empirischen Teils dieser Diplomarbeit nur VAR-Prozesse simuliert werden, sollen auch nur deren Eigenschaften im kommenden Abschnitt eingehender erläutert werden.

1.1 Vektorielle Autoregressive-(VAR-)Prozesse

Der Prozess

$$\mathbf{x}_t = \mathbf{c} + \Phi_1 \mathbf{x}_{t-1} + \Phi_2 \mathbf{x}_{t-2} + \dots + \Phi_p \mathbf{x}_{t-p} + \boldsymbol{\epsilon}_t \quad (1.1)$$

wird als vektorieller (m -dimensionaler) autoregressiver Prozess der Ordnung p (VAR(p)-Prozess) bezeichnet. Darin sind: $\Phi_i (i = 1, 2, \dots, p)$ ($m \times m$)-Parametermatrizen mit $\Phi_p \neq \mathbf{0}$; \mathbf{c} ein Konstantenvektor (der ohne Beschränkung der Allgemeinheit häufig gleich einem Nullvektor gesetzt wird); $\boldsymbol{\epsilon}_t$ vektorielles weißes Rauschen als aktueller Zufallseinfluss, der den aktuellen Prozess zusammen mit dem Prozess zu früheren Zeitpunkten bestimmt.

Terminologisch stellt der Begriff *Autoregression* eine Analogie zur Multiplen Regression in dem Sinne her, dass die Prozessrealisationen zu vorangegangenen Zeitpunkten die Prädiktoren darstellen, während der Prozess selbst das Kriterium beschreibt. Ein autoregressiver Prozess zeichnet sich durch ein unendliches Gedächtnis aus, da er zu jedem Zeitpunkt² t auf sich selbst „zurückgreift“ (Stier, 2001, S. 44). McCleary und

¹Von einigen Autoren wird für die Angabe der Dimension n gewählt. In dieser Arbeit wird wie von Rinne und Specht (2002) aber bewusst m verwendet, um eine mögliche Verwechslung mit der Notation der Stichprobengröße zu vermeiden.

²Allgemeine Anmerkung zur Notation: im Rahmen der Diplomarbeit steht t für den *Zeitpunkt*

Hay (1980, S. 55f.) beschreiben autoregressive Prozesse als unendliche Summe von gewichteten vergangenen Zufallseinflüssen (Zufallsschocks bzw. **Random Shocks**), von denen jeder einzelne den zukünftigen Verlauf des autoregressiven Prozesses beeinflusst. Seine Wirkung werde jedoch mit der Zeit immer kleiner.

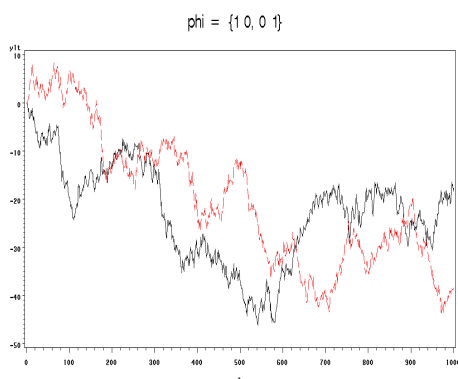


Abbildung 1.1: Bivariater Prozess mit 2 (unabhängigen) Komponentenprozessen; T=1000

Der Fehlerterm ϵ_t in multivariaten Modellen ist eine vektorielle Verallgemeinerung des Fehlerterms ϵ_t in univariaten Modellen. In der autokorrelierten Fehlerstruktur eines univariaten $AR(p)$ -Prozesses (vgl. Werner, 2005, S. 75) wird der Fehlerterm im Zeitraum t über die vorangegangenen Fehlerterme und einen **Random Shock** mit der *White-Noise*-Verteilung $WN(0, \sigma^2)$ bestimmt, d.h. die Zufallsschocks sind *independently identically distributed (iid)*, was von den meisten Autoren mit *identisch und unabhängig voneinander normalverteilt* übersetzt wird (vgl. z.B. Schmitz, 1989, S. 43, bzw. Thome, 1997, S. 204). Weitere Ausführungen zu **White Noise** sind auf Seite 22 zu finden.

Analog zu einem univariaten $AR(p)$ -Prozess kann ein $VAR(p)$ -Prozess in Backshift-Operator-Schreibweise³ folgendermaßen geschrieben werden:

$$\Phi(B)\mathbf{x}_t = \mathbf{c} + \epsilon_t \quad (1.2)$$

Das zugehörige Matrix-Lag-Polynom mit der $(m \times m)$ -Einheitsmatrix lautet:

$$\Phi(B) := \mathbf{I}_m - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p \quad (1.3)$$

Dabei beschreibt die Gleichung

$$\mathbf{x}_t = (\mathbf{I}_2 - \Phi B)\mathbf{x}_t \quad \text{oder} \quad \mathbf{c} + \Phi \mathbf{x}_{t-1} + \epsilon_t = \mathbf{c} + \epsilon_t$$

und tritt deshalb auch als Index auf, während T für die *Gesamtmenge* im Sinne von *Anzahl der Messzeitpunkte* verwendet wird.

³Erläuterungen zum Backward-Shift-Operator sind im Anhang A, S. 139, zu finden.

den einfachsten Fall eines AR(1)-Prozesses der Dimension $m = 2$, also einen bivariaten VAR(1)-Prozess. Häufig wird bei der (2×2) -Parametermatrix Φ auf das Suffix 1 für den AR-Grad verzichtet.

Ausgeschrieben ist der Prozess wie üblich folgendermaßen darstellbar:

$$\begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

oder

$$x_{1t} = c_1 + \phi_{11}x_{1,t-1} + \phi_{12}x_{2,t-1} + \epsilon_{1t}$$

$$x_{2t} = c_2 + \phi_{21}x_{1,t-1} + \phi_{22}x_{2,t-1} + \epsilon_{2t}$$

Ist der VAR(p)-Prozess aus Gleichung 1.1, S. 16, mit

$$E(\epsilon_t \epsilon'_{t-h}) = \begin{cases} \Sigma, & h = 0 \\ \mathbf{0}, & h \neq 0 \end{cases}$$

stationär⁴, dann gilt:

$$E(\mathbf{x}_t) = \boldsymbol{\mu} = \mathbf{c} + \Phi_1\boldsymbol{\mu} + \Phi_2\boldsymbol{\mu} + \dots + \Phi_p\boldsymbol{\mu}$$

d.h. es ist:

$$\boldsymbol{\mu} = [\mathbf{I}_m - \Phi_1 - \Phi_2 - \dots - \Phi_p]^{-1} \mathbf{c}$$

und der Prozess kann in der Form

$$\mathbf{x}_t - \boldsymbol{\mu} = \Phi_1(\mathbf{x}_{t-1} - \boldsymbol{\mu}) + \dots + \Phi_p(\mathbf{x}_{t-p} - \boldsymbol{\mu}) + \epsilon_t$$

geschrieben werden. Wie bei univariaten AR-Prozessen wird aus Gründen der Einfachheit häufig $\boldsymbol{\mu} = \mathbf{0}$ gesetzt.

Da ein multivariater Prozess mehrere Einzelprozesse mit einschließt, müssen im Gegensatz zum univariaten Fall nicht nur die „inneren“ zeitlichen Abhängigkeiten eines Prozesses x_t - ausgedrückt durch Autokorrelations- bzw. Autokovarianzfunktionen -, sondern auch die Abhängigkeiten oder Interaktionen zwischen den einzelnen Komponentenprozessen mithilfe von Kreuzkorrelations- bzw. Kreuzkovarianz-Funktionen berücksichtigt werden.

⁴Erklärende Erläuterungen sind im Exkurs zur Stationarität unter 2.1, S. 21f., zu finden.

In Abhängigkeit davon, welche der Elemente der Φ -Matrix Null oder verschieden von Null sind, lassen sich verschiedene zeitliche Beziehungen zwischen x_{1t} und x_{2t} ablesen. Stier (2001, S. 67f.) listet folgende Spezialfälle auf, die bei der Betrachtung der Matrizen, die im empirischen Teil der Arbeit Anwendung finden, hilfreich sein können (vgl. Auflistung der simulierten Matrizen, s. S. 88f.).

1. $\Phi = \mathbf{0}$: \mathbf{x}_t ist ein White-Noise-Prozess, weil alle $\phi_{ij} = 0$.
2. Φ als Diagonalmatrix: $\phi_{12} = \phi_{21} = 0$ mit ϕ_{11} und $\phi_{22} \neq 0$. Es gilt also:

$$\begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} \phi_{11} & 0 \\ 0 & \phi_{22} \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

Die beiden univariaten AR(1)-Prozesse stehen in keiner asynchronen Beziehung, dennoch können x_{1t} und x_{2t} kontemporär zum Zeitpunkt t korreliert sein, falls die Kovarianzmatrix von ϵ_t keine Diagonalmatrix ist, d.h. ϵ_{1t} und ϵ_{2t} korreliert sind.

3. Φ als obere Dreiecksmatrix, sodass gilt:

$$\begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ 0 & \phi_{22} \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

Es wird deutlich, dass x_{1t} durch $x_{2,t-1}$ beeinflusst wird, während x_{2t} nur von sich selbst abhängig ist, was eine einseitige/kausale Abhängigkeit impliziert. Somit ist x_{2t} als Input- und x_{1t} als Output-Prozess zu begreifen. Es erscheint der Hinweis notwendig, dass man den vektoriellen Prozess in diesem Fall in die Gruppe der Transferfunktions- oder (V)ARMAX-Modelle einordnen kann, die sowohl den univariaten als auch den multivariaten Modellen zuzuordnen sind, da die Matrix Φ die Form einer oberen Dreiecksgestalt aufweist und somit die Variablenabhängigkeit nur in einer Wirkrichtung besteht. Diese Modelle mit einseitigen oder kausalen Abhängigkeiten weisen aus diesem Grund eine Verwandtschaft zu den Regressionsmodellen auf.

4. Φ als untere Dreiecksmatrix und damit die Umkehrung von 3., in der x_{1t} als

Input und x_{2t} als Output angesehen werden kann:

$$\begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} \phi_{11} & 0 \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

5. Die Elemente der Nebendiagonalen von Φ sind von Null verschieden, während die Hauptdiagonalwerte beliebig sind. Man spricht in diesem Fall von einer Rückkopplung/Feedback zwischen x_{1t} und x_{2t} .

1.2 Voraussetzungen multivariater stochastischer Prozesse

Zu den Voraussetzungen multivariater Prozesse scheinen vorab einige allgemeine Anmerkungen vor dem Hintergrund notwendig, dass die erforderlichen statistischen Voraussetzungstests weit weniger entwickelt sind als für univariate Modelle. Man kann dennoch festhalten, dass es nicht ausreicht, die Voraussetzungstests nur auf die einzelnen Variablen anzuwenden. Sofern nicht ausschließlich Kointegration untersucht werden soll und damit - wie später verdeutlicht wird (vgl. 3, S. 32ff.) - der Stellenwert der Stationaritätsannahme relativiert ist, ist im multivariaten Fall z.B. eine Stationaritätsprüfung (vgl. die Erläuterungen im Exkurs 2.1, S. 21) lediglich anhand einzelner Variablen *nicht* ausreichend, da die (Kreuz-)Kovarianz zwischen den Variablen zu berücksichtigen ist. Werner (2005, S. 125ff.) verdeutlicht die Notwendigkeit, die Ansätze zum Unit-Root-Test (vgl. Abschnitt 2.5.2, S. 29) für die Testung der vektoriellen ARMA-Modelle auf die entsprechenden **Matrizenpolynome** und damit auf mehrfache **Eigenwerte** (vgl. Abschnitt 2.5.1, S. 28) zu generalisieren und zu fordern, dass die Nullstellen des AR- bzw. MA-Polynoms außerhalb, bzw. dass die Nullstellen der **Charakteristischen Gleichung** (vgl. Erklärungen im Abschnitt 2.5.1, S. 28) innerhalb des Einheitskreises liegen. Wie später anhand der für den empirischen Teil generierten Matrizen deutlich wird, sind diese Methoden v.a. deshalb zur Beurteilung der Stationarität heranzuziehen, da in der Φ -Matrix der VAR-Modelle sowohl einzelne Werte als auch Zeilensummen > 1 auftreten können, ohne dass der Prozess bereits explodiert⁵. Der folgende theoretische Exkurs liefert

⁵Wie einzelne Simulationen als Vorbereitung der Diplomarbeit gezeigt haben, ist die Tatsache, ob ein multivariater Prozess explodiert oder nicht, an den Eigenwerten der Modellmatrix abzulesen,

die notwendigen Definitionen.

2 Theoretischer Exkurs

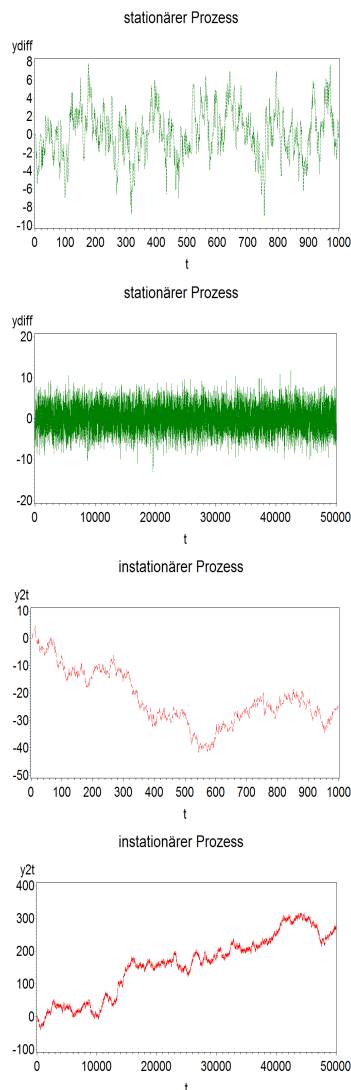


Abbildung 2.1: Stationärer Prozess (grün), instationärer Prozess (rot); jeweils für $T = 1000$ und $T = 50000$ zur Veranschaulichung des langfristigen Verlaufs stationärer und instationärer Prozesse.

2.1 Stationarität

Stationarität und Instationarität spielen bei der Darstellung von Kointegration eine entscheidende Rolle. Die Handhabung von Instationarität im Rahmen des Koin-

wird aber aus Gründen des Umfangs im Rahmen dieser Arbeit nicht ausführlicher beschrieben.

Um Ko-Integration zu definieren, ist eine formale Definition von Integration erforderlich. Aus diesem Grund werden in diesem Exkurs bereits aus der univariaten Zeitreihenanalyse bekannte Konzepte - wie das Konzept der Stationarität (2.1) und das damit unmittelbar verknüpfte Konzept der integrierten Prozesse (2.2) - skizziert. Darüber hinaus sind zur angemessenen Darstellung des Kointegrationskonzeptes Vorbemerkungen zur Differenzenbildung (2.3), die Beschreibung von Random-Walk-Prozessen (2.4) und die Darstellung der Überprüfungsmöglichkeiten von Stationarität (2.5) notwendig. Zwar gehören grundsätzliche Erläuterungen zum Problem der Scheinregression eigentlich inhaltlich in diesen theoretischen Exkurs, da Scheinregression aber in unmittelbarem Zusammenhang mit Kointegration zu diskutieren ist, wird erst in Abschnitt 3.2, S. 44ff., in das Thema eingeführt.

tegrationskonzepts erscheint nämlich im Vergleich zum univariaten Fall - wie im Verlauf der Arbeit deutlich wird - in einem völlig anderen Licht.

Im Allgemeinen wird zwischen **strenger** und **schwacher Stationarität** unterschieden. Strenge Stationarität ist dann gegeben, wenn die Verteilungsfunktion gegenüber einer zeitlichen Verschiebung invariant ist, sodass der „Beobachtungszeitpunkt unwesentlich für das Beobachtungsergebnis“ (Rinne & Specht, 2002, S. 160) ist, weil *alle* statistischen Kennwerte über die Zeit hinweg keiner Schwankung unterliegen dürfen.

In Abgrenzung davon besteht die notwendige Bedingung für schwache Stationarität nur in der zeitlichen Stabilität der ersten beiden Momente Mittelwert und Varianz/Kovarianz des erzeugenden Prozesses, d.h. Erwartungswert (Niveau) und Varianz der realisierbaren Werte sind für alle Beobachtungszeitpunkte konstant. In der Regel wird schwache Stationarität als ausreichend angesehen.

Der einfachste stationäre Prozess wird als **weißes Rauschen (White Noise)** bezeichnet. Weißes Rauschen ist eine Folge von Zufallsvariablen, die nicht nur paarweise unkorreliert sind und den Erwartungswert Null und die Varianz σ^2 haben, sondern auch paarweise unabhängig und identisch verteilt sind (*independently identically distributed (iid)*); vgl. Werner, 2005, S. 11, 75).

Vektorielles Rauschen ist durch den Prozess

$$\boldsymbol{\epsilon}_t = \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \cdot \\ \cdot \\ \cdot \\ \epsilon_{mt} \end{pmatrix} \text{ mit } E(\boldsymbol{\epsilon}_t) = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

und der Kovarianzmatrixfunktion

$$E(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}'_{t-h}) = \begin{cases} \Sigma, h = 0 \\ \mathbf{0}, h \neq 0 \end{cases}$$

darstellbar.

Bei einem vektoriiellen stationären Prozess sind folglich alle beteiligten univariaten Einzelprozesse stationär.

Zum Schluss der Hinweis, dass Stationarität zwar eine notwendige aber nicht hinreichende Voraussetzung für die Schätzung der Prozessparameter aus den empirisch ermittelten Daten einer Reihe ist. Hinzu kommt die Voraussetzung der **Ergodizität**, die nach Schlittgen und Streitberg (1994, S. 231) für die meisten in Frage kommenden stationären Zeitreihen erfüllt ist.

2.2 Integrierte Prozesse

Allgemein sind (instationäre) Prozesse dann als **integriert** zu bezeichnen, wenn sie durch Differenzenbildung (vgl. Erläuterungen im nächsten Abschnitt 2.3, S. 25) in stationäre Prozesse transformiert werden können (Stier, 2001, S. 50). In diesem Exkurs sei formal-definitiorisch festgehalten, dass eine Zeitreihe x_t ohne deterministische Komponenten mit der Ordnung d integriert ist, d.h. $x_t \sim I(d)$, wenn sie nach d -facher Differenzenbildung stationär (und invertierbar) ist (Cromwell, Hannan, Laby & Terraza, 1994, S. 17 ; Engle & Granger, 1987, S. 252). Wenn x_t beispielsweise wie in vielen Anwendungsfällen $I(1)$, also vom Grad 1 integriert ist, dann ist die Reihe nach einmaliger Differenzenbildung⁶ stationär oder $I(0)$. Anschaulich steht der Integrationsgrad also für die Anzahl der Differenzenbildungen, die notwendig sind, um den Prozess stationär zu machen. In der Literatur bezeichnet man mit integrierten Prozessen meist instationäre Prozesse. Es wird aber deutlich, dass stationäre Prozesse per definitionem ebenfalls integriert sind: sie haben den Integrationsgrad $d = 0$. $I(d)$ -Prozesse besitzen grundsätzlich d Einheitswurzeln (siehe unter 2.5.2, S. 29f.). Das Tripel (p, \mathbf{d}, q) charakterisiert einen (vektoriellen) ARIMA-Prozess vollständig.

⁶Erläuterungen zur Differenzenbildung im Abschnitt 2.3, S. 25ff.

2.2.1 $I(1)$ - und $I(0)$ -Prozesse

Die wesentlichen Unterschiede zwischen $I(1)$ - und $I(0)$ -Prozessen sind durch folgende Charakteristika beschreibbar (Banerjee, Dolado, Galbraith & Hendry, 1993; Engle & Granger, 1987): $I(0)$ -Prozesse besitzen eine endliche Varianz und Innovationen haben nur kurzfristige Effekte auf die Reihe. $I(1)$ -Prozesse weisen für $t \rightarrow \infty$ dagegen eine unendliche Varianz auf, wobei die integrierte Zeitreihe die Effekte der Vergangenheit akkumuliert und Innovationen eine dauerhafte Wirkung haben. Vor diesem Hintergrund erscheint auch logisch, dass eine integrierte Reihe instationär sein muss, weil ihre weitere Entwicklung von allen Einflüssen der Vergangenheit abhängig ist und nicht an einen Mittelwert, zu dem sie letztendlich zurückkehren könnte, gebunden ist.

Zorn⁷ (2001, Kursinformation vom 4. Oktober) fasst die Beziehungen zwischen $I(1)$ - und $I(0)$ -Prozessen folgendermaßen zusammen und unterstreicht dadurch - wie später deutlich wird - die Besonderheit des Kointegrationskonzepts als *stationäre* Linearkombination *instationärer* Prozesse:

1. wenn x_{1t} $I(0)$ ist, ist auch jede Linearkombination $\beta_0 + \beta_1 x_{1t}$ von x_{1t} $I(0)$,
2. dasselbe gilt für $I(1)$ -Zeitreihen, d.h. die Transformation $\beta_0 + \beta_1 x_{1t}$ verändert die $I(1)$ -Reihe nicht,
3. wenn zwei Reihen x_{1t} und x_{2t} $I(0)$ sind, dann ist jede Kombination von $\beta_1 x_{1t} + \beta_2 x_{2t}$ ebenfalls $I(0)$,
4. wenn x_{1t} $I(0)$ und x_{2t} $I(1)$ ist, dann ist die Kombination $\beta_1 x_{1t} + \beta_2 x_{2t}$ $I(1)$, d.h. Integration dominiert Stationarität.

⁷Die deutsche Übersetzung ist aus Werner (2005, S. 134) unter Anpassung der Notation an die Schreibweise dieser Diplomarbeit entnommen.

2.3 Differenzenbildung

Differenzenbildung ist ursprünglich - vor der Einführung des Kointegrationskonzepts - das Verfahren der Wahl, um vorhandene Drifts⁸ so zu eliminieren, dass nach Differenzenbildung ein stationärer Prozess resultiert.

Es gilt für die erste Differenz ($d = 1$):

$$\Delta x_t = x_t - x_{t-1} = (1 - B)x_t$$

Unter Umständen muss die Differenzreihe Δx_t nochmals differenziert werden, bevor Stationarität erreicht wird. Die mehrfache Differenzenbildung vom Grad d ist durch das Exponieren des Differenzenoperators darstellbar, d.h. für $d = 2$:

$$\Delta^2 x_t = [(x_t - x_{t-1}) - (x_{t-1} - x_{t-2})] = (1 - B)^2 x_t.$$

Wie umstritten das Konzept der Differenzenbildung ist und dass im multivariaten Kontext diese Methode mithilfe des Kointegrationskonzepts nicht mehr notwendig ist, da instationäre Prozesse modelliert werden können, deutet Lütkepohl (1991, S. 351) an. Er beschreibt, dass bei einem VAR(p)-Prozess, dessen Komponenten alle $I(d)$ sind, die Differenzierung jeder einzelnen Komponente interessante Merkmale der Beziehung zwischen den ursprünglichen Variablen verzerre und damit die VAR-Analyse nach der Differenzenbildung nicht mehr wie üblich durchgeführt werden könne. Bei Systemen mit kointegrierten Variablen sei Differenzenbildung inadäquat. Dies steht offensichtlich im deutlichen Gegensatz zu univariaten Zeitreihen, bei denen keine andere Möglichkeit der Handhabung von Instationarität besteht. Thome (1997, S. 219) geht weiter:

⁸In vielen Quellen wird terminologisch nicht zwischen *Trend* und *Drift* differenziert und meist in diesem Zusammenhang der Begriff *Trend* verwendet. In Abschnitt 2.4, S. 26f., wird erläutert, warum diese begriffliche Unterscheidung aber aufgrund ihrer inhaltlichen Relevanz notwendig ist. Deshalb wird im Rahmen der Diplomarbeit im entsprechenden Fall entweder direkt der Begriff *Drift* verwendet oder in Anlehnung an die Quellen der Begriff *Trend* übernommen, aber durch Anführungszeichen kenntlich gemacht, dass die genaue Verwendung des Begriffs noch zu diskutieren ist.

Die Differenzenbildung führt letztlich zu einer Fehlspezifikation, wenn es darum geht, die strukturelle Beziehung zwischen zwei oder mehr Zeitreihen zu modellieren, die „kointegriert“ sind.

Murray (1994, S. 37) wird sehr deutlich, indem er die Differenzenbildung als „mistaken way“ zur Untersuchung von (linearen) Zusammenhängen zwischen potentiell kointegrierten Variablen darstellt. Dies muss unmittelbar einleuchten, denn durch die Nivellierung des langfristigen Effekts bei der Differenzenbildung wird ein *per se* *anderes* Modell interpretiert.

2.4 Random-Walk-Prozesse

Da Random-Walk-Prozesse der Gruppe differenzenstationärer Prozesse, deren langfristige Entwicklung durch stochastische Drifts bestimmt ist, zuzuordnen sind, erscheint es notwendig, kurz die Charakteristika differenzen-(DS-) und trendstationärer (TS-)Prozesse zu unterscheiden.

DS-Prozesse gehören zu den stochastischen Drift-Modellen und beschreiben instationäre, vom Grad d integrierte Prozesse, die nach d -maliger Differenzenbildung stationär sind. Im Vergleich dazu gehören TS-Prozesse zur Gruppe der deterministischen Trend-Modelle und stehen für ebenfalls instationäre Prozesse, die aber nach einer Polynomsubtraktion stationär sind (vgl. auch DeJong, Nankervis, Savin & Whiteman, 1992b). Zufallseinflüsse können den Prozessverlauf dabei nicht nachhaltig beeinflussen, da bei einem deterministischen (linearen) Trend die Zuwächse konstant sind, während bei einem stochastischen Trend der Erwartungswert der Zuwächse konstant ist (Wolters, 2000, S. 33). Das heißt, dass die Verläufe beider Prozesse durchaus sehr ähnlich sein können. Änderungen des Verlaufs haben jedoch bei beiden Prozessen unterschiedliche Ursachen. Die Richtung von DS-Prozessen entwickelt sich zufällig, während TS-Prozesse einem deterministischen Trend folgen. Für diese Arbeit ist außerdem der Sachverhalt relevant, dass TS-Prozesse bei konstanter und endlicher Varianz keine Einheitswurzeln haben, während DS-Prozesse d Einheitswurzeln besitzen und ihre Varianz für $t \rightarrow \infty$ gegen unendlich strebt. Weiterhin ist wichtig, dass Zufallsschocks bzw. sogenannte *Innovationen* einen dauerhaften Einfluss auf

das Prozessniveau haben, während deren Wirkung bei TS-Prozessen auf denjenigen Zeitpunkt beschränkt bleibt, in dem die Innovationen auftreten. Rinne und Specht (2002, S. 356ff.) und Stadnytska (2004) legen einen ausführlichen Vergleich vor.

Die für den empirischen Teil dieser Arbeit relevanten Random-Walk-Prozesse beschreiben ARIMA(0,1,0)-Prozesse und stellen einen speziellen und auch praktisch relevanten Fall von Instationarität (Stier, 2001, S. 46) dar. Formal sind Random-Walk-Prozesse autoregressive Prozesse 1. Ordnung mit $\phi = 1$. Das autoregressive Polynom $(1 - \phi B) = (1 - B)$ hat die Wurzel (Nullstelle) $B = 1$. Genauere Erläuterungen hierzu sind unter 2.5.2, S. 29ff., zu finden. Random Walks beschreiben damit sogenannte Einheitswurzel- oder Unit-Root-Prozesse. Sie können durch Differenzbildung (vgl. Abschnitt 2.3, S. 25) in stationäre Prozesse überführt werden. Es gilt für einfache Random-Walk-Prozesse:

$$x_t = x_{t-1} + \epsilon_t, \quad \text{d.h.} \quad x_t - x_{t-1} = \epsilon_t$$

Darin steht ϵ_t für stationäres weißes Rauschen. Da die Varianz mit der Zeit wächst, ist der Prozess instationär. Der Prozess pendelt unregelmäßig hin und her und kann für längere Zeit auch eine dominant aufsteigende oder abfallende Richtung verfolgen, was aus folgender Formulierung deutlich wird: Im Gegensatz zu einem stationären Prozess hat der Random Walk nur eine schwache Affinität zu seinem Erwartungswert, wie aus graphischen Darstellungen leicht ersichtlich ist. Theoretisch wiederholt er erst im Unendlichen irgendeinen Wert, den er zuvor schon einmal realisiert hat (Mills, 1990, S. 99), d.h. die Zeit für einen wiederholten „Nulldurchgang“ ist unendlich (∞). Die ständigen Auf- und Abwärtsbewegungen mit ihren unregelmäßigen, vielfachen Richtungswechseln stellen keine deterministischen, sondern stochastische Bewegungen dar, deren Korrespondenz über mehrere Zeitreihen nicht determiniert, sondern empirisch offen ist.

Murray (1994, S. 37) vergleicht den Entwicklungsverlauf von Random Walks mit dem „drunkard’s walk“, also mit dem Fortbewegungspfad eines Betrunkenen. Dieser Vergleich ist deshalb stimmig, weil Random-Walk-Prozesse als DS-Prozesse wie erläutert einen instationären *und* zufälligen Verlauf zeigen.

2.5 Überprüfung der Stationarität

2.5.1 Lag-Polynom und Charakteristische Gleichung

Wie in Abschnitt 2.4, S. 26ff., beschrieben, haben Random-Walk-Modelle eine Nullstelle von Eins und weisen damit eine Einheitswurzel auf. Dieses Konzept der Einheitswurzeln und der zugehörigen Testverfahren zur Identifikation dieser Wurzeln wird im folgenden genauer beleuchtet, da es im empirischen Teil für die Generierung von Random-Walk-Prozessen eine bedeutende Rolle spielen wird.

Wie aus der univariaten Zeitreihenanalyse bekannt ist, werden zur Identifikation von Einheitswurzeln das Lag-Polynom oder die Charakteristische Gleichung⁹, die zueinander in einer inversen Beziehung stehen, herangezogen (Werner, 2005, S. 101).

Ein VAR(p)-Prozess ist dann als stationär zu bezeichnen, wenn alle Wurzeln von

$$|\mathbf{I}_m - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p|$$

außerhalb des Einheitskreises liegen (Werner, 2005, S. 125). $|\mathbf{A}|$ bezeichnet darin die Determinante der Matrix \mathbf{A} und \mathbf{I}_m die $(m \times m)$ -Einheitsmatrix.

Das Matrix-Lag-Polynom für einen bivariaten VAR(1)-Prozess lautet abgeleitet aus Gleichung 1.3, S. 17

$$\mathbf{I}_2 - \Phi_1 B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} B$$

Wenn man den Operator B durch eine Variable z ersetzt, ergibt sich die Determinante $|\mathbf{I}_2 - \Phi_1 z|$.

Werner (2005, S. 126) zeigt, wie sich inverse Lösungen aus dem Ansatz zum Charakteristischen Polynom der Matrix $\mathbf{A} - \lambda \mathbf{I}$ ergeben, sofern das Lag-Polynom in $\Phi_1^{-1}(\mathbf{I}_2 - \Phi_1(\lambda))$ umgeformt werde.

⁹Die Lösung der Charakteristischen Gleichung entspricht der Faktorisierung des Lag-Polynoms und wird deshalb in der Regel herangezogen.

2.5.2 Unit-Root-Tests

Einheitswurzeltests (Unit-Root-Tests) dienen der Unterscheidung stationärer und instationärer Prozesse, wobei im Fall der Instationarität wie beschrieben nochmals eine Differenzierung zwischen DS- und TS-Prozessen nötig ist (vgl. Abschnitt 2.4, S. 26f.). Im Gegensatz zu TS-Prozessen weisen nur DS-Prozesse Einheitswurzeln auf, die aber in empirischen Zeitreihen von den Einheitswurzeltests nicht in jedem Fall eindeutig identifiziert werden. Im Rahmen dieser Diplomarbeit sind Random-Walk-Prozesse und damit DS-Prozesse von Bedeutung.

Die Standard-Einheitswurzeltests wurden von Dickey und Fuller ausgearbeitet (Dickey & Fuller, 1979; Dickey & Fuller, 1981; Fuller, 1976).

Zur Durchführung des Tests in der einfachen Version verwenden Dickey und Fuller drei elementare Schätzgleichungen, deren Parameter nach dem üblichen OLS-Regressionsverfahren geschätzt werden:

$$\begin{aligned}
 x_t &= \phi x_{t-1} + \nu_t && \rightarrow \text{Random Walk} \\
 x_t &= \mu + \phi x_{t-1} + \nu_t && \rightarrow \text{Random Walk} + \text{Konstante} \\
 x_t &= \mu + \beta t + \phi x_{t-1} + \nu_t && \rightarrow \text{Random Walk} + \text{Konstante} \\
 &&& + \text{deterministischer Trend}
 \end{aligned}$$

Es gilt die Annahme, dass die Residuen unabhängig und identisch normalverteilt sind $\nu_t \sim iid N(0, \sigma^2)$.

Der Dickey-Fuller-Test existiert für unabhängige Modell-Residuen („einfacher“ DF-Test) und für autokorrelierte Residuen („erweiterter/augmented“ ADF-Test). Die Nullhypothese postuliert in beiden Fällen, dass der vorliegende Prozess differenzstationär ist ($H_0 : \phi = 1$), die Alternativhypothese dagegen, dass er stationär oder trendstationär ist ($H_1 : \phi < 1$).

Das grundsätzliche Problem dieser Tests besteht in der Tatsache, dass jede endliche Reihe von T Realisationen, die tatsächlich von einem Einheitswurzelprozess generiert wurde, sich ebensogut durch ein stationäres Prozessmodell darstellen lässt, in dem die entscheidenden Koeffizienten eine Wurzel in der Nähe von 1 implizieren. Es gibt also immer lokale Alternativen, denen gegenüber der Test praktisch keine

Trennschärfe hat (was im Grunde für alle statistischen Tests gilt). Bei kurzen Zeitreihen (Daumenregel: $T \leq 100$) ist die Trennschärfe aber auch gegenüber größeren Abweichungen von der Nullhypothese so gering, dass Kritiker die Einheitswurzeltests in solchen Fällen grundsätzlich nicht für sinnvoll einsetzbar halten (DeJong, Nankervis, Savin & Whiteman, 1992a).

Während etwa bei einem t -Test mit geringer Power, der folglich kleine Unterschiede nicht nachweisen kann, die geringe Power in dem Sinne möglicherweise zu tolerieren ist, da die Unterschiede ohnehin gering sind, ist die Unterscheidung zwischen DS oder TS inhaltsrelevant. Die geringe Power der Unit-Root-Tests kann folglich dramatische Konsequenzen implizieren. Eine ausführliche Darstellung ist in der aktuellen Arbeit von Stadnytska (2004), in Stier (2001, S. 286ff.) und bei Lanne und Lütkepohl (2001) zu finden.

Kointegration

in Theorie und Simulation

If each element of a vector of time series \mathbf{x}_t first achieves stationarity after differencing, but a linear combination $\boldsymbol{\alpha}'\mathbf{x}_t$ is already stationary, the time series \mathbf{x}_t are said to be co-integrated with co-integrating vector¹⁰ $\boldsymbol{\alpha}$. (Engle & Granger, 1987; Abstract)

The components of the vector \mathbf{x}_t are said to be *co-integrated of order d, b* , denoted $\mathbf{x}_t \sim CI(d, b)$, if (i) all components of \mathbf{x}_t are $I(d)$; (ii) there exists a vector $\boldsymbol{\alpha} (\neq \mathbf{0})$ so that $\mathbf{z}_t = \boldsymbol{\alpha}'\mathbf{x}_t \sim I(d - b), b > 0$. The vector $\boldsymbol{\alpha}$ is called *co-integrating vector*. (Engle & Granger, 1987, S. 253)

Diese Definition von Engle und Granger (1987, S. 251 und 253) stammt aus dem Artikel, der hauptsächlich das Kointegrationskonzept begründete. Seitdem wurde das Konzept in vielerlei Hinsicht erforscht. Seine Definition ist deshalb in der Fachliteratur nicht immer eindeutig und unterscheidet sich teilweise sogar in wesentlichen Punkten. Die theoretische Ausgangssituation ist also nicht einheitlich und klar einzugrenzen. Um vorab die für den empirischen Teil relevanten Informationen zur Verfügung zu stellen, erscheint es bei der folgenden definitorischen Darstellung des Kointegrationskonzepts aus diesem Grund sinnvoll, sowohl die gemeinsame Essenz der einschlägigsten Definitionen herauszukristallisieren und darzustellen als auch die Unterschiede zwischen den Definitionen zum Gegenstand der Betrachtung zu machen. Die „Moderatorvariable“, in der sich die Definitionen hauptsächlich unterscheiden, ist die definitorische Enge bzw. Weite bezüglich der notwendigen Bedingungen des Kointegrationskonzepts. Ziel der theoretischen Abhandlung ist es, Kointegration definitorisch einzugrenzen, um das Konzept so scharf zu definieren, dass es test- und interpretierbar sein kann. Deshalb lassen sich aus der Betrachtung der Unterschiede und der daraus teilweise auch resultierenden offenen Fragen viele wichtige

¹⁰Anmerkung zur Notation: Im weiteren Verlauf der Arbeit ist der Kointegrationsvektor mit $\boldsymbol{\beta}$ notiert.

Impulse für den empirischen Teil der Arbeit ableiten. Allerdings lassen sich nicht alle „Impulse“ im Rahmen dieser Arbeit untersuchen. Dies sei weiteren Publikationen vorbehalten.

Zu Beginn ist es notwendig, den expliziten Hinweis zu geben, dass im empirischen Teil nur Fälle mit *einer* Kointegrationsbeziehung betrachtet werden, da eine Betrachtung von Fällen mit mehr als einer Kointegrationsbeziehung und damit mehr als zwei Komponentenprozessen den Rahmen dieser Arbeit sprengen würde. Banerjee, Dolado, Galbraith und Hendry (1993, S. 138) machen deutlich, dass viele Herangehensweisen aus dem bivariaten Fall auf andere Fälle übertragbar sind. Somit wird hier hauptsächlich der bivariate Fall als einfachster multivariater Prozess dargestellt, aus dessen Betrachtung sich wiederum Hypothesen für komplexere multivariate Fälle ableiten lassen, die in einer anderen Studie untersucht werden können. Zwar findet im Rahmen dieser Arbeit nur die Kointegration von $I(1)$ -Variablen Anwendung, doch soll auch darauf verwiesen werden, dass es Kointegration von $I(2)$ -Variablen geben kann (vgl. Engle & Yoo, 1991, S. 251; Holtemöller, 2002; Johansen, 1995a).

3 Kointegration aus theoretischer Perspektive

Wir sollten uns des Wertes der Theorie bewusst sein. Ein Geschäftsmann hat einmal festgestellt, „es sei nichts so praktisch wie eine gute Theorie“.
(Lewin, 1982, S. 217)

3.1 Definitionen in der einschlägigen Fachliteratur

Wie aus der univariaten Zeitreihenanalyse bekannt, ist Stationarität eine grundlegende Voraussetzung für die Modellierung von ARMA-Prozessen. Auch im multivariaten Fall wäre diese Voraussetzung zu erwarten. Das Kointegrationskonzept stellt einen Ausweg aus dem Dilemma dar, da vor dem Hintergrund der Stationaritätsforderung Zeitreihen bereinigt werden müssten und damit wichtige Information verlorenginge (siehe 2.3, S. 25ff.). Wie Stadnytska (2004) zeigt, werden Entscheidun-

gen über DS- oder TS-Prozesse (vgl. 2.4, S. 26ff.) häufig nicht korrekt gefällt, was zu Fehlschlüssen führt.

Mithilfe des Kointegrationsansatzes können instationäre Prozesse und somit auch die in ihnen enthaltene Information multivariat modelliert werden. Nach Stier (2001, S. 316) beschreibt Kointegration ein „stabiles System“, in dem Abweichungen vom langfristigen Gleichgewicht nur vorübergehender Natur sind. Das System könne zwar kurzfristig vom Gleichgewicht abweichen, da die Einzelprozesse ständigen entwicklungsbedingten Veränderungen unterworfen seien, kehre aber immer wieder zu diesem Gleichgewicht, das die Einzelprozesse miteinander verbinde und als Linearkombination zu schreiben sei, zurück (vgl. auch Hamilton, 1994, S. 572).

Murray (1994, S. 37) nutzt zur Illustration des Kointegrationsansatzes wie bereits zur Erklärung von Random-Walk-Prozessen (siehe unter 2.4, 26ff.) das Bild vom Betrunknen und seinem Hund. Beide für sich genommen bewegen sich wie Random Walks fort. Trotz der Nichtstationarität der beiden Pfade, würde man aber sagen: „if you find the drunkard, the dog is unlikely to be very far away“, d.h. dass die Distanz zwischen den beiden Bewegungspfaden stationär ist und von Kointegration gesprochen werden kann. Kointegration ist ein probabilistisches Konzept, da der Hund nicht an der Leine und somit der Abstand zwischen den beiden nicht fixiert ist. Trotz der Instationarität der Einzelpfade ist der Abstand für den kointegrierten Fall aber eine stationäre Zufallsvariable.

Kointegration impliziert Granger-Kausalität (Thome, 1997, S. 219), d.h. dass bei zwei kointegrierten Reihen mindestens eine Granger-kausal für die andere ist. Genauere Erläuterungen zu diesem Kausalitätskonzept sind im Anhang B, S. 140f., zu finden.

Weitgefasste Definition des Kointegrationskonzepts Das Konzept der Kointegration ist deshalb so bedeutungsvoll, weil es ermöglicht, ein Gleichgewicht zwischen zwei oder mehreren Reihen, die jede für sich genommen instationär sind, zu beschreiben. Es stellt eine Möglichkeit dar, den bekannten Begriff des Gleichgewichts mit Hilfsmitteln der Zeitreihenanalyse zu operationalisieren (Stier, 2001, S. 316). Die Momente (Mittelwert, Varianz und Kovarianz) der Einzelreihen können also über die

Zeit hinweg variieren und gleichzeitig existiert eine Linearkombination dieser Reihen mit zeitinvarianten Eigenschaften, die den Gleichgewichtszusammenhang beschreibt. Aus dieser Umschreibung des Konzeptes nach Banerjee, Dolado, Galbraith und Hendry (1993, S. 136ff.) wäre nach dem allgemeinen Stationaritätsverständnis (vgl. 2.1, S. 21f.) zwingend zu schließen, dass diese Linearkombination durch Stationarität charakterisiert ist. Banerjee, Dolado, Galbraith und Hendry (1993, S. 145) berufen sich zwar auf Engle und Granger (1987), halten die Definition von Kointegration dennoch sehr allgemein (Banerjee, Dolado, Galbraith & Hendry, 1993, S. 137):

However, a linear combination of series may have a lower order of integration than any one of them has individually. In this case the variables are said to be co-integrated.

Das heißt folglich, dass nach Meinung der Autoren *zwei einzelne Zeitreihen dann kointegriert sind, wenn eine Linearkombination dieser Variablen einen geringeren Grad an Integration aufweist als die Einzelreihen selbst.*

Sie machen in der Fußnote¹¹ deutlich, dass Kointegration in ihrer Eigenschaft als Gleichgewichtsbeziehung das Vorliegen einer stationären Linearkombination mit dem Integrationsgrad Null impliziere und beziehen sich dabei auf die Definition von Engle und Granger (1987), die Stationarität der „Abweichung“ bzw. Linearkombination („stationarity of the deviation“; Banerjee, Dolado, Galbraith & Hendry, 1993, S. 7) $x_{1t} - \beta x_{2t}$ fordert. Sie heben dennoch hervor, dass ihrer Meinung nach per definitionem jede Reduktion des Integrationsgrades ausreiche, damit Variablen als kointegriert bezeichnet werden dürfen.

If $\{x_{1t} - \beta x_{2t}\}$ has a constant unconditional mean but is not stationary, then we may still want to say that an equilibrium relationship holds; the series will not, however, fit the strict Engle-Granger definition of co-

¹¹Banerjee, Dolado, Galbraith und Hendry, 1993, S. 137, Fußnote 3: „When regarding a co-integrating combination as an 'equilibrium relationship', it is natural to expect this combination to be integrated of order zero. However, definitionally, any reduction in the order of integration - say from d to $d - b$ (where $b > 0$) - is sufficient for the variables to be called 'co-integrated'.“

integration, which requires that some linear combination be stationary.
(Banerjee, Dolado, Galbraith & Hendry, 1993, S. 7)

Die Autoren führen dabei weiter aus:

When $b > 0$ a linear relation exists between the variables which is integrated of lower order than either of the variables themselves, but which may none the less not be $I(0)$. In the latter case ($a - b > 0$), the variables may deviate from the linear relationship by an ever-growing amount, and so it is not the kind of relationship that we have been calling a long-run equilibrium. Nevertheless, variables that are $CI(a, b)$ for $b > 0$ do contain some information about the long-run behaviour of the series involved¹².
(Banerjee, Dolado, Galbraith & Hendry, 1993, S. 7)

Folglich gehen die Autoren in der letzten Bemerkung davon aus, dass auch eine Kointegrationsbeziehung, die durch eine instationäre Linearkombination charakterisiert ist ($b > 0$), Informationen über das langfristige Verhalten der Zeitreihe enthält. Dies scheint zunächst akzeptierbar, allerdings muss die genauere Modellierbarkeit einer solchen Kointegrationsbeziehung in Frage gestellt werden.

Mehrheitlich geteilte Ausgangsdefinition und deren Weiterentwicklungen

Der sehr allgemeine Ansatz, der eine Reduktion des Integrationsgrades als ausreichend ansieht, wird von der Mehrheit der Autoren nicht geteilt, denn das besondere Charakteristikum von Kointegration, das das Konzept als Teil der Forschung an Einheitswurzeln zum „most vigorous development in econometrics in recent years“ (Hatanaka, 1996, Vorwort) gemacht hat, besteht gerade in der Stationarität der Linearkombination. In diesem Sinne wäre die Reduktion des Integrationsgrades zwar eine notwendige, nicht aber eine hinreichende Bedingung für das Vorliegen von Kointegration. Wenn x_{1t} und x_{2t} $I(1)$ sind, gilt nach den herkömmlichen Regeln, dass

¹²Die Notation wird später eingeführt. a beschreibt dabei den Integrationsgrad der Variablen und b die Reduktion desselben, sodass der Integrationsgrad der Linearkombination durch $a - b$ beschreibbar ist.

Tabelle 3.1: Grundidee des Kointegrationskonzepts: VAR(1)-Prozesse; Modellmatrix Φ_1 , zugehörige Eigenwerte (EW) und Dickey-Fuller-Tests (DF) der VARMAX-Prozedur über Komponentenprozesse (black, red) und Linearkombination (green) für T=1000.

VAR(1)-Prozesse					
Modellmatrix			DF	DF	DF
ϕ_{11}	ϕ_{12}		$Pr < Tau$	$Pr < Tau$	$Pr < Tau$
ϕ_{21}	ϕ_{22}	EW	(black)	(red)	(green)
1	0	1	0,52	0,798	none
0	1	1			
0,95	0,05	1	0,5993	0,6123	<,0001
0,1	0,9	0,85			
0,8	0,2	1	0,692	0,7202	<,0001
0,1	0,9	0,7			
0,6	0,4	1	0,7201	0,7665	<,0001
0,1	0,9	0,5			
0,2	0,8	1	0,7199	0,7992	<,0001
0,1	0,9	0			

jede Linearkombination auch $I(1)$ sein wird (vgl. 2.2.1, S. 24f.). Wenn allerdings Kointegration gegeben ist, beschreibt die Linearkombination $x_{1t} - \beta x_{2t}$ einen $I(0)$ -Prozess (vgl. Granger, 1988, S. 556f.). Dies hat die außerordentlich bedeutende und bahnbrechende praktische Implikation, dass bei instationären Prozessen eine Prognose dann möglich ist, wenn diese kointegriert sind, weil die Stationarität ihrer Linearkombination eine Vorhersage möglich macht.

Die in Tabelle 3.1 aufgeführten VAR(1)-Prozesse¹³ und deren graphische Illustrationen (vgl. Abbildung 3.1, S. 37) verdeutlichen den dem Kointegrationskonzept zugrundeliegenden Gedanken im Sinne der mehrheitlich geteilten Ausgangsdefinition, die in der vorangestellten Definition auf S. 31 zusammengefasst ist.

Nach Thome (1997, S. 213) ist nur dann eine solche stationäre Linearkombination nichtstationärer Prozesse möglich, wenn die „Trend“-Bewegungen¹⁴ der einzelnen Reihen miteinander korrespondieren, wenn also eine Änderung des lokalen „Trends“ in x_{1t} mit einer mehr oder weniger schnellen Anpassung des lokalen „Trends“ in x_{2t} einhergeht. Die gemeinsamen stochastischen „Trends“ („Common Trends“), also die gemeinsamen $I(1)$ -Komponenten (Zorn, 2001, S. 1f.) der instationären Varia-

¹³In der rechten Spalte der Tabelle ist zuoberst deshalb „none“ vermerkt, da aufgrund der Unabhängigkeit der Komponentenprozesse keine Kointegration vorliegen kann, weshalb eine durch den Kointegrationsvektor gewichtete Linearkombination nicht definiert ist.

¹⁴Erläuterungen zur Handhabung der terminologischen Differenzierung zwischen *Trend* und *Drift* im Rahmen dieser Arbeit sind der Fußnote auf S. 25 zu entnehmen.

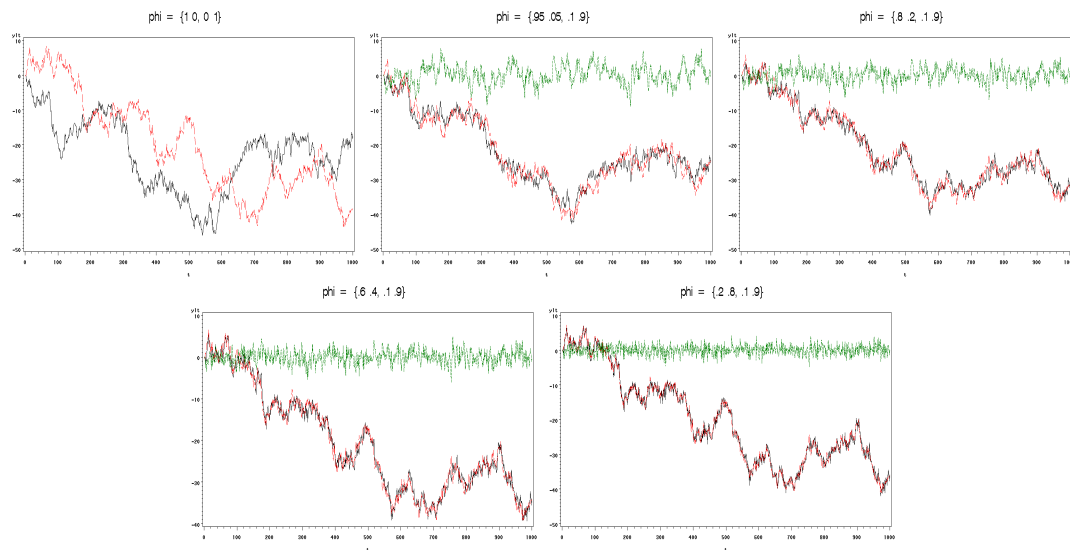


Abbildung 3.1: Grundidee der Kointegrationsmethodologie: demonstriert durch multivariate Prozesse mit zunehmender Abhängigkeit der Einzelreihen, beginnend mit einem Modell mit völlig unabhängigen Komponentenprozessen (links oben). Modelleigenschaften in Tabelle 3.1, S. 36.

blen, werden durch das Bilden einer Linearkombination kompensiert (vgl. Rinne & Specht, 2002; Stock & Watson, 1988). Wenn keine Kointegration gegeben ist, dann enthält das System zwei stochastische „Trends“. Wenn die beiden Variablen kointegriert sind, enthält das System durch den gemeinsamen stochastischen „Trend“ eine Kointegrationsbeziehung (Wolters, 2000, S. 35).

Engle und Granger (1987, S. 254) weisen explizit darauf hin, dass die Reduktion des Integrationsgrads die Besonderheit der - wie später zu zeigen sein wird - interpretier- und testbaren (vgl. 3.5, S. 60ff.) Beziehung ausmache. Wenn die Einzelprozesse bereits stationär seien, sei die Stationarität der Linearkombination trivial und nicht mehr dem Konzept der Kointegration zuzuordnen. Banerjee, Dolado, Galbraith und Hendry (1993, S. 4-5) verdeutlichen, dass bei Stationarität der Einzelreihen auch jede Relation $x_{1t} - \beta x_{2t}$ stationär und damit im statistischen Gleichgewicht sei, und zwar unabhängig davon, ob es ein β gebe, das ein echtes Gleichgewicht beschreibe.

Stier (2001, S. 315f.) legt ähnlich wie Engle und Granger (1987; vgl. Definition auf S. 31) Bedingungen für Kointegration fest. Es sei *notwendig* für die Kointegration von x_{1t} und x_{2t} , dass sie den gleichen Integrationsgrad besäßen. Hinreichend ist diese Bedingung seiner Meinung nach aber noch nicht. Kointegration liege nur

dann vor, wenn außer der gleichen Kointegrationsordnung d eine Linearkombination $\beta_1 x_{1t} + \beta_2 x_{2t}$ mit dem Integrationsgrad $d - b$ existiere, d.h.: $x_{1t}, x_{2t} \sim CI(d, b)$ mit $d \geq b > 0$ und (β_1, β_2) als Kointegrationsvektor. Das erste Element des Kointegrationsvektors ist in der Regel auf Eins normiert¹⁵ (vgl. Erläuterungen zur Normierung des Kointegrationsvektors auf S. 39f.). Die Funktion des Kointegrationsvektors wird ausführlich in Abschnitt 3.4.1, S. 57ff., illustriert.

Allgemein ist ein vektorieller Prozess $\mathbf{x}'_t = (x_{1t} \ x_{2t} \ \dots \ x_{mt})$ nach Stier (2001, S. 315f.)¹⁶ dann als kointegriert zu bezeichnen, wenn jedes x_{it} ($i = 1, 2, \dots, m$) vom Grad d integriert ist und wenn ein Vektor $\beta' = (\beta_1 \ \beta_2 \ \dots \ \beta_m)$ existiert, sodass gilt $\beta' \mathbf{x}_t \sim I(d, b)$.

Anders als Banerjee, Dolado, Galbraith und Hendry (1993) heben Rinne und Specht (2002, S. 540) und Stier (2001) hervor, dass der Fall $d = b = 1$, d.h. dass alle Einzelprozesse vom Grad 1 integriert sind, mit $\mathbf{x}_t \sim CI(1, 1)$ in der Praxis am bedeutendsten sei. Für diesen Fall beschreibt die Linearkombination auch einen $I(0)$ -Prozess. Im Artikel von Engle und Granger (1987, S. 251), der unter anderen das Kointegrationskonzept begründete, legen sich die Autoren bezüglich dieser Frage fest, indem sie fordern, dass nur dann von Kointegration gesprochen werden dürfe, wenn eine stationäre Linearkombination¹⁷ $\beta' \mathbf{x}_t$ existiere. Andere Autoren (Hamilton, 1994, S. 571ff.; Rinne & Specht, 2002, S. 541; Stewart und Gill, 1998, S. 243) schließen sich

¹⁵Auf diese Normierung kann zurückgeführt werden, dass in den zitierten Quellen für das zweite Element des Kointegrationsvektors sowohl β_2 als auch einfach β geschrieben wird (vgl. Werner, 2005, S. 132). Da der Fettdruck β aber immer unmissverständlich kenntlich macht, wann der Kointegrationsvektor und nicht sein zweites Element β gemeint ist, sind im Rahmen der Diplomarbeit in Anlehnung an die jeweiligen Quellen beide Notationen zulässig.

¹⁶Stier verwendet für die Angabe der Dimension die Notation n , die in dieser Arbeit durch m ersetzt wird, um Verwechslungen mit der Notation der Stichprobengröße zu vermeiden (vgl. Festlegung der Schreibweise auf S. 16).

¹⁷Kointegration beschreibt eine Linearkombination instationärer Variablen. Es ist nicht ausgeschlossen, dass es auch nichtlineare Kombinationen instationärer Variablen gibt, die ein langfristiges Gleichgewicht beschreiben. Eine Theorie der nichtlinearen Kointegration steht nach Rinne und Specht (2002, S.541) allerdings noch aus. Hinweise finden sich bei Granger (1991, S. 284ff.). Bei Boucher (2003) kann ein Anwendungsbeispiel für nichtlineare Kointegration nachgelesen werden.

dieser Meinung an, wobei Hamilton (1994, S. 571) aber bei seiner Definition von Kointegration die Integrationsgrade dezidiert festlegt. Eine vektorielle Zeitreihe \mathbf{x}_t sei dann kointegriert, wenn jede Einzelreihe $I(1)$ - und damit nichtstationär mit einer Einheitswurzel - und die Linearkombination $\beta' \mathbf{x}_t$ mit $\beta \neq \mathbf{0}$ stationär, d.h. $I(0)$, sei. Die Abweichungen vom Gleichgewicht besitzen eine beschränkte Variation (Wolters, 2000, S. 33) und „in the long run“ bewegen sich die untransformierten (vgl. Zorn, 2001, S. 17) Variablen also auf das Gleichgewicht $\beta' \mathbf{x}_t = 0$ zu, denn es gilt: $x_{1t} - \beta x_{2t} = \epsilon_t$, wobei $\epsilon_t \sim I(d - b = 0)$, mit $b > 0$ und $d \geq 0$. Nur unter dieser Bedingung ist der Begriff des Gleichgewichts gerechtfertigt (vgl. Rinne & Specht, 2002, S. 539). Die Variable ϵ_t beschreibt das Ausmaß, mit dem das System \mathbf{x}_t vom Gleichgewicht abweicht und wird deshalb Gleichgewichtsfehler oder „equilibrium error“ (Cromwell, Hannan, Labys & Terraza, 1994, S. 18; Wolters, 2000, S. 33) oder „Gleichgewichtsstörung“ (Rinne & Specht, 2002, S. 539) genannt. Die notwendigen Bedingungen einer *interpretier- und testbaren* Kointegrationsbeziehung, die sich aus den beschriebenen Definitionen ableiten lassen, sind in der folgenden Tabelle zusammengefasst:

Notwendige Bedingungen (NB) einer interpretier- und testbaren Kointegrationsbeziehung
NB 1: instationäre Einzelreihen mit identischem Integrationsgrad $[I(d)]$
NB 2: geringerer Integrationsgrad der Linearkombination $[I(d - b), b > 0]$
NB 3: Stationarität der Linearkombination $[I(0)] \Rightarrow$ „Gleichgewicht“

Zur Normierung des Kointegrationsvektors, die auch im empirischen Teil Beachtung finden wird, scheinen einige allgemeine Erläuterungen notwendig, die sich auch auf vektorielle Prozesse mit mehr als zwei Variablen übertragen lassen. Die Funktion des Kointegrationsvektors β als Gewichtevektor der Linearkombination (Murray, 1994, S. 37; Thome, 1997, S. 219) ist es, aus den m nichtstationären Prozessen *einen* stationären Prozess zu generieren. Wie oben erläutert, sind zwei $I(1)$ -Prozesse x_{1t} und x_{2t} mit dem Kointegrationsvektor $\beta' = (1 \quad -\beta)$ kointegriert,

weil die Linearkombination $x_{1t} - \beta x_{2t}$ stationär ist. Allgemein existieren für m Einzelreihen genau h unabhängige Kointegrationsvektoren, die für $h < m$ in einer $(h \times m)$ -Kointegrationsmatrix¹⁸ zusammengefasst werden können. h wird auch (Ko-)Integrationsrang genannt (Rinne & Specht, 2002, S. 541; Werner, 2005, S.133). Das erste Element des Kointegrationsvektors ist in der Regel auf Eins normiert (vgl. Hamilton, 1994, S. 574; Stier, 2001, S. 316). Diese Normierung dient dem Zweck der Eindeutigkeit, denn wenn $\beta' \mathbf{x}_t$ stationär ist, dann ist auch ein Produkt $c \beta' \mathbf{x}_t$ aus $\beta' \mathbf{x}_t$ und einem beliebigen Skalar c ($\neq 0$) stationär. Abgesehen von der Normierung gibt es *bei nur zwei Prozessen* ($m = 2$) *nur einen Kointegrationsvektor*. Bei mehr als zwei Prozessen ist dies nicht mehr der Fall. Ausführlichere Erläuterungen hierzu und zur Normierung für $m > 2$ finden sich bei Rinne und Specht (2002, S. 541).

Auseinanderdriftende Zeitpfade der Komponentenprozesse Im Gegensatz zu den meisten Autoren, die auseinanderdriftende Zeitpfade der Komponentenprozesse *nicht* explizit in die Definition von Kointegration mit einschließen, relativieren Rinne und Specht (2002, S. 542f.) die Behauptung, dass kointegrierte Variablen Zeitpfade aufweisen, die *nicht* auseinander driften.

Dabei muss terminologisch klar gestellt werden, dass sich in diesem Zusammenhang das Adjektiv „auseinanderdriftend“ nur auf den *graphischen* Eindruck bezieht und dass kointegrierte Reihen sich statistisch auch im Rahmen dieser Definition in einer Gleichgewichtsbeziehung befinden.

Auch wenn andere Autoren möglicherweise diesen Ansatz in ihre Definition von Kointegration implizit mit eingeschlossen haben mögen, weisen Rinne und Specht (2002) das Kointegrationskonzept mit dieser Aussage aus. Mit diesem Hinweis bewegen sie sich aber weiterhin im Rahmen der oben und von anderen Autoren beschriebenen Grundidee von Kointegration (vgl. zusammenfassende Tabelle auf Seite 39).

Es wird oft behauptet, dass kointegrierte Variablen Zeitpfade aufweisen,

¹⁸Da in dieser Arbeit bivariate Modelle mit nur einer Kointegrationsbeziehung und damit nur einem Kointegrationsvektor im Mittelpunkt stehen, wird auf die genaue Darstellung der Kointegrationsmatrix an dieser Stelle verzichtet. Sie ist z.B. bei Werner (2005, S.133) dargestellt.

die nicht auseinander driften. Das trifft aber nur dann zu, wenn - etwa bei $m = 2$ Variablen - der Kointegrationsvektor durch $\beta' = (1, -1)$ gegeben ist. (Rinne & Specht, 2002, S. 542)

Wenn man davon ausgehe, dass $(x_{1t}, x_{2t}) \sim CI(1, 1)$ mit dem Kointegrationsvektor $\beta' = (1 \quad -\beta_2)$ vorliege, dann gelte $x_{1t} - \beta_2 x_{2t} \sim I(0)$. Wenn „im Sonderfall“ (Rinne & Specht, 2002, S. 542) die Reihen x_{1t} und x_{2t} nicht auseinander driften würden, dann müsse ihre Differenz $x_{1t} - x_{2t} \sim I(0)$ sein, d.h. $\beta_2 = 1$.

Die Autoren leiten weiter her (S. 542f.)

Im allgemeinen Fall ($\beta_2 \neq 1$) ist die gewöhnliche Differenz $Y_{1t} - Y_{2t} = Y_{1t} - \beta_2 Y_{2t} - (1 - \beta_2)Y_{2t}$ instationär, denn sie ist eine Linearkombination einer stationären Variablen ($Y_{1t} - \beta_2 Y_{2t}$) und einer $I(1)$ -Variablen $(1 - \beta_2)Y_{2t}$.¹⁹

Folglich - so machen die Autoren deutlich - würden sich die beiden kointegrierten Variablen mit wachsender Zeit voneinander entfernen und die Differenz hänge vom Niveau einer der beiden Variablen ab.

VAR(1)-Prozesse mit auseinanderdriftenden Zeitpfaden der Komponentenprozesse sind in Abbildung 3.3, S. 57, graphisch veranschaulicht.

Gleichheit des Integrationsgrads der Komponentenprozesse In der Literatur herrscht - wie bereits auch bei der Definition von Banerjee, Dolado, Galbraith und Hendry (1993) angedeutet - Uneinigkeit bezüglich der Frage nach der Gleichheit der Integrationsgrade der Einzelprozesse. Engle und Granger (1987, S. 253) formulieren in ihrer Definition von Kointegration als eine notwendige Voraussetzung „all components of \mathbf{x}_t are $I(d)$ “, d.h. dass die Einzelprozesse denselben Integrationsgrad aufweisen müssen. Neben anderen Autoren (vgl. oben: Cromwell, Hannan, Labys & Terraza, 1994, S. 17; Hamilton, 1994; Murray, 1994, S. 37; Rinne & Specht, 2002, S. 540) äußert Thome (1997, S. 213) sich hierzu in Form einer *notwendigen*

¹⁹In der Schreibweise der Diplomarbeit werden die in diesem Zitat benannten Komponentenprozesse Y_{1t} und Y_{2t} mit x_{1t} und x_{2t} notiert. Erläuterungen zur Notation des zweiten Elements des Kointegrationsvektors finden sich auf Seite 38 (Fußnote).

Bedingung, dass die Ausgangsreihen integrierte Prozesse der gleichen Ordnungsstufe d sein müssten und somit durch d -fache Differenzenbildung jeweils in stationäre Reihen transformiert werden könnten.

Die beteiligten Zeitreihen sind nicht nur „integriert“, sondern integriert mit demselben Ordnungsgrad. (Thome, 1997, S. 219)

Wenn die Reihen von unterschiedlicher Ordnung integriert seien, könnten sie keinen langfristigen linearen strukturellen Zusammenhang aufweisen. Auch Stier (2001, S. 315) hebt aufs Deutlichste hervor, dass Prozesse unterschiedlichen Integrationsgrades „**nie kointegriert**“ seien, „während Prozesse gleichen Integrationsgrades kointegriert sein können, aber nicht kointegriert sein müssen“ (Hervorhebungen im Original). Er argumentiert ähnlich wie Thome (1997), dass bei Einzelprozessen unterschiedlichen Integrationsgrades die Differenzen der Einzelreihen beispielsweise über die Zeit hinweg ständig zunehmen könnten und nicht wie für kointegrierte vektorielle Prozesse mit Variablen gleichen Integrationsgrades charakteristisch weder ständig größer noch kleiner würden. Da der empirische Teil der Arbeit mithilfe von SAS bestritten wird, erscheint es notwendig, darauf hinzuweisen, dass SAS seine Definition von Kointegration explizit an die Definition von Engle und Granger (1987) anlehnt (vgl. SAS/ETS Software: Changes and Enhancements, 2001).

Lütkepohl (1991, S. 352ff.) grenzt sich in Bezug auf die Frage nach der Gleichheit der Integrationsgrade der Ausgangsprozesse als notwendige Bedingung für Kointegration deutlich von den oben angeführten Autoren ab und schließt die Möglichkeit nicht aus, dass Reihen unterschiedlicher Integrationsgrade kointegriert sein können, und verweist direkt auf die Möglichkeit einer Kointegrationsbeziehung zwischen einer stationären und einer instationären Variablen. Ähnliche Hinweise auf ungleiche Integrationsgrade in kointegrierten Systemen geben auch Johansen (1995b, S. 35ff.) und Pagan (2004d, S. 10). Alle genannten Autoren führen in den zitierten Quellen diese Hinweise allerdings nicht weiter aus und wenden in ihren Beispielen diese Konzepterweiterung auch nicht an.

Da dieser definitorische Unterschied im empirischen Teil der Arbeit aus Gründen des Umfangs nicht aufgegriffen werden kann, soll auf seine ausführliche Darstellung an dieser Stelle verzichtet werden. Einzelne von mir durchgeführte Simulationen geben Hinweise auf die Existenz solcher Kointegrationsbeziehungen und sollen in einer nachfolgenden Publikation vertieft werden.

Die Begründer des Kointegrationskonzepts - Engle und Granger - schreiben:

Often researchers would like to include in the cointegrating relationship variables which are not $I(1)$. In principle, inclusion of a stationary variable is prohibited but should not affect the remaining coefficients (assuming it is not the dependent variable). It appears that it also should not affect the asymptotic critical values of the test statistics. More problematical are dummy variables for regime changes or data corrections or seasonality. (Engle & Granger, 1991, S. 14)

Banerjee, Dolado, Galbraith und Hendry (1993, S. 138) führen Gründe an, warum das Kointegrationskonzept als so zentral für die Modellierung mit integrierten Variablen und die Untersuchung der langfristigen Beziehungen zwischen diesen Variablen zu betrachten ist. Zwei der genannten Gründe sollen hier aufgeführt werden, um zu den nachfolgenden Abschnitten überzuleiten:

- Vor dem Hintergrund von Kointegration als Gleichgewicht könne wirkliche versus scheinbare Regression (spurious regression, vgl. 3.2.1, S. 44ff.) gegenübergestellt werden. Regressionen über Zeitreihen mit instationären Variablen seien nur und ausschließlich dann sinnvoll, wenn diese Variablen kointegriert seien. Ein Kointegrationstest (vgl. weitere Erläuterungen dazu in Abschnitt 3.5, S. 60ff.) könne somit als nützliche Methode angesehen werden, um wirkliche Regressionen von Scheinregressionen zu unterscheiden.
- Kointegrierte Variablen besäßen darüber hinaus noch die wichtige Eigenschaft, dass ein „Set“ kointegrierter Variablen neben anderen Repräsentationen auch eine Fehler-Korrektur-Darstellung habe, sodass der Zusammenhang durch die

Integration eines Terms im Modell, der für die Abweichung der beobachteten Werte vom langfristigen Gleichgewicht stehe, beschrieben werden könne.

Die beiden Punkte, also die Betrachtung von Scheinregression vor dem Hintergrund der Kointegration und die Fehlerkorrekturdarstellung von kointegrierten Prozessen, werden in den folgenden beiden Abschnitten (3.2 und 3.3) weiter beleuchtet, da sie auch im empirischen Teil von Bedeutung sein werden.

3.2 Kointegration und Scheinregression

3.2.1 Abhängigkeitskonzept der Scheinregression

Bereits 1926 konstatierte Yule (S. 2), dass in der Regression einer Zeitreihenvariablen auf eine andere Zeitreihenvariable eine statistisch signifikante Beziehung entdeckt werden kann, die sachlogisch nicht begründbar ist. Als humoristisches Beispiel führen Rinne und Specht (2002, S. 539) den gleichzeitigen Rückgang der Geburtenzahlen und der Zahl nistender Störche in Europa oder das Steigen der Pastorengelälter und Spirituosenpreise in Deutschland an. Aus diesen Gegebenheiten würde niemand schließen, dass der Klapperstorch die Babys bringe oder die Pastoren vorwiegend Alkoholiker seien.

Dieses von Granger und Newbold (1974, 2001) als *Scheinregression* oder *Spurious Regression* bezeichnete Phänomen ist vor dem Hintergrund des Kointegrationsansatzes aus einer neuen Perspektive zu diskutieren, sodass es an dieser Stelle notwendig erscheint, seine Grundzüge zu skizzieren.

Scheinregression beschreibt die Tatsache, dass in einem Regressionsmodell für nicht-stationäre Zeitreihen auch dann ein signifikanter Steigungskoeffizient geschätzt werden kann, wenn keinerlei strukturelle (kausale) Beziehung zwischen den Reihen besteht.

Banerjee, Dolado, Galbraith und Hendry (1993, S. 74ff.) konstatieren, dass bei Simulationsstudien mit $T = 100$ Realisationen die korrekte Nullhypothese in 75% aller Fälle bei einem nominellen Signifikanzniveau von 5% zurückzuweisen war. Die Ablehnungsquote nahm dabei entgegen der Erwartung mit größerem Stichprobenumfang

nicht ab, sondern weiter zu. Eine höhere Integrationsordnung der Prozesse führte zu einem nochmaligen Anstieg dieser Quote, wobei die Korrelationskoeffizienten stark gegen $r = \pm 1$ tendierten. Die Ergebnisse konnten im Rahmen dieser Diplomarbeit durch Simulationen mit 500 Wiederholungen und in 20er Schritten variierten Messzeitpunkten (bis $T = 300$) validiert werden. Dabei lag der Prozentsatz der signifikanten (0.01) Steigungsparameter zunächst bei 20 Messzeitpunkten bei 33.2 %, stieg dann aber schnell bei 80 Messzeitpunkten auf 67 % an und stabilisierte sich schließlich mit nur noch geringem Anstieg mit 76.6 % bei $T = 160$ und 83.4 % bei $T = 300$. Im Rahmen dieser Simulationen wurde zusätzlich zwischen den beiden Signifikanzniveaus von 0.05 (s) und 0.01 (ss) differenziert. Es wird deutlich, dass sich die von den Autoren gemachten Angaben auf den Prozentsatz *hochsignifikanter* (0.01) Schein-Regressionen zweier unabhängiger Zufallsprozesse beziehen müssen. Eine detaillierte Darstellung der Replikationsergebnisse liegt bei der Autorin vor.

In einem persönlichen, von Phillips (1997, S. 262) geführten Interview beschreibt Granger, dass er bei einer Präsentation dieser Ergebnisse an der *London School of Economics* auf völligen Unglauben und den Verdacht eines Programmierfehlers gestoßen sei.

Es sind deshalb bei naiver Anwendung des üblichen Regressions- und Trendbereinigungsverfahrens erhebliche Fehlschlüsse zu erwarten, weil in der realen Forschungspraxis a priori nicht bekannt ist, ob zwei Reihen strukturell miteinander verbunden sind.

Wenn die Parameter des Regressionsmodells nicht mit den Rohdaten, sondern mit den DS-Reihen, d.h. mit Zeitreihen nach Differenzenbildung, geschätzt werden, kann das Problem der Spurious Regression bei integrierten Prozessen vermieden werden. Nach Thome (1997) vermeidet die Differenzenbildung zwar das Problem der beschriebenen Scheinkausalität, gleichzeitig ermöglicht sie aber den Fehlschluss der scheinbaren Nicht-Kausalität. Differenzenbildung zerstört wie bereits beschrieben (vgl. Abschnitt 2.3, S. 25ff.) wichtige Information, da bei der Verwendung bereinigter Variablen nur das kurzfristige Verhalten sichtbar gemacht werden kann, da mögliche langfristige Gleichgewichtsbeziehungen und Anpassungsdynamiken zwischen den

integrierten Prozessen nicht mehr beobachtbar sind. Scheinregression bringt also Verzerrungseffekte mit sich, wenn in Regressionsanalysen Trendkomponenten nicht adäquat berücksichtigt werden. Das Kointegrationskonzept stellt eine innovative Modellierungsstrategie zur Vermeidung dieses Problems dar. Darüber hinaus wird im Verlauf der Arbeit deutlich, dass auch die mit dem Kointegrationskonzept unmittelbar verbundenen Fehlerkorrekturmodelle einen Ausweg aus dem Dilemma darstellen, weil sie gleichzeitig sowohl die Dynamiken kurzfristiger als auch langfristiger Anpassungsprozesse modellieren können.

Werner (2005, S. 128f.) skizziert die theoretischen Erklärungen der experimentellen Befunde von Granger und Newbold (1974) durch Phillips (1986). Phillips (1986, S. 318/319) fasst dabei zusammen:

Low values of the Durbin-Watson²⁰ statistic DW and moderate values of the coefficient of determination R^2 are therefore to be expected in spurious regressions (...).

Die Beispiele von Granger und Newbold (1974) legen nahe, dann von Scheinregression auszugehen, wenn der multiple Korrelationskoeffizient R^2 größer als der Wert der Durbin-Watson-Statistik ist. Banerjee, Dolado, Galbraith und Hendry (1993, S. 139) verweisen auf „a high value of R^2 and a low value of the Durbin-Watson statistic.“ Rinne und Specht (2002, S. 539) formulieren:

Eine Warnung vor Scheinregression ist ein hohes Bestimmtheitsmaß R^2 und ein Durbin-Watson-Koeffizient von nahezu Null, d.h. hohe positive Autokorrelation erster Ordnung der Regressionsresiduen.

Dass die Operationalisierung von Scheinregression über R^2 und die Durbin-Watson-Statistik nicht in allen Fällen anwendbar ist, wird nachfolgend in einer eigenen Simulation angedeutet.

Drobetz (2003) führt das Phänomen der Scheinregression als Grund für die Entwicklung des Konzepts der Kointegration an. Dies wird auch von der *Royal Swedish*

²⁰Zur Erläuterung dieser Größe: wie aus der univariaten Zeitreihenanalyse bekannt, können positive oder negative Autokorrelationen erster Ordnung anhand des Durbin-Watson (DW)-Tests überprüft werden (vgl. Werner, 2005, S. 75f.).

Academy of Sciences und der *Bank of Sweden* anlässlich der Verleihung des Nobelpreises an C.W.J. Granger „für Methoden zur Analyse ökonomischer Zeitreihen mit gemeinsamen veränderlichen Trends (Kointegration)“ in dieser Weise dargestellt (*Bank of Sweden & The Royal Swedish Academy of Sciences*, 2003a; 2003b, S. 4).

3.2.2 Kointegration vs. Scheinregression

In die Klärung des Zusammenhangs der beiden *Abhängigkeitskonzepte* Scheinregression und Kointegration kann das folgende Zitat allgemein einführen.

Wenn die beiden Variablen nicht kointegriert sind, entstehen die Probleme der Scheinregression, und die OLS-Schätzer produzieren nichtstationäre Residuen. (Thome, 1997, S. 214f.)

Die herkömmliche Methode der Differenzenbildung (vgl. 2.3, S. 25ff.) zur Vermeidung von Scheinregression hat einen Informationsverlust zur Folge. Wenn allerdings keine Differenzenbildung durchgeführt wird, kann Scheinregression nur dann vermieden werden, wenn die Instationarität der Reihen angemessen modelliert wird. Dies wird durch das Kointegrationskonzept gewährleistet. Kointegration könnte also als eine Modellierungs-Strategie unter Ausschluss von Scheinregressionen verstanden werden.

Maddala und Kim (1998) unterstreichen diesen Sachverhalt nochmals vor dem Hintergrund der Logik der Scheinregression. Von Scheinregression ist per definitionem nur dann auszugehen, wenn unabhängige $I(1)$ -Einzelprozesse gegeben sind - in allen anderen Fällen ist die Regression nicht scheinbar, sondern real existent. Sobald die Einzelprozesse abhängig voneinander sind und damit also Einflüsse, die wechselseitig sein können, vorliegen, kann von Kointegration ausgegangen werden. In diesem Fall ist aber wiederum Scheinregression auszuschließen und von faktisch begründeter Regression auszugehen, weil der Zusammenhang durch die Kointegration erklärbar ist. Kointegration und Scheinregression sind somit als gegenläufige Abhängigkeitskonzepte aufzufassen.

In Bezug auf die dargestellte Faustregel von Granger und Newbold (1974), dass $R^2 > DW$ Scheinregression indiziert, die auch nochmals von Banerjee, Dolado,

Tabelle 3.2: Für die folgenden VAR(1)-Prozesse ermittelte R^2 - und DW -Statistiken. $T = 20000$, 20 Wiederholungen.

Unabhängige Einzelprozesse				Rückgekoppelte Einzelprozesse				Nicht-rückgekoppelte Einzelprozesse			
<i>keine Kointegration</i>				<i>Kointegration</i>				<i>Kointegration</i>			
Modellmatrix				Modellmatrix				Modellmatrix			
ϕ_{11}	ϕ_{12}	R^2	DW	ϕ_{11}	ϕ_{12}	R^2	DW	ϕ_{11}	ϕ_{12}	R^2	DW
ϕ_{21}	ϕ_{22}			ϕ_{21}	ϕ_{22}			ϕ_{21}	ϕ_{22}		
1	0	0,35359	0,00045	0,95	0,05	0,997755	0,29895	0	0,1	0,97308	1,99955
0	1			0,1	0,9			0	1		
				0,2	0,8	0,99836	3,4005	0	0,5	0,998605	2,0022
				0,9	0,1			0	1		
								0	0,95	0,999405	2,0041
								0	1		
								0	-0,95	0,99941	1,99855
								0	1		
								0	1	0,99945	2,00415
								0	1		

Galbraith und Hendry (1993, S. 81) abgeleitet wurde²¹, kann man vermuten, dass sich im Fall von Kointegration ein mittlerer DW-Wert von 2 ergeben müsste.

Eigene Simulationen ergeben, dass sich diese Faustregel auf den *direkten* Vergleich der beiden unterschiedlich standardisierten Statistiken bezieht. Unter Verwendung des nicht-rückgekoppelten Kointegrationsmodells von Hamilton (1994, S. 571; vgl. auch bei Stier, 2001, S. 316) zeigt eine im Rahmen der Vorbereitung der Diplomarbeit von mir durchgeführte Simulation für $T = 20000$ und 20 Wiederholungen bei vorgegebenen unabhängigen $I(1)$ -Prozessen ein mittleres R^2 von .35 und ein mittleres DW von .0004. Diese Werte verschieben sich bei abhängigen $I(1)$ -Prozessen im Mittel zu .99 und 2.00. Bei rückgekoppelten Modellen lässt sich diese Beziehung hingegen weniger deutlich beobachten. In Tabelle 3.2, S. 48, sind die beschriebenen Ergebnisse dargestellt.

²¹S. 81: „These authors suggested treating any regression for which $R^2 > DW$ as one that is likely to be spurious. This could be interpreted as a sign of lack of any equilibrium relationship among the variables in the regression, which in turn implies a non-stationary error-term and so very strong autocorrelation in the regression residuals.“

Insgesamt scheint folglich - v.a. bezogen auf nicht-rückgekoppelte Modelle - die Aussage von Hassler (2003, S. 3) bezüglich der Gegenläufigkeit von Scheinregression und Kointegration bestätigt:

Denn genau dann, wenn Kointegration vorliegt, besteht bei einer Regression von Niveaus integrierter Variablen nicht die Gefahr einer Scheinregression.

3.3 Fehlerkorrekturmodelle

Kointegration impliziert die Möglichkeit, die strukturelle Beziehung zwischen den beteiligten Variablen als Fehlerkorrekturmodelle (Error Correction Models ((V)ECM)) darstellen zu können. Nach Stier (2001, S. 318) sind Fehler-Korrektur-Modelle für Anwendungen wohl die wichtigste *Darstellungsform kointegrierter Prozesse*. Wie oben beschrieben, stellen Kointegrationsmodelle und auch ihre Fehlerkorrekturrepräsentation eine Strategie dar, mithilfe derer sich die Instationarität von Zeitreihen positiv handhaben lässt. Mithilfe dieser Modelle ist es möglich, simultan eine eventuell vorhandene langfristige Gleichgewichtsbeziehung zwischen zwei oder mehr Variablen und die kurzfristige Anpassungsdynamik, mit der eine Störung des Gleichgewichts vom System bearbeitet wird, zu modellieren. Fehlerkorrekturmodelle stellen den Anpassungsmechanismus dar, dass instationäre Variablen sich aufgrund von Gleichgewichtsabweichungen aus früheren Perioden an eine gleichgewichtige Beziehung anpassen, wenn Kointegration gegeben ist:

In der folgenden VAR-Darstellung (vgl. Stier, 2001, S. 318) erscheinen die Niveaus der beiden Komponentenprozesse in der stationären Linearkombination, die die Kointegrationsbeziehung der beiden Prozesse ausdrückt:

$$\begin{pmatrix} \Delta x_{1t} \\ \Delta x_{2t} \end{pmatrix} = \begin{pmatrix} -1 & \beta \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

Die Fehler-Korrektur-Form eines kointegrierten Systems ergibt sich allgemein (spezifischere Beschreibung im nächsten Abschnitt 3.4) aus dieser VAR-Darstellung, indem

die neue Variable $z_t := x_{1t} - \beta x_{2t}$ eingeführt wird und somit gilt:

$$\begin{pmatrix} \Delta x_{1t} \\ \Delta x_{2t} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} z_{t-1} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} \quad (3.1)$$

Das Fehlerkorrekturmodell modelliert also die stationäre Linearkombination der kointegrierten Einzelprozesse und nicht direkt die $I(1)$ -Prozesse selbst. Durch diese Eigenschaft der Stationarität (vgl. 2.1. S. 21ff.) in der Fehlerkorrekturform ist eine Prognose möglich.

Dabei repräsentiert die Variable z_t die (kurzfristigen) Abweichungen der kointegrierten Prozesse x_{1t} und x_{2t} von ihrem langfristigen Gleichgewichtszustand $x_{1t} = \beta x_{2t}$. z_{t-1} erhält beispielsweise dann ein negatives Vorzeichen, wenn $\beta x_{2,t-1} > x_{1,t-1}$ ist. In diesem Fall korrigiert das kointegrierte System diese Abweichung vom Gleichgewichtspfad oder den „Fehler“ durch eine positive Veränderung von x_{1t} im Zeitpunkt t , d.h. dann ist $\Delta x_{1t} > 0$. Wenn dagegen z_{t-1} positiv ist, dann ist im nächsten Zeitabschnitt eine negative Korrektur von x_{1t} zu beobachten, sodass $\Delta x_{1t} < 0$. Wenn sich das System im Zeitpunkt $t - 1$ im Gleichgewicht befindet, dann gilt: $z_{t-1} = 0$ und x_{1t} wird nicht weiter modifiziert, sodass gilt: $\Delta x_{1t} = 0$. Wie deutlich wird, hat dieser Korrektur- bzw. Anpassungsmechanismus zur Benennung des kointegrierten Systems als „Fehler-Korrektur-Repräsentation“ geführt. Engle und Granger (1987) beschreiben als Grundidee von Fehlerkorrekturmodellen die Tatsache, dass ein Teil des Ungleichgewichts („disequilibrium“, S. 254) aus einer Periode in der nächsten Periode korrigiert werde. In einem Zwei-Variablen-System würde beispielsweise ein Fehlerkorrekturmodell die Veränderung in einer Variablen sowohl mit Störungen des Gleichgewichts in der Vergangenheit als auch mit Veränderungen in beiden Variablen in der Vergangenheit verbinden. Somit modelliert das Fehlerkorrekturmodell jede Art der Anpassung an ein neues Gleichgewicht.

In Abgrenzung von der Kointegrationsgleichung macht Thome (1997, S. 216) deutlich, dass diese sich auf Gleichgewichtsniveaus beziehe und somit als statistische Regressionsgleichung keine Angaben über den zeitlichen Verlauf der Anpassungsreaktionen, mit denen ein temporär entstandenes Ungleichgewicht reduziert werde, mache. Sie mache auch keine Aussage darüber, ob die „Reequilibrierungsprozesse“,

d.h. die Rückkehrbewegung in Richtung des Gleichgewichtsniveaus, durch kurzfristige Effekte überlagert würden. Diese dynamische Spezifikation erfolge mithilfe der Fehlerkorrekturmodelle.

Granger unterstreicht in seiner „Nobel Lecture“ (Bank of Sweden / The Royal Swedish Academy of Sciences, 2003a) die praktische Bedeutung von Fehlerkorrekturmodellen:

The error-correction model has been particularly important in making the idea of cointegration practically useful.

3.4 Engle-Granger-Repräsentationstheorem

Der Zusammenhang zwischen Fehlerkorrekturmodellen und Kointegration wurde zuerst durch Granger (1981) hervorgehoben (vgl. Engle & Granger, 1987, S. 255ff.).

Das Engle-Granger-Repräsentationstheorem besagt, dass zu einem System kointegrierter $I(1)$ -Variablen ein Fehlerkorrekturmodell existiert, das die Kurzfristedynamik beschreibt und dessen Mechanismus die Abweichungen zwischen den Variablen begrenzt, **und dass es umgekehrt** zu gegebenem Fehlerkorrekturmodell für $I(1)$ -Variablen eine diese verbindende Kointegrationsbeziehung gibt. (Rinne & Specht, 2002, S. 546; Hervorhebungen nicht im Original)

Engle und Granger (1987) beschreiben eine allgemeine Herleitung für die Umformung von m -dimensionalen $\text{VAR}(p)$ -Modellen in $\text{VECM}(p)$ -Modelle, die hier dargestellt werden soll, da ihre anschließende Vereinfachung für den bivariaten Fall im empirischen Teil Anwendung finden wird.

Werner (2005, S. 137f.) hat die einzelnen Schritte der Herleitung elaboriert und aus Gründen der Anschaulichkeit soll seine Darstellung übernommen werden. Im SAS-VARMAX-Manual (SAS/ETS Software: Changes and Enhancements, 2001, S.113f.) wird diese Herleitung für die autoregressive Repräsentation in ähnlicher Weise, aber nicht in dieser Vollständigkeit dargestellt.

Ausgehend vom bekannten VAR(p)-Modell aus Gleichung 1.2, S. 17,

$$\Phi(B)\mathbf{x}_t = \boldsymbol{\epsilon}_t \quad (3.2)$$

mit den fixen anfänglichen Werten $\mathbf{x}_{-p+1}, \dots, \mathbf{x}_0$ und $\boldsymbol{\epsilon}_t \sim N(0, \Sigma)$ kann der AR-Operator $\Phi(B)$ folgendermaßen umgeschrieben werden:

$$\Phi(B) = \Phi^*(B)(1 - B) + \Phi(1)B \quad (3.3)$$

Darin sind

$$\Phi^*(B) = \mathbf{I}_m - \sum_{i=1}^{p-1} \Phi_i^* B^i \quad \text{und} \quad \Phi_i^* = - \sum_{j=i+1}^p \Phi_j \quad (3.4)$$

Nach dem SAS-VARMAX-Manual (SAS/ETS Software: Changes and Enhancements, 2001, S.113) und Werner (2005, S. 137) gilt in der Fehlerkorrekturgleichung außerdem:

$$-\Phi(1) = -\mathbf{I}_m + \Phi_1 + \Phi_2 + \dots + \Phi_p = \boldsymbol{\alpha}\boldsymbol{\beta}' = \boldsymbol{\Pi} \quad (3.5)$$

Diese Gleichung ergibt sich per definitionem aus dem aus dem Grundlagenabschnitt 1.3, S. 17, bekannten Matrix-Lag-Polynom, das $\Phi(B)$ definiert:

$$\Phi(B) := \mathbf{I}_m - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p$$

Darin beschreibt \mathbf{I}_m eine ($m \times m$)-Einheitsmatrix.

Für $B = 1$ kann umgeschrieben werden:

$$-\Phi(1) = -\mathbf{I}_m + \Phi_1 + \Phi_2 + \dots + \Phi_p$$

Engle und Granger (1987) weisen durch eine Herleitung auf S. 255/256 ihres Artikels nach, dass außerdem gilt:

$$-\Phi(1) = \boldsymbol{\alpha}\boldsymbol{\beta}'$$

Dieser Nachweis ist in der im Anhang E (S. 172) zusätzlich von mir durchgeführten Herleitung von Gl. 3.2 zu Gl. 3.3, S. 52, in Anlehnung an Engle und Granger (1987) enthalten.

Der Anpassungskoeffizient $\boldsymbol{\alpha}$ wird erstmals in der Fehlerkorrektur-Darstellung eingeführt, $\boldsymbol{\beta}$ ist der bereits beschriebene Kointegrationsvektor. Auf die Funktion des

neu eingeführten Parameters α wird im folgenden Abschnitt 3.4.1, S. 54, eingegangen. In Abschnitt 3.4.2, S. 59, wird die Rückrechnung aus dem Fehlerkorrekturmodell auf $\alpha\beta'$ für nur eine Kointegrationsbeziehung vereinfacht. Durch diese Vereinfachung können die Populationswerte von α und β aus der Φ_1 -Matrix errechnet werden, was zur Beurteilung der Schätzgüte durch den Vergleich von Schätz- und Populationswerten im empirischen Teil notwendig ist.

Das Vektor-Fehlerkorrekturmodell (VECM) ergibt sich unter Verwendung der Gleichung 3.5 und durch Einsetzen der Gleichung 3.3 in Gleichung 3.2:

$$\begin{aligned} [\Phi^*(B)(1-B) + \Phi(1)B]\mathbf{x}_t &= \epsilon_t \\ [\Phi^*(B)(1-B)]\mathbf{x}_t &= -\Phi(1)B\mathbf{x}_t + \epsilon_t \quad | \text{vgl. Gl. 3.5} \\ &= \alpha\beta'\mathbf{x}_{t-1} + \epsilon_t \quad | \text{vgl. Gl. 3.4} \end{aligned}$$

Das VECM lautet folglich:

$$[\Phi^*(B)(1-B)]\mathbf{x}_t = \alpha\beta'\mathbf{x}_{t-1} + \epsilon_t \quad (3.6)$$

Durch die beiden folgenden Herleitungsschritte kann das Fehlerkorrekturmodell auch wie in Gleichung 3.7 geschrieben werden:

Aus Gl. 3.6 folgt:

$$[(\mathbf{I}_m - \sum_{i=1}^{p-1} \Phi_i^* B^i)(1-B)]\mathbf{x}_t = \alpha\beta'\mathbf{x}_{t-1} + \epsilon_t$$

Es lässt sich umformen:

$$\mathbf{I}_m(1-B)\mathbf{x}_t - \sum_{i=1}^{p-1} \Phi_i^* B^i(1-B)\mathbf{x}_t = \alpha\beta'\mathbf{x}_{t-1} + \epsilon_t$$

Nach Ausmultiplikation auf der linken Seite der Gleichung und Neuordnung der Elemente ergibt sich eine andere Darstellung des Fehlerkorrekturmodells:

$$\Delta\mathbf{x}_t = \alpha\beta'\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \Phi_i^* \Delta\mathbf{x}_{t-i} + \epsilon_t \quad (3.7)$$

Diese Gleichung wird in vielen Fällen um einen deterministischen Ausdruck ϑ erweitert, der aus einer Konstanten, einem linearen Trend, saisonalen Kodiervariablen oder nichtstochastischen Regressoren bestehen kann. In dieser Gleichung wird $\alpha\beta'$ durch Π repräsentiert (vgl. SAS/ETS Software: Changes and Enhancements, 2001,

S.114).

$$\Delta \mathbf{x}_t = \mathbf{\Pi} \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \mathbf{\Phi}_i^* \Delta \mathbf{x}_{t-i} + \boldsymbol{\vartheta} + \boldsymbol{\epsilon}_t \quad (3.8)$$

Allgemein lässt sich für die Rücktransformation eines VECM(p)-Modells in ein VAR(p)-Modell schreiben:

$$\mathbf{x}_t = \boldsymbol{\vartheta} + (\mathbf{I}_m + \mathbf{\Pi} + \mathbf{\Phi}_1^*) \mathbf{x}_{t-1} + \sum_{i=2}^{p-1} (\mathbf{\Phi}_i^* - \mathbf{\Phi}_{i-1}^*) \mathbf{x}_{t-1} - \mathbf{\Phi}_{p-1}^* \mathbf{x}_{t-p} + \boldsymbol{\epsilon}_t \quad (3.9)$$

Wie beim Dickey-Fuller-Test (vgl. Abschnitt 2.5.2, S. 29f.) die Berücksichtigung einer Konstanten und/oder eines Trends die Signifikanzschranken ändert, so sind beim Ableiten eines VECM(p)- aus einem VAR(p)-Modell die Konstanten bzw. Trends ebenso besonders zu beachten. Deterministische kointegrierte Beziehungen aus dem VAR(p)-Modell treten in der VECM(p)-Repräsentation nicht mehr auf. Stochastische Kointegrationsbeziehungen aus der VAR(p)-Form dagegen sind im VECM(p)-Modell in Form von deterministischen Zusammenhängen als Fehlerkorrekturwert oder unabhängiger Ausdruck repräsentiert (vgl. Werner, 2005, S. 138).

3.4.1 Interpretation der Funktion des Anpassungskoeffizienten α und des Kointegrationsvektors β

Funktion des Anpassungskoeffizienten α Im Fehlerkorrekturmodell wird erstmals α eingeführt, das insofern eine wichtige Rolle spielt, als es als sogenannter Anpassungskoeffizient angibt, wie schnell der Prozess sein Gleichgewicht wiederherstellt, während der bereits eingeführte Koeffizient β den langfristigen Gleichgewichtszustand („long-run parameters“, vgl. SAS/ETS Software: Changes and Enhancements, 2001, S.113f., S. 114) beschreibt. Im empirischen Teil der Arbeit wird das VECM unter anderem zur Schätzung dieser beiden Parameter dienen.

Die Funktion des Anpassungskoeffizienten wird in der Literatur kaum beschrieben und interpretiert. Rinne und Specht (2002, S. 545f.) machen hier eine Ausnahme:

Es drückt $|\alpha_i|$ die Anpassungsgeschwindigkeit aus, und α_i heißt Anpassungskoeffizient. (...)

Natürlich muss mindestens einer der Anpassungskoeffizienten von Null

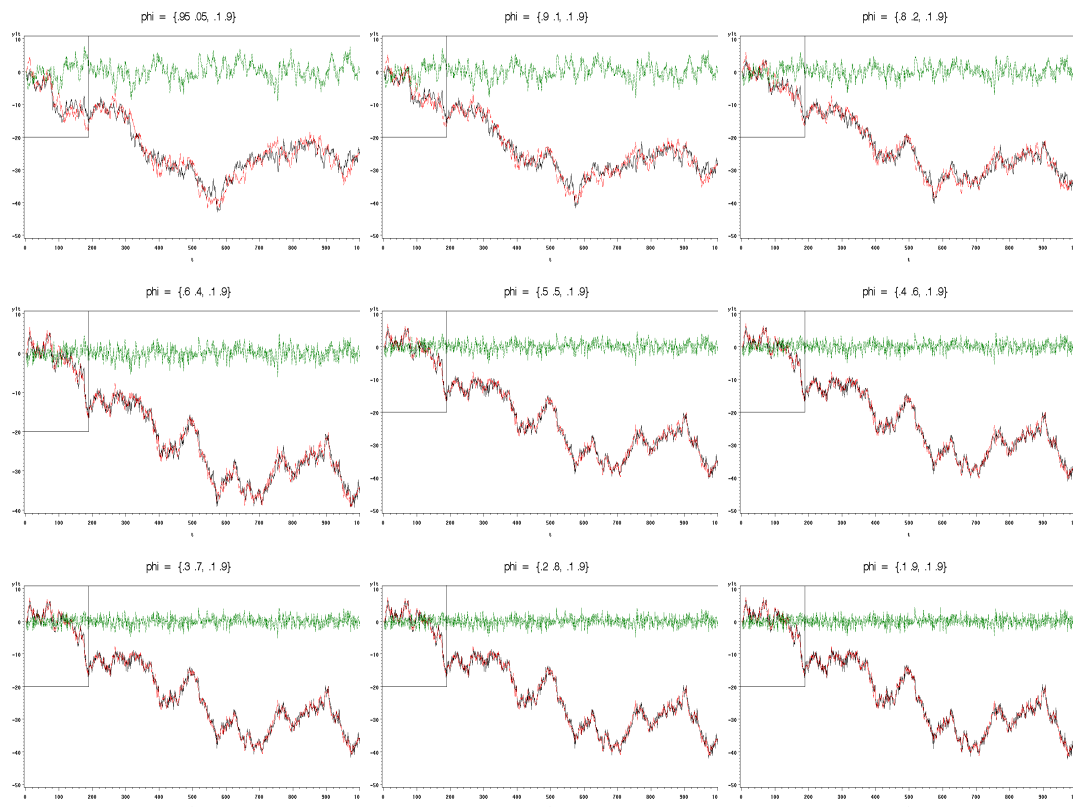


Abbildung 3.2: Graphische Veranschaulichung der Funktion des Anpassungskoeffizienten. Modelleigenschaften in Tabelle 3.3, S. 56. $|\alpha_1|$ von links oben nach rechts unten größer werdend.

verschieden sein, damit nach einer Gleichgewichtsstörung die Rückkehr zum Gleichgewicht möglich ist. (Rinne & Specht, 2002, S. 545f.)

Eigene Simulationen verdeutlichen die Funktion des Anpassungskoeffizienten durch unterschiedliche Variationen von α bei Konstanthaltung von β durch Generierung zugehöriger Φ_1 -Matrizen. Die Graphiken in Abbildung 3.2, S. 55, die die in Tabelle 3.3, S. 56, aufgeführten Prozesse illustrieren, zeigen, wie direkt $|\alpha_1|$ oder $|\alpha_2|$ die Anpassungsgeschwindigkeit determinieren²². Die Abbildung enthält eine Abfolge von Graphiken zur Veranschaulichung der sich verändernden Anpassungsgeschwindigkeit durch Variation von $|\alpha_i|$ (vgl. obiges Zitat von Rinne & Specht, 2002, S. 545f.). Um die Aufmerksamkeit der Leserin bzw. des Lesers auf den Bereich in den Graphiken zu lenken, in dem durch die Variation des Anpassungskoeffizienten eine Veränderung

²²Bei der Interpretation der Funktion des Anpassungskoeffizienten spielt das Vorzeichen von α_1 bzw. α_2 keine Rolle, weshalb der Betrag ($|\alpha_1|$; $|\alpha_2|$) zu schreiben ist.

Tabelle 3.3: Funktion des Anpassungskoeffizienten: VAR(1)-Prozesse; Modellmatrix Φ_1 , zugehörige Eigenwerte (EW) und Populationsparameter α_1 und α_2 . Dickey-Fuller-Tests (DF) der VARMAX-Prozedur über Komponentenprozesse (black, red) und Linearkombination (green) für T=1000. Dabei $\beta_1 = 1$ und $\beta_2 = -1$. Veranschaulichung in Abbildung 3.2, S. 55.

VAR(1)-Prozesse								
Modellmatrix		EW	α_1	α_2	DF	DF	DF	
ϕ_{11}	ϕ_{12}				$Pr < Tau$	$Pr < Tau$	$Pr < Tau$	
ϕ_{21}	ϕ_{22}			(black)	(red)	(green)		
0,95	0,05	1	-0,05	0,1	0,5993	0,6123	<,0001	
0,1	0,9	0,85						
0,9	0,1	1	-0,1	0,1	0,6461	0,6645	<,0001	
0,1	0,9	0,8						
0,8	0,2	1	-0,2	0,1	0,692	0,7202	<,0001	
0,1	0,9	0,7						
0,6	0,4	1	-0,4	0,1	0,7201	0,7665	<,0001	
0,1	0,9	0,5						
0,5	0,5	1	-0,5	0,1	0,7232	0,7785	<,0001	
0,1	0,9	0,4						
0,4	0,6	1	-0,6	0,1	0,7234	0,7873	<,0001	
0,1	0,9	0,3						
0,3	0,7	1	-0,7	0,1	0,722	0,7939	<,0001	
0,1	0,9	0,2						
0,2	0,8	1	-0,8	0,1	0,7199	0,7992	<,0001	
0,1	0,9	0,1						
0,1	0,9	1	-0,9	0,1	0,7173	0,8035	<,0001	
0,1	0,9	0						

der Geschwindigkeit der Anpassung beobachtet werden kann, wurde dieser Bereich jeweils links oben durch ein Rechteck markiert. Dies ist bei dieser Abbildung vor allem deshalb notwendig, weil die Funktion des Anpassungskoeffizienten graphisch nicht so überdeutlich sichtbar wird, wie bei der Veranschaulichung der Funktion des Kointegrationsvektors im nachfolgenden Abschnitt.

Um den Einfluss der Größe des Anpassungskoeffizienten auf das Verhalten der Komponentenprozesse (schwarz und rot) und deren Linearkombination (grün) in einem VAR(1)-Prozess illustrieren zu können, ist es notwendig, dass

eines der beiden Elemente des Anpassungskoeffizienten systematisch in seiner numerischen Größe variiert wird. Dies ist nur möglich, wenn das andere Element konstant gehalten wird. Deshalb wird α_1 so variiert, dass sein Betrag systematisch zunimmt, während α_2 konstant bleibt. Im jeweils links oben markierten Rechteck in Abbildung 3.2, S. 55, lässt sich über die Variationen von $|\alpha_1|$ hinweg eine Annäherung an die stationäre Linearkombination beobachten: je größer $|\alpha_1|$, desto stärker die Anpassung der beiden Komponentenprozesse an das langfristige Gleichgewicht. Diese Entwicklung ist vor allem beim Vergleich der ersten und der letzten Graphik in Abbildung 3.2 zu sehen. Die Größe des Anpassungskoeffizienten determiniert folglich zumin-

dest für den Fall nicht-auseinanderdriftender Zeitpfade der Komponentenprozesse den graphisch beobachtbaren Abstand zwischen den beiden Komponentenprozessen und ihrer Linearkombination.

Funktion des Kointegrationsvektors β Die Funktion des Kointegrationsvektors wird in der Literatur kaum interpretiert. Dieser Abschnitt soll die Rolle von β für die langfristige Gleichgewichtsbeziehung kointegrierter Prozesse anhand von Graphiken veranschaulichen.

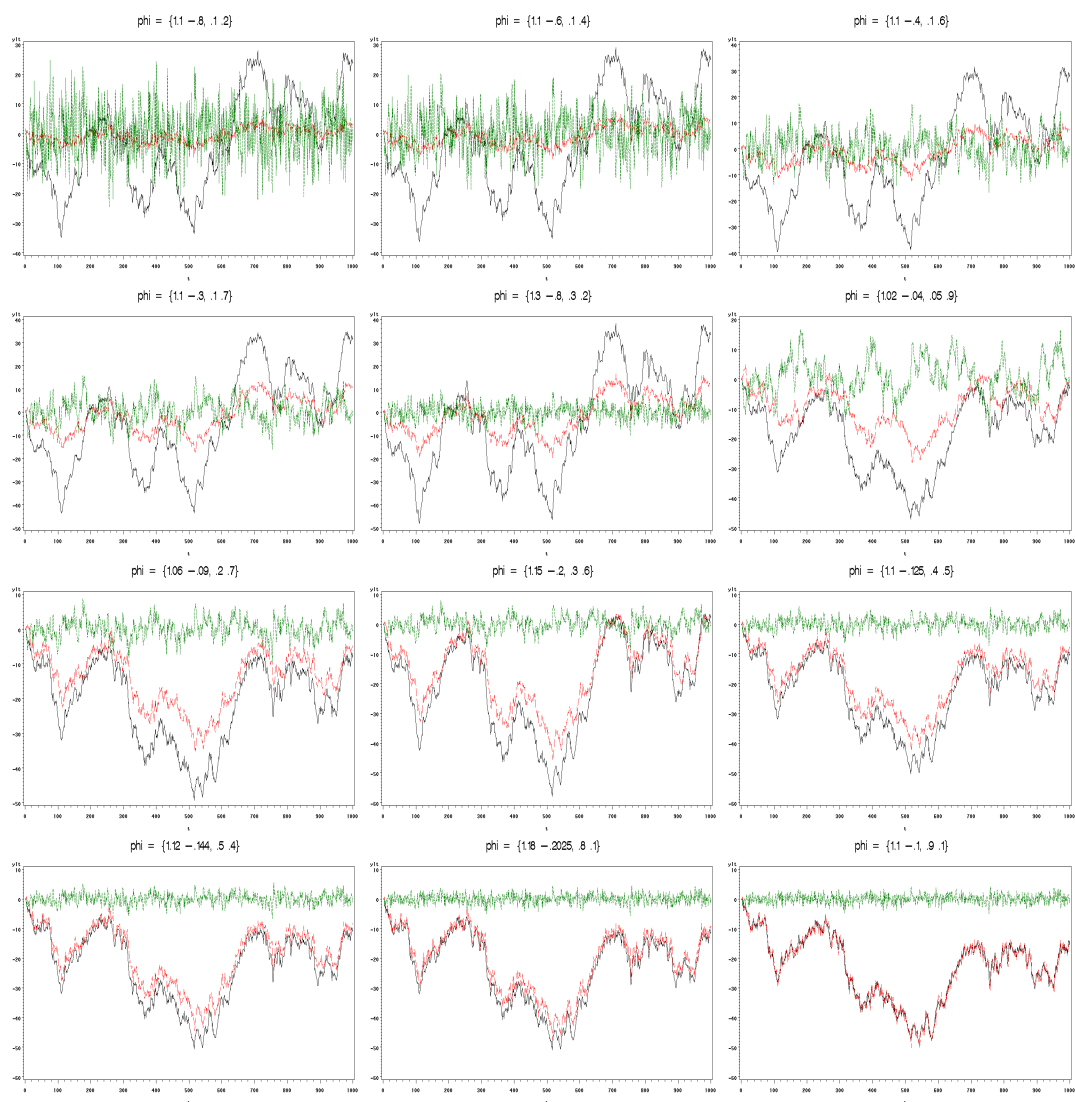


Abbildung 3.3: Graphische Veranschaulichung der Funktion des Kointegrationsvektors. Modelleigenschaften in Tabelle 3.4, S. 58. Zunehmende Annäherung von $\beta' = (1 \quad -8)$ an $\beta' = (1 \quad -1)$ von links oben nach rechts unten.

Tabelle 3.4: Funktion des Kointegrationsvektors: VAR(1)-Prozesse; Modellmatrix Φ_1 , zugehörige Eigenwerte (EW) und Populationsparameter β_1 und β_2 . Dickey-Fuller-Tests (DF) der VARMAX-Prozedur über Komponentenprozesse (black, red) und Linearkombination (green) für T=1000. Veranschaulichung in Abbildung 3.3, S. 57.

VAR(1)-Prozesse									
$\beta_2 < 0$									
Modellmatrix		EW	β_1	β_2	DF	DF	DF		
ϕ_{11}	ϕ_{12}				$Pr < Tau$	$Pr < Tau$	$Pr < Tau$	(black)	(red)
ϕ_{21}	ϕ_{22}								
1,1	-0,8	0,3	1	-8	0,1816	<,0001	<,0001		
0,1	0,2	1							
1,1	-0,6	0,5	1	-6	0,2143	<,0001	<,0001		
0,1	0,4	1							
1,1	-0,4	0,7	1	-4	0,3055	0,0026	<,0001		
0,1	0,6	1							
1,1	-0,3	0,8	1	-3	0,4044	0,0429	<,0001		
0,1	0,7	1							
1,3	-0,8	0,5	1	-2,6	0,1962	0,05	<,0001		
0,3	0,2	1							
1,02	-0,04	1	1	-2	0,3857	0,182	<,0001		
0,05	0,9	0,92							
1,06	-0,09	1	1	-1,5	0,4092	0,2839	<,0001		
0,2	0,7	0,76							
1,15	-0,2	1	1	-1,3	0,3437	0,2789	<,0001		
0,3	0,6	0,75							
1,1	-0,125	1	1	-1,25	0,4173	0,3356	<,0001		
0,4	0,5	0,6							
1,12	-0,144	1	1	-1,2	0,4141	0,3397	<,0001		
0,5	0,4	0,52							
1,18	-0,2025	1	1	-1,125	0,3987	0,3381	<,0001		
0,8	0,1	0,28							
1,1	-0,1	1	1	-1	0,4633	0,4065	<,0001		
0,9	0,1	0,2							

Der Kointegrationsvektor determiniert allgemein die Nähe der Einzelreihen (schwarz und rot) zueinander und damit auch, ob die Zeitpfade der Komponentenprozesse auseinanderdriften oder nicht. Je größer nämlich die numerische Annäherung an $\beta' = (1 \quad -1)$, desto weniger driften die Zeitpfade auseinander, wie anhand der zu den in Tabelle 3.4, S. 58, aufgeführten VAR(1)-Prozessen gehörigen Graphiken demonstriert werden kann, die in Abbildung 3.3, S. 57, dargestellt sind.

β_1 ist aufgrund der Normierung 1, $\beta_2 < 0$ wurde systematisch variiert.

Wenn man das Verhalten der Einzelreihen zueinander im Verlauf der Abbildung 3.3, S. 57, in seiner Entwicklung bei zunehmender Annäherung von $\beta' = (1 \quad -8)$ an $\beta' = (1 \quad -1)$ von Graphik zu Graphik verfolgt, kann man sehen, dass sich die Komponentenprozesse im Verlauf des Schaubilds allmählich annähern.

$\beta' = (1 \quad -1)$ steht gemäß Rinne und Specht (2002, S. 542; vgl. Zitat auf S. 41) für den Fall nicht-auseinanderdriftender Zeitpfade der Einzelreihen. Je mehr β_2 von -1 abweicht, desto mehr driften die Zeitpfade der Komponentenprozesse auseinander.

Dies gilt auch für Abweichungen in den positiven Bereich. Wenn $\beta_2 > 0$, erfüllt der Kointegrationsvektor dieselben Funktionen, die Zeitpfade der Einzelreihen verhalten sich aber graphisch wie „gespiegelt“, wie eigene Simulationen zeigen konnten. Ebenso wurden Simulationen zur Funktion des Kointegrationsvektors für nicht-rückgekoppelte Modelle durchgeführt und graphisch veranschaulicht. Die Funktion des Kointegrationsvektors unterscheidet sich für diese Fälle aber nicht von den rückgekoppelten Modellen. Die Ergebnisdarstellung dieser Simulationen soll deshalb aus Gründen des Umfangs nicht in diese Diplomarbeit aufgenommen werden.

Im empirischen Teil der Diplomarbeit wird sich zeigen, dass die Unterscheidungen zwischen VAR(1)-Prozessen mit auseinanderdriftenden und nicht-auseinanderdriftenden Zeitpfaden das Verhalten von Test- und Schätzverfahren deutlich beeinflusst. Deshalb ist es für die Forschungspraxis von Interesse, dass der Kointegrationsvektor bezüglich der Frage, ob die Einzelreihen auseinanderdriften oder nicht, eine Indikatorfunktion einnimmt.

3.4.2 Schätzung von Anpassungskoeffizient α und Kointegrationsvektor β mithilfe des für *eine* Kointegrationsbeziehung vereinfachten Fehlerkorrekturmodells

Das VECM soll im empirischen Teil zur Beurteilung der Schätzgüte von $\alpha\beta'$ durch den Vergleich von Schätz- und Populationswerten genutzt werden, und aus diesem Grund ist es sinnvoll, die Rückrechnung aus dem VECM auf die inhaltlich so bedeutenden Parameter $\alpha\beta'$ für den bivariaten VAR(1)-Fall mit folglich nur einer Kointegrationsbeziehung ($m = 2$; $h = 1$) zu vereinfachen.

Der Teil der Fehlerkorrekturmodell-Gleichung (vgl. Gleichung 3.6, S. 53, bzw. Gleichung 3.7, S. 53), der $\alpha\beta'$ spezifiziert (vgl. SAS/ETS Software: Changes and Enhancements, 2001, S.113f.; Werner, 2005, S. 137), wurde bereits durch Gleichung 3.5, S.

52, beschrieben, soll aber hier zum besseren Überblick nochmals aufgeführt werden:

$$-\Phi(1) = -\mathbf{I}_m + \Phi_1 + \Phi_2 + \dots + \Phi_p = \alpha\beta'$$

Es kann für den Fall $p = 1$ folgendermaßen vereinfacht werden:

$$-\mathbf{I}_{m=2} + \Phi_{p=1} = \alpha\beta'$$

Folglich kann beim Vorliegen von nur einer Kointegrationsbeziehung mithilfe einer vereinfachten Gleichung vom VECM auf $\alpha\beta'$ rückgerechnet werden.

$$\alpha\beta' = \Phi_1 - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (3.10)$$

Die eben dargestellte Vereinfachung hat den Vorteil, dass von vorgegebenen Φ_1 -Matrizen auf die Populationswerte von α und β rückgerechnet werden kann. Das ist in dieser Arbeit deshalb sinnvoll, weil erstens dadurch gewährleistet ist, dass zur Veranschaulichung Matrizen herangezogen werden, deren Elemente nur wenige Dezimalstellen aufweisen, und zweitens - und das ist entscheidend - die Matrizen mit einer Einheitswurzel generiert werden. Dieses Vorgehen erlaubt also die Operationalisierung von Kointegration über theoretische Populationsparameter unabhängig von Testverfahren, von denen im folgenden Abschnitt die einschlägigsten dargestellt werden sollen. Die Matrizen und mit ihnen die Populationswerte von α und β werden im Rahmen der *Monte-Carlo-Simulationen* (vgl. Abschnitt 4.2.1, S. 81ff.) als wahre Parameter bei der Modellbildung vorgegeben und auf Basis der generierten Daten können z.B. Schätzungen durchgeführt werden, sodass durch den Vergleich von Parameter und Schätzwert die Güte der Schätzung beurteilt werden kann.

3.5 Testverfahren

Aus der Liste der Kointegrationstests von Cromwell, Hannan, Labys und Terraza (1994, S. 7) sollen an dieser Stelle die für den empirischen Teil der Arbeit relevanten Verfahren dargestellt werden.

Kointegrationstests sind als Tests auf Einheitswurzeln in einer geschätzten Zeitreihe zu verstehen. Es existieren verschiedene Tests, die bestimmen sollen, ob ein Vektor

integrierter Zeitreihen kointegriert ist. Für alle Testverfahren ist es zunächst notwendig, dass jede univariate Zeitreihe $I(1)$ ist. Somit können die klassischen Einheitswurzeltests (2.5.2, S. 29f.) Anwendung finden. Wenn diese Voraussetzung gegeben ist, können die multivariaten Tests durchgeführt werden.

Wie im empirischen Teil noch ausführlicher beschrieben, werden im Rahmen dieser Arbeit die Komponentenprozesse also vorab mithilfe von Dickey-Fuller-Tests auf ihre Instationarität getestet, um sicherzustellen, dass gemäß der dargestellten Definition Kointegration als stationäre Linearkombination von instationären Ausgangsprozessen vorliegt.

3.5.1 Dickey-Fuller-Kointegrationstest

Wenn man den Kointegrationsvektor β als Koeffizienten der Kointegrationsbeziehung $x_{1t} - \beta x_{2t}$ kennt und nur eine Kointegrationsbeziehung existiert und somit eine der beteiligten Variablen als abhängige Variable angesehen werden kann, dann reduziert sich der Test auf Kointegration auf einen Unit-Root-Test (vgl. unter 2.5.2, S. 29f.) für die bekannte Linearkombination und der Kointegrationsvektor kann mittels OLS-Regression geschätzt werden (Thome, 1997, S. 219). Andernfalls sind wie in Abschnitt 3.5.2, S. 62ff., beschrieben Maximum-Likelihood-Methoden anzuwenden. Ist die H_0 der Nichtstationarität der Linearkombination von $I(1)$ -Zeitreihen verworfen, liegt Kointegration vor. Wenn der Kointegrationsvektor β dagegen nicht bekannt ist, so schlagen Engle und Granger (1987) ein zweistufiges Vorgehen vor, das von Drobetz (2003, S. 16) als „Engle-Granger-Test auf Kointegration“ bezeichnet wird.

Hassler (2001, S. 21) nennt diese Vorgehensweise Dickey-Fuller-Kointegrationstest. Die Bezeichnung wird in dieser Arbeit übernommen, da sie den Kern des Verfahrens treffend beschreibt.

Dabei wird wie folgt vorgegangen:

1. *Kointegrierende Regression*, d.h. Schätzung der Kointegrationsbeziehung mit-

hilfe der OLS-Schätzung der entsprechenden linearen Regressionsgleichung:

$$x_{1t} = \hat{\beta}x_{2t} + \hat{\epsilon}_t$$

Drobetz (2003) macht deutlich, dass - wie in vielen anderen Quellen ebenso betont (Engle & Granger, 1987; Phillips & Durlauf, 1986; Stock, 1987 u.a.) - die OLS-Schätzung bei Vorliegen von Kointegration trotz Instationarität der Einzelreihen konsistent ist. Es liegt sogar **Superkonsistenz** vor, d.h. dass die Schätzwerte nicht wie bei Regressionen mit stationären Variablen mit der Wurzel des Stichprobenumfangs gegen ihre wahren Parameterwerte konvergieren, sondern mit der Rate T^{-1} . Die Verteilungen sind nicht mehr normal und vor allem in kleinen Stichproben können starke Verzerrungen auftreten.

2. Anwendung des Dickey-Fuller t -Tests für das OLS-Residuum ϵ_t

Wird die Nullhypothese der Instationarität der Linearkombination \hat{z}_t verworfen - enthält das Residuum also keine Einheitswurzel - kann von der Existenz einer Kointegrationsbeziehung ausgegangen werden. Das Hypothesenpaar lautet folglich (Hassler, 2001, S. 22):

$$H_0 : \text{keine Kointegration} \rightarrow z_t \text{ ist } I(1)$$

$$H_1 : \text{Kointegration} \rightarrow z_t \text{ ist stationär}$$

3.5.2 Maximum-Likelihood-Analyse kointegrierter Systeme:

Das Johansen-Testverfahren

Mit dem Dickey-Fuller-Kointegrationstest (vgl. 3.5.1) kann nur eine einzige Kointegrationsbeziehung ermittelt werden. Bei m Variablen sind aber allgemein bis zu $h = m - 1$ Kointegrationsbeziehungen vorstellbar. Neben der Untersuchung der Frage, ob Kointegration vorliegt oder nicht, sollten Testverfahren auch Aufschluss darüber geben können, ob mehrere oder gegebenenfalls wieviele Kointegrationsbeziehungen vorliegen. Es gilt jedoch anzumerken, dass der Johansen-Test auch im Fall von nur einer Kointegrationsbeziehung anwendbar ist und im empirischen Teil auch für diesen Fall Anwendung finden wird.

Cromwell, Hannan, Labys und Terraza (1994, S. 29) machen deutlich, dass Kointegrationstests für den multivariaten Fall gegenüber dem bivariaten Fall zwei zusätz-

liche Probleme aufweisen. Man müsse entscheiden, wie die Kointegrationsbeziehung normalisiert werden solle, d.h. welche Variable relevant sei, um als eine Linearkombination zwischen den anderen Variablen zum Ausdruck zu kommen. Außerdem existiere mehr als eine Kointegrationsbeziehung.

Engle und Yoo (1987) umgehen das erste Problem, indem sie berichten, dass nur kleine Unterschiede in den Ergebnissen durch unterschiedliche Normalisierungen verursacht würden. Normalisierung stelle für sie nur dann ein Problem dar, wenn keine „natürliche“ Normalisierung existiere. Deswegen führen sie den Kointegrationstest für den multivariaten Fall in derselben Weise durch wie für den bivariaten Fall, mit der Ausnahme, dass die kritischen Werte für die Dickey-Fuller-Tests mithilfe von Monte-Carlo-Simulationen²³ ermittelt werden. Diese basieren auf bis zu 5 Variablen in der Kointegrationsregression.

Als Alternative zu diesem Ansatz entwickelte Johansen (1988) ein auf Maximum-Likelihood-Techniken basierendes Verfahren, das beiden Problemen Beachtung schenkt. Johansen und Juselius (1990) entwickelten das Verfahren zum Testen von Hypothesen über den Kointegrationsrang und zum Schätzen der Kointegrationsvektoren weiter. Dieser Full-Information Maximum-Likelihood-Approach (FIML) (Stier, 2001, S. 329), der auf dem aus der multivariaten Statistik bekannten Konzept der **Kanonischen Korrelation** (Hassler, 1994, S. 571ff.; Stier, 2001, S. 329) basiert, erlaubt mehrere Kointegrationsbeziehungen, ist ohne eine Koeffizientennormierung durchführbar und kann nachweisen, wieviele Kointegrationsbeziehungen vorliegen.

Hamilton (1994, S. 635ff.) beschreibt die Lösung des Optimierungsproblems durch Johansen (Johansen, 1988; Johansen, 1991) in drei Schritten, die dem FIML-Ansatz folgen.

1. *OLS-Schätzung von Hilfsgleichungen*
2. *Kanonische Korrelationen (Berechnung von Eigenwerten und Eigenvektoren)*
3. *Maximum-Likelihood-Schätzungen der Modellparameter*

²³zu finden in: Engle & Yoo (1987), Appendix Tabelle A.3

Das Maximum-Likelihood-Verfahren beruht auf der Auswertung der Eigenwerte einer Matrix und ermittelt somit die „stationärste“ (Drobetz, 2003, S. 19) Kointegrationsbeziehung. Das unterscheidet dieses Verfahren von der Dickey-Fuller-Methode, die auf einer OLS-Schätzung basierend die Kointegrationsbeziehung mit der minimalen Varianz ausmacht. Im einfachsten Fall bildet ein VAR(1)-Modell für die m $I(1)$ -Niveauvariablen den Ausgangspunkt für das Johansen-Verfahren. Der Rang einer Matrix wird durch die Anzahl von Null verschiedenen Eigenwerten bestimmt. Das Verfahren basiert somit auf einem Test für die Anzahl der von Null verschiedenen Eigenwerte. Die Zahl der Kointegrationsbeziehungen bestimmt den Rang r .

Aus dem oben dargestellten dreistufigen FIML-Ansatz lassen sich zwei Signifikanztests - der Spur-Test und der λ -max-Test - ableiten.

Im Spurkriterium wird die Differenz der Log-Likelihood-Funktion gebildet, die die beiden Hypothesen LLF_{H_1} und LLF_{H_0} enthält. Die H_1 postuliert, dass m Kointegrationsbeziehungen vorliegen, wogegen unter H_0 angenommen wird, dass $h < m$ Kointegrationsbeziehungen vorliegen.

$$LLF_{H_1} - LLF_{H_0} = -\frac{T}{2} \sum_{i=h+1}^m \log(1 - \hat{\lambda}_i)$$

Die Verdoppelung dieses Likelihood-Quotienten entspricht der Spur eines mehrdimensionalen Wiener Prozesses²⁴ und liefert die Teststatistik zur Überprüfung der Hypothese, dass höchstens h Kointegrationsvektoren mit $m - h$ Einheitswurzeln vorliegen.

Mit dem Spur-Test (Trace-Test) werden die folgenden Hypothesen getestet:

$$H_0 : r \leq h$$

gegen

$$H_1 : r > h$$

Die entsprechende Trace-Teststatistik (Tr) lautet:

$$Tr = \lambda_{trace} = -T \sum_{i=h+1}^m \log(1 - \hat{\lambda}_i), h = 0, 1, 2, \dots, m - 1$$

²⁴Erläuterungen zu Wiener Prozessen befinden sich im Anhang C.

mit T als Stichprobengröße, m als Anzahl der Variablen und $\hat{\lambda}_i$ als Eigenwerten. Die Eigenwerte werden sequentiell in absteigender Folge geordnet, d.h. $\hat{\lambda}_1 > \hat{\lambda}_2 > \dots > \hat{\lambda}_m$. Beim Vorliegen von h Kointegrationsbeziehungen ist λ_h der letzte signifikante Test, der die H_0 der $m - h$ Einheitswurzeln zurückweisen kann, denn λ_{h+1} enthält bereits eine Einheitswurzel. Die Trace-Teststatistik wird für jedes h von $h = 0$ bis $h = m - 1$ berechnet und entsprechend sinkt sie mit steigendem h . Wenn sich beide Hypothesen auf stationäre Prozesse beziehen würden, könnte man erwarten, dass die Teststatistik asymptotisch χ^2 verteilt ist. Es existieren Testwerte mit anderen Verteilungen, die bei Johansen und Juselius (1990) tabelliert sind. Stier (2001, S. 334) gibt den Hinweis, dass Johansen (1988) eine relativ einfache χ^2 -Approximation für die Verteilung des Spur-Tests entwickelt habe. Von manchen Autoren, wie z.B. Cromwell, Hannan, Labys und Terraza (1994, S. 29f.), die in ihrer Zusammenfassung des Johansen-Verfahrens von einer χ^2 -Verteilung ausgehen, wird in Bezug auf diesen Sachverhalt nicht differenziert.

Da im empirischen Teil Simulationen mithilfe von SAS durchgeführt werden, ist es notwendig zu überprüfen, wie der Johansen-Test in SAS implementiert ist. Das VARMAX-Manual (SAS/ETS Software: Changes and Enhancements, 2001, S. 53) gibt Aufschluss darüber, dass dieselbe Teststatistik wie oben beschrieben verwendet wird und auch $r = 0$ gegen $r > 0$ und $r = 1$ gegen $r > 1$ getestet werden.

Zur Veranschaulichung ist nachfolgend ein typischer SAS-Ausdruck eingefügt.

Cointegration Rank Test

H_0:	H_1:			Critical	Drift	DriftIn
Rank=r	Rank>r	Eigenvalue	Trace	Value	InECM	Process
0	0	0.7739	74.12	12.21	NOINT	Constant
1	1	0.0254	1.26	4.14		

Die Zahl der Kointegrationsbeziehungen bestimmt den Rang r . SAS/ETS Software: Changes and Enhancements (2001, S. 54) gibt zur Interpretation des Outputs dabei folgende Hilfestellung:

H_0 is the null hypothesis and H_1 is the alternative hypothesis. The first

row tests $r = 0$ against $r > 0$; the second row tests $r = 1$ against $r > 1$.
 (...) By default, the critical values at 5% significance level are used for testing. You can compare the test statistics and critical values in each row; there is one cointegrated process.

Da im empirischen Teil der Arbeit nur VAR(1)-Prozesse, d.h. multivariate Prozesse mit nur zwei Komponentenprozessen simuliert werden, wird folglich nur die erste Reihe des Ausdrucks relevant sein, da in diesen Modellen nur maximal eine Kointegrationsbeziehung gegeben sein kann.

Die Tatsache, dass im Fall der Beibehaltung der Nullhypothese von Einheitswurzeln ausgegangen und auf Instationarität geschlossen wird, legt nahe, dass dieser Test als multivariates Analogon des Dickey-Fuller-Tests (vgl. 2.5.2, S. 29f.) zu verstehen ist.

Als zweites Beispiel für einen Maximum-Likelihood-Kointegrations-Test sei auf den λ -max-Test oder den Test für den größten Eigenwert λ^{max} verwiesen. Erläuterungen zu diesem Test sind bei Stier (2001, S. 335) zu finden. Da aufgrund der offensichtlichen Gemeinsamkeiten ähnliche Ergebnisse wie beim Johansen-Test zu erwarten sind, wird dieses Verfahren im empirischen Teil der Arbeit keine Verwendung finden.

Die Verteilungen der Teststatistiken des Trace- und des λ -max-Tests sind leicht zu simulieren, da sie nur von der Dimension m des zugrundeliegenden Prozesses \mathbf{x}_t abhängen (Banerjee, Dolado, Galbraith & Hendry, 1993, S. 268).

Beide Tests überprüfen die Anzahl der in einem kointegrierten System potentiell vorhandenen Kointegrationsbeziehungen und werden deshalb als Kointegrationsrangtests (Cointegration Rank Test) bezeichnet.

Stier (2001, S. 335) macht deutlich, dass über die Effizienz der Johansen-Kointegrationstests wenig bekannt ist. Banerjee, Dolado, Galbraith und Hendry (1993, S. 278) bezeichnen dies als „lack of knowledge“.

Von den beiden Johansen-Testverfahren sind die Tests von Stock und Watson (1988) abzugrenzen, die auf gemeinsame „Trends“ („Common Trends“) testen. Die „Common Trend“-Darstellung geht vom Moving-Average-Ansatz aus. Man spricht von „common trends“, da Random Walks sich lokal wie „Trends“ verhalten. Sowohl Beobach-

tung als auch Theorie legen nahe, dass kointegrierte Reihen die gleichen stochastischen „Trends“ beinhalten. Wenn jede der m Reihen von der Ordnung 1 integriert ist und gleichzeitig aber alle gemeinsam durch $k \leq m$ stochastische „Trends“ charakterisiert werden können, hat die vektorielle Repräsentation dieser Reihen k Einheitswurzeln und $m - k$ distinkte stationäre Linearkombinationen. Die Testverfahren können alternativ als Tests auf die Anzahl der gemeinsamen „Trends“, der linear unabhängigen Kointegrationsvektoren oder der autoregressiven Einheitswurzeln des vektoriellen Prozesses betrachtet werden. Beide der vorgeschlagenen Testverfahren sind asymptotisch ähnlich.

Einzelne Simulationen haben gleiche Ergebnisse wie die Johansen-Testverfahren geliefert, weshalb im empirischen Teil aus Gründen des Umfangs der vorliegenden Arbeit auf systematische Simulationen der Stock-Watson-Tests verzichtet wird.

3.6 Weiterentwicklungen des Kointegrationskonzepts:

Ausblick

Aus Gründen der Vollständigkeit und zur Verdeutlichung des breiten Forschungsspektrums, das durch das Integrations- und Kointegrationskonzept aufgespannt wird, sei hier kurz auf einige Weiterentwicklungen verwiesen, von denen das Konzept der fraktionalen Kointegration (vgl. Abschnitt 3.6.1) wohl für die Psychologie - zumindest nach aktuellem Forschungsstand - die größte Relevanz besitzt. Anschließend wird auf die Konzepte der saisonalen und der Multi-Kointegration verwiesen (vgl. Abschnitt 3.6.2). Dabei werden die Themenbereiche nur angerissen, da sie für den empirischen Teil nicht von unmittelbarer Bedeutung sind.

3.6.1 ARFIMA-Modelle und fraktionale Kointegration

ARFIMA-Modelle, d.h. fraktional integrierte Modelle (**A**uto **R**egressive **F**raktionally **I**ntegrated **M**oving **A**verage), erregen die Aufmerksamkeit des psychologischen Forschers vor allem deshalb, weil sie in jüngster Zeit Anwendung bei vielen psychologischen Fragestellungen gefunden haben. Granger und Joyeux (1980) und Hosking

(1981) regen in Abgrenzung von den Eigenschaften herkömmlich integrierter Modelle mit einem diskreten Integrationsgrad d (vgl. Abschnitt 2.2, S. 23ff.) erstmals an, dass Werte des Integrationsgrades d , die nicht ganzzahlig seien, auch sinnvolle Modelle implizieren, und damit Zeitreihen „zwischen klassischer Stationarität und der hier bislang unterstellten Integriertheit“ (Hassler, 2003, S. 4) angesiedelt werden könnten. Diese Zeitreihen können zwar stationär sein, weisen aber als „Modelle mit langem Gedächtnis“ eine sehr starke Persistenz und eine extrem weit in die Vergangenheit zurückreichende Autokorrelation auf (vgl. Hassler, 2003, S. 4). Die herkömmlichen ARIMA-Modelle erfassen dagegen das „Short-Term“-Gedächtnis. In einer endlichen Stichprobe kann dieses Langzeitgedächtnis bzw. diese starke Persistenz durch ARMA-Modelle mit hoher Parameterzahl modelliert werden. Im Sinne des „Sparsamkeitsprinzips“ (Rinne & Specht, 2002, S. 310) stellt fraktionale Integration jedoch eine sparsamere und modelltheoretisch einfachere Lösung dar und ist deshalb vorzuziehen. Allgemein sind ARFIMA-Modelle dadurch eine Erweiterung von ARIMA-Modellen, dass der Differenzenparameter d in diesen Modellen jede beliebige reelle Zahl sein kann und nicht mehr auf den nur ganzzahligen (vgl. Maddala & Kim, 1998, S. 297) und positiven Zahlenraum beschränkt ist (vgl. Rinne & Specht, 2002, S. 309; Werner, 2005, S. 145). Fraktionale Integration ist folglich durch den fraktionalen Differenzenoperator $|d| < 1$ definiert. In der Literatur wird der Einfluss des *fraktionalen* Differenzenparameters d auf die Stationarität und das Gedächtnis eines Prozesses nicht einheitlich beschrieben. Im Rahmen der Beschäftigung mit fraktionaler Integration habe ich die unterschiedlichen Kategorisierungen von d und die von den jeweiligen Autoren resultierenden Implikationen für (In-)Stationarität aus unterschiedlichen Quellen zusammengetragen und verglichen. Eine ausführliche Darstellung dieses Vergleichs sei einer anderen Publikation vorbehalten, da er für den empirischen Teil keine unmittelbare Relevanz besitzt, dennoch kann der Hinweis gegeben werden, dass die in der Literatur formulierten Stationaritätsgrenzen für d weiter untersucht werden müssen. Grundsätzlich ist (In-)Stationarität nämlich als ein Charakteristikum einer Zeitreihe aufzufassen, das objektiv beurteilt werden kann und aus diesem Grund ist es indiziert, die in der Literatur angegebenen Sta-

tionaritätsgrenzen graphisch und durch entsprechende Testverfahren zu validieren. Werner (2005, S. 145) fasst zusammen, dass der Prozess für $d < 1$ mittelwertsneutral (*mean reverting*), aber nicht für alle $(d < 1)$ -Werte kovarianzstationär sei, d.h. dass der Prozess - und das ist besonders für Simulationen zur fraktionalen Kointegration interessant - für $d > .5$ instationär, für $d < .5$ aber stationär und invertierbar sei. Bezüglich der Gedächtnisfunktion sei weiterhin bemerkenswert, dass für $0 < d < .5$ eine langfristige (vgl. auch Maddala & Kim, 1998, S. 297), für $-.5 < d < 0$ eine mittelfristige und für $d = 0$ eine kurzfristige Abhängigkeit modellierbar sei (vgl. auch Barkoulas, Baum & Oguz, 1996, S. 5).

Relevanz fraktional (ko-)integrierter Prozesse für die Psychologie Die hohe Anwendungsrelevanz univariater fraktional integrierter Zeitreihen lässt vermuten, dass das Konzept der fraktionalen Kointegration ebenfalls praktische Implikationen für die psychologische Anwendung haben könnte, auch wenn es möglicherweise zu diesem Zweck noch nicht hinreichend erforscht sein dürfte. Um dieses mögliche Anwendungsspektrum anzudeuten, seien die in jüngster Zeit erschlossenen psychologischen Anwendungsfelder für fraktional integrierte Zeitreihen angedeutet. Farrell, Wagenmakers und Ratcliff (2004, S.2) betonen, dass aufgrund der Entdeckung von $1/f$ -Noisekomponenten in psychologischen Zeitreihen (vgl. u.a. Van Orden, Holden & Turvey, 2003; Van Orden, Moreno & Holden, 2002), die auch als *pink* oder *flicker noise* bezeichnet werden und ein besonders anwendungsrelevantes Beispiel fraktional-integrierter Prozesse darstellen, durch die Arbeitsgruppe um David Gilden (Gilden, 1997, 2001; Gilden, Thornton & Mallon, 1995; Gilden & Wilson, 1995a, 1995b), die Zeitreihenanalyse in der Psychologie in jüngster Zeit zusätzliche Impulse erhalten habe. Pink Noise ist in den zeitabhängigen Fluktuationen unterschiedlichster Phänomene zu beobachten, so z.B. in der durch das EEG aufgezeichneten Hirnaktivität oder im Bereich von Musik und Sprache. Pink Noise findet für die psychologische Anwendung besondere Beachtung, da laut Gilden (2001) und Werner (2005, S. 144f.) das, was bisher als unerklärte Varianz in psychologischen Experimenten verstanden worden sei und bis zu 90% der Gesamtvarianz ausmache, zu einem beachtlichen Teil

auf Pink Noise zurückführbar sei. Einige Studien weisen auf die Existenz von $1/f$ -Noise in grundsätzlichen kognitiven Fertigkeiten, wie in Wahlentscheidungsaufgaben und Schätzexperimenten untersucht, hin.

So beobachtet die Arbeitsgruppe um Gilden (vgl. Gilden, 2001; Gilden, Thornton, Mallon, 1995; Gilden & Wilson, 1995a, 1995b) $1/f$ -Noise in den Leistungen der Versuchspersonen bei Aufgaben zur mentalen Rotation, lexigraphischen Entscheidung, beschleunigten visuellen Suche, Signalentdeckung, zum Absoluturteil und bei Schätzungen von Distanzen, Größe und Kraft. Gilden (1997) untersuchte z.B. auch in einem Reaktionszeitexperiment die Fluktuation der für elementare Entscheidungen benötigten Zeit mithilfe dieses Ansatzes. Vor allem bei Zeitschätzungsexperimenten scheint $1/f$ -Noise eine Rolle zu spielen (Gilden, 2001; Gilden, Thornton & Mallon, 1995; Wagenmakers, Farrell & Ratcliff (in press)). Außerdem lässt sich ein $1/f$ -Noise ähnliches Rauschen in Zeitreihen aus einfachen Reaktionszeitexperimenten (vgl. Van Orden, Holden & Turvey, 2003; Wagenmakers, Farrell & Ratcliff (in press); Gilden, Thornton & Mallon, 1995) und Wortbenennungsexperimenten feststellen (Van Orden, Holden & Turvey, 2003). Auch in Studien zur Psychomotorik ist $1/f$ -Noise eine wichtige Erklärungsvariable. Genauere Erläuterungen hierzu finden sich bei Farrell, Wagenmakers und Ratcliff (2004, S.5). Van Orden, Holden und Turvey (2003) betonen, dass Anwendungsmöglichkeiten von $1/f$ -Noise durchaus nicht nur auf Experimente zu einfachen repetitiven Verhaltensweisen oder Verhaltensbeobachtungen über den Zeitverlauf eines Experiments hinweg (z.B. 1000 Beobachtungen während eines einstündigen Experiments) beschränkt seien, sondern beispielsweise auch in ökologisch valideren Settings oder über eine Vielzahl von Zeitspannen hinweg von Bedeutung seien. So verweisen sie z.B. auf eine Untersuchung von Delignieres, Fortes und Ninot (in press), in der die zeitabhängige Dynamik des Selbstwertgefühls erforscht wird. Peterson und Leckman (1998) fanden heraus, dass sich die zeitliche Dynamik der Tic-Dauer beim Tourette-Syndrom einem $1/f$ -Noise-Muster annähert. Ähnlich zeigt eine Studie von Gottschalk, Bauer und Whybrow (1995), dass $1/f$ -Noise in den tagesabhängigen Fluktuationen der selbstgemessenen Stimmung bei bipolaren Patienten über einen Zeitraum von einem bis zu zweieinhalb Jahren nach-

weisbar ist. Neben Schiepek (2003), der eine Erforschung klinischer Fälle mithilfe von dynamischen Systemen empfiehlt, macht die Studie von Yoshinaga, Miyazima und Mitake (2000) deutlich, dass Fingerklopfmuster bei behinderten und gesunden Erwachsenen sich durch das Ausmaß von $1/f$ -Noise in den resultierenden Zeitreihen unterscheiden könnten. Nicht zuletzt Ward (2002) weist darauf hin, dass diese Befunde über $1/f$ -Noise in einer Vielzahl psychologischer Zeitreihen die Entwicklung psychologischer Theorien inspirieren könnten.

Auch wenn die psychologische Anwendung von zeitreihenanalytischen Modellierungen im Vergleich zur Ökonometrie beispielsweise keine sehr lange Geschichte aufzuweisen hat (vgl. Farrell, Wagenmakers & Ratcliff, 2004, S.6), so eröffnen die dargestellten Studien dennoch Perspektiven auf ein weites noch zu erschließendes psychologisches Forschungsgebiet. Sie wurden an dieser Stelle deshalb so ausführlich beschrieben, weil die Existenz von fraktional integrierten Prozessen in der Psychologie darauf hinweist, dass psychologische Phänomene ebenfalls durch fraktional kointegrierte Prozesse modellierbar sein müssten.

Fraktionale Kointegration Barkoulas, Baum und Oguz (1996, S.3) formulieren, dass dann fraktionale Kointegration vorliege, wenn bei einem gegebenen vektoriellem Prozess \mathbf{x}_t die Linearkombination $\beta' \mathbf{x}_t$ fraktional integriert, d.h. $I(d - b)$ mit $b > 0$, sei. Nach Brunetti und Gilbert (1999) wird dabei vorausgesetzt, dass die kointegrierten Komponentenprozesse den gleichen Integrationsgrad haben, was beim Vorliegen empirischer Daten zu überprüfen sei. Die Linearkombination ist dabei mittelwertsneutral (mean reverting) aber nicht notwendigerweise kovarianzstationär und beschreibt wie bei der herkömmlichen Kointegration eine langfristige Gleichgewichtsbeziehung, die aber laut Werner (2005, S. 146) zur Stabilisierung einen langen Zeitraum benötige, was darauf zurückzuführen sei, dass der kumulative Einfluss der Random Shocks möglicherweise erst im Unendlichen gegen Null konvergiere. Ebenso wie herkömmliche Kointegration ist das fraktionale Pendant in der Literatur nicht einheitlich definiert, weshalb eine ausführliche theoretische Darstellung, die Informationen aus verschiedenen Quellen zusammenträgt und vergleicht, noch aussteht.

Dueker und Startz (1994, 1995) und Baillie und Bollerslev (1994) gehen in ihren Artikeln ausführlich auf fraktional kointegrierte Modelle ein. Die letztgenannten Autoren schildern den Fall, dass die Linearkombination zweier $I(1)$ -Reihen nicht exakt einen $I(0)$ -Prozess beschreibt, sondern für $0 < d < 1$ vom Grad d integriert ist. Maddala und Kim (1998, S. 302) stellen allerdings in Frage, dass dieser Fall tatsächlich die Bedingungen fraktionaler Kointegration erfüllt. Sie weisen auf eine weiter gefasste Definition fraktionaler Kointegration von Dueker und Startz (1995) hin, die argumentieren, dass fraktionale Kointegration dann vorliege, wenn eine Linearkombination von zwei $I(d)$ -Prozessen existiere, die $I(d - b)$ für $b \geq 0$ sei. Sie sind der Meinung, dass durch diese Definition die kontinuierliche Messung von Kointegration mehr Informationen bereitstelle als dies in der einfachen $I(0)/I(1)$ -Kategorisierung möglich sei.

Abhandlungen über den Kointegrationsrang fraktional integrierter Prozesse sind bei Breitung und Hassler (2000) zu finden. Auch für fraktional kointegrierte Zeitreihen existieren Fehlerkorrekturdarstellungen, die bei Dittmann (2004) ausführlich dargestellt sind.

In verschiedenen Studien hat sich gezeigt, dass die fraktionale Kointegration in manchen Fällen gegenüber der ganzzahligen Kointegration vorzuziehen ist. So fanden im ökonomischen Bereich Lardic und Mignon (2004), dass die fraktionale Kointegration bei der Abbildung kurz- und langfristiger Zinsdaten flexibler ist als die herkömmliche Kointegration. Krämer (1997) stellt dar, wie dies auf die Kointegration von Aktienkursen zu übertragen ist. Ein interessantes Anwendungsbeispiel liefern Caporale und Gil-Alana (2002), indem sie in ihrer Studie über den Zusammenhang zwischen Arbeitslosigkeit und Input-Preisen explizit die Modellierung durch fraktionale Kointegration als adäquater für diese Themenstellung beschreiben, weil sie die langfristigen Gedächtniseffekte und die langsame Rückkehr zum Gleichgewicht abbilden könne.

Das *Journal of Econometrics* (1996) widmet einen gesamten Band, nämlich das von Baillie und King herausgegebene *Volume 73* mit dem Titel „Fractional Differencing and Long Memory Processes“, der Auseinandersetzung mit unterschiedlichen

Themenbereichen, die für die Modellierung fraktional integrierter Prozesse relevant sind. Der Band enthält unterschiedliche Artikel namhafter Autoren. Zur weiteren Vertiefung und Auseinandersetzung mit fraktional integrierten und kointegrierten Prozessen sind die enthaltenen Artikel zu empfehlen.

3.6.2 Saisonale Kointegration und Multi-Kointegration

Neben der dargestellten Erweiterung in Form der fraktionalen Kointegration war Granger noch an einer Reihe anderer Erweiterungen beteiligt, von denen hier exemplarisch nur zwei genannt werden sollen. Ihre genaue theoretische und empirische Analyse würde den Rahmen dieser Arbeit sprengen. Hassler (2003, S. 4) beschreibt saisonale Zeitreihen bei Quartalsdaten beispielsweise als solche bei denen die Differenzbildung zum Vorquartal noch keine Stationarität liefere und es erforderlich sei, Jahresdifferenzen zum Quartal des Vorjahres zu bilden, um Stationarität zu erhalten. In diesem Fall spricht man von *saisonal integrierten* Zeitreihen, zu denen auch das Analogon der *saisonalen Kointegration* existiert (vgl. Hylleberg, Engle, Granger & Yeo, 1990). *Multi-Kointegration* beschreibt sozusagen eine Verschränkung verschiedener Kointegrationsbeziehungen ineinander (vgl. Granger & Lee, 1990 in Engle & Granger, 1991). So nennt wiederum Hassler (2003, S. 4) das Beispiel von Konsum und Einkommen, bei dem die Reihen integriert seien, die Ersparnis der Differenz aber stationär sei und somit das Gegenwartsvermögen als kumulierte Ersparnis der Vergangenheit definitionsgemäß wieder integriert sein müsse. Dadurch bestehe statistisch die Möglichkeit einer zweiten plausiblen Kointegrationsbeziehung, nämlich zwischen den drei Variablen Konsum, Einkommen und Vermögen.

3.7 Beispiele für die Anwendung des Kointegrationskonzepts

Dieser Abschnitt, der die theoretischen Ausführungen abschließt, widmet sich der Darstellung von Beispielen der empirischen Anwendung des Kointegrationskonzepts und stellt somit sowohl eine Überleitung zum folgenden empirischen Teil als auch einen Ausblick auf mögliche Forschungsaktivitäten dar.

Wie in der Einleitung (S. 10ff.) dargestellt, dominieren in der Psychologie eher quer-

schnittliche als prozessbezogene Analyseverfahren. Darin liegt sicherlich auch die Tatsache begründet, dass in den psychologischen Datenbanken, wie z.B. *PsycINFO*, kaum Anwendungsbeispiele für die jüngsten Entwicklungen im Rahmen zeitreihenanalytischer Verfahren wie das Kointegrationskonzept zu finden sind. Die nachfolgend dargestellten, u.a. soziologischen, damit bereits sozialwissenschaftlichen und somit der Psychologie näheren Beispiele zu meist gesellschaftlichen Themen können dennoch einen Einblick in die Möglichkeiten der Anwendung des Kointegrationskonzepts geben. Gleichzeitig spiegeln sie den Bedarf wieder, vor dem Hintergrund der in der Einleitung dargestellten Anwendungsrelevanz der multivariaten Modellierung in stationärer Zeitreihen durch kointegrierte Modelle für die psychologische Forschung, aktiv in diesem Sinne längsschnittliche Forschung zu betreiben, um der „Realität“ von prozessbezogenen psychologischen miteinander verbundenen Phänomenen durch eine angemessene Modellbildung gerecht zu werden.

Ökonometrische Publikationen zur Anwendung des Kointegrationskonzepts sind in großer Anzahl zu finden. Deshalb sollen im Rahmen dieses Abschnittes nur die nicht bzw. nur teilweise ökonometrischen Anwendungen aus aktuellen Studien dargestellt werden.

Eine Studie von Bremmer und Kesselring (2004) beschreibt die Modellierung von kointegrierten Prozessen zur Klärung des Zusammenhangs von Scheidungsrate und Arbeitskraftbeteiligung von Frauen. Unter Anwendung kointegrativer Verfahren werden die Kausalitätsrichtungen zwischen den folgenden vier Variablen bestimmt: ansteigende Scheidungsrate in den Vereinigten Staaten in den letzten 40 Jahren, wachsende weibliche Arbeitskraftbeteiligung, wachsendes Einkommen von Frauen und signifikante Reduktion der Geburtenrate bei verheirateten Frauen. In ähnlicher Weise stellt McNown (2003) ein kointegriertes Modell zur Klärung des Zusammenhangs zwischen der altersabhängigen Geburtenrate und der weiblichen Arbeitskraftbeteiligung auf. Weitere Variablen sind: Einkommen der Frauen, Arbeitslosenraten, Bildung und Einkommen der Männer. Die Kointegrationsbeziehung zwischen Fruchtbarkeit und Arbeitsverhalten von Frauen auf dem Arbeitsmarkt zeigt sich dabei robust über beide Altersgruppen und andere Variablen. Luiz (2001) klärt durch ko-

integrierte Modelle den Zusammenhang zwischen Kriminalitätsdynamik und ihren ökonomischen Determinanten. Eine Studie von Witt und Witte (2000) nutzt das Kointegrationskonzept zur Abbildung der Beziehung von Kriminalität, Inhaftierung und Arbeitskraftbeteiligung von Männern. Des weiteren beschreiben Darrat und Al-Yousif (1999) den langfristigen Zusammenhang zwischen Bevölkerungs- und Wirtschaftswachstum in Entwicklungsländern. O'Brien (1999) analysiert die Konvergenz bzw. Divergenz „of serious crime arrest rates“ von Männern und Frauen über 35 Jahre hinweg durch die Bildung kointegrierter Modelle. Eine Studie von Dupagne und Green (1996) nutzt Kointegrationstechniken zur Untersuchung von „consumer mass media expenditures in Belgium“. In der psychologischen Datenbank *PsycINFO*²⁵ konnten chronologisch rückwärts bis 1996 zum Suchwort „cointegration“ nur die beschriebenen Artikel von Luiz (2001), Witt und Witte (2000), O'Brien (1999) und Dupagne und Green (1996) gefunden werden. Die außerdem benannten Studien von Darrat und Al-Yousif (1999), Bremmer und Kesselring (2004) und McNown (2003) sind einer Recherche in der Datenbank *ProQuest*²⁶ entnommen.

Im Abschnitt 3.6.1, S. 69ff., wurden bereits viele psychologische Anwendungsbeispiele für univariate fraktional integrierte Zeitreihen dargestellt und die dort referenzierten Quellen weisen auf ein großes Anwendungsfeld hin.

Wie in der Einleitung (S. 10ff.) angedeutet, ist davon auszugehen, dass eine Vielzahl psychologischer Phänomene *instationär* ist, ihr Verlauf in der gemeinsamen Betrachtung mit anderen Phänomenen, d.h. *multivariat*, besser prognostiziert werden kann und *Granger-kausal* miteinander verbunden und durch eine *stationäre Linear-*

²⁵Der Zugang zu dieser Datenbank für psychologische Fachrecherche ist über die Homepage der Universitätsbibliothek der Universität Heidelberg <http://www.ub.uni-heidelberg.de> möglich. Die Datenbank der *American Psychological Association* wertet unter vierteljährlicher Aktualisierung ca. 1.300 internationale Zeitschriften, sowie Bücher und Buchkapitel aus den Gebieten Psychologie, Psychiatrie, Soziologie, Erziehungswissenschaften und medizinische Psychologie aus (Beschreibung entnommen aus <http://www.ub.uni-heidelberg.de/helios/epubl/info/daba/psyclit.html>).

²⁶Der Zugang zu dieser fächerübergreifenden Datenbank steht mir aufgrund meines Online-Stipendiums bei *e-fellows.net* unter der Internetadresse <http://www.e-fellows.net> offen. Diesem Stipendienggeber sei an dieser Stelle für die Recherchemöglichkeiten und das Stipendium herzlich gedankt.

kombination darstellbar und damit kointegriert sind. Wenn man beispielsweise die von Hautzinger, Baumgartner, Neßhöver und Schmitz (1987) durchgeführte und in der *Zeitschrift für Klinische Psychologie* publizierte „Zeitreihenanalyse kognitiver Verhaltenstherapie bei depressiven Patienten“ betrachtet, kann man ohne Zweifel zusätzliche Faktoren ableiten, die in diesem Zusammenhang eine Rolle spielen und möglicherweise mit den dort beschriebenen Variablen in einer langfristigen Gleichgewichtsbeziehung stehen. Ebenso könnte möglicherweise die von Bürgy (1998) beschriebene zeitreihenanalytische Modellierung psychologischer Regulationsvorgänge, die neben anderen Anwendungsbeispielen auch in Werner (2005) beschrieben ist, multivariat erweitert und ein langfristiger Zusammenhang identifiziert werden. Aber auch die vielfältigen durch fraktional integrierte Modelle abbildbaren psychologischen Phänomene, die v.a. in jüngster Zeit an Bedeutung gewinnen (vgl. Abschnitt 3.6.1, S. 69ff.), stehen ohne Frage in (Wechsel-)Beziehungen zu anderen Phänomenen. Auch wenn die „Ausbeute“ an Anwendungsartikeln zur Kointegration psychologischer Phänomene zu diesem Zeitpunkt aufgrund der Neuheit des Konzepts noch relativ gering ist, kann die Diplomarbeit als informativer Impuls verstanden werden. In diesem Sinne ist aus den psychologischen Anwendungsartikeln zu univariaten Zeitreihen, die in größerer Anzahl vertreten sind, ebenfalls ein Informationsgewinn für die Anwendung kointegrierter Modelle zu ziehen, da unter den benannten Bedingungen durch die Vernetzung der psychologischen Einzelreihen die Ableitung von Kointegrationsbeziehungen und damit eine noch genauere Abbildung der Realität als Funktion kointegrierter Modelle möglich sein kann.

4 Kointegration als empirisches Konzept: Simulationen

4.1 Problemstellung

Ein maßgebliches Ziel der theoretischen Gegenüberstellungen der unterschiedlichen Definitionen von Kointegration (vgl. Abschnitt 3, S. 32ff.) ist es, notwendige Bedingungen für das Vorliegen von Kointegration zu finden, um Kointegration als ein Konzept definieren zu können, das *interpretier- und testbar* ist. Deshalb ist es wichtig, einige in den Definitionen enthaltene Prämissen mithilfe von Simulationen zu überprüfen. Zu diesem Zweck können die folgenden Hypothesen formuliert werden, die nach Themenbereichen geordnet sind.

4.1.1 Fragestellung und Hypothesen

Hypothesen zur Klärung des Kointegrationskonzepts Die erste Hypothese wird als *Hypothese 0* bezeichnet, da diese Fragestellung in der Literatur kaum Beachtung findet und möglicherweise von vielen Autoren als trivial betrachtet wird. Neben Stier (2001, S. 315), der Kointegration als „interessanten Spezialfall“ bezeichnet, erwecken auch die anderen bereits bei der Definition von Kointegration (vgl. Kapitel 3, S. 32ff.) zitierten Autoren den Eindruck, dass Kointegration eher ein besonderes, nicht allzu häufig auftretendes Phänomen der multivariaten Zeitreihenanalyse ist und unter Umständen nicht gegeben sein kann. Der Nobelpreisträger C.W.J. Granger formuliert in seiner „Nobel Lecture“ am 8. Dezember 2003 zum Thema *Time Series Analysis, Cointegration, and Applications* (Bank of Sweden / The Royal Swedish Academy of Sciences, 2003a) die folgende eher uneindeutige Aussage:

For cointegration, a pair of integrated, or smooth series, must have the property that a linear combination of them is stationary. Most pairs of integrated series will not have the property, so that cointegration should be considered as a surprise when it occurs. In practice, many pairs of macroeconomic series seem to have the property, as is suggested by eco-

nomic theory.

Einzelne vorläufige Simulationen im Rahmen dieser Diplomarbeit geben Anlass zur Formulierung der folgenden Hypothese, durch deren Überprüfung eine strukturierte Vorlage für inhaltliche Interpretationen generiert werden kann.

- Hypothese 0:

Für VAR(1)-Prozesse gilt, dass in Abhängigkeit von der Anzahl der Messzeitpunkte bereits bei minimalster Abhängigkeit der Komponentenprozesse Kointegration vorliegt.

Wie bereits dargestellt (vgl. 3.1, S. 40f.) äußern nur wenige Autoren, wie z.B. Rinne und Specht (2002, S. 542f.), explizit, dass auseinanderdriftende Zeitpfade nicht im Widerspruch zum Kointegrationskonzept stehen:

Es wird oft behauptet, dass kointegrierte Variablen Zeitpfade aufweisen, die nicht auseinander driften. Das trifft aber nur dann zu, wenn - etwa bei $m = 2$ Variablen - der Kointegrationsvektor durch $\beta' = (1, -1)$ gegeben ist. (Rinne & Specht, 2002, S. 542)

Dieser Sachverhalt, der durch die im theoretischen Teil bereits erfolgte graphische Darstellung kointegrierter Prozesse für $\beta_2 \neq \pm 1$ visualisiert werden konnte (vgl. Abbildung 3.3, S. 57), soll durch die Überprüfung der folgenden Hypothese empirisch fundiert werden.

- Hypothese 1a:

Für VAR(1)-Prozesse gilt, dass auch bei anhand von Graphiken beobachtbaren auseinanderdriftenden Zeitpfaden durch das Vorliegen zweier instationärer Komponentenprozesse und einer stationären Linearkombination von Kointegration zu sprechen ist.

Obwohl wie in der oben formulierten Hypothese vermutlich sowohl nicht-auseinanderdriftende als auch auseinanderdriftende Zeitpfade dem Kointegrationskonzept zuzuordnen sind, habe ich in vorangegangenen Simulationen beobachtet, dass der Johansen-Kointegrationstest bezüglich der beiden Verlaufstypen unterscheidet:

- Hypothese 1b:

Für VAR(1)-Prozesse gilt, dass der Johansen-Test bei nicht-auseinanderdriftenden Zeitpfaden bei identischer Anzahl der Messzeitpunkte und vergleichbarer Abhängigkeit der Komponentenprozesse erst ab einem höheren Grad der Abhängigkeit der Komponentenprozesse („langsamer“) Kointegration indiziert als bei auseinanderdriftenden Einzelprozessen.

Dem Dickey-Fuller-Kointegrationstest und dem Johansen-Kointegrationstest liegen unterschiedliche Teststatistiken zugrunde (vgl. 3.5, S. 60ff.). Deshalb sollen durch die folgende Hypothese die beiden Verfahren miteinander verglichen werden:

- Hypothese 1c:

Für VAR(1)-Prozesse gilt, dass der DF-Kointegrationstest und der Johansen-Kointegrationstest sich bei identischer Anzahl der Messzeitpunkte und vergleichbarer Abhängigkeit der Komponentenprozesse bezüglich der diagnostizierten Häufigkeit von Kointegrationsbeziehungen unterscheiden.

Für die im empirischen Teil generierten vektoriellen Prozesse mit nur einer Kointegrationsbeziehung lassen sich die Hypothesen 0 und 1a jeweils sowohl mit dem Dickey-Fuller-Kointegrationstest (vgl. 3.5.1, S. 61ff.) als auch mit dem Johansen-Kointegrationstest (vgl. 3.5.2, S. 62ff.) überprüfen, wobei für die Klärung von Hypothese 1a hauptsächlich der DF-Kointegrationstest herangezogen wird, weil er sowohl die Einzelreihen als auch die Linearkombination auf (In-)Stationarität untersucht. Hypothese 1b bezieht sich auf das Testverfahren von Johansen und Hypothese 1c auf den Vergleich der beiden Verfahren.

Hypothesen zur Schätzung von α und β Durch die in Abschnitt 3.4.2, S. 59f., geleistete Vereinfachung der Rückrechnung aus dem VECM auf α und β für den bivariaten VAR(1)-Fall ist eine Beurteilung der Güte der Schätzung der beiden Parameter möglich, die durch die folgenden Hypothesen geleitet sein soll, und deren Überprüfung zur Interpretation der bisher nur selten inhaltlich beleuchteten Parameter beitragen kann. Simulationen im Rahmen der Vorbereitung der Diplomarbeit lassen vermuten, dass sich die Güte der Schätzung von α und β unterscheidet. In der

Literatur sind nur wenige Anhaltspunkte zur Qualität der Schätzungen zu finden. Deshalb scheint die Formulierung gerichteter Hypothesen in diesem Zusammenhang nicht sinnvoll. Aufgrund der Betrachtung der Prozessverläufe und der mit α und β zu erfassenden Sachverhalte würde man vermuten, dass die Güte der Schätzung des Kointegrationsvektors ähnlich hoch ausfällt wie die Qualität der Schätzung von ARMA-Parametern (vgl. Werner, 2005, S. 165ff.).

- Hypothese 2a:

Für VAR(1)-Prozesse gilt, dass die Güte der Schätzung des Kointegrationsvektors β über das Fehlerkorrekturmodell sehr hoch ist, d.h. im üblichen Range der Schätzungen von ARMA-Parametern liegt, nämlich zwischen einem MSE von 0,2 und 0,002 (vgl. Werner, 2005, S. 165ff.).

- Hypothese 2b:

Für VAR(1)-Prozesse gilt, dass die Güte der Schätzung des Anpassungskoeffizienten α sich von der Güte der Schätzung des Kointegrationsvektors β bei gleicher Anzahl der Messzeitpunkte und vergleichbarer Abhängigkeit der Komponentenprozesse unterscheidet.

- Hypothese 2c:

Für VAR(1)-Prozesse gilt, dass die Schätzgüte von α vs. β sich in Abhängigkeit vom Verlauf der Zeitpfade der Komponentenprozesse (auseinanderdriftend vs. nicht-auseinanderdriftend) bei gleicher Anzahl der Messzeitpunkte und vergleichbarer Abhängigkeit der Komponentenprozesse unterscheidet.

- Hypothese 2d:

Für VAR(1)-Prozesse gilt, dass die Schätzgüte von α vs. β sich in Abhängigkeit von der variierten Abhängigkeit der Komponentenprozesse über die Nebendiagonalen bei gleicher Anzahl der Messzeitpunkte unterscheidet.

4.2 Methode

Wie aus den Grundlagen der Statistik bekannt ist, kennt man die Populationsparameter in der empirischen Forschungspraxis in der Regel nicht. Nur mithilfe von

Simulationen ist es möglich, die Populationsparameter der Zeitreihenmodelle zu kennen. Die Kenntnis der wahren Parameterwerte wiederum ist Voraussetzung für die empirische Bearbeitung der vorgestellten Problemstellung. Während man sich beispielsweise in der empirischen Praxis auf die Güte bestimmter Schätz-Verfahren verlassen muss, kann man unter Zuhilfenahme von Simulationen durch die Kenntnis der wahren Werte Schätz- und Populationswert miteinander vergleichen. Aus diesem Grund ist die **Monte-Carlo-Simulation** (vgl. Hammersley & Handscomb, 1965; Milligan, 1981; Rubinstein, 1981) die Methode der Wahl und soll im folgenden dargestellt werden. Darüber hinaus werden die Vorgehensweise bei der Datenerzeugung und der Untersuchungsablauf beschrieben.

4.2.1 Monte-Carlo-Simulation

Die Monte-Carlo-Methode ist auf die Arbeit von Metropolis und Ulam (1949) zurückzuführen und ist zu verschiedenen Forschungszwecken einsetzbar. Die notwendigen Daten werden dabei mithilfe des Computers simuliert und nicht an realen Versuchspersonen erhoben. Diese Vorgehensweise ermöglicht die Kenntnis über Modelleigenschaften wie Parameterzahl und -ausprägung. Bei der Erhebung von Daten an realen Objekten ist diese Kenntnis nicht gegeben. Für die methodologische Forschung und damit die Forschung an einer Methode, die aufgrund ihrer Bewährtheit dann in der Forschung am realen Objekt Anwendung finden kann, ist das Vorgehen deshalb hilfreich, weil nur auf diesem Weg durch Kenntnis der Population die Güte statistischer Verfahren überprüft werden kann. Deshalb ist die Methode für die vorliegende Arbeit unersetzlich. Kalos und Whitlock (1986) und Kleijnen (1987) geben Hinweise zur praktischen Durchführung.

Stichprobenkennwerte müssen gewisse Gütekriterien erfüllen, wenn sie Schätzungen der Populationsparameter sein sollen (vgl. Werner, 1997, S. 23ff.). Als eine Erweiterung der sogenannten Fisher-Kriterien für gute Schätzer, nämlich *Erwartungstreue*, *Konsistenz*, relative statistische *Effizienz* und *Suffizienz* (vgl. Bortz, 1999, S. 95ff.; Fisher, 1925; Hays, 1994, S. 211ff.), stellt Werner (1997, S. 24) den Mittleren Quadratischen Fehler (MSE: Mean Square Error) heraus, der zwei der vier Fisher-Kriterien,

nämlich Erwartungstreue und Effizienz, in sich vereint. Eine Schätzung ist dann erwartungstreu, wenn der *Bias* Null beträgt. Der Bias entspricht dem Erwartungswert der quadrierten Differenz $[E(b) - \beta]$ ²⁷. Gleichzeitig ist es für die Beurteilung eines Schätzers aber auch wichtig zu wissen, wie sehr die Schätzungen streuen. Der Schätzer mit der höheren statistischen Effizienz und damit der Schätzer mit der kleineren Varianz und größerer mittlerer Abweichung vom Populationswert ist dabei gegenüber einem alternativen Schätzer mit geringerer mittlerer Abweichung vom Populationswert zu präferieren. Die relative statistische Effizienz eines Schätzers gegenüber einem alternativen Schätzer errechnet sich aus dem Quotienten der Varianzen der beiden Schätzer.

Der mittlere quadratische Fehler schließlich erfasst beide Eigenschaften zugleich, nämlich Erwartungstreue und Varianz. Es gibt Schätzer, die einen (deutlichen, wenn auch geringen) Bias durch eine (sehr viel) geringere Varianz „wettmachen“ können, sodass der mittlere quadratische Fehler (MSE) gegenüber dem erwartungstreuen Schätzer kleiner ist. (Werner, 1997, S. 24; Hervorhebungen im Original)

Der MSE wird nachfolgend als sinnvolles Kriterium zur Beurteilung der Schätzgüte von α und β genutzt und deshalb zu Beginn des empirischen Teils der Arbeit an dieser Stelle beschrieben. Er lässt sich wie folgt berechnen:

$$\text{MSE}(b) = E(b - \beta)^2 = \text{var}(b) + E[E(b) - \beta]^2 = \text{var}(b) + \text{Bias} \quad (4.1)$$

Eine detaillierte Ableitung ist bei Werner (1997, S. 23) zu finden.

Ein Vorteil der Monte-Carlo-Methode für die Untersuchung der aktuellen Problemstellung liegt in der Nutzung von Zufallsgeneratoren, die in die programmierten Zeitreihen x_{1t} und x_{2t} integriert werden und so den Zufallsanteil modellieren. In der vorliegenden Arbeit werden für jede Komponentenreihe andere Zufallszahlen verwendet, um sicherzustellen, dass die Prozesse nicht deshalb kointegriert sind, weil ihre Zufallsanteile auf dieselbe Zufallszahl zurückzuführen sind.

²⁷Gemäß der üblichen Notation steht b für einen beliebigen Stichprobenkennwert, der griechische Buchstabe β für den dazugehörigen Populationswert.

Ein Vorteil von Simulationen zur *Forschung an der Methode* besteht darin, dass die Fehler *nicht* unbekannt sind. Damit wird vermieden, dass die in der Regel vorhandenen unerklärten Variablen, die nicht im Modell spezifiziert wurden, in der Fehlergröße subsumiert werden.

Im Anschluss an die Simulation werden die generierten Daten den für die Fragestellung relevanten zeitreihenanalytischen Verfahren, so beispielsweise dem Johansen-Test, unterzogen. Die wahren Parameter werden bereits bei der Modellbildung vorgegeben und sind somit bekannt. Somit kann die Güte der Schätzung durch den Vergleich von Schätz- und Parameterwerten - hier für α und β - überprüft werden. Mithilfe von Monte-Carlo-Simulationen kann folglich beleuchtet werden, was unter optimalen Bedingungen geschieht.

Die simulierten Daten können unter Vorgabe derselben Zufallszahlen repliziert werden. Rubinstein (1981) verdeutlicht, dass nur der Startwert des Zufallszahlengenerators und der Zufallsgenerator selbst bekannt sein müssen.

Um die Monte-Carlo-Methode als wesentliches Instrument der methodologischen Forschung zu illustrieren, kann man eine Parallele zur allgemeinspsychologischen Szenario-Forschung ziehen: auch hier wird ein Modell zur Erzeugung von Daten generiert, in dem anschließend Komponenten variiert werden, um beispielsweise das Verhalten der Versuchsperson (analog: Test- und Schätzverfahren) in diesen unterschiedlichen Varianten zu beobachten.

Die Wichtigkeit der Monte-Carlo-Methode für die empirische Forschung wird von Bortz (1993, S. 130) unterstrichen, der hervorhebt, dass sie Entscheidungen darüber erleichtert, „unter welchen Umständen ein bestimmter Test eingesetzt oder nicht eingesetzt werden kann.“

4.2.2 Datenerzeugung

In allen Simulationen zur Bearbeitung der Hypothesen beschreiben die Einzelreihen Random-Walk-Prozesse (vgl. 2.4, S. 26ff.). Es erscheint notwendig, nochmals anzumerken, dass die beiden Random-Walk-Prozesse im Fall von Kointegration voneinan-

der abhängig sind, da sie einen gemeinsamen stochastischen „Trend“ aufweisen und somit derselbe Zufalls-Prozess in beiden wirksam ist. Damit dies gegeben ist, sind Φ_1 -Matrizen mit einer Einheitswurzel bzw. einem Eigenwert von 1 zu generieren, wie folgendermaßen gezeigt werden kann:

Gegeben ist die Φ_1 -Matrix:

$$\Phi_1 = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}$$

Gemäß dem Ansatz zum Charakteristischen Polynom der Matrix $\Phi_1 - \lambda \mathbf{I}$

$$\Phi_1 - \lambda \mathbf{I} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \phi_{11} - \lambda & \phi_{12} \\ \phi_{21} & \phi_{22} - \lambda \end{pmatrix}$$

ist die *Determinante* abzuleiten:

$$|\Phi_1 - \lambda \mathbf{I}| = (\phi_{11} - \lambda)(\phi_{22} - \lambda) - \phi_{12}\phi_{21}$$

Homogene Gleichungssysteme der Ordnung $n \times n$ haben nur dann eine von Null verschiedene Lösung, wenn „die Determinante der Matrix der Koeffizienten gleich Null ist“ (Werner, 1997, S. 40), d.h.:

$$(\phi_{11} - \lambda)(\phi_{22} - \lambda) - \phi_{12}\phi_{21} = 0$$

$$\phi_{11}\phi_{22} - \phi_{11}\lambda - \phi_{22}\lambda + \lambda^2 - \phi_{12}\phi_{21} = 0$$

$$\lambda^2 + (-\phi_{22} - \phi_{11})\lambda + \phi_{11}\phi_{22} - \phi_{12}\phi_{21} = 0$$

Ziel ist es, Φ_1 -Elemente für λ_1 oder $\lambda_2 = 1$ zu bestimmen²⁸. Unter Auflösung der Gleichung für den Fall $\lambda_1 = 1$ erhält man:

$$1 - \phi_{22} - \phi_{11} + \phi_{11}\phi_{22} - \phi_{12}\phi_{21} = 0$$

Um einen nichtstationären VAR(1)-Prozess - bestehend aus zwei Random-Walk-Einzelreihen - zu erzeugen, muss die folgende Bedingung erfüllt sein:

$$\phi_{11} + \phi_{22} - \phi_{11}\phi_{22} + \phi_{12}\phi_{21} = 1$$

²⁸Anmerkung: Da es sich um eine quadratische Gleichung handelt, ergeben sich für λ zwei Lösungen: λ_1 und λ_2 .

Diese Gleichung wird zur Generierung der Φ_1 -Matrizen zur Sicherstellung von Random-Walk-Einzelreihen im vektoriellen Modell in SAS IML programmiert. Dabei werden drei Matrixelemente gesetzt und das vierte Element wird durch das Programm errechnet. Zur Kontrolle wird anschließend die generierte Φ_1 -Matrix mithilfe der `eigval`-Option von IML oder der `vtroot`-Option von VARMAX (vgl. SAS/ETS Software: Changes and Enhancements, 2001) auf eine Einheitswurzel überprüft (vgl. Anhang D1, S. 146).

4.2.3 Programmierung in SAS

Alle Zeitreihen werden im Rahmen dieser Arbeit mit dem Statistikprogramm SAS - Release 8.2 programmiert. Zur Programmierung in SAS sei auf einige hilfreiche Literatur verwiesen, damit die im Anhang aufgeführten Programme bei Bedarf leicht nachvollzogen werden können. Eine allgemeine, leicht verständliche Einführung findet sich bei Ortseifen (1997). Wie in SAS/Language (1990) können in den Benutzerhandbüchern von SAS/ETS (1999) die notwendigen Prozedurschritte nochmals im Einzelnen nachvollzogen werden. Brocklebank und Dickey (1986) führen in die SAS-Programmierung von Zeitreihen ein. Angaben zur Programmierung multivariater Zeitreihen mithilfe des Prozedurschrittes `proc varmax` können in SAS/ETS Software (2001) nachgelesen werden. Des weiteren existiert hilfreiche Literatur zur Erstellung von Graphiken in SAS (SAS/GRAPH, 1991), zur Programmierung mit Makros (SAS, 1990) und zur Programmierung mit IML (SAS/IML, 1990).

Die für die Simulationen verwendeten Programme sind in Anhang D, S. 141ff., aufgeführt. Zu Beginn werden dort in einem Überblick die wesentlichen Programm- und Prozedurschritte erläutert, um es der Leserin bzw. dem Leser zu ermöglichen, die Programme nachzuvollziehen und dadurch gegebenenfalls Hilfestellungen für das eigene Programmieren in SAS zu erhalten.

4.2.4 Untersuchungsablauf

Der Untersuchungsablauf des empirischen Teils kann in Orientierung an der chronologischen Abfolge der Hypothesen in zwei große inhaltliche Bereiche und damit

in zwei große Simulationsblöcke unterteilt werden, wobei der erste Teil wieder in zwei inhaltliche Unterabschnitte unterteilt werden kann. Zunächst werden die Testverfahren und mit ihnen die Operationalisierung von Kointegration durch systematische Simulationen nach Vorgabe der Hypothesen untersucht. Dabei ist zwischen dem Testverfahren von Johansen und dem Dickey-Fuller-Kointegrationstest zu unterscheiden, da für beide Verfahren eigene Simulationen über die jeweils gleichen Modelle durchgeführt werden. Anschließend folgen systematische Simulationen zur Güte der Parameterschätzungen von Kointegrationsvektor und Anpassungskoeffizient. Durch die systematische Generierung von Matrizen und damit von unterschiedlichen Modellen, die v.a. den Kategorien *nicht-auseinanderdriftend*, d.h. $\beta_2 = \pm 1$, vs. *auseinanderdriftend*, d.h. $\beta_2 \neq \pm 1$, zugeordnet werden können, ist es möglich, abschließend allgemeine Aussagen über diese beiden Modellkategorien vor dem Hintergrund der Hypothesen zu treffen.

Testverfahren: Dickey-Fuller- und Johansen-Kointegrationstest Hypothese 0 und Hypothese 1a dienen v.a. der Validierung der im theoretischen Teil formulierten Annahmen zur Häufigkeit des Vorliegens von Kointegrationsbeziehungen und zu Kointegration von Prozessen mit auseinanderdriftenden Zeitpfaden. Hypothese 1b und Hypothese 1c haben direkte praktische Implikationen: ihre Überprüfung macht es möglich, die Anwendung der beiden Verfahren in ihren Vor- und Nachteilen abzuwägen, v.a. was das Verhalten der Testverfahren bezüglich der zu wählenden Anzahl der Messzeitpunkte, beim Vorliegen von auseinanderdriftenden Zeitpfaden und in Abhängigkeit von der Stärke des Zusammenhangs der Einzelreihen²⁹ des multivariaten Prozesses betrifft.

²⁹Die Stärke des Zusammenhangs wird im Rahmen der Diplomarbeit zur besseren Kategorisierung und zum besseren Vergleich der einzelnen VAR(1)-Prozesse zusätzlich dargestellt und entspricht der Summe der Nebendiagonalelemente, d.h. $|\phi_{12}| + |\phi_{21}|$. Dies erfasst damit die Stärke des Zusammenhangs bzw. die Höhe der globalen Abhängigkeit zwischen den Komponentenprozessen, nicht aber die Richtung des kausalen oder Gewichtung des möglicherweise wechselseitigen Zusammenhangs.

Güte der Parameterschätzungen von Kointegrationsvektor und Anpassungskoeffizient über das Fehlerkorrekturmodell Die Einflussgröße, die sicherlich die Beurteilung der Hypothesen moderiert, ist die Anzahl der Messzeitpunkte. Für den psychologischen Anwender eröffnet sich durch die nachfolgend vorgelegten Simulationsergebnisse zu den Hypothesen 2a bis 2d die Möglichkeit, im praktischen Fall zu entscheiden, wieviele Messzeitpunkte für eine reliable Schätzung (operationalisiert durch den MSE) von Kointegrationsvektor und Anpassungskoeffizient notwendig sind. Eine verlässliche Schätzung ist Voraussetzung für eine angemessene Abbildung der Realität. Ähnlich wie bereits im univariaten Fall geschehen und u.a. bei Stadnytska (2004) beschrieben, wäre es als Folgeuntersuchung der Diplomarbeit außerdem möglich, für den multivariaten Anwendungsfall eine Mindestanzahl an Messzeitpunkten für die unterschiedlichen VAR(1)-Modelle zu identifizieren, ab der stabile Schätzungen möglich sind, und dies noch systematischer als im Rahmen dieser Arbeit möglich zu dokumentieren. In Hypothese 2b bis 2d werden Vergleiche der Schätzgüte von α und β unter unterschiedlichen Modellbedingungen bei gleicher Anzahl der Messzeitpunkte durchgeführt.

4.2.5 Übersicht über die im Rahmen der Simulationen verwendeten Matrizen

In den folgenden Tabellen sind die im Rahmen der Simulationen verwendeten Matrizen und Eigenschaften der VAR(1)-Prozesse, die durch die Matrizen definiert sind, aufgeführt. Dabei sind in Tabelle 4.1 VAR(1)-Prozesse mit nicht-auseinanderdriftenden Zeitpfaden der Einzelreihen beschrieben, während in Tabelle 4.2 VAR(1)-Prozesse mit auseinanderdriftenden Zeitpfaden der Komponentenprozesse repräsentiert sind. Die Matrizen sind wie in den Tabellen benannt in unterschiedlicher Form variiert. Alle durch die in den Tabellen benannten Matrizen charakterisierten und simulierten Prozesse wurden auch graphisch visualisiert und ursprünglich in den Anhang der Arbeit aufgenommen. Aus Gründen des Umfangs muss darauf im Rahmen der Diplomarbeit verzichtet werden. Exemplarisch finden sich im Verlauf der Arbeit einige graphische Visualisierungen zum besseren Verständnis des Kointegra-

tionskonzepts.

Tabelle 4.1: Simulierte VAR(1)-Prozesse mit **nicht-auseinanderdriftenden** Zeitpfaden der Komponentenprozesse.

nicht-auseinanderdriftend <i>Variation minimaler Abhängigkeiten</i> Hypothese 0 $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} > 0$ $\phi_{11} + \phi_{12} = 1$ $\phi_{21} + \phi_{22} = 1$ ϕ_{21} und ϕ_{22} konstant Modell- matrix ϕ_{11} ϕ_{12} ϕ_{21} ϕ_{22} EW Det Abh β_2						nicht-auseinanderdriftend alle Hypothesen $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} > 0$ $\phi_{11} + \phi_{12} = 1$ $\phi_{21} + \phi_{22} = 1$ ϕ_{21} und ϕ_{22} konstant Modell- matrix ϕ_{11} ϕ_{12} ϕ_{21} ϕ_{22} EW Det Abh β_2					
1	0	1	1	0	-1	0,95	0,05	1	0,85	0,15	-1
0	1	1				0,1	0,9	0,85			
0,995	0,005	1	-0,005	0,105	-1	0,9	0,1	1	0,8	0,2	-1
0,1	0,9	0,895				0,1	0,9	0,8			
0,99	0,01	1	-0,01	0,11	-1	0,85	0,15	1	0,75	0,25	-1
0,1	0,9	0,89				0,1	0,9	0,75			
0,985	0,015	1	-0,015	0,115	-1	0,8	0,2	1	0,7	0,3	-1
0,1	0,9	0,885				0,1	0,9	0,7			
0,98	0,02	1	-0,02	0,12	-1	0,75	0,25	1	0,65	0,35	-1
0,1	0,9	0,88				0,1	0,9	0,65			
0,975	0,025	1	0,875	0,125	-1	0,7	0,3	1	0,6	0,4	-1
0,1	0,9	0,875				0,1	0,9	0,6			
						0,65	0,35	1	0,55	0,45	-1
						0,1	0,9	0,55			
						0,6	0,4	1	0,5	0,5	-1
						0,1	0,9	0,5			
						0,55	0,45	1	0,45	0,55	-1
						0,1	0,9	0,45			
						0,5	0,5	1	0,4	0,6	-1
						0,1	0,9	0,4			
						0,4	0,6	1	0,3	0,7	-1
						0,1	0,9	0,3			
						0,3	0,7	1	0,2	0,8	-1
						0,1	0,9	0,2			
						0,2	0,8	1	0,1	0,9	-1
						0,1	0,9	0,1			
						0,1	0,9	1	0	1	-1
						0,1	0,9	0			

Auch mag es der Leserin bzw. dem Leser auffallen, dass die Elemente der zweiten Zeile der aufgelisteten Modellmatrizen konstant sind. Zusätzlich wurden bei der

Tabelle 4.2: Simulierte VAR(1)-Prozesse mit **auseinanderdriftenden** Zeitpfaden der Komponentenprozesse.

auseinanderdriftend alle Hypothesen $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} \in \mathbb{R}$ $\phi_{11} + \phi_{12} < 1$ $\phi_{21} + \phi_{22} < 1$ ϕ_{21} und ϕ_{22} konstant Modell- matrix ϕ_{11} ϕ_{12} ϕ_{21} ϕ_{22} EW Det Abh β_2						auseinanderdriftend alle Hypothesen $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} > 0$ $\phi_{11} + \phi_{12} > 1$ $\phi_{21} + \phi_{22} < 1$ ϕ_{21} und ϕ_{22} konstant Modell- matrix ϕ_{11} ϕ_{12} ϕ_{21} ϕ_{22} EW Det Abh β_2					
1,01	-0,05	1	0,51	0,15	-5	0,99	0,05	1	0,49	0,15	-5
0,1	0,5	0,51				0,1	0,5	0,49			
1,02	-0,1	1	0,52	0,2	-5	0,98	0,1	1	0,48	0,2	-5
0,1	0,5	0,52				0,1	0,5	0,48			
1,03	-0,15	1	0,53	0,25	-5	0,97	0,15	1	0,47	0,25	-5
0,1	0,5	0,53				0,1	0,5	0,47			
1,04	-0,2	1	0,54	0,3	-5	0,96	0,2	1	0,46	0,3	-5
0,1	0,5	0,54				0,1	0,5	0,46			
1,05	-0,25	1	0,55	0,35	-5	0,95	0,25	1	0,45	0,35	-5
0,1	0,5	0,55				0,1	0,5	0,45			
1,06	-0,3	1	0,56	0,4	-5	0,94	0,3	1	0,44	0,4	-5
0,1	0,5	0,56				0,1	0,5	0,44			
1,07	-0,35	1	0,57	0,45	-5	0,93	0,35	1	0,43	0,45	-5
0,1	0,5	0,57				0,1	0,5	0,43			
1,08	-0,4	1	0,58	0,5	-5	0,92	0,4	1	0,42	0,5	-5
0,1	0,5	0,58				0,1	0,5	0,42			
1,09	-0,45	1	0,59	0,55	-5	0,91	0,45	1	0,41	0,55	-5
0,1	0,5	0,59				0,1	0,5	0,41			
1,1	-0,5	1	0,6	0,6	-5	0,9	0,5	1	0,4	0,6	-5
0,1	0,5	0,6				0,1	0,5	0,4			
1,12	-0,6	1	0,62	0,7	-5	0,88	0,6	1	0,38	0,7	-5
0,1	0,5	0,62				0,1	0,5	0,38			
1,14	-0,7	1	0,64	0,8	-5	0,86	0,7	1	0,36	0,8	-5
0,1	0,5	0,64				0,1	0,5	0,36			
1,16	-0,8	1	0,66	0,9	-5	0,84	0,8	1	0,34	0,9	-5
0,1	0,5	0,66				0,1	0,5	0,34			
1,18	-0,9	1	0,68	1	-5	0,82	0,9	1	0,32	1	-5
0,1	0,5	0,68				0,1	0,5	0,32			
1,2	-1	1	0,7	1,1	-5	0,8	1	1	0,3	1,1	-5
0,1	0,5	0,7				0,1	0,5	0,3			

Erstellung der Arbeit auch Matrizen generiert, in denen die Elemente der zweiten Zeile variieren. Die anhand dieser Matrizen simulierten Prozesse wurden denselben

Test- und Schätzverfahren unterzogen. Auf den ersten Blick ergeben sich dabei keine stark abweichenden Ergebnisse. Mögliche Unterschiede sollen in einer Folgepublikation beleuchtet werden. Dieses Datenmaterial kann auf Nachfrage jederzeit gerne eingesehen werden.

In allen Matrizen sind die Elemente der Nebendiagonalen variiert, da ihre numerische Größe die Abhängigkeit der Komponentenprozesse voneinander bestimmt. Dies ist auch deutlich graphisch zu sehen. Um zu zeigen, dass die Variation der Hauptdiagonalen keinen Einfluss auf die Abhängigkeit der Einzelreihen voneinander hat, wurde dies mit systematischen Simulationen und den dazugehörigen Graphiken überprüft und nachgewiesen.

In den Tabellen entspricht die Notation [Abh] der Summe der Elemente in der Nebendiagonalen, d.h. $|\phi_{12}| + |\phi_{21}|$, und erfasst damit die Höhe der globalen Abhängigkeit zwischen den Komponentenprozessen, nicht aber die Richtung des kausalen oder Gewichtung des wechselseitigen Zusammenhangs. Der hier verwendete Begriff der Abhängigkeit ist also keinem bekannten Abhängigkeitsmaß zuzuordnen, sondern lediglich eine Operationalisierung der Stärke des Zusammenhangs der Komponentenprozesse zu Interpretationszwecken. $[\Phi_1]$ steht für die Modellmatrix, [EW] für Eigenwerte und [Det] für Determinante von Φ_1 bei systematisch variiertes (ansteigender) Abhängigkeit der Komponentenprozesse. Zusätzlich ist der Kointegrationsvektor β aufgeführt ($\beta_1 = 1$; $[\beta_2]$ variiert).

Außerdem ist der Hinweis zu geben, dass das Verhalten der Testverfahren bei *variieren Abhängigkeiten im minimalen Wertebereich* ($< 0,15$) im Rahmen dieser Diplomarbeit nur für Modelle mit nicht-auseinanderdriftenden Zeitpfaden demonstriert wird.

5 Interpretation der Ergebnisse und Überprüfung der Hypothesen

Die vollständige Darstellung der Ergebnis-Tabellen hinsichtlich der in den Hypothesen formulierten Fragestellungen würde diesen Teil der Ergebnisbeschreibung und -interpretation unübersichtlich machen. Um der Interpretation der Ergebnisse dennoch nicht nur anhand der Graphiken folgen zu können, wurden die Ergebnistabellen reduziert und die für die Ergebnisinterpretation besonders relevanten Daten sind in Form von Tabellenausschnitten hier aufgenommen. In den Tabellen sind die jeweils im Text erwähnten Werte **fett** markiert, um ein sinnvolles Nachvollziehen der Interpretationsschritte zu ermöglichen. Bei Interesse an weiteren zusätzlich erhobenen Statistiken können die Gesamt-Ergebnistabellen auf Anfrage gern zur Verfügung gestellt werden.

Zum Verständnis dieser hier aufgenommenen Tabellenausschnitte seien noch folgende Hinweise bezüglich der verwendeten Abkürzungen gegeben. Allgemein gilt, dass alle Simulationen über 200 Wiederholungen durchgeführt wurden. Die Modellmatrix $[\Phi_1]$, die den multivariaten Prozess durch ihre Elemente $[\phi_{11}]$, $[\phi_{12}]$, $[\phi_{21}]$ und $[\phi_{22}]$ definiert, ist in allen Tabellen zu finden. [T] beschreibt die systematisch variierte Anzahl der Messzeitpunkte. Das Abhängigkeitsmaß für die Stärke des Zusammenhangs zwischen den Komponentenprozessen wird mit [Abh] notiert (vgl. Erläuterungen auf S. 90). Nicht-Signifikanz bzw. Signifikanz auf dem 0,05-Niveau wird durch die Abkürzung [n; s] angegeben, während die Häufigkeit der (Nicht-)Signifikanzen durch $[x/200]$ und deren prozentualer Anteil entsprechend durch [%] abgebildet werden. Im Gegensatz zum Johansen-Test erfolgt dies beim Dickey-Fuller-Test jeweils getrennt für die Komponentenprozesse und die [Linearkombination]. In den Tabellen zur Schätzung von Anpassungskoeffizient und Kointegrationsvektor sind zusätzlich die folgenden Abkürzungen relevant: die mittleren quadratischen Fehler $[MSE_{\tilde{\beta}_2}]$, $[MSE_{\tilde{\alpha}_1}]$, $[MSE_{\tilde{\alpha}_2}]$ und $[MSE_{\tilde{\alpha}_{ges}}]$. $MSE_{\tilde{\alpha}_{ges}}$ beschreibt das arithmetische Mittel aus $MSE_{\tilde{\alpha}_1}$ und $MSE_{\tilde{\alpha}_2}$.

Aufgrund der im theoretischen Teil erläuterten Normierung von β_1 (vgl. S. 39) muss

angemerkt werden, dass der MSE von β_1 nicht zur Beurteilung der Schätzgüte herangezogen werden kann, da er einen festgelegten Wert beschreibt. Daraus folgt, dass die Bildung eines Gesamt-MSE für den Kointegrationsvektor analog zur Bildung des Gesamt-MSE für den Anpassungskoeffizienten ($MSE_{\alpha_{ges}}$) *nicht* vorgenommen wird. Es leuchtet unmittelbar ein, dass der MSE von β_2 durch die Normierung von $\beta_1 = 1$ nicht unbeeinflusst bleibt und deshalb eigentlich nicht ausschließlich eine Aussage zur Güte der Schätzung von β_2 macht, sondern - ähnlich wie der $MSE_{\alpha_{ges}}$ - ein globales Qualitätsurteil über die Güte der Schätzung des Kointegrationsvektors im Allgemeinen abgibt. Deshalb sind keine Einzelvergleiche, etwa zwischen $MSE_{\hat{\beta}_2}$ und $MSE_{\hat{\alpha}_2}$, möglich, sondern *der Vergleich der Schätzgüte von Kointegrationsvektor und Anpassungskoeffizient kann nur durch den Gesamt-Vergleich von $MSE_{\hat{\beta}_2}$ und $MSE_{\alpha_{ges}}$ erfolgen.*

Die aus den Tabellen entnommenen und in der nachfolgenden Ergebnisdarstellung interpretierten Simulationsergebnisse sind zusätzlich in Abbildungen veranschaulicht. Bei der Betrachtung der Schaubilder zu den Testverfahren (Dickey-Fuller- und Johansen-Kointegrationstest; vgl. Abschnitte 5.1.1 und 5.1.2, S. 93ff.) sollte beachtet werden, dass aus Gründen der Übersichtlichkeit nicht alle in den Simulationen durchgeführten Messzeitpunktevariationen (von $T = 50$ bis $T = 350$ (Dickey-Fuller-Test) bzw. $T = 500$ (Johansen-Test) in 50er-Schritten) in die Graphiken aufgenommen wurden. Die in den Graphiken dargestellte höchste Anzahl der Messzeitpunkte beschreibt das T , ab dem in 100% aller Fälle vom jeweiligen Testverfahren eine stationäre Linearkombination (Dickey-Fuller-Test) bzw. eine signifikante Kointegrationsbeziehung (Johansen-Test) indiziert wird. Bei der graphischen Veranschaulichung der Ergebnisse zu den Parameterschätzungen von Anpassungskoeffizient und Kointegrationsvektor (vgl. Abschnitt 5.3, S. 107ff.) wird nur der Ausschnitt der insgesamt simulierten Messzeitpunktevariationen (von $T = 50$ bis $T = 500$ in 50er-Schritten) mit in die Abbildungen aufgenommen, in dem sich der mittlere quadratische Fehler (MSE) über die T -Variationen hinweg wesentlich und sichtbar unterscheidet. Das T , ab dem sich die MSE-Ergebnisse überlagern, ist in den Graphiken jeweils das höchste T . In diesem Sinne kann das in den Schaubildern aufgeführte höchste T als „Schwel-

lenwert“ aufgefasst werden. Beim Vergleich der Schaubilder untereinander sei das unterschiedliche Skalenspektrum beachtet, da es sich in vielen Fällen nur aufgrund einzelner Ausreißerwerte vergrößert und bei Nichtbeachtung dieses Sachverhalts ein verfälschter Eindruck von den Ergebnissen entstehen kann. Die in den Abbildungen auftretende Abkürzung $[Abh_{glob}]$ steht für das in den Tabellen mit $[Abh]$ notierte Maß für die Abhängigkeit der Einzelreihen.

5.1 Testverfahren: Ergebnisse

5.1.1 Dickey-Fuller-Kointegrationstest

Bei der Analyse der Tabellen 5.1 und 5.2, S. 94f., in denen die Ergebnisse zum Dickey-Fuller-Kointegrationstest zusammengefasst sind, lassen sich folgende Muster in Abhängigkeit von der Anzahl der Messzeitpunkte und der Abhängigkeit der Komponentenprozesse erkennen. Die Diagnose des Tests bezüglich der Instationarität der Einzelreihen soll dabei im Rahmen dieser Arbeit außer Acht gelassen werden, da die Instationarität einerseits in den meisten Fällen ohnehin gegeben ist und andererseits im Fall ihres Nicht-Vorliegens nicht klar bestimmt werden kann, ob die dann diagnostizierte Stationarität artifiziell und ein durch die Höhe des Kointegrationsvektors hergestelltes Artefakt ist (vgl. die Darstellungen zum Einfluss des Kointegrationsvektors unter 3.4.1, S. 57ff. und zur Kontroverse über die Stationarität einer der Einzelreihen eines kointegrierten Systems auf S. 42f.).

In Tabelle 5.1, S. 94, in der u.a. die Ergebnisse für die Variation minimaler Abhängigkeiten der Komponentenreihen dargestellt sind, erklärt sich von selbst, dass beim erstgenannten Fall mit zwei unabhängigen Einzelreihen (Nebendiagonalwerte von Null) keine Stationaritätsaussage über die Linearkombination getroffen werden kann, da aufgrund der Unabhängigkeit der Komponentenprozesse keine Kointegration gegeben ist und folglich auch keine durch den Kointegrationsvektor gewichtete Linearkombination existiert. Der Dickey-Fuller-Test gibt in SAS unter Angabe einer Fehlermeldung keine Stationaritätsaussage aus und ist damit in gewisser Hinsicht in der Lage, Nicht-Kointegration anzuzeigen, die dann gegeben ist, wenn die instationären Einzelprozesse unabhängig sind (\rightarrow none). Wenn man die Tabelle in der

Tabelle 5.1: Dickey-Fuller-Kointegrationstest für VAR(1)-Prozesse mit **nicht-auseinanderdriftenden** Zeitpfaden der Komponentenprozesse.

nicht-auseinanderdriftend Dickey-Fuller-Kointegrationstest <i>Variation minimaler Abhängigkeiten</i> für Hypothese 0 $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} > 0$ $\phi_{11} + \phi_{12} = 1; \phi_{21} + \phi_{22} = 1$ ϕ_{21} und ϕ_{22} konstant						
Modellmatrix			Linearkombination			
ϕ_{11}	ϕ_{12}	T	Abh	Sign. n; s	Häufg. x/200	Proz. %
1	0	50	0	none	none	none
0	1	100		none	none	none
		150		none	none	none
		200		none	none	none
0,995	0,005	50	0,105	n	134	67
0,1	0,9	100		s	66	33
		150		n	35	17,5
		200		s	165	82,5
		150		n	5	2,5
		200		s	195	97,5
0,975	0,025	50	0,125	n	118	59
0,1	0,9	100		s	82	41
		150		n	15	7,5
		200		s	185	92,5
		150		n	2	1
		200		s	198	99
		200		s	200	100

nicht-auseinanderdriftend Dickey-Fuller-Kointegrationstest <i>alle Hypothesen</i> $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} > 0$ $\phi_{11} + \phi_{12} = 1; \phi_{21} + \phi_{22} = 1$ ϕ_{21} und ϕ_{22} konstant						
Modellmatrix			Linearkombination			
ϕ_{11}	ϕ_{12}	T	Abh	Sign. n; s	Häufg. x/200	Proz. %
0,95	0,05	50	0,15	n	92	46
0,1	0,9	100		s	108	54
		150		n	5	2,5
		200		s	195	97,5
		150		s	200	100
0,9	0,1	50	0,2	n	47	23,5
0,1	0,9	100		s	153	76,5
		150		s	200	100
		200		s	200	100
0,65	0,35	50	0,45	n	1	0,5
0,1	0,9	100		s	199	99,5
		150		s	200	100
		200		s	200	100
0,6	0,4	50	0,5	s	200	100
0,1	0,9	100		s	200	100
		150		s	200	100

linken Spalte (ebenfalls Tab. 5.1, S. 94) zu **nicht-auseinanderdriftenden Zeitpfaden** weiter betrachtet, kann man feststellen, dass sich in Abhängigkeit von T und den Abhängigkeiten der Einzelreihen ein Muster abzeichnet: im Abhängigkeitspektrum von 0,105 bis 0,125 wird ab $T = 200$ in 100% der Fälle eine stationäre Linearkombination diagnostiziert. Bei einer Abhängigkeit von 0,15 (rechte Spalte der Tabelle) wird bereits bei $T = 150$ in 100% der Fälle Stationarität angezeigt, was sich weiterhin steigert. Im Abhängigkeitsspektrum von 0,2 bis 0,45 wird bereits ab

$T = 100$ und bei allen höheren Abhängigkeiten wird bereits bei $T = 50$ in 100% der Fälle Stationarität bescheinigt.

Tabelle 5.2: Dickey-Fuller-Kointegrationstest für VAR(1)-Prozesse mit **auseinanderdriftenden** Zeitpfaden der Komponentenprozesse.

auseinanderdriftend Dickey-Fuller-Kointegrationstest $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} \in \mathfrak{R}$ $\phi_{11} + \phi_{12} < 1$; $\phi_{21} + \phi_{22} < 1$ ϕ_{21} und ϕ_{22} konstant							auseinanderdriftend Dickey-Fuller-Kointegrationstest $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} > 0$ $\phi_{11} + \phi_{12} > 1$; $\phi_{21} + \phi_{22} < 1$ ϕ_{21} und ϕ_{22} konstant						
Modellmatrix				Linearkombination			Modellmatrix				Linearkombination		
ϕ_{11}	ϕ_{12}	T	Abh	Sign. n; s	Häufg. x/200	Proz. %	ϕ_{11}	ϕ_{12}	T	Abh	Sign. n; s	Häufg. x/200	Proz. %
ϕ_{21}	ϕ_{22}						ϕ_{21}	ϕ_{22}					
1,01	-0,05	50	0,15	s	200	100	0,99	0,05	50	0,15	s	200	100
0,1	0,5						0,1	0,5					
		100		s	200	100			100		s	200	100
1,02	-0,1	50	0,2	s	200	100	0,98	0,1	50	0,2	s	200	100
0,1	0,5						0,1	0,5					
		100		s	200	100			100		s	200	100
1,07	-0,35	50	0,45	s	200	100	0,93	0,35	50	0,45	s	200	100
0,1	0,5						0,1	0,5					
		100		s	200	100			100		s	200	100
1,08	-0,4	50	0,5	n	1	0,5	0,92	0,4	50	0,5	s	200	100
0,1	0,5			s	199	99,5	0,1	0,5					
		100		s	200	100			100		s	200	100

Bei **auseinanderdriftenden Zeitpfaden**, die in Tabelle 5.2, S. 95, dargestellt sind, verhält sich der Test anders: in der Regel wird immer schon ab $T = 50$, unabhängig von der Abhängigkeit der Komponentenprozesse, von einer stationären Linearkombination ausgegangen. Nur in sehr wenigen Ausnahmen wird ab einer höheren Anzahl der Messzeitpunkte Stationarität angezeigt, dies erfolgt aber immer spätestens ab $T = 100$. Dies gilt sowohl für die auseinanderdriftenden Modelltypen mit der Summe der Elemente in der ersten Zeile der Modellmatrix kleiner (links) als auch größer (rechts) 1.

Die Ergebnisse können neben den benannten Tabellen auch anhand der Abbildung 5.1, S. 96, verifiziert werden. Die Abbildung zeigt ebenfalls, dass ab $T = 200$ bereits bei einer Abhängigkeit von 0,105 in 100% der Fälle Kointegration angezeigt wird (Bild links oben). Außerdem ist zu erkennen, dass ab einer Abhängigkeit von 0,15

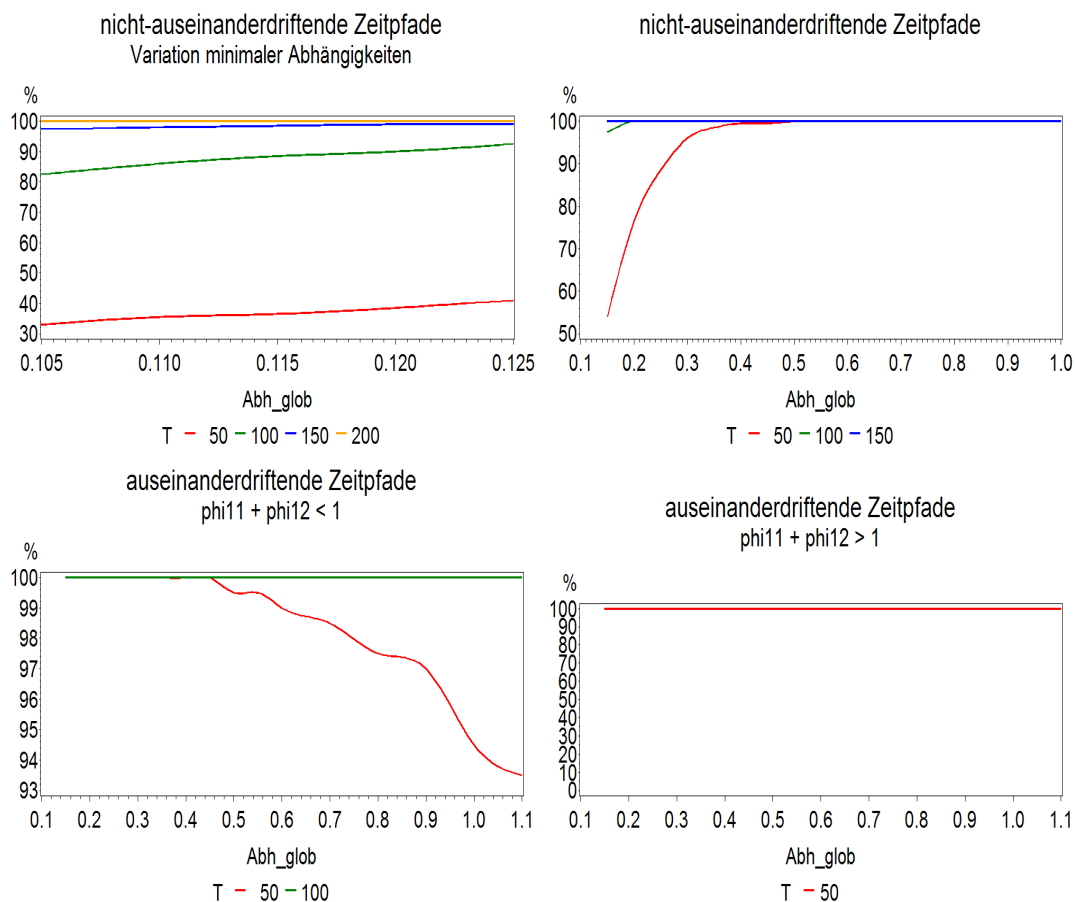


Abbildung 5.1: Ergebnisse des **Dickey-Fuller-Kointegrationstests**: prozentualer Anteil stationärer Linearkombinationen.

immer weniger Messzeitpunkte notwendig sind, um in 100% der Fälle Kointegration zu diagnostizieren (Bild rechts oben). Gleichzeitig ist zu sehen, dass die Teststärke bei auseinanderdriftenden Zeitpfaden der Einzelreihen deutlich höher ist: abgesehen von einer leicht erkennbaren Instabilität bei $T = 50$ (Bild links unten), wird unabhängig vom Grad der Abhängigkeit in 100% der Fälle Kointegration indiziert (auch Bild rechts unten).

5.1.2 Johansen-Kointegrationstest

Auch bei der Betrachtung der Tabellen 5.3 und 5.4, S. 97ff., die die Ergebnisse des *Cointegration Rank Tests* von Johansen abbilden, sind die nachfolgend beschriebenen Muster beobachtbar, wiederum getrennt für Modelle mit auseinanderdriftenden und nicht-auseinanderdriftenden Zeitpfaden. Tabelle 5.3, S. 97, zeigt an erster Stelle

Tabelle 5.3: Johansen-Testung auf eine Kointegrationsbeziehung ($H_0 : r = 0$ vs. $H_1 : r > 0$) der VAR(1)-Prozesse mit **nicht-auseinanderdriftenden** Zeitpfaden der Komponentenprozesse.
 ⇒ Fortsetzung auf der folgenden Seite ⇒

nicht-auseinanderdriftend Johansen-Testverfahren <i>Variation minimaler Abhängigkeiten für Hypothese 0</i> $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} > 0$ $\phi_{11} + \phi_{12} = 1; \phi_{21} + \phi_{22} = 1$ ϕ_{21} und ϕ_{22} konstant						
Modell- matrix		T	Abh	Sign. n; s	Häufg. x/200	Proz. Anteil %
ϕ_{11} ϕ_{21}	ϕ_{12} ϕ_{22}					
1 0	0 1	50	0	n s	189 11	94,5 5,5
		100		n s	189 11	94,5 5,5
		150		n s	191 9	95,5 4,5
		200		n s	193 7	96,5 3,5
		250		n s	192 8	96 4
0,995 0,1	0,005 0,9	50	0,105	n s	152 48	76 24
		100		n s	83 117	41,5 58,5
		150		n s	19 181	9,5 90,5
		200		n s	1 199	0,5 99,5
		250		s	200	100

nicht-auseinanderdriftend Johansen-Testverfahren <i>alle Hypothesen</i> $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} > 0$ $\phi_{11} + \phi_{12} = 1; \phi_{21} + \phi_{22} = 1$ ϕ_{21} und ϕ_{22} konstant						
Modell- matrix		T	Abh	Sign. n; s	Häufg. x/200	Proz. Anteil %
ϕ_{11} ϕ_{21}	ϕ_{12} ϕ_{22}					
0,95 0,1	0,05 0,9	50	0,15	n s	169 31	84,5 15,5
		100		n s	99 101	49,5 50,5
		150		n s	22 178	11 89
		200		n s	2 198	1 99
		250		s	200	100
0,9 0,1	0,1 0,9	50	0,2	n s	149 51	74,5 25,5
		100		n s	58 142	29 71
		150		n s	6 194	3 97
		200		s	200	100
		250		s	200	100
0,85 0,1	0,15 0,9	50	0,25	n s	128 72	64 36
		100 ¹		n s	9 190	4,5 95,5
		150		s	200	100
		200		s	200	100
		250		s	200	100

die Ergebnisse für die VAR(1)-Matrix mit zwei unabhängigen Komponentenprozessen (beide Nebendiagonalwerte von Null). Hier zeigt der Johansen-Test in maximal

Fortsetzung Tabelle 5.3, S. 97						
Modellmatrix		T	Abh	n; s	Häufg. x/200	Proz. Anteil %
ϕ_{11}	ϕ_{12}					
ϕ_{21}	ϕ_{22}					
0,8 0,1	0,2 0,9	50	0,3	n	93	46,5
				s	107	53,5
		100		s	200	100
		150		s	200	100
		200		s	200	100
		250	s	200	100	
0,55 0,1	0,45 0,9	50	0,55	n	1	0,5
				s	199	99,5
		100		s	200	100
		150		s	200	100
		200		s	200	100
		250	s	200	100	
0,5 0,1	0,5 0,9	50	0,6	s	200	100
		100		s	200	100
		150		s	200	100
		200		s	200	100
		250		s	200	100

5,5% (bei $T = 50$ und $T = 100$) und damit in einem verschwindend geringen Anteil der Fälle Kointegration an und ist so offensichtlich in der Lage, den Fall zu erkennen, in dem keine Kointegration vorliegt, nämlich bei exakter Nullabhängigkeit der Komponenten. Sobald die graphisch **nicht-auseinanderdriftenden Komponenten** aber minimal abhängig sind (0,105), lässt sich das Verhalten folgendermaßen charakterisieren: im Abhängigkeitsspektrum von 0,105 bis 0,15 (einschließlich der Betrachtung der rechten Spalte) wird ab $T = 250$ in 100% der Fälle Kointegration angezeigt, für die Abhängigkeit von 0,2 geschieht dies bereits ab $T = 200$, für 0,25 ab $T = 150$ und im Abhängigkeitsbereich von 0,3 bis 0,55 ist dies ab $T = 100$ beobachtbar.

Ab der Abhängigkeit von 0,6 stabilisiert sich dieses Verhalten bei $T = 50$.

Bei VAR(1)-Prozessen mit **auseinanderdriftenden Zeitpfaden der Komponentenprozesse** ist wie auch schon beim Dickey-Fuller-Kointegrationstest ein anderes Verhalten zu beobachten. In Tabelle 5.4, S. 99, wird für Modelle, bei denen die Elemente der ersten Zeile der Modellmatrix in der Summe kleiner als 1 sind (linke Spalte der Tabelle), im Abhängigkeitsbereich von 0,15 bis 0,35 ab $T = 100$ in 100% der Fälle Kointegration indiziert.

Dies stabilisiert sich aber schon sehr schnell ab einer Abhängigkeit von 0,4 auf $T = 50$. Falls die Elemente der ersten Zeile der Modellmatrix in der Summe größer als 1 sind, sieht es ähnlich aus (rechte Spalte der Tabelle). Der Test stabilisiert sich bei einer Abhängigkeit von 0,55 bei $T = 50$ auf 100% der Fälle.

Tabelle 5.4: Johansen-Testung auf eine Kointegrationsbeziehung ($H_0 : r = 0$ vs. $H_1 : r > 0$) der VAR(1)-Prozesse mit **auseinanderdriftenden** Zeitpfaden der Komponentenprozesse.

auseinanderdriftend Johansen-Testverfahren $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} \in \mathbb{R}$ $\phi_{11} + \phi_{12} < 1; \phi_{21} + \phi_{22} < 1$ ϕ_{21} und ϕ_{22} konstant						
Modellmatrix		T	Abh.	Sign.	Häufg.	Proz. Anteil
ϕ_{11}	ϕ_{12}			n, s	x/200	%
ϕ_{21}	ϕ_{22}					
1,01	-0,05	50	0,15	n	14	7
0,1	0,5			s	186	93
		100		s	200	100
1,05	-0,25	50	0,35	n	2	1
0,1	0,5			s	198	99
		100		s	200	100
1,06	-0,3	50	0,4	s	200	100
0,1	0,5					
		100		s	200	100

auseinanderdriftend Johansen-Testverfahren $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} > 0$ $\phi_{11} + \phi_{12} > 1; \phi_{21} + \phi_{22} < 1$ ϕ_{21} und ϕ_{22} konstant						
Modellmatrix		T	Abh.	Sign.	Häufg.	Proz. Anteil
ϕ_{11}	ϕ_{12}			n; s	x/200	%
ϕ_{21}	ϕ_{22}					
0,99	0,05	50	0,15	n	22	11
0,1	0,5			s	178	89
		100		s	200	100
0,95	0,25	50	0,35	n	12	6
0,1	0,5			s	188	94
		100		s	200	100
0,91	0,45	50	0,55	s	200	100
0,1	0,5					
		100		s	200	100

Die Ergebnisse sind zusätzlich zu den beschriebenen Tabellen in Abbildung 5.2, S. 100, veranschaulicht. Die Abbildung zeigt, dass der Johansen-Test bei variierten Abhängigkeiten im minimalen Spektrum eine größere Anzahl an Messzeitpunkten benötigt, um in 100 % der Fälle Kointegration anzuzeigen (Bild links oben) als bei höheren Abhängigkeiten (ab 0,15; Bild rechts oben). Im Gegensatz dazu zeigt die Abbildung, dass der Johansen-Test für auseinanderdriftende Zeitpfade der Einzelreihen abgesehen von leichten Schwankungen bei $T = 50$ viel häufiger Kointegration anzeigt (Bilder unten links und rechts).

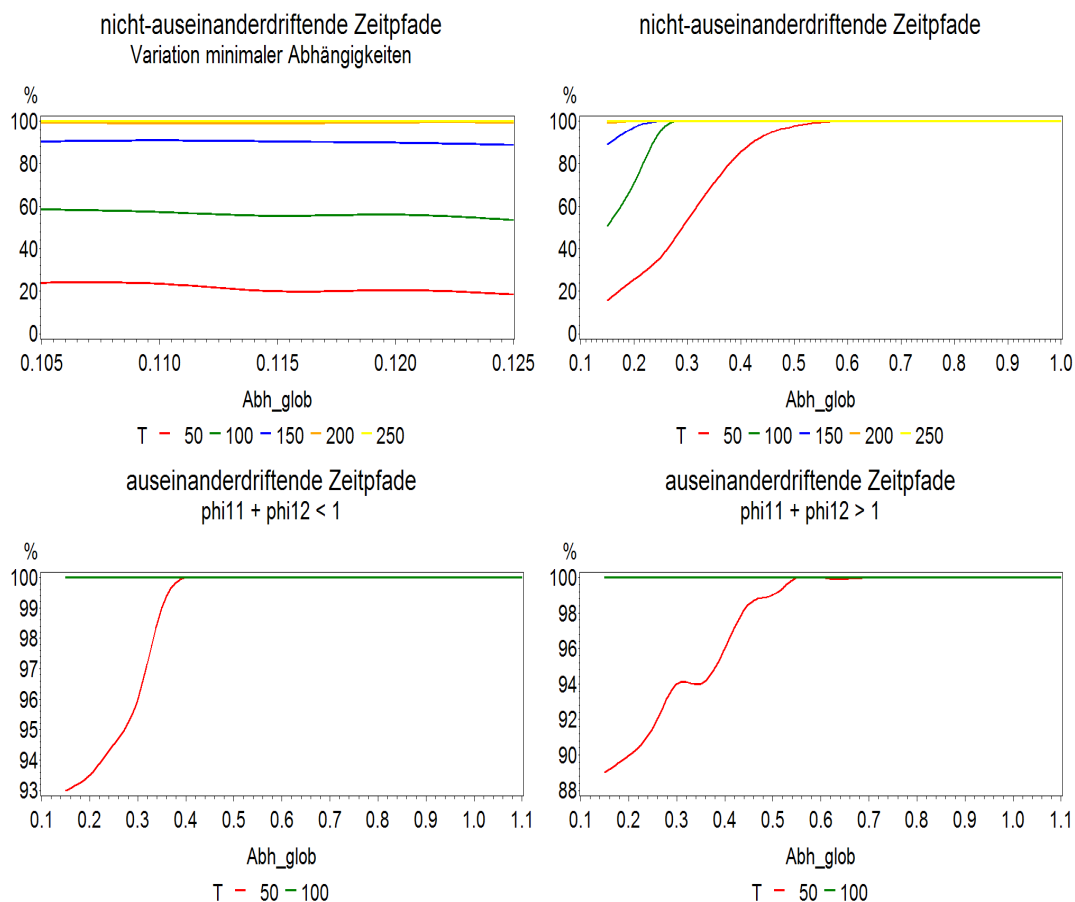


Abbildung 5.2: Ergebnisse des **Johansen-Kointegrationstests**: prozentualer Anteil einer signifikanten Kointegrationsbeziehung.

5.2 Überprüfung der Hypothesen zu den Testverfahren

Hypothese 0 (S. 78), die postuliert, dass in Abhängigkeit von der Anzahl der Messzeitpunkte bereits bei minimalster Abhängigkeit der Komponentenprozesse Kointegration vorliegt, kann als bestätigt gelten, zumindest wenn man die Operationalisierungen von Kointegration durch die beiden verwendeten Testverfahren, den Dickey-Fuller- und den Johansen-Kointegrationstest, betrachtet. Dabei ist anzumerken, dass beide Verfahren bereits bei minimalster Abhängigkeit der Komponentenprozesse Kointegration anzeigen, aber dennoch in der Lage sind, den Fall unabhängiger Komponentenprozesse als Nicht-Kointegration zu erkennen. Aber bereits bei einer Abhängigkeit von 0,105 wird ab $T = 200$ (Dickey-Fuller-Test) bzw. $T = 250$ (Johansen-Test) in 100% der Fälle Kointegration angezeigt. Dieser Befund kann

deswegen als relevant für das Verständnis des Kointegrationskonzepts betrachtet werden, weil in den meisten der im theoretischen Teil beschriebenen „klassischen“ Quellen zum Kointegrationskonzept der Eindruck erweckt wird, dass Kointegration ein ausgesprochen besonderes und nicht sehr häufiges Phänomen sei. Es wäre gesondert zu klären, warum die Autoren dies vermuten. Möglicherweise gehen sie implizit davon aus, dass die meisten Prozesse voneinander unabhängig sind. Dies dürfte aber für Prozessverläufe genauso selten sein wie für querschnittliche Betrachtungen. Querschnittsanalysen haben nämlich gezeigt, dass die meisten Variablen abhängig sind ohne allerdings nachweisen zu können, wie hoch genau diese Abhängigkeit ist. Da allerdings über die Häufigkeit der Abhängigkeit von $I(1)$ -Prozessen wenig bekannt ist, wäre es vorschnell, aus der Bestätigung der Hypothese 0 den Schluss zu ziehen, dass die meisten $I(1)$ -Prozesse *nicht* unkorreliert sind, auch wenn die Ergebnisse vermuten lassen, dass die Wahrscheinlichkeit, kointegrierte Modelle bilden zu können, hoch ist. Außerdem mag der entstehende Eindruck der Seltenheit des Vorliegens von Kointegrationsbeziehungen darin begründet liegen, dass die Verfahren der Kointegrationsmethodologie in vielen inhaltlichen Bereichen noch nicht „erprobt“ sind und deshalb in manchen Forschungsbereichen noch nicht ausreichend Erfahrungswissen, d.h. empirisches Wissen, über die Häufigkeit des Vorliegens von Kointegration existiert. Insgesamt kann die Bestätigung dieser Hypothese durch die Variation minimaler Abhängigkeiten der Komponentenprozesse bei nicht-auseinanderdriftenden Zeitpfaden der Einzelreihen zu weiteren noch systematischeren Untersuchungen dieses Sachverhalts auch für Modelle mit auseinanderdriftenden Zeitpfaden der Komponentenprozesse anregen, um dem psychologischen Forscher noch fundierteres Wissen zur Verfügung zu stellen, ob Kointegration ein eher seltenes oder eher häufiges Phänomen ist.

Dieser Sachverhalt wird sicherlich durch die Tatsache gefördert, dass Kointegration bahnbrechende Neuerungen für die Forschung mit sich gebracht hat und diese Besonderheit fälschlicherweise intuitiv mit Seltenheit des Auftretens gleichgesetzt wird. Die Überprüfung von Hypothese 0 zeigt aber, dass bereits bei minimalster Abhängigkeit der Einzelreihen von einer über die hier verwendeten Testverfahren

operationalisierten Kointegrationsbeziehung (für VAR(1)-Modelle) auszugehen ist.

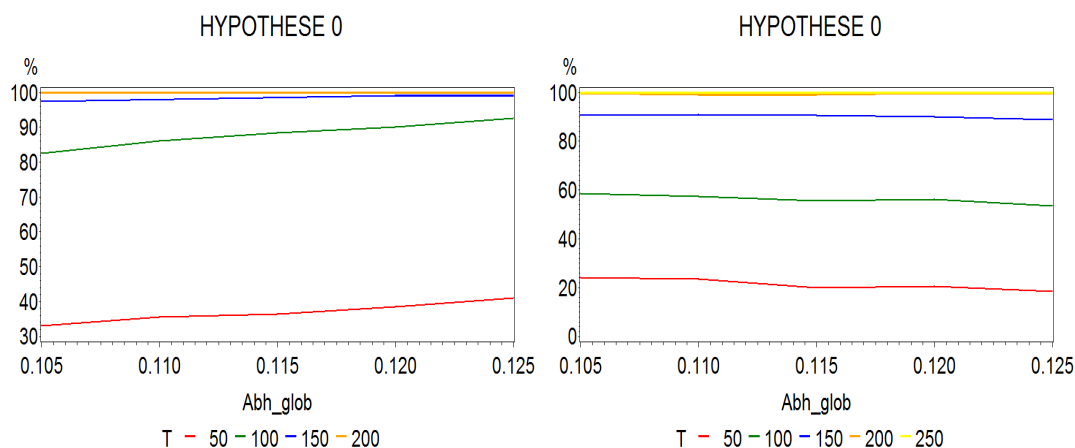


Abbildung 5.3: HYPOTHESE 0 (Variation minimaler Abhängigkeiten für nicht-auseinanderdriftende Zeitpfade) mittels Dickey-Fuller-Kointegrationstest (links: prozentualer Anteil stationärer Linearkombinationen) bzw. Johansen-Kointegrationstest (rechts: prozentualer Anteil einer signifikanten Kointegrationsbeziehung).

Die Befunde sind in Abbildung 5.3, S. 102, illustriert. Die Abbildung zeigt nochmals das beschriebene Verhalten der beiden Testverfahren - DF-Test und Johansen-Test - bei variierten Abhängigkeiten im Spektrum von 0,105 bis 0,125. Dabei werden v.a. die Unterschiede zwischen den beiden Verfahren deutlich: der Dickey-Fuller-Test (linkes Bild) weist insofern eine höhere Teststärke auf, als er auch im minimalen Abhängigkeitsspektrum häufiger Kointegration anzeigt als der Johansen-Kointegrationstest (rechtes Bild).

Hypothese 1a (S. 78) ist eindeutig bestätigt. Damit ist nachgewiesen, dass das nur in wenigen Quellen und in jüngster Zeit beschriebene Phänomen von kointegrierten Einzelreihen trotz auseinanderdriftender Zeitpfade dem Kointegrationskonzept deshalb zuzuordnen ist, weil es die beschriebenen notwendigen und von den meisten Autoren geteilten Bedingungen für Kointegration erfüllt (vgl. zusammenfassende Tabelle auf S. 39): die Gleichgewichtsbeziehung zwischen zwei instationären Reihen, deren Zeitpfade auseinanderdriften, ist durch eine stationäre Linearkombination charakterisierbar, was das Vorliegen von Kointegration impliziert. In allen Fällen (bis auf den Fall der exakten Nullabhängigkeit der Komponentenprozesse; s. Erläuterungen zu Hypothese 0) weisen sowohl der Dickey-Fuller-Test als auch der

Johansen-Test auf die Kointegration von Einzelreihen mit auseinanderdriftenden³⁰ Zeitpfaden hin. Bei der Besprechung der nachfolgenden Hypothesen wird sich zeigen, dass die Kategorisierung in „nicht-auseinanderdriftende Zeitpfade“ ($\beta_2 = \pm 1$) und „auseinanderdriftende Zeitpfade“ ($\beta_2 \neq \pm 1$), die im Rahmen dieses empirischen Teils vorgenommen wird, vonnöten ist, da sich sowohl beide Testverfahren als auch die Güte der Parameterschätzungen von Kointegrationsvektor und Anpassungskoeffizient für auseinanderdriftende und nicht-auseinanderdriftende Zeitpfade in ihrem Verhalten unterscheiden. Für den Anwendungsforscher impliziert dies, dass er aus der Betrachtung des Kointegrationsvektors Rückschlüsse auf den Verlauf der Zeitpfade ziehen kann und daraus im nächsten Schritt Implikationen für die Güte der Test- und Schätzverfahren ableitbar sind. Infolgedessen wird er in der Lage sein, eine angemessene Entscheidung über die Anzahl der Messzeitpunkte zu treffen, zumal durch die Diplomarbeit ein relativ breites Spektrum an Modelloptionen abgedeckt ist, die als Orientierungshilfen dienen können.

Da davon auszugehen ist, dass in der Forschungsrealität in den meisten Fällen β_2 nicht exakt -1 bzw. $+1$ ist, erhöht der durch die Bestätigung der Hypothese mögliche Einbezug der auseinanderdriftenden Fälle zusätzlich die in der Einleitung (S. 10ff.) angesprochene Abbildung eines breiteren Realitätsspektrums durch kointegrierte Modelle: sie sind in der Lage, die Realität psychologischer Phänomene möglichst genau abzubilden (instationär, multivariat *und* auseinanderdriftende Zeitpfade).

Eine Veranschaulichung dieser Ergebnisse findet sich in Abbildung 5.4, S. 104. Absichtlich sind hier nur die Messzeitpunktevariationen (T) aus dem niedrigeren Bereich dargestellt, da ihre Betrachtung für einen Großteil der psychologischen Forschungspraxis höhere Relevanz besitzt. Es zeigt sich deutlich, dass VAR(1)-Prozesse

³⁰Im Rahmen der Arbeit wurden zwei unterschiedliche Typen von kointegrierten Modellen mit auseinanderdriftenden Zeitpfaden simuliert, nämlich Modelle mit $\phi_{11} + \phi_{12} < 1$ und $\phi_{11} + \phi_{12} > 1$. Wie die Ergebnisdarstellung zu Test- und Schätzverfahren zeigt, unterscheiden sich die Verfahren in Bezug auf die beiden Modelltypen kaum. Aus diesem Grund bezieht sich die Hypothesenüberprüfung - falls nicht anders gekennzeichnet - immer auf den erstgenannten Typ. Mögliche feinere Unterschiede zwischen diesen beiden Typen sind in einer Nachfolgeuntersuchung zu dieser Diplomarbeit zu klären.

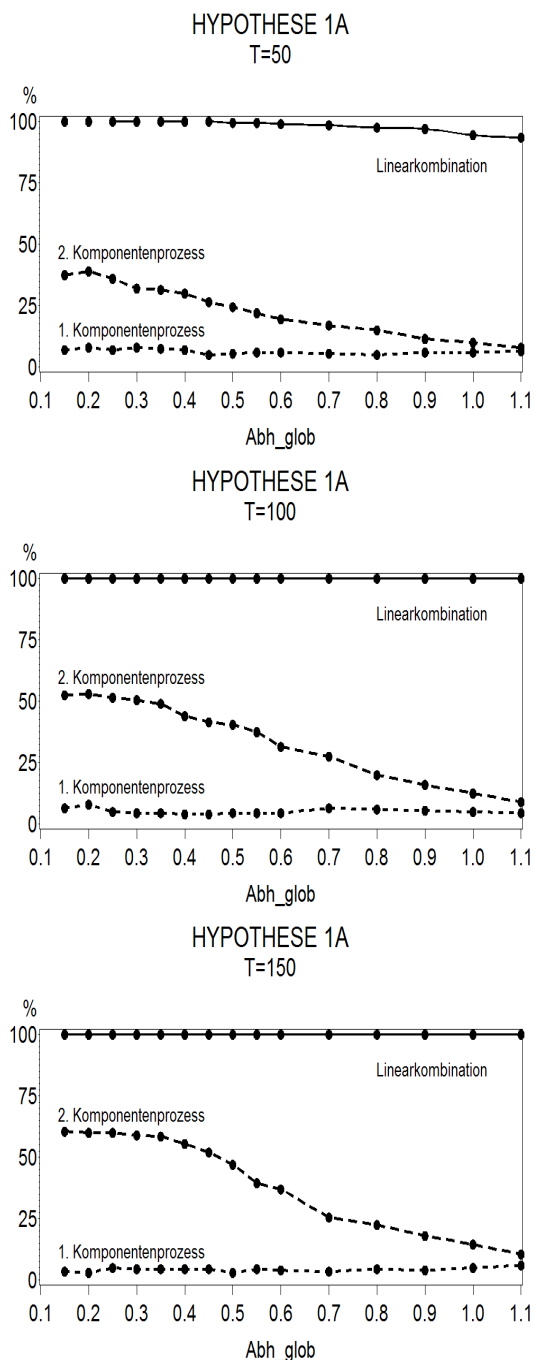


Abbildung 5.4: HYPOTHESE 1A (prozentualer Anteil stationärer Komponentenprozesse und stationärer Linearkombinationen) mittels Dickey-Fuller-Kointegrationstest für auseinanderdriftende Zeitpfade (Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$).

mit graphisch auseinanderdriftenden Zeitpfaden umso häufiger kointegriert sind, je höher sowohl die Anzahl der Messzeitpunkte (variiert von oben nach unten) als auch der Grad der Abhängigkeit der Komponentenprozesse ist: der Dickey-Fuller-Test zeigt zunehmend in 100% der Fälle stationäre Linearkombinationen an, während gleichzeitig in immer mehr Fällen instationäre Komponentenprozesse diagnostiziert werden und somit der Anteil fälschlicherweise als stationär deklarerter Komponentenprozesse sinkt.

Hypothese 1b (S. 79), die angibt, dass das Johansen-Testverfahren bei nicht-auseinanderdriftenden Zeitpfaden bei identischer Anzahl der Messzeitpunkte und vergleichbarer Abhängigkeit der Komponentenprozesse erst ab einem höheren Grad der Abhängigkeit der Komponentenprozesse („langsamer“) Kointegration indiziert als bei auseinanderdriftenden Einzelprozessen, ist ebenfalls bestätigt. Ihre Validierung untermauert die Wichtigkeit der Unterscheidung von Prozessen mit auseinanderdriftenden und nicht-auseinanderdriftenden Zeitpfaden.

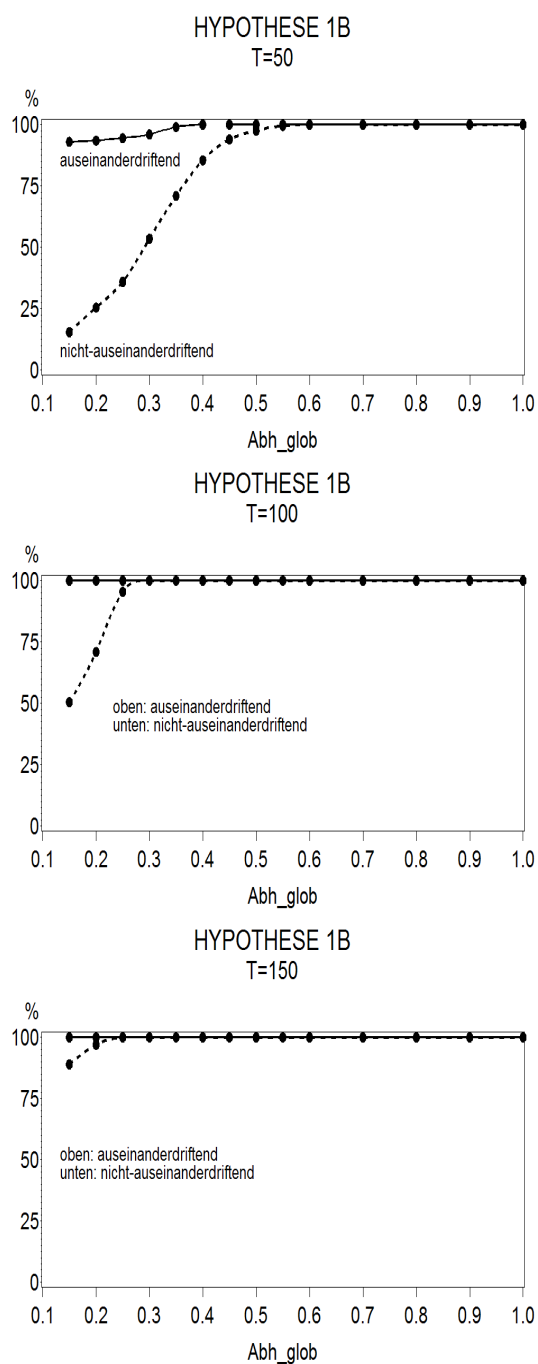


Abbildung 5.5: HYPOTHESE 1B: Vergleich des Verhaltens des Johansen-Testverfahrens (prozentualer Anteil einer signifikanten Kointegrationsbeziehung) für Modelle mit nicht-auseinanderdriftenden vs. auseinanderdriftenden Zeitpfaden (Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$).

Man kann nämlich sehen, dass nicht nur der Johansen-Kointegrationstest, sondern auch der Dickey-Fuller-Test, der an dieser Stelle aber nicht genauer betrachtet werden soll, bei den beiden Verlaufstypen der Zeitpfade der Komponentenprozesse ein unterschiedliches Verhalten zeigen. Zunächst mag es so anmuten, als würde sich der Johansen-Test bei auseinanderdriftenden Zeitpfaden der Einzelreihen nicht sehr differenziert verhalten, indem er bereits bei einer Abhängigkeit von 0,4 schon bei $T = 50$ Messzeitpunkten in 100% der Fälle Kointegration anzeigt. Aus statistischer Perspektive ist dies allerdings keinesfalls negativ zu bewerten: wenn keine falsch positiven Ergebnisse resultieren, ist der Test nämlich hervorragend, d.h. er weist eine hohe Teststärke auf. Bei nicht-auseinanderdriftenden Zeitpfaden diagnostiziert der Johansen-Test bei der gleichen geringen Anzahl der Messzeitpunkte von 50 sehr viel „später“ das Vorliegen einer Kointegrationsbeziehung. Erst ab einer Abhängigkeit von 0,6 wird in 100% der Fälle Kointegration angezeigt.

Abbildung 5.5, S. 105, weist ebenso dieses Verhalten des Johansen-Kointegrationstests graphisch nach: es zeigt sich, dass der beschriebene Unterschied für Modelle mit

auseinanderdriftenden vs. nicht-auseinanderdriftenden Zeitpfaden mit zunehmender Anzahl der Messzeitpunkte (von oben nach unten) immer kleiner wird.

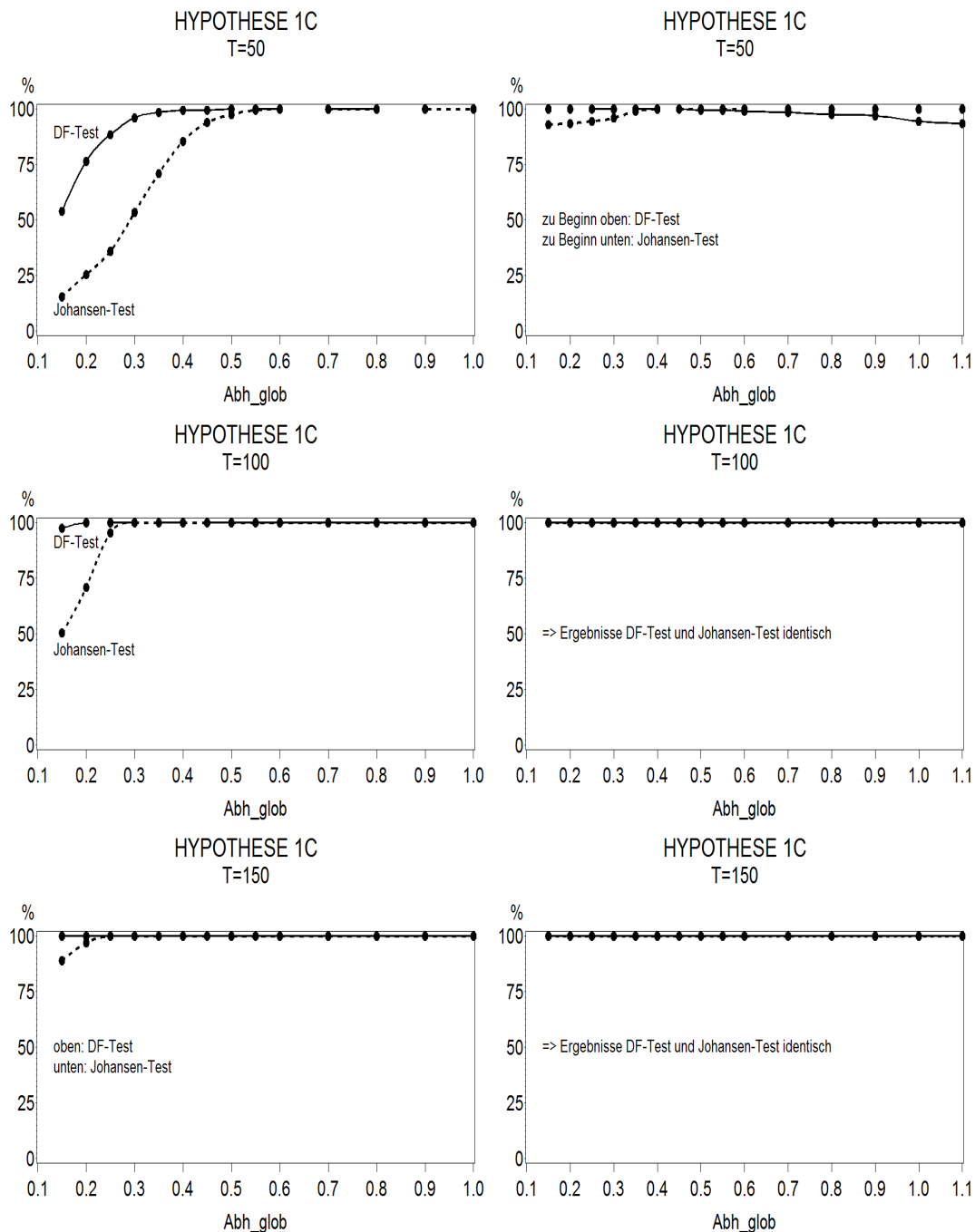


Abbildung 5.6: HYPOTHESE 1C: Vergleich des Verhaltens des Dickey-Fuller- vs. Johansen-Kointegrationstests für Modelle mit *nicht*-auseinanderdriftenden Zeitpfaden (links) und Modelle mit auseinanderdriftenden Zeitpfaden (rechts; Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$); prozentualer Anteil stationärer Linearkombinationen (DF-Test) bzw. einer signifikanten Kointegrationsbeziehung (Johansen-Test).

Hypothese 1c (S. 79) kann als bedingt bestätigt gelten, wenn auch die Unterschiede nicht sehr deutlich sind. Die Hypothese nimmt an, dass der DF-Kointegrationstest und der Johansen-Kointegrationstest sich bei identischer Anzahl der Messzeitpunkte und vergleichbarer Abhängigkeit der Komponentenprozesse bezüglich der diagnostizierten Häufigkeit von kointegrierten Beziehungen unterscheiden. Dieser Unterschied ist allerdings nicht sehr deutlich. Der Johansen-Test zeigt ungefähr dieselbe Tendenz wie der Dickey-Fuller-Kointegrationstest, allerdings immer etwas „später“, um in der Terminologie „Geschwindigkeit“ zu bleiben, nämlich - zumindest was Modelle mit nicht-auseinanderdriftenden Zeitpfaden betrifft - ungefähr um einen Abhängigkeitsunterschied von 0,1 oder 0,2 versetzt. Man könnte also schlussfolgern, dass der Dickey-Fuller-Test eine minimal höhere Teststärke aufweist.

Für nicht-auseinanderdriftende (linke Spalte) und auseinanderdriftende (rechte Spalte) Zeitpfade können die Befunde anhand der Abbildung 5.6, S. 106, nachvollzogen werden. Wenn man die linke und die rechte Spalte der Abbildung vergleicht, kann man deutlich sehen, dass die in der Hypothese postulierten Unterschiede bei Modellen mit auseinanderriftenden Zeitpfaden der Einzelreihen geringer ausfallen.

5.3 Güte der Parameterschätzungen von Kointegrationsvektor und Anpassungskoeffizient: Ergebnisse

Im Zentrum der Betrachtung zur Beurteilung der Güte der Parameterschätzungen stehen zunächst die Tabellen 5.5, S. 108f., als repräsentative „Stichprobe“ für multivariate Prozesse mit nicht-auseinanderdriftenden Zeitpfaden und die Tabellen 5.6, S. 110f., und 5.7, S. 112, als Vertretergruppe der VAR(1)-Prozesse mit auseinanderdriftenden Zeitpfaden. Die Tabellen beschreiben die mittleren quadratischen Fehler (MSE) für β_2 , α_1 , α_2 und auch den Gesamt-MSE für α . Wie auf S. 92 durch die Normierung von $\beta_1 = 1$ begründet wurde, sollte nochmals darauf hingewiesen werden, dass der Vergleich der Schätzgüte von Kointegrationsvektor und Anpassungskoeffizient nur durch den Gesamt-Vergleich von $\text{MSE}_{\tilde{\beta}_2}$ und $\text{MSE}_{\alpha_{ges}}$ erfolgen kann.

Insgesamt wird vor allem der Bereich der geringen Messzeitpunkte betrachtet, da die-

Tabelle 5.5: MSEs für α und β ; VAR(1)-Prozesse mit **nicht-auseinanderdriftenden** Zeitpfaden der Komponentenprozesse. \Rightarrow Fortsetzung auf der nächsten Seite \Rightarrow

nicht-auseinanderdriftend $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} > 0$ $\phi_{11} + \phi_{12} = 1; \phi_{21} + \phi_{22} = 1$ ϕ_{21} und ϕ_{22} konstant								
Modellmatrix		Abh	T	MSE	MSE	MSE	MSE	α_{ges}
ϕ_{11}	ϕ_{12}			β_2	α_1	α_2	α_{ges}	
0,95 0,1	0,05 0,9	0,15	50	29,4836	0,0076	0,006	0,0068	
			100	52,7924	0,0026	0,0023	0,0024	
			150	0,0995	0,0015	0,0012	0,0013	
			200	0,0141	0,0009	0,0008	0,0009	
			250	0,0082	0,0006	0,0007	0,0007	
			300	0,0063	0,0005	0,0005	0,0005	
			350	0,0045	0,0004	0,0004	0,0004	
			400	0,0035	0,0004	0,0004	0,0004	
			450	0,0029	0,0003	0,0003	0,0003	
			500	0,0025	0,0003	0,0003	0,0003	
0,75 0,1	0,25 0,9	0,35	50	0,029	0,0099	0,0074	0,0086	
			100	0,0065	0,0033	0,0033	0,0032	
			150	0,0041	0,0023	0,0022	0,0023	
			200	0,0021	0,0015	0,0016	0,0015	
			250	0,0014	0,0011	0,0013	0,0012	
			300	0,0009	0,0009	0,0011	0,001	
			350	0,0006	0,0007	0,0009	0,0008	
			400	0,0006	0,0007	0,0007	0,0007	
			450	0,0004	0,0006	0,0007	0,0006	
			500	0,0004	0,0005	0,0006	0,0006	
0,55 0,1	0,45 0,9	0,55	50	0,0049	0,0107	0,0092	0,0099	
			100	0,0016	0,0041	0,0044	0,0042	
			150	0,0011	0,003	0,0032	0,0031	
			200	0,0006	0,0019	0,0022	0,002	
			250	0,0004	0,0015	0,0018	0,0016	
			300	0,0003	0,0012	0,0016	0,0014	
			350	0,0002	0,001	0,0014	0,0012	
			400	0,0002	0,0009	0,0011	0,001	
			450	0,0002	0,0008	0,001	0,0009	
			500	0,0001	0,0007	0,0009	0,0008	

ser für die psychologische Anwendung Relevanz besitzt und beurteilt werden kann, welche Mindestanzahl von Messzeitpunkten bei welcher Stärke des Zusammenhangs der Komponentenprozesse gegeben sein muss, um von reliablen Schätzungen ausgehen zu können.

In Tabelle 5.5, S. 108, für **nicht-auseinanderdriftende Zeitpfade** wird zunächst der $MSE_{\hat{\beta}_2}$ betrachtet. Bei geringer Abhängigkeit der Komponentenprozesse (0,15) und bei geringer Anzahl der Messzeitpunkte ($T = 50$ bzw. 100) ist der MSE zunächst ausgesprochen hoch (bis 52,7924), nimmt aber dann sehr schnell ab und beträgt bei einer Abhängigkeit von 0,35 und $T = 50$ nur noch 0,029. Diese Tendenz nimmt mit zunehmender Abhängigkeit zu (0,0049 bei $Abh = 0,55$) und liegt bei einer sehr hohen Abhängigkeit von 0,9 bei 0,0013.

Das heißt, dass die Güte der Schätzung von β_2 sich bei einer geringen Anzahl der Messzeitpunkte ($T = 50$) mit zunehmender Abhängigkeit der Komponentenprozesse sehr schnell verbessert. Bei Erhöhung der Messzeitpunkte verstärkt sich dieses

Fortsetzung Tabelle 5.5, S. 108							
ϕ_{11}	ϕ_{12}			MSE	MSE	MSE	MSE
ϕ_{21}	ϕ_{22}	Abh	T	$\tilde{\beta}_2$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_{ges}$
0,2	0,8	0,9	50	0,0013	0,0117	0,0112	0,0114
			100	0,0005	0,0047	0,0058	0,0052
			150	0,0003	0,0034	0,0039	0,0037
			200	0,0002	0,0023	0,0027	0,0025
			250	0,0001	0,0019	0,0022	0,002
			300	0,0001	0,0015	0,002	0,0017
			350	0,0001	0,0012	0,0017	0,0015
			400	0,0001	0,0011	0,0013	0,0012
			450	0	0,001	0,0013	0,0011
			500	0	0,0009	0,0011	0,001

Verhalten selbstredend (MSE von 0 bei Abhängigkeit von 0,9 und $T = 500$). Die Schätzung von α_1 ist schon bei sehr viel geringerer Abhängigkeit der Komponentenprozesse und geringerer Anzahl der Messzeitpunkte ($T = 50$) vergleichsweise ausgezeichnet.

Bei einer Abhängigkeit von 0,15 liegt der MSE bei $T = 50$ bei 0,0076, was einen überdeutlichen Unterschied zu den beschriebenen Ergebnissen von β_2 demonstriert. Allerdings steigt der MSE für α_1 mit zunehmender Abhängigkeit der Komponentenprozesse an. Er liegt bei einer Abhängigkeit von 0,55 und $T = 50$ bei 0,0107 und nimmt bis zu einer Abhängigkeit von 0,9 weiter zu bis 0,0117. Diese Tendenz ist für den Bereich der hohen Messzeitpunkte zwar vorhanden, aber schwächt sich etwas ab (vgl. auch Abbildung 5.7, S. 113).

Ein ähnliches Verhalten ist für α_2 zu beobachten. Auch hier ist bei geringer Abhängigkeit von 0,15 im drastischen Vergleich zu β_2 ein verhältnismäßig geringer MSE von 0,006 (für $T = 50$) zu sehen, der bei einer Abhängigkeit von 0,55 aber bereits auf eine Abhängigkeit von 0,0092 angestiegen ist und bei einer Abhängigkeit von 0,9 bei 0,0112 liegt.

Damit zeichnet sich dieselbe Tendenz für das arithmetische Mittel, den $MSE_{\tilde{\alpha}_{ges}}$, ab. Die Schätzung des Anpassungskoeffizienten wird mit zunehmender Abhängigkeit der Komponentenprozesse bei gleicher Anzahl der Messzeitpunkte ($T = 50$) geringfügig schlechter und erreicht bei einer hohen Abhängigkeit der Komponentenprozesse von 0,9 einen Wert von 0,0114. Dies ist ein Unterschied im Vergleich zu einem $MSE_{\tilde{\beta}_2}$ bei gleicher Abhängigkeit und gleicher Anzahl der Messzeitpunkte von 0,0013.

Für Prozesse mit **auseinanderdriftenden Zeitpfaden der Komponentenprozesse** sind sowohl die Fälle für $\phi_{11} + \phi_{12} < 1$ (Tabelle 5.6, S. 110) als auch für

Tabelle 5.6: MSEs für α und β ; VAR(1)-Prozesse mit **auseinanderdriftenden** Zeitpfaden der Komponentenprozesse ($\phi_{11} + \phi_{12} < 1$). \Rightarrow Fortsetzung auf der nächsten Seite \Rightarrow

Modellmatrix		Abh		T	MSE β_2	MSE α_1	MSE α_2	MSE α_{ges}
auseinanderdriftend $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} \in \mathfrak{R}$ $\phi_{11} + \phi_{12} < 1; \phi_{21} + \phi_{22} < 1$ ϕ_{21} und ϕ_{22} konstant								
1,01	-0,05	0,15	50	2,6876	0,0009	0,0026	0,0018	
0,1	0,5		100	1,5291	0,0003	0,0006	0,0005	
			150	0,5948	0,0002	0,0004	0,0003	
			200	0,2018	0,0001	0,0002	0,0002	
			250	0,0971	0,0001	0,0002	0,0001	
			300	0,0977	0,0001	0,0002	0,0001	
			350	0,0948	0,0001	0,0001	0,0001	
			400	0,0644	0,0001	0,0001	0,0001	
			450	0,0634	0	0,0001	0,0001	
			500	0,0395	0	0,0001	0,0001	
1,06	-0,3	0,4	50	4,5485	0,0008	0,0012	0,001	
0,1	0,5		100	0,2296	0,0003	0,0004	0,0003	
			150	0,1586	0,0002	0,0003	0,0002	
			200	0,0723	0,0001	0,0002	0,0002	
			250	0,0656	0,0001	0,0001	0,0001	
			300	0,0514	0,0001	0,0001	0,0001	
			350	0,0366	0,0001	0,0001	0,0001	
			400	0,0249	0,0001	0,0001	0,0001	
			450	0,024	0	0,0001	0,0001	
			500	0,0185	0	0,0001	0	
1,07	-0,35	0,45	50	1,3537	0,0007	0,001	0,0009	
0,1	0,5		100	0,219	0,0003	0,0004	0,0003	
			150	0,1144	0,0002	0,0003	0,0002	
			200	0,0643	0,0001	0,0002	0,0001	
			250	0,0561	0,0001	0,0001	0,0001	
			300	0,0441	0,0001	0,0001	0,0001	
			350	0,0323	0,0001	0,0001	0,0001	
			400	0,0194	0,0001	0,0001	0,0001	
			450	0,0177	0	0,0001	0	
			500	0,0135	0	0,0001	0	

$\phi_{11} + \phi_{12} > 1$ (Tabelle 5.7, S. 112) dargestellt. Bei den Prozessen mit auseinanderdriftenden Zeitpfaden der Einzelreihen sind im Allgemeinen andere Ergebnisse als bei Prozessen mit nicht-auseinanderdriftenden Zeitpfaden zu beobachten. Der MSE von β_2 schwankt bei den erstgenannten Modellen (Tabelle 5.6, S. 110) zunächst für $T = 50$ im Abhängigkeitsspektrum von 0,15 bis 0,4 stark, stabilisiert sich aber ab einer Abhängigkeit von 0,45 (1,3537) und wird zunehmend besser mit einem Wert von 0,1695 bei einer Abhängigkeit von 0,9. Trotz dieser Verbesserung ist die Güte der Schätzung des Kointegrationsvektors offensichtlich bei auseinanderdriftenden Zeitpfaden bedeutend schlechter als bei nicht-auseinanderdriftenden Zeitpfaden. Bei einer Abhängigkeit von 0,9 beträgt der MSE für β_2 dort 0,0013.

Der MSE für α_1 verschlechtert sich zwar nicht mit zunehmender Abhängigkeit wie im Fall von nicht-auseinanderdriftenden Zeitpfaden, er weist aber gleichzeitig auch

Fortsetzung Tabelle 5.6, S. 110							
ϕ_{11}	ϕ_{12}	Abh	T	MSE	MSE	MSE	MSE
ϕ_{21}	ϕ_{22}			$\tilde{\beta}_2$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\alpha_{\tilde{ges}}$
1,09	-0,45	0,55	50	0,9193	0,0007	0,0008	0,0008
0,1	0,5		100	0,2659	0,0003	0,0003	0,0003
			150	0,0663	0,0002	0,0002	0,0002
			200	0,0405	0,0001	0,0002	0,0001
			250	0,0285	0,0001	0,0001	0,0001
			300	0,0231	0,0001	0,0001	0,0001
			350	0,0169	0,0001	0,0001	0,0001
			400	0,0121	0,0001	0,0001	0,0001
			450	0,011	0	0,0001	0
			500	0,0087	0	0	0
1,16	-0,8	0,9	50	0,1695	0,0007	0,0006	0,0006
0,1	0,5		100	0,032	0,0002	0,0003	0,0002
			150	0,0109	0,0001	0,0002	0,0002
			200	0,0068	0,0001	0,0001	0,0001
			250	0,0044	0,0001	0,0001	0,0001
			300	0,0038	0,0001	0,0001	0,0001
			350	0,0027	0	0,0001	0,0001
			400	0,0019	0	0	0
			450	0,0016	0	0	0
			500	0,0013	0	0	0

keine deutlichen Änderungen auf, d.h. es kann keine wirkliche Aussage darüber gemacht werden, in welcher Weise und in Abhängigkeit von welchen Faktoren sich die Schätzgüte von α_1 bei auseinanderdriftenden Zeitpfaden verändert. Der MSE schwankt relativ bald, nämlich ab einer Abhängigkeit von 0,45, um 0,0007.

Auch bei der durch den MSE operationalisierten Schätzgüte von α_2 sind keine sehr deutlichen Veränderungen, aber eine leichte Tendenz zu einer Verbesserung der Schätzgüte mit zunehmender Abhängigkeit der Komponentenprozesse zu beobachten (MSE= 0,0008 bei Abh= 0,55 und 0,0006 bei Abh=0,9 für jeweils $T = 50$). Vor diesem Hintergrund muss sich auch der $MSE_{\alpha_{\tilde{ges}}}$ ähnlich entwickeln. Er stabilisiert sich bei einer Abhängigkeit von 0,9 bei 0,0006 für $T = 50$.

Beim Vergleich der Tabellen zu auseinanderdriftenden Zeitpfaden (5.6, S. 110, und 5.7, S. 112) kann man feststellen, dass die Schätzgüte bei den eben beschriebenen Modellen mit auseinanderdriftenden Zeitpfaden für $\phi_{11} + \phi_{12} < 1$ (Tabelle 5.6, S. 110) mit dem Verhalten der Schätzgüte bei den außerdem noch simulierten Modellen für $\phi_{11} + \phi_{12} > 1$ (Tabelle 5.7, S. 112) vergleichbar ist. Dies kann anhand der Tabelle 5.7, S. 112, verifiziert werden. Deshalb wäre es redundant, diese Ergebnisse nochmals zu beschreiben. Die beschriebenen Modelle aus Tabelle 5.6, S. 110, werden im weiteren Verlauf der Arbeit als repräsentativ für die Gruppe der Modelle mit aus-

Tabelle 5.7: MSEs für α und β ; VAR(1)-Prozesse mit **auseinanderdriftenden** Zeitpfaden der Komponentenprozesse ($\phi_{11} + \phi_{12} > 1$).

Modellmatrix		Abh		T	MSE $\tilde{\beta}_2$	MSE $\tilde{\alpha}_1$	MSE $\tilde{\alpha}_2$	MSE α_{ges}
auseinanderdriftend $\phi_{11}, \phi_{12}, \phi_{21}$ und $\phi_{22} > 0$ $\phi_{11} + \phi_{12} > 1; \phi_{21} + \phi_{22} < 1$ ϕ_{21} und ϕ_{22} konstant								
0,99 0,1	0,05 0,5	0,15		50	25,5931	0,001	0,0018	0,0014
				100	1,3692	0,0003	0,0006	0,0005
				150	0,4347	0,0002	0,0004	0,0003
				200	0,1731	0,0001	0,0002	0,0002
				250	0,0951	0,0001	0,0002	0,0001
				300	0,0988	0,0001	0,0002	0,0001
				350	0,093	0,0001	0,0001	0,0001
				400	0,0615	0,0001	0,0001	0,0001
				450	0,0538	0,0001	0,0001	0,0001
				500	0,0436	0	0,0001	0,0001
0,93 0,1	0,35 0,5	0,45		50	13,6293	0,0012	0,0012	0,0012
				100	2,0471	0,0005	0,0005	0,0005
				150	0,3242	0,0002	0,0003	0,0003
				200	0,1161	0,0002	0,0002	0,0002
				250	0,0671	0,0001	0,0002	0,0001
				300	0,0502	0,0001	0,0001	0,0001
				350	0,0431	0,0001	0,0001	0,0001
				400	0,0362	0,0001	0,0001	0,0001
				450	0,0287	0,0001	0,0001	0,0001
				500	0,0254	0,0001	0,0001	0,0001
0,84 0,1	0,8 0,5	0,9		50	1,0165	0,0012	0,0008	0,001
				100	0,1087	0,0004	0,0004	0,0004
				150	0,052	0,0003	0,0003	0,0003
				200	0,0287	0,0002	0,0002	0,0002
				250	0,0192	0,0001	0,0002	0,0001
				300	0,0151	0,0001	0,0001	0,0001
				350	0,0111	0,0001	0,0001	0,0001
				400	0,009	0,0001	0,0001	0,0001
				450	0,0068	0,0001	0,0001	0,0001
				500	0,0051	0,0001	0,0001	0,0001

einanderdriftenden Zeitpfaden angesehen, da ihr Verhalten auch in Bezug auf die Testverfahren repräsentativ ist, wie bereits gezeigt werden konnte (vgl. Abschnitte 5.1.1, S. 93ff., und 5.1.2, S. 96ff.).

Zusätzlich zu den beschriebenen Tabellen sind die Ergebnisse anhand der Abbildungen 5.7 und 5.8 auf den Seiten 113 und 114 nachvollziehbar. Wie auch schon bei der graphischen Darstellung der Ergebnisse zu den Testverfahren deutlich wurde, unterscheiden sich die Ergebnisse auch bei den Parameterschätzungen für die beiden Modelltypen auseinanderdriftender Zeitpfade mit $\phi_{11} + \phi_{12} < 1$ vs. > 1 nicht wesentlich, weshalb hier nur die Ergebnisse für die erstgenannte Modellkategorie mit $\phi_{11} + \phi_{12} < 1$ graphisch illustriert werden.

In Abbildung 5.7, S. 113, die die Qualität der Schätzung von Kointegrationsvektor und Anpassungskoeffizient bei Einzelreihen mit nicht-auseinanderdriftenden Zeitpfaden graphisch zusammenfasst, fällt deutlich auf, dass der MSE für den Anpassungs-

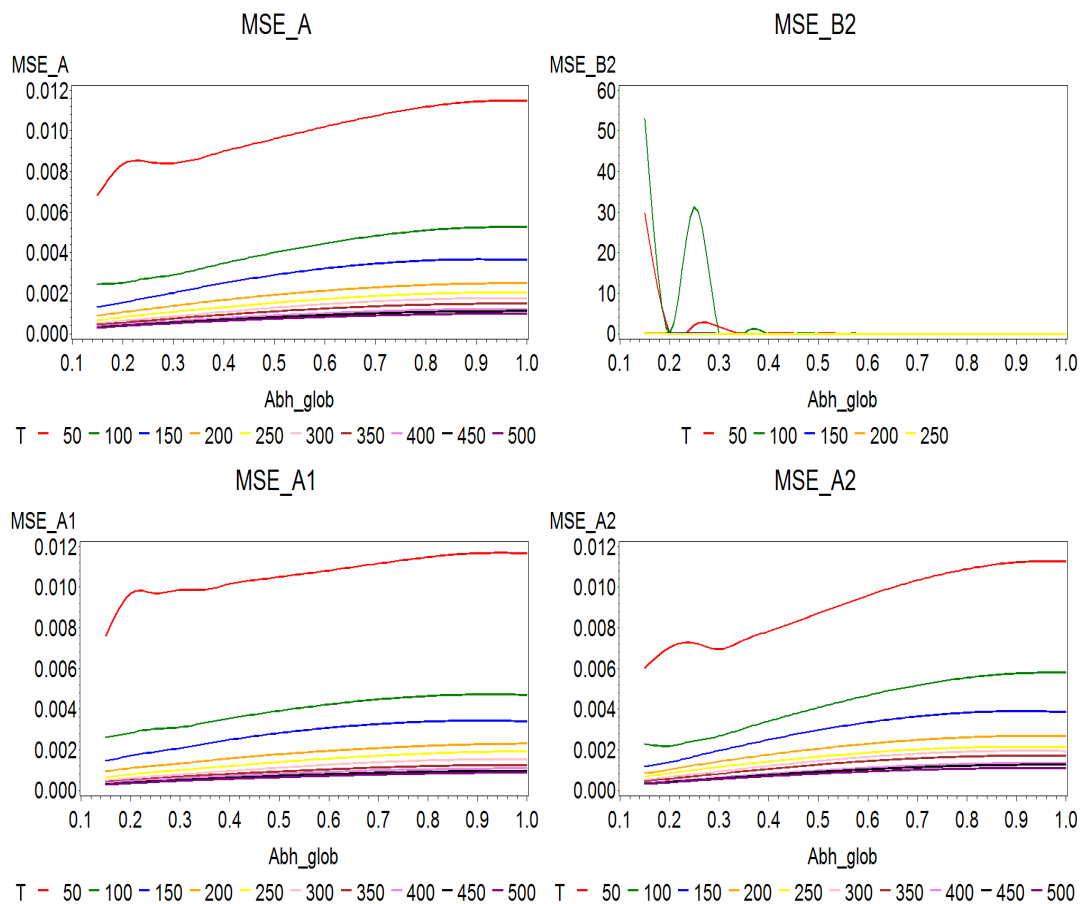


Abbildung 5.7: Ergebnisse der Parameterschätzungen von Kointegrationsvektor und Anpassungskoeffizient für *nicht*-auseinanderdriftende Zeitpfade.

koeffizienten mit zunehmendem Grad der Abhängigkeit der Komponentenprozesse ansteigt, während der MSE für den Kointegrationsvektor abnimmt.

Neben dieser Abbildung stimmt auch die Veranschaulichung der Ergebnisse zur Schätzgüte im Fall von auseinanderdriftenden Zeitpfaden der Komponentenprozesse in Abbildung 5.8, S. 114, mit den in den Tabellen dargestellten Ergebnissen überein: im Fall auseinanderdriftender Zeitpfade der Einzelreihen steigt die Güte der Schätzung des Anpassungskoeffizienten mit zunehmender Abhängigkeit der Komponentenprozesse. Dies trifft auch auf die Qualität der Schätzung des Kointegrationsvektors zu. Die Entwicklung seiner Schätzgüte ist allerdings aufgrund eines Ausreißers nicht so deutlich graphisch zu sehen.

Wenn man die beschriebenen unterschiedlichen Funktionen von Anpassungskoeffizi-

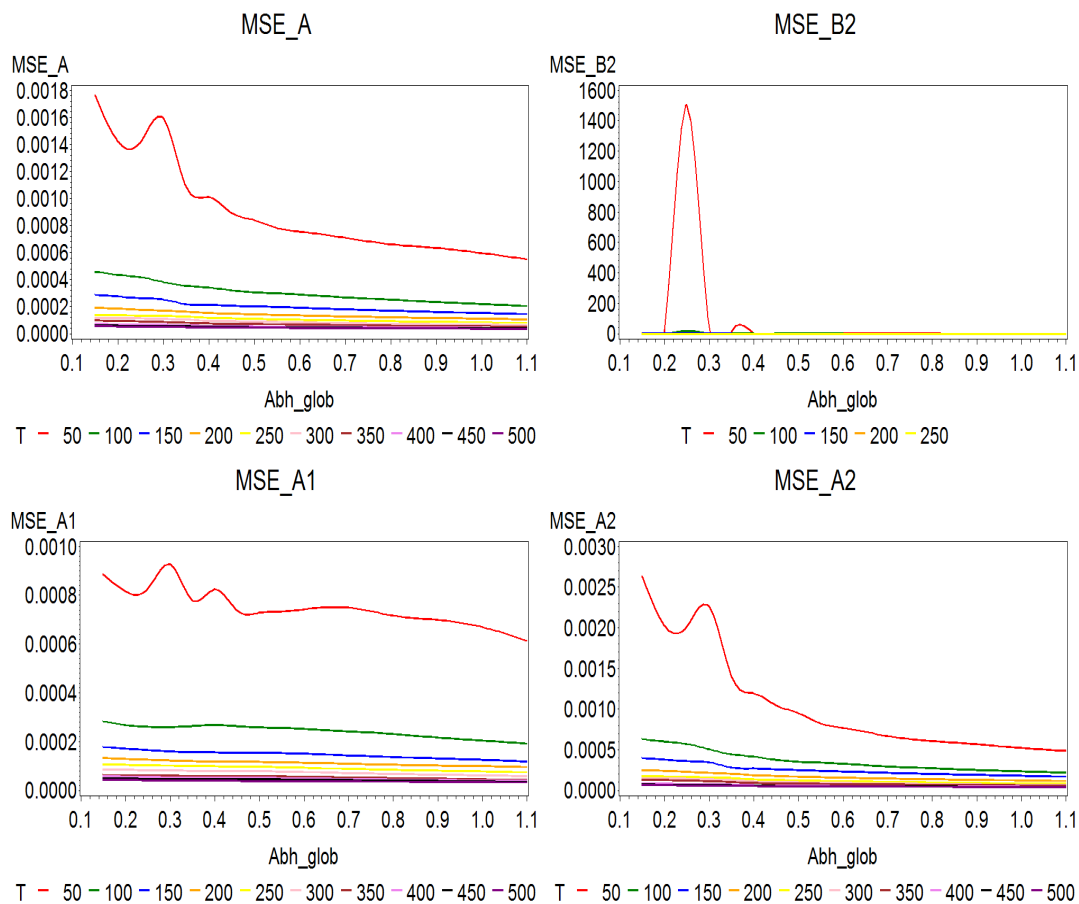


Abbildung 5.8: Ergebnisse der Parameterschätzungen von Kointegrationsvektor und Anpassungskoeffizient für auseinanderdriftende Zeitpfade ($\phi_{11} + \phi_{12} < 1$).

ent und Kointegrationsvektor bedenkt (vgl. Abschnitt 3.4.1, S. 54ff.), die an der angegebenen Stelle auch graphisch veranschaulicht sind, können diese Ergebnisse zum Verhalten der Schätzverfahren für die beiden Parameter unmittelbar einleuchten. Wie ausführlich dargestellt wurde, bestimmt die Größe des Kointegrationsvektors den Abstand zwischen den Einzelreihen (schwarz und rot) und die Größe des Anpassungskoeffizienten den Abstand der beiden Reihen zur Linearkombination (grün), also die „Anpassung“ an das Gleichgewicht. Zwar sind in der Literatur aufgrund der existierenden Testverfahren keine Abstufungen von Kointegration beschrieben, sodass man von „stärker kointegrierten“ oder „weniger stark kointegrierten“ Prozessen sprechen könnte, dennoch haben die Testverfahren gezeigt (vgl. Abschnitte 5.1.1, S. 93ff., und 5.1.2, S. 96ff.), dass bei zunehmender Abhängigkeit der Komponentenprozesse in immer mehr Fällen Kointegration indiziert wird. So könnte man weiter

denken, dass der Kointegrationsvektor, der den Abstand zwischen den Einzelreihen bestimmt, zusätzlich angibt, wie stark die Prozesse kointegriert sind. Folglich leuchtet ein, warum der Kointegrationsvektor umso besser und stabiler geschätzt wird, je abhängiger die Komponentenprozesse sind. Mit zunehmender Stabilität der Schätzungen ist auch von einer zunehmenden Präzision der Prognose auszugehen. Diese Schlussfolgerungen sind in Nachfolgestudien genauer zu untersuchen. Die Betrachtung der Güte der Schätzverfahren hat also weitreichende Implikationen. Die Funktion des Anpassungskoeffizienten (Abstand der Reihen zum Gleichgewicht) dagegen steht nicht in unmittelbarem Zusammenhang mit der Abhängigkeit der Komponentenprozesse und deshalb ist erklärbar, dass die Schätzverfahren für den Anpassungskoeffizienten mit zunehmender Abhängigkeit der Komponentenprozesse kein eindeutiges Muster erkennen lassen.

5.4 Überprüfung der Hypothesen zu den Parameterschätzungen

Hypothese 2a (S. 80) kann in ihrer Bestätigung oder Nicht-Bestätigung nur beurteilt werden, wenn man zwischen auseinanderdriftenden und nicht-auseinanderdriftenden Zeitpfaden und dabei jeweils nochmals zwischen dem Grad der Abhängigkeit der Komponentenprozesse unterscheidet. Die Hypothese nimmt an, dass die Güte der Schätzung des Kointegrationsvektors über das Fehlerkorrekturmodell sehr hoch ist, d.h. im üblichen Range der Schätzungen von ARMA-Parametern liegt, nämlich zwischen einem MSE von 0,2 und 0,002. Da β_1 immer auf Eins normiert wird, sind keine Schätzbeurteilungen zu β_1 möglich, weshalb die Güte der Schätzung des Kointegrationsvektors nur auf Basis der Ergebnisse zu β_2 beurteilt werden kann (vgl. Erläuterungen auf S. 92). Für nicht-auseinanderdriftende Zeitpfade ist die Güte der Schätzung bei geringer Abhängigkeit der Komponentenprozesse zunächst sehr niedrig, sie wird aber bei der Erhöhung der Abhängigkeit der Komponentenprozesse sehr schnell viel besser und erreicht bei einer sehr hohen Abhängigkeit (0,9) ein relativ hohes Niveau: MSE von 0,0013 bei $T = 50$ und 0,0001 ab $T = 250$. Man kann also sagen, dass die Güte der Schätzung des Kointegrationsvektors bei nicht-auseinan-

derdriftenden Zeitpfaden dann sehr hoch ist, wenn die Einzelprozesse eine mittlere bis hohe Abhängigkeit aufweisen.

Bei auseinanderdriftenden Zeitpfaden ist die Schätzung von β_2 bei niedrigem T deutlich schlechter. Nach anfänglicher Schwankung stabilisiert sie sich zwar ab einer Abhängigkeit von 0,45 (MSE: 1,3537), der MSE ist aber bei einer hohen Abhängigkeit (0,9) im Vergleich zu nicht-auseinanderdriftenden Zeitpfaden sehr hoch (0,1695) und damit ist die Schätzung von β_2 bei auseinanderdriftenden Zeitpfaden im Vergleich zu nicht-auseinanderdriftenden Zeitpfaden schlechter.

Diese Befunde können graphisch anhand von Abbildung 5.9, S. 116, verifiziert werden.

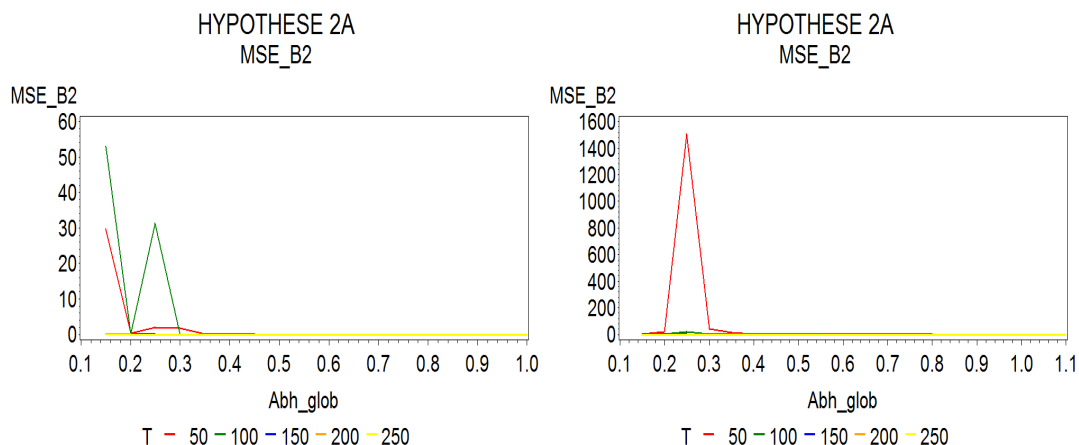


Abbildung 5.9: HYPOTHESE 2A zur Schätzung des Kointegrationsvektors für VAR(1)-Modelle mit *nicht*-auseinanderdriftenden (links) und auseinanderdriftenden Zeitpfaden (rechts; Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$).

Aufgrund der Ausreißer sind die anhand der Tabellen 5.5 bis 5.7, S. 108ff., beschriebenen Unterschiede zwischen Modellen mit nicht-auseinanderdriftenden vs. auseinanderdriftenden Zeitpfaden der Komponentenprozesse graphisch nicht so deutlich zu erkennen.

Solange noch kein Vergleich mit der Schätzgüte eines anderen Parameters, etwa mit der des Anpassungskoeffizienten (vgl. Hypothese 2b), angestellt wird, könnte man aus den Tabellen und Schaubildern schlussfolgern, dass die Qualität der Schätzung des Kointegrationsvektors im Allgemeinen als „gut“ bezeichnet werden kann. In Be-

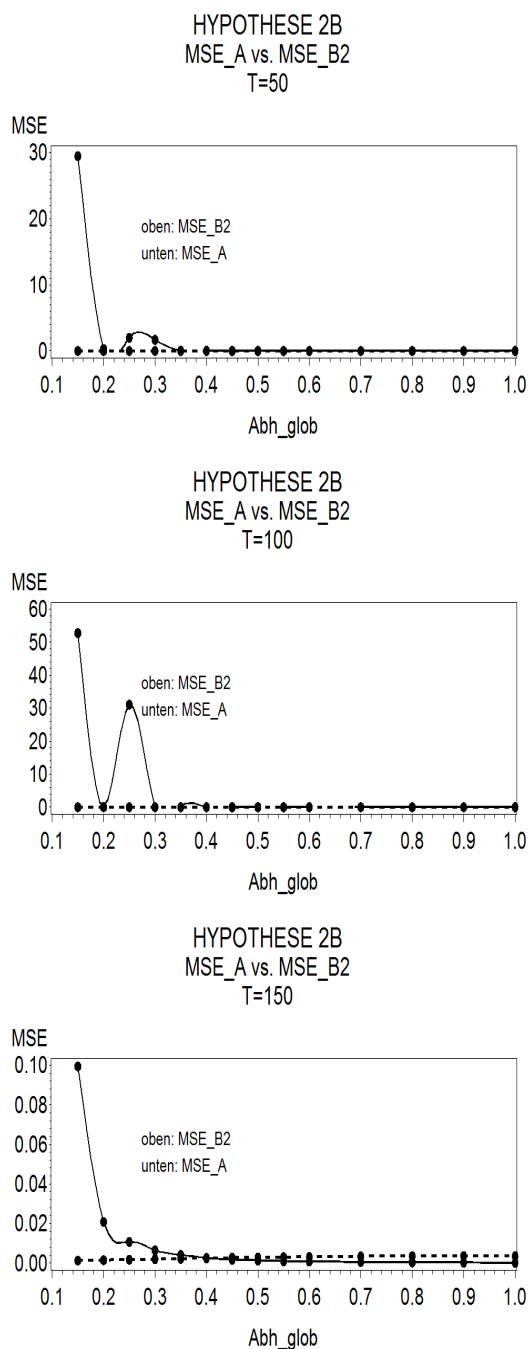


Abbildung 5.10: HYPOTHESE 2B: übergeordneter Vergleich der Schätzgüte von Kointegrationsvektor vs. Anpassungskoeffizient für Modelle mit *nicht*-auseinanderdriftenden Zeitpfaden.

Sie postuliert, dass die Güte der Schätzung des Kointegrationsvektors sich bei gleicher Anzahl der Messzeitpunkte und vergleichbarer Abhängigkeit der Komponentenprozesse von der Güte der Schätzung des Anpassungskoeffizienten unterscheidet.

zug auf den in der Hypothese angestellten Vergleich mit der Güte der Schätzung von ARMA-Parametern (vgl. Werner, 2005, S. 165ff.), deren üblicher Range zwischen einem MSE von 0,2 und 0,002 lokalisiert werden kann, muss allerdings differenziert werden. In diesem Sinne kann die Güte der Schätzung des Kointegrationsvektors für nicht-auseinanderdriftende Zeitpfade bei sehr geringen Abhängigkeiten der Komponentenprozesse ($< 0,35$) erst ab $T = 150$, bei allen höheren Abhängigkeiten bereits bei $T = 50$ als „sehr gut“ angesehen werden. Wenn man diesen Vergleich aber mit der Güte der Schätzung von ARMA-Parametern als Gütekriterium bei auseinanderdriftenden Zeitpfaden heranzieht, ist das Qualitätsurteil „sehr gut“ nicht angemessen: die Schätzungen fallen erst bei einer sehr hohen Abhängigkeit von 0,9 bei $T = 50$ in den Range zwischen einem MSE von 0,2 und 0,002.

Hypothese 2b (S. 80) kann auch nur beurteilt werden, wenn zwischen auseinanderdriftenden und nicht-auseinanderdriftenden Zeitpfaden unterschieden wird.

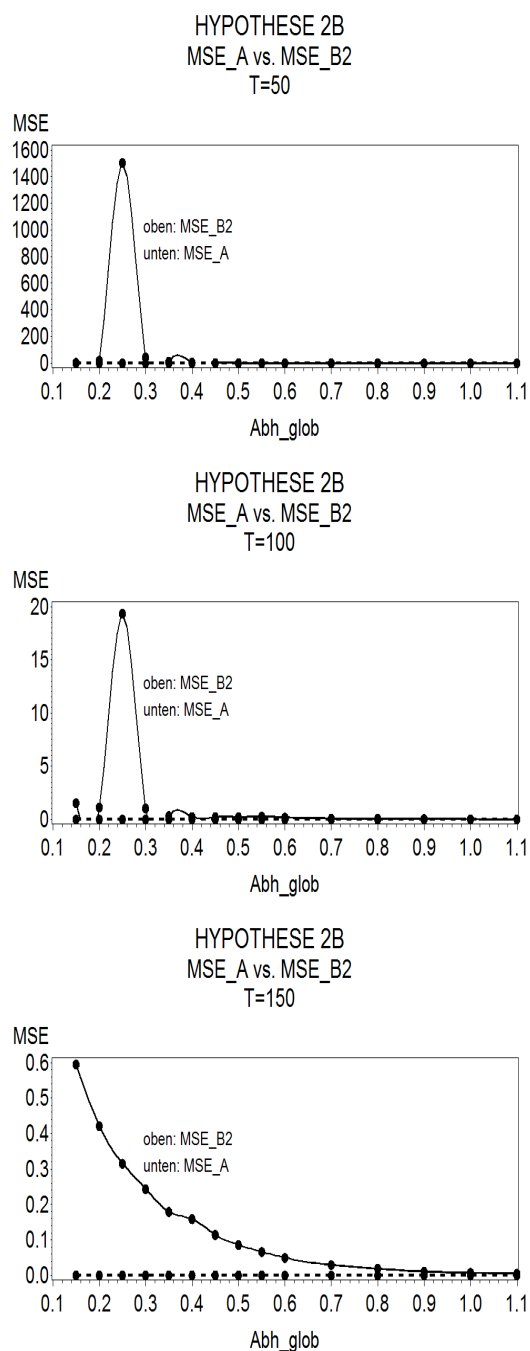


Abbildung 5.11: HYPOTHESE 2B: übergeordneter Vergleich der Schätzgüte von Kointegrationsvektor vs. Anpassungskoeffizient für Modelle mit auseinanderdriftenden Zeitpfaden (Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$).

Insgesamt weisen die Ergebnisse einen solchen Unterschied nach, der ohne Zweifel bedeutend ist: die Güte der Schätzungen der beiden Parameter verhält sich in ihrer Tendenz bei nicht-auseinanderdriftenden Zeitpfaden gegenläufig. Diese Gegenläufigkeit könnte v.a. deshalb einleuchten, weil Kointegrationsvektor und Anpassungskoeffizient als Produkt ins Fehlerkorrekturmodell eingehen (vgl. Gleichung 3.5, S. 52). Die Schätzung von α_1 ist bei einer sehr geringen Abhängigkeit der Komponentenprozesse (0,15) bei $T = 50$ ausgezeichnet (MSE: 0,0076). Wider Erwarten verschlechtert sich die Schätzung von α_1 aber mit zunehmender Abhängigkeit der Komponentenprozesse geringfügig und erreicht bei einer sehr hohen Abhängigkeit von 0,9 einen MSE von 0,0117. Die Güte der Schätzung von α_2 verhält sich in ihrem Verlauf ähnlich. Damit fällt die Güte der Schätzung des Anpassungskoeffizienten für hohe Abhängigkeiten zwar nicht schlecht, aber geringer aus als für den Kointegrationsvektor.

Für auseinanderdriftende Zeitpfade ist die Schätzung des Anpassungskoeffizienten insgesamt besser als die des Kointegrationsvektors.

Diese Ergebnisse sind in Abbildung 5.10, S. 117, veranschaulicht, die sich auf ko-

integrierte Modelle beziehen, deren Zeitpfade nicht auseinanderdriften. Zusätzlich werden die Befunde für Modelle mit auseinanderdriftenden Zeitpfaden in Abbildung 5.11, S. 118 dargestellt. Sowohl die Abbildung für Modelle mit nicht-auseinander-

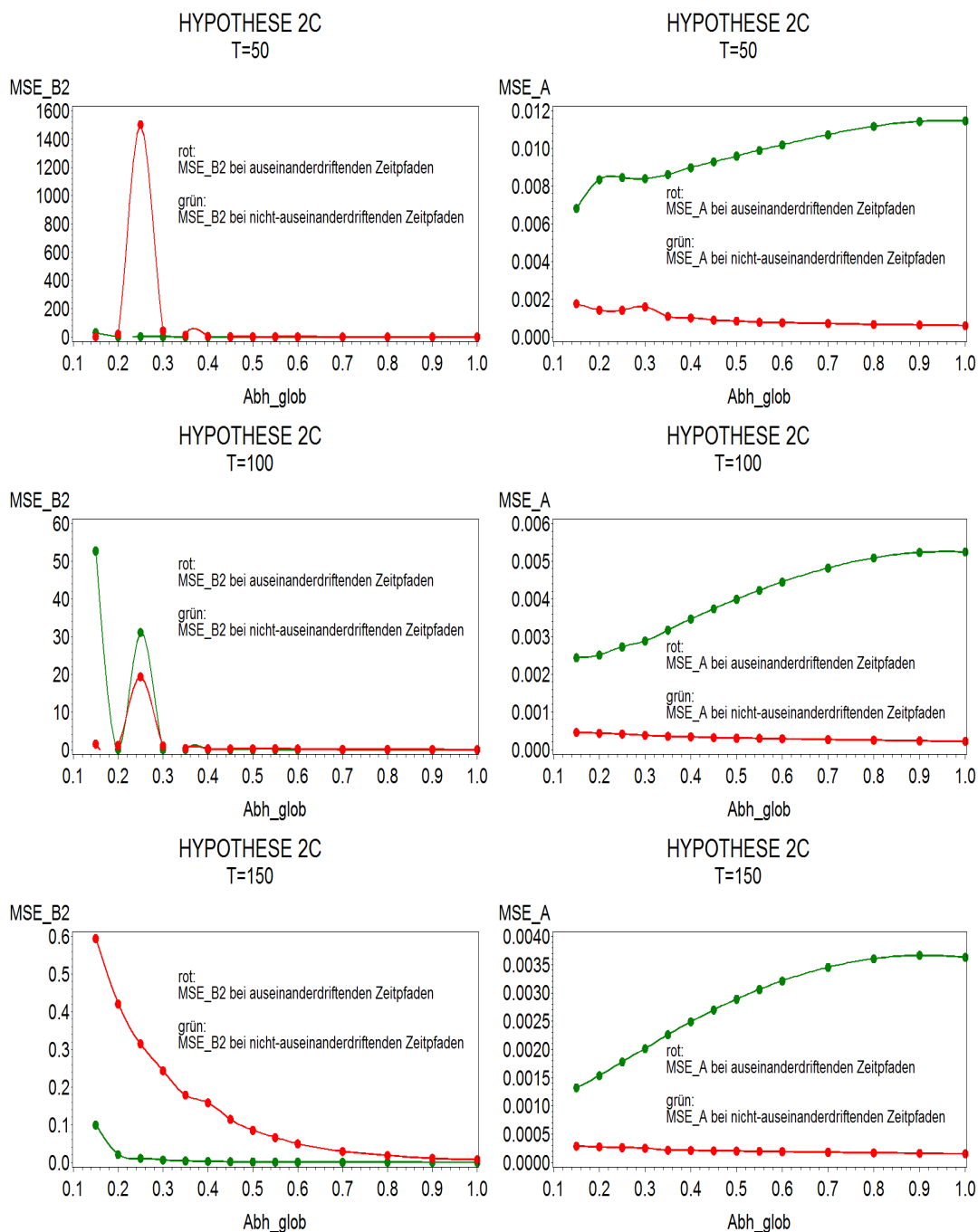


Abbildung 5.12: HYPOTHESE 2C: übergeordneter Vergleich der Schätzgüte von Kointegrationsvektor (links) vs. Anpassungskoeffizient (rechts) für Modelle mit nicht-auseinanderdriftenden (grün) vs. auseinanderdriftenden (rot) Zeitpfaden (Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$).

derdriftenden als auch mit auseinanderdriftenden Zeitpfaden zeigt die beschriebene Unterschiedlichkeit der Schätzgüte von Anpassungskoeffizient und Kointegrationsvektor.

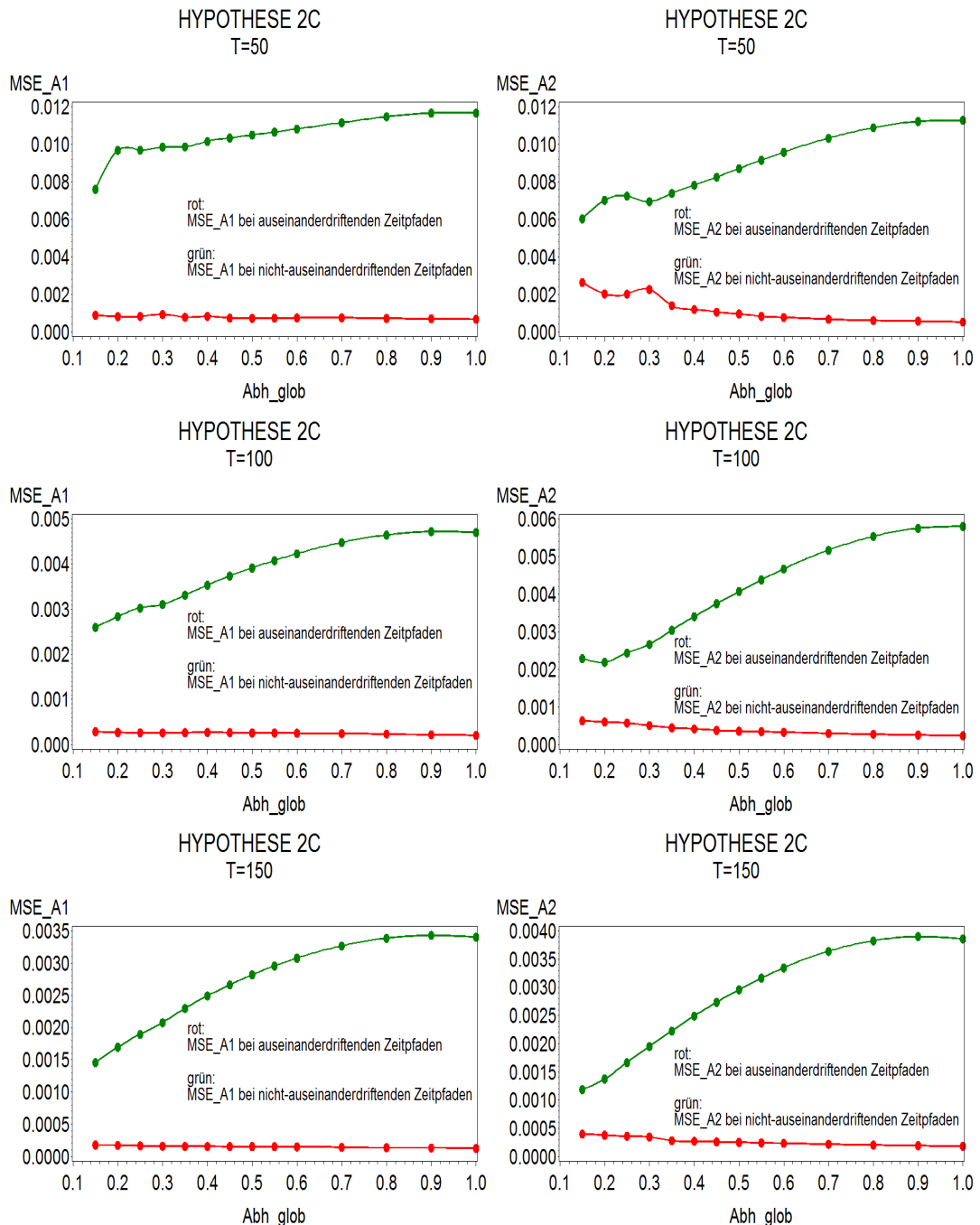


Abbildung 5.13: HYPOTHESE 2C: Einzelvergleiche der Schätzgüte von α_1 (links) vs. α_2 (rechts) für Modelle mit nicht-auseinanderdriftenden (grün) vs. auseinanderdriftenden (rot) Zeitpfaden (Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$).

Hypothese 2c (S. 80) ist durch die Beschreibungen zu den Hypothesen 2a und b bereits bestätigt. Das bedeutet, dass die Schätzgüte von Anpassungskoeffizient vs. Kointegrationsvektor sich in Abhängigkeit vom Verlauf der Zeitpfade der Einzelreihen bei gleicher Anzahl der Messzeitpunkte und vergleichbarer Abhängigkeit der Komponentenprozesse unterscheidet. Damit ist nachgewiesen, dass bei der Forschung mit kointegrierten Modellen unbedingt zwischen Modellen mit auseinanderdriftenden und nicht-auseinanderdriftenden Zeitpfaden unterschieden werden muss, um die Schätzgüte von Kointegrationsvektor und Anpassungskoeffizient angemessen beurteilen zu können.

Dies wird auch durch die Abbildungen 5.12 und 5.13, S. 119f., veranschaulicht. Die grünen (nicht-auseinanderdriftende Zeitpfade) und die roten (auseinanderdriftende

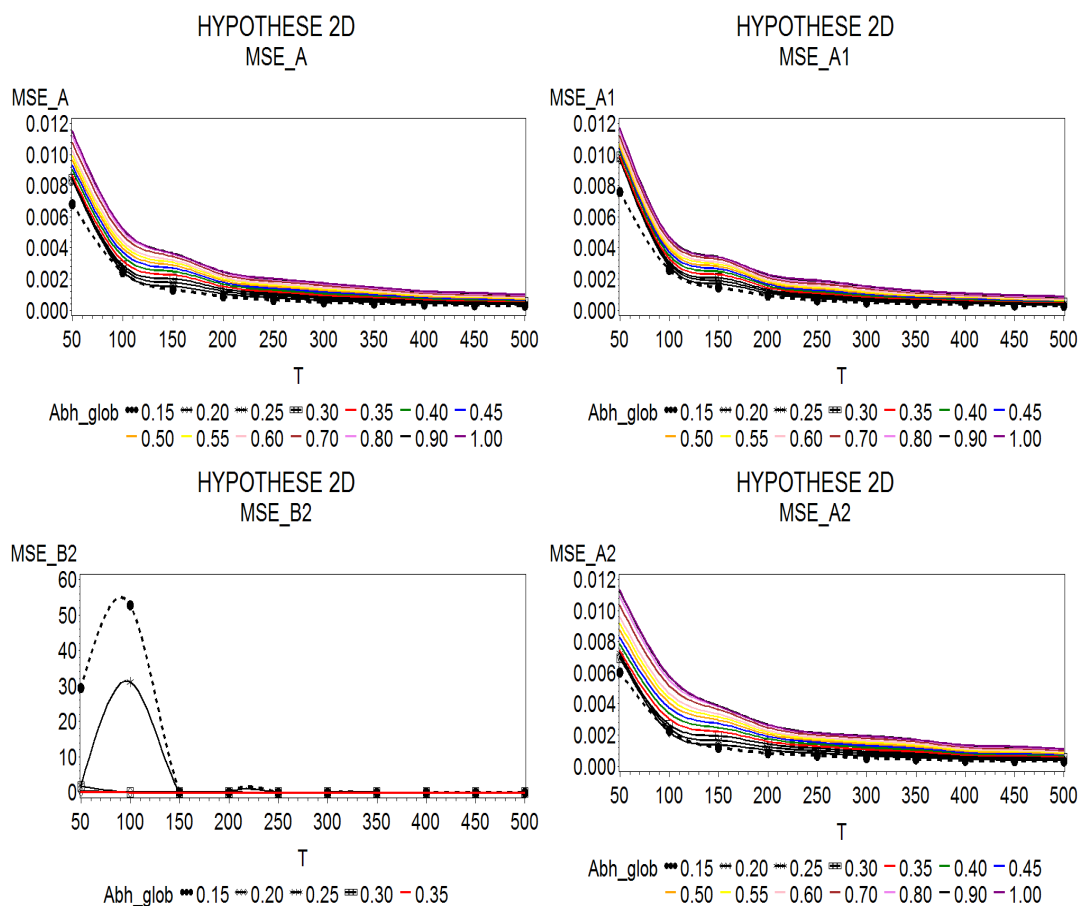


Abbildung 5.14: HYPOTHESE 2D zur Schätzung von Kointegrationsvektor (links) und Anpassungskoeffizient (rechts) für Modelle mit *nicht*-auseinanderdriftenden Zeitpfaden; unterschieden nach der variierten Abhängigkeit der Komponentenprozesse.

Zeitpfade) weichen bei den Graphiken zum Anpassungskoeffizienten viel deutlicher voneinander ab, als bei den Graphiken zum Kointegrationsvektor.

Hypothese 2d (S. 80), die postuliert, dass die Schätzgüte von Anpassungskoeffizient vs. Kointegrationsvektor sich in Abhängigkeit von der variierten Abhängigkeit der Komponentenprozesse über die Nebendiagonalen bei gleicher Anzahl der Messzeitpunkte unterscheidet, ist ebenfalls durch die Beschreibungen zu den vorangegangenen Hypothesen bereits bestätigt. Die Güte der Schätzung des Kointegrationsvektors nimmt bei nicht-auseinanderdriftenden Zeitpfaden mit zunehmender Abhängigkeit der Komponentenprozesse zu, während die Schätzgüte des Anpassungskoeffizienten sich gleichzeitig mit zunehmender Abhängigkeit verschlechtert.

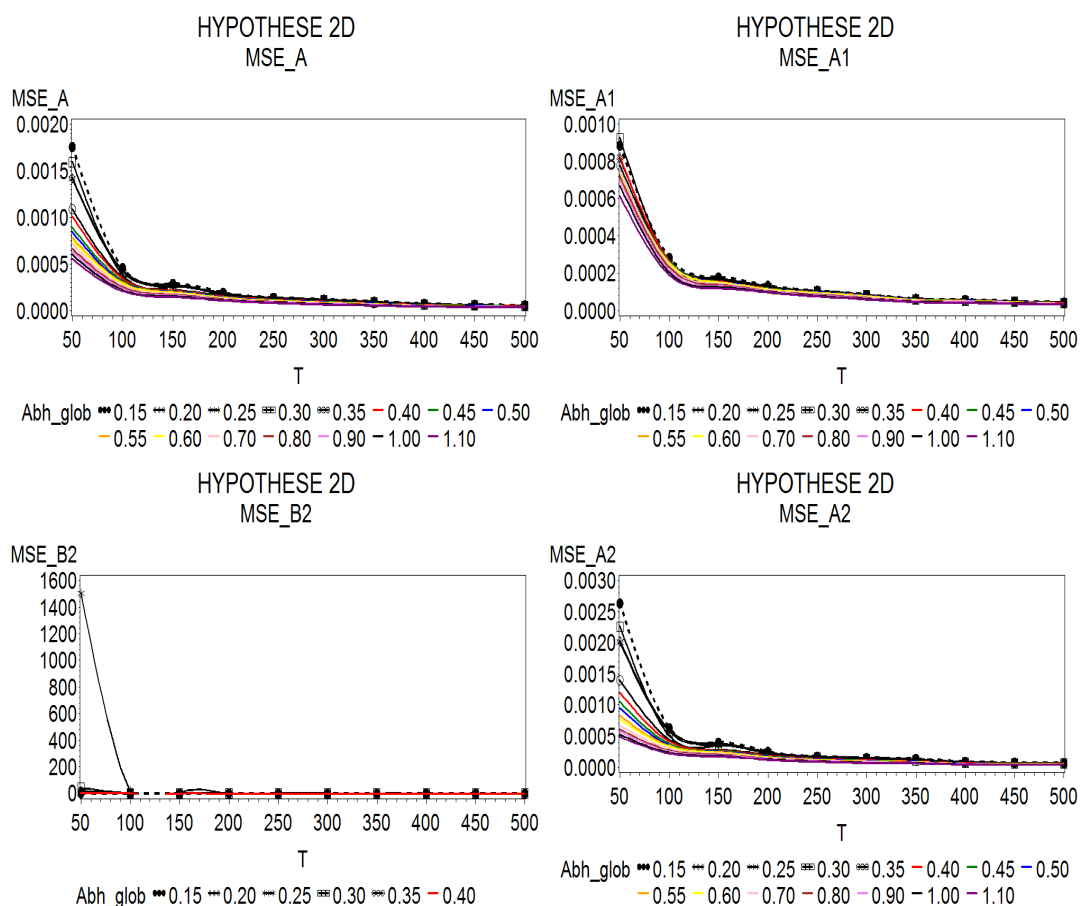


Abbildung 5.15: HYPOTHESE 2D zur Schätzung von Kointegrationsvektor (links) und Anpassungskoeffizient (rechts) für Modelle mit auseinanderdriftenden Zeitpfaden (Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$); unterschieden nach der variierten Abhängigkeit der Komponentenprozesse.

Bei auseinanderdriftenden Zeitpfaden ist das Umgekehrte der Fall: während der Kointegrationsvektor mit zunehmender Abhängigkeit der Einzelreihen schlechter geschätzt wird, verbessert sich gleichzeitig die Schätzungsgüte des Anpassungskoeffizienten.

Wie sich dieses Verhalten bei Modellen mit auseinanderdriftenden Zeitpfaden verändert, wird durch den Vergleich der Abbildungen 5.14, S. 121, und 5.15, S. 122, veranschaulicht. Bei auseinanderdriftenden Zeitpfaden können in Bezug auf die Schätzungsgüte des Kointegrationsvektors keine wesentlichen Veränderungen im Vergleich zum nicht-auseinanderdriftenden Fall mithilfe der Graphiken beobachtet werden. Für die Schätzungsgüte des Anpassungskoeffizienten erscheinen die Graphiken bei Modellen mit auseinanderdriftenden Zeitpfaden der Einzelreihen wie umgekehrt, d.h. es deutet sich die Tendenz an, dass die Schätzungen umso schlechter sind, je geringer der Grad der Abhängigkeit der Komponentenprozesse ist.

In den Schaubildern sind nur die Abhängigkeitsgrade zu sehen, die sich nicht überschneiden, d.h. der höchste im Schaubild verzeichnete Abhängigkeitsgrad ist derjenige, ab dem sich die Simulationsergebnisse für alle höheren Abhängigkeitsgrade nicht wesentlich unterscheiden.

Diskussion und Einordnung der Ergebnisse

Auch wenn zeitreihenanalytische und damit prozessbezogene Verfahren in der psychologischen Forschungspraxis zum Zweck der Abbildung des *dynamischen Systems Mensch* in seinen Eigenschaften und Verhaltensweisen wie in der Einleitung angedeutet und von vielen Autoren bemängelt bisher zu selten Anwendung finden, so ist doch in jüngster Zeit eine vermehrte Anwendung zeitbezogener Verfahren - zumindest aus dem Bereich der univariaten Zeitreihenanalyse - zu beobachten. Die Psychologie bietet ein großes Spektrum an zeitbezogenem Datenmaterial, das eines angemessenen Auswertungsverfahrens bedarf. Die Liste der beschriebenen Anwendungsbeispiele kann beliebig weiter um Datenbeispiele - wie z.B. physiologische Daten (EEG-Daten, Daten zu Blutzuckerfluktuation, Herzfrequenz, Hauttemperatur, Muskelspannung etc.) oder durch Daten zu Blickbewegung, Stimmungsfluktuation oder psychologischen Regulationsvorgängen, wie z.B. der Beanspruchungsregulation - ergänzt werden, bei denen auch die Erhebung einer ausreichenden Anzahl an Messzeitpunkten möglich ist. Dies lässt auf „interdisziplinäre“ Kooperationsprojekte hoffen: beispielsweise mit der Biologischen, der Allgemeinen, der Entwicklungspsychologie, der psychologischen Kognitionsforschung und mit Unternehmen.

Nicht zuletzt der empirische Teil der Diplomarbeit zeigt, dass der Forschungsaufwand bei der multivariaten Modellierung psychologischer Zeit-Phänomene aufgrund der Komplexität der Modelle höher als im univariaten Fall ist, was die Befürchtung aufkommen lassen könnte, dass die Anwendungsattraktivität multivariater Zeitreihenmodelle für die psychologische Forschung dadurch sinkt bzw. erst gar nicht richtig in Gang kommt. Es muss jedoch - wie schon in der Einleitung beschrieben - ohne Zweifel einleuchten, dass multivariate *kointegrierte* Modelle bisher nicht realisierte Modellierungsoptionen - nämlich die direkte Abbildung instationärer Prozesse ohne vorgeschaltete Transformation, die das „Langzeitgedächtnis“ eliminiert - eröffnen und damit wiederum höchst anwendungsattraktiv sind. Hinzu kommt, dass koin-

tegrierte Modelle ein ausgesprochen hohes Modellierungspotential aufweisen, d.h. sie können das Realitätsspektrum sehr viel genauer abbilden als bei herkömmlichen zeitreihenanalytischen Modellen, da viele psychologische Phänomene einen instationären Verlauf zeigen und in Beziehung zu anderen Phänomenen stehen. Außerdem sind - wie im Rahmen der Simulationen deutlich gezeigt werden konnte - Kointegrationsmodelle auch bei auseinanderdriftenden Zeitpfaden der Komponentenprozesse möglich, was zusätzlich die Modellierungsvielfalt erhöht. Es zeigt sich insgesamt, dass die Kointegrationsmethodologie ein außerordentlich flexibles Verfahren darstellt, was wohl unter anderem auch zur Verleihung des Nobelpreises an C.W.J. Granger geführt hat.

Um die Kointegrationsmethodologie, die ihren Ursprung in der Ökonometrie hat, aber unmittelbare Relevanz für die Psychologie besitzt, erschließend darzustellen, hat die Diplomarbeit es sich zunächst zum Ziel gemacht, nach Darstellung der notwendigen zeitreihenanalytischen Grundlagen die theoretischen Implikationen des Konzepts schrittweise zu erarbeiten und damit die Basis für das Verständnis und die Anwendung kointegrierter Modelle zu verdeutlichen. Wichtig ist in diesem Zusammenhang sicherlich der explizite Einbezug von Modellen mit auseinanderdriftenden Zeitpfaden in das Kointegrationskonzept, der auch weitreichende empirische Implikationen mit sich gebracht hat, wie im Rahmen der Simulationen gezeigt werden konnte. Darüber hinaus sind die theoretische Einführung in die Testverfahren als Operationalisierung von Kointegration und die Darstellung des gegenläufigen Zusammenhangs von Scheinregression und Kointegration bedeutende Schritte zur Erschließung des Ansatzes. Wie sich in den Simulationen zeigt, ergeben sich interessante und nicht immer erwartungskonforme Ergebnisse für die Güte der Parameterschätzungen von Kointegrationsvektor und Anpassungskoeffizient über das Fehlerkorrekturmodell. Die Beurteilung der Güte ist allerdings nur durch den Vergleich mit den Populationsparametern möglich. Deshalb findet sich im theoretischen Teil neben der Erläuterung der für diesen Sachverhalt relevanten Fehlerkorrekturmodelle die fundierte Vereinfachung der Rückrechnung auf die Populationsparameter für bivariate VAR(1)-Modelle. Aus den praktischen Anwendungsbeispielen für fraktionale

Integration sind Anwendungsfelder für fraktionale Kointegration erschließbar.

Aber auch die Tatsache, dass der empirische Beleg erbracht werden konnte, dass Prozesse mit auseinanderdriftenden Zeitpfaden durchaus kointegriert sein können, erweitert das Wissen über kointegrierte Modelle deutlich, weil bezüglich Schätzungen und Testverfahren Modelle mit auseinanderdriftenden und nicht-auseinanderdriftenden Zeitpfaden zwingend unterschieden werden müssen und diese Unterscheidung für den psychologischen Forscher damit unmittelbare Relevanz besitzen muss.

Die empirische Analyse des Verhaltens der betrachteten Testverfahren mit zunehmender Abhängigkeit der Komponentenprozesse - unterschieden für die Anzahl der Messzeitpunkte und auseinanderdriftende und nicht-auseinanderdriftende Zeitpfade - ermöglicht durch die tabellarische Darstellung der Ergebnisse, durch die eine Vielzahl der in der Realität vorkommenden Modelltypen abgedeckt werden kann, praktische Entscheidungen bei der Wahl des Testverfahrens, der Anzahl der Messzeitpunkte usw.

Nicht alle Modelltypen, zu denen Simulationen durchgeführt wurden, können im Rahmen dieser Arbeit aufgeführt werden, doch das zusätzliche Wissen um ihre aus den Simulationen abgeleiteten Eigenschaften hat die Interpretation der beschriebenen Modelle bereichert und Impulse für weitere Untersuchungen gegeben, mithilfe derer ein noch größeres Modellspektrum abgedeckt werden kann.

Darüber hinaus lassen die Ergebnisse zur Güte der Schätzungen von Kointegrationsvektor und Anpassungskoeffizient Muster erkennen und geben deswegen direkte Handlungsrichtlinien für die Wahl der Messzeitpunkte bei gegebener Abhängigkeit der Komponentenprozesse vor. Stabile Schätzungen dieser beiden Parameter sind nämlich unbedingt notwendig, da die beiden Größen den Kern der Prozesseigenschaften kointegrierter Systeme abbilden: langfristige Gleichgewichtsbeziehung und dynamische Anpassungsprozesse.

Aus eben diesem Grund ist die genaue Beschäftigung mit diesen Parametern und ihrer Funktion unbedingt notwendig, nicht zuletzt auch deshalb, weil beispielsweise am Kointegrationsvektor abgelesen werden kann, ob die Zeitpfade der Einzelpro-

zesse auseinanderdriften oder nicht. Dieses Wissen ist eine entscheidende Prämisse für die Wahl der richtigen Moderatorvariablen, wie z.B. die Wahl der Anzahl der Messzeitpunkte. Von der angemessenen Wahl dieser Einflussgrößen wiederum hängt das Verhalten von Test- und Schätzverfahren ab.

Aufgrund der Wichtigkeit von Kointegrationsvektor und Anpassungskoeffizient wurde im Rahmen der Diplomarbeit deren Funktion interpretiert und durch graphische Veranschaulichungen untermauert.

Zur weiteren Erforschung des Kointegrationskonzepts ist ohne Zweifel die Interpretation von Kointegrationsvektor und Anpassungskoeffizient, die beide das Konzept maßgeblich konstituieren und charakterisieren, zunehmend in den Mittelpunkt der psychologisch-methodologischen Betrachtung zu rücken.

Die vorliegende Arbeit kann sowohl aus theoretischer als auch aus empirischer Perspektive als „Informations-Basis“ für die weiterführende Untersuchung verwandter Konzepte, wie beispielsweise von methodologischen Ansätzen im Zusammenhang mit fraktionaler, saisonaler oder Multi-Kointegration, und für deren empirische Validierung, z.B. durch Reanalysen, nützlich sein oder allgemein als Leitfaden in der psychologische Forschungspraxis dienen.

Literatur

Baillie, R.T. & Bollerslev, T. (1994). Cointegration, fractional cointegration, and exchange rate dynamics. *Journal of Finance*, 49, 737–745.

Baillie, R.T. & King, M.L. (Hrsg.) (1996). Annals of econometrics: fractional differencing and long memory processes. *Journal of Econometrics*, Vol. 73. The whole volume contains only articles on the topic of fractional differencing and long memory processes.

Banerjee, A., Dolado, J.J., Galbraith, J.W. & Hendry, D.F. (1993). *Co-integration, error correction, and the econometric analysis of non-stationary data*. Oxford: Oxford University Press.

Bank of Sweden / The Royal Swedish Academy of Sciences. The Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel 2003. (2003a). <http://www.nobel.se/economics/laureates/2003/index.html>.

Bank of Sweden / The Royal Swedish Academy of Sciences. Time-series econometrics: cointegration and autoregressive conditional heteroskedasticity. (2003b). <http://www.nobel.se/economics/laureates/2003/index.html>. Advanced information on the Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel.

Barkoulas, J.T., Baum, Ch. F. & Oguz, G.S. Fractional dynamics in a system of long term international interest rates. (1996). Boston College Working Papers in Economics 315., Boston College Department of Economics <http://fmwww.bc.edu/EC-P/WP315.pdf>. Außerdem erschienen unter dem Titel: *Fractional cointegration analysis of long term international interest rates*, *International Journal of Finance* (1997), 9:2, pp. 586-606.

Bortz, J. (1993). *Statistik für Sozialwissenschaftler* (4. Aufl. ed.). Heidelberg: Springer-Verlag.

Bortz, J. (1999). *Statistik für Sozialwissenschaftler* (5. Aufl. ed.). Heidelberg: Springer-Verlag.

Boucher, C. (2003). Testing for rational bubbles with time varying risk premium and non-linear cointegration: evidence from the US and French stock markets. Technical Report, Paris: Universität Paris-Nord, France.

Box, G.E.P. & Jenkins, G.M. (1976). *Time series analysis: forecasting and control (rev. ed.)*. San Francisco: Holden-Day.

Breitung, J. & Hassler, U. (2000). Inference on the cointegration rank in fractionally integrated processes. Technical Report: Humboldt University Berlin / Free University Berlin.

Bremmer, D. & Kesselring, R. (2004). Divorce and female labor force participation: evidence from time-series data and cointegration. *Atlantic Economic Journal*, 32,3, 175–190.

Bürgy, R. (1998). *Die Zeitreihenanalytische Modellierung psychologischer Regulationsvorgänge: zur Anwendung von ARIMA-Modellen in der Psychologie*. Hamburg: Kovač.

Brocklebank, J.C. & Dickey, D.A. (1986). *SAS System for forecasting time series*. Cary, NC, USA: SAS Institute Inc.

Brunetti, C. & Gilbert, Ch.L. Bivariate FIGARCH and fractional cointegration. (1999). <http://www.econ.qmul.ac.uk/papers/docs/wp408.pdf>. Edinburgh: University of Edinburgh.

Buchner, A. & Funke, J. (1993). Finite state automata: Dynamic task environments in problem solving research. *Quarterly Journal of Experimental Psychology*, 46A, 83–118.

Caporale, G.M. & Gil-Alana, L.A. (2002). Unemployment and input prices: a fractional cointegration approach. *Applied Economic Letters*, 9(6), 347–51.

Cromwell, J.B., Hannan, M.J., Labys, W.C. & Terraza, M. (1994). *Multivariate tests for time series models*. # 100. Thousand Oaks, London: Sage Publications.

- Darrat, A.F. & Al-Yousif, Y.K. (1999). On the long-run relationship between population and economic growth: Some time series evidence for developing countries. *Eastern Economic Journal*, 25,3, 301–306.
- DeJong, D., Nankervis, J.C., Savin, N.E. & Whiteman, C.H. (1992a). The power problems of unit root tests in time series with autoregressive errors. *Journal of Econometrics*, 53, 323–343.
- DeJong, D., Nankervis, J.C., Savin, N.E. & Whiteman, C.H. (1992b). Integration versus trend stationarity in time series. *Econometrica*, 60, 423–433.
- Delignieres, D., Fortes, M. & Ninot, G. The fractal dynamics of self-esteem and physical self. *Nonlinear dynamics in Psychology and Life Sciences* (in press).
- Dickey, D.A. & Fuller, W.A. (1979). Distribution of estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 427–431.
- Dickey, D.A. & Fuller, W.A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49, 1057–1072.
- Dittmann, I. (2004). Error correction models for fractionally cointegrated time series. *Journal of Time Series Analysis*, 25, 27–32(6).
- Drobetz, W. (2003). Nichtstationarität und Kointegration: Eine Einführung. Technical Report, Basel: Universität Basel (WWZ), Abteilung Finanzmarkttheorie.
- Dueker, M. & Startz, R. Maximum-Likelihood estimation of fractional cointegration with an application to U.S. and Canadian bond rates. (1994). <http://research.stlouisfed.org/wp/1994/94-027.pdf>. St. Louis: Federal Reserve Bank.
- Dueker, M. & Startz, R. (Working Paper, October 1995). Maximum likelihood estimation of fractional cointegration with an application to the short end of the yield curve. Technical Report: University of Washington.

Dupagne, M. & Green, R.J. (1996). Revisiting the principle of relative constancy: Consumer mass media expenditures in Belgium. *Communication Research*, 23(5), 612–635.

Engle, R. & Yoo, S. (1991). Cointegrated economic time series: an overview with new results. In Engle, R. & Granger, C.W.J. (Eds.), *Long-run economic relationships. Readings in cointegration*, pp. 237–265. New York: Oxford University Press.

Engle, R.F. & Granger, C.W.J. (1991). *Long-run economic relationships. Readings in cointegration*. New York: Oxford University Press.

Engle, R.F. & Granger, W.J. (1987). Co-Integration and error correction: representation, estimation and testing. *Econometrica*, 55(2), 251–276.

Engle, R.F. & Yoo, B.S. (1987). Forecasting and testing in co-integrated systems. *Journal of Econometrics*, 35, 143–159.

Erdfelder, E. & Funke, J. (Hrsg.). (2004). *Allgemeine Psychologie und deduktivistische Methodologie*. Göttingen: Vandenhoeck & Ruprecht.

Farrell, S., Wagenmakers, E.-J. & Ratcliff, R. ARFIMA time series modeling of serial correlations in human performance. (2004). <http://eis.bris.ac.uk/~pssaf/-arfimapmsubmitted.pdf>. Clifton Bristol: University of Bristol.

Fisher, R.A. (1925). Theory of statistical estimation. *Proceedings of the Cambridge Philosophical Society*, 22, 700–725.

Fuller, W.A. (1976). *Introduction to statistical time series*. New York: Wiley.

Gilden, D.L. (1997). Fluctuations in the time required for elementary decisions. *Psychological Science*, 8, 296–301.

Gilden, D.L. (2001). Cognitive emissions of 1/f noise. *Psychological Review*, 108, 33–56.

Gilden, D.L., Thornton, T. & Mallon, M.W. (1995). 1/f noise in human cognition. *Science*, 267, 1837–1839.

- Gilden, D.L. & Wilson, S.G (1995a). On the nature of streaks in signal detection. *Cognitive Psychology*, 28, 17–64.
- Gilden, D.L. & Wilson, S.G. (1995b). Streaks in skilled performance. *Psychonomic Bulletin & Review*, 2, 260–265.
- Gottschalk, A., Bauer, M.S. & Whybrow, P.C. (1995). Evidence of chaotic mood variation in bipolar disorder. *Archives of General Psychiatry*, 52, 947–959.
- Granger, C.W.J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37, 424–438.
- Granger, C.W.J. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, 16, 121–130.
- Granger, C.W.J. (1988). Causality, cointegration, and control. *Journal of Economic Dynamics and Control*, 12, 551–559.
- Granger, C.W.J. (1991). Some recent generalizations of cointegration and the analysis of long-run relationships. In Engle, R. & Granger, C.W.J. (Eds.), *Long-run economic relationships. Readings in cointegration*, pp. 277–288. New York: Oxford University Press.
- Granger, C.W.J. & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1, 15–29.
- Granger, C.W.J. & Lee, T.W. (1990). Multicointegration. In Fomby, Th. B. & Rhodes, G.F. (Eds.), *Co-integration, spurious regressions, and unit-roots: advances in econometrics, Vol. 8*, pp. 71–84: JAI Press.
- Granger, C.W.J. & Newbold, P. (1974). Spurious regression in econometrics. *Journal of Econometrics*, 2, 111–120.
- Granger, C.W.J & Newbold, P. (2001). Spurious regression in econometrics. In Ghysels, E., Swanson, N.R. & Watson, M. W. (Eds.), *Essays in econometrics: Volume 2, Causality, integration and cointegration, and long memory: collected papers*

of Clive W. J. Granger (*Econometric Society Monographs*), pp. 109–118. New York: Cambridge University Press.

Hamilton, J.D. (1994). *Time series analysis*. Princeton, N.J.: Princeton University Press.

Hammersley, J.M. & Handscomb, D.C. (1965). *Monte Carlo methods*. London: Methuen.

Hartley, B. & Hawkes, T.O. (1970). *Rings, modules and linear algebra*. London: Chapman Hall.

Hassler, U. (1994). Einheitswurzeltests - Ein Überblick. *Allgemeines Statistisches Archiv*, 78, 207–228.

Hassler, U. (2001). Leitfaden zum Testen und Schätzen von Kointegration. Technical Report: FU Berlin.

Hassler, U. (2003). Zeitabhängige Volatilität und instationäre Zeitreihen: Zum Nobelpreis an Robert F. Engle und Clive W.J. Granger. *Wirtschaftsdienst: Wissenschaft für die Praxis*, 12, 1–6.

Hatanaka, M. (1996). *Time series based econometrics. Unit roots and cointegration*. Oxford: Oxford University Press.

Hautzinger, M., Baumgartner, P., Neßhöver, W. & Schmitz, B. (1987). Zeitreihenanalyse kognitiver Verhaltenstherapie bei depressiven Patienten. *Zeitschrift für Klinische Psychologie*, 16(3), 256–263.

Hays, W. (1994). *Statistics* (5. Aufl. ed.). Orlando, Florida: Harcourt Brace.

Holtemöller, O. (2002). Money and prices: an $I(2)$ analysis for the Euro area. Technical Report, Berlin: Humboldt-Universität, Sonderforschungsbereich 373.

Hosking, J.R.M. (1981). Fractional differencing. *Biometrika*, 68, 165–176.

Hylleberg, S., Engle, R.F., Granger, C.W.J. & Yeo, B.S. (1990). Seasonal integration and co-integration. *Journal of Econometrics*, 44, 215–238.

- Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12, 231–254.
- Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59, 1551–1580.
- Johansen, S. (1995a). A statistical analysis of cointegration for $I(2)$ variables. *Econometric Theory*, 11, 25–59.
- Johansen, S. (1995b). *Likelihood-based inference in cointegrated vector autoregressive models*. New York: Oxford University Press.
- Johansen, S. & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration - with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, 52, 169–210.
- Kalos, M.H. & Whitlock, P.A. (1986). *Monte Carlo methods*. New York: Wiley.
- Kleijnen, J.P.C. (1987). *Statistical tools for simulation practitioners*. New York: Marcel Dekker.
- Krämer, W. (1997). Kointegration von Aktienkursen. Technical Report: Fachbereich Statistik, Universität Dortmund.
- Lanne, M. & Lütkepohl, H. (2001). Unit root tests for time series with level shifts: a comparison of different proposals. Technical Report: Humboldt University Berlin.
- Lardic, S. & Mignon, V. (2004). Fractional cointegration and the term structure. *Empirical Economics*, 181, 1–14.
- Lewin, K. (1982). *Feldtheorie (= Kurt-Lewin-Werkausgabe, Band 4)*. Bern: Huber. Herausgegeben von Carl-Friedrich Graumann.
- Lütkepohl, H. (1991). *Introduction to multiple time series analysis*. New York: Springer.

Luiz, J.-M. (2001). Temporal association, the dynamics of crime, and their economic determinants: A time series econometric model of South Africa. *Social Indicators Research*, 53(1), 33–61.

Maddala, G. & Kim, I.-M. (1998). *Unit roots, cointegration and structural change*. Cambridge: Cambridge University Press.

McCleary, R. & Hay, R.A. (1980). *Applied time series analysis for the social sciences*. London: Sage Publications.

McNown, R. (2003). A cointegration model of age-specific fertility and female labor supply in the United States. *Southern Economic Journal*, 70(2), 344–358.

Metropolis, N. & Ulam, S. (1949). The Monte Carlo method. *J. Am. Statist. Assoc.*, 44, 335.

Milligan, G.W. (1981). A review of Monte Carlo tests of cluster analysis. *Multivariate behavior research*, 16, 379–407.

Mills, T.C. (1990). *Time series techniques for economists*. Cambridge: Cambridge University Press.

Murray, M.P. (1994). A drunk and her dog: an illustration of cointegration and error correction. *The American Statistician*, 48, No. 1, 37–39.

O'Brien, R.M. (1999). Measuring the convergence/divergence of „serious crime“ arrest rates for males and females: 1960-1995. *Journal of Quantitative Criminology*, 15(1), 97–114.

Ortseifen, C. (1997). *Der SAS Kurs. Eine leicht verständliche Einführung*. Bonn: Internat. Thomson Publ.

Pagan, A. (2004d). Economics 607: applied macro-econometrics. Spring 2004. Topic 11: Co-integration. Technical Report: John Hopkins University.

Peterson, B.S. & Leckman, J.F. (1998). The temporal dynamics of the tics in Gilles de la Tourette syndrome. *Biological Psychiatry*, 44, 1337–1348.

Phillips, P.C.B. (1986). Understanding spurious regression in econometrics. *Journal of Econometrics*, 33, 311–340.

Phillips, P.C.B. (1997). The ET interview: Professor Clive Granger. *Econometric Theory*, 13, 253–303.

Phillips, P.C.B. & Durlauf, S.N. (1986). Multiple time series regressions with integrated processes. *Review of Economic Studies*, LIII, 473–495.

Pöppe, C. (2003). Nobelpreis für Wirtschaftswissenschaften - Zeitreihen in der Ökonometrie. *Spektrum der Wissenschaft*, Dezember 2003, S. 23 ff.

Rinne, H. & Specht, K. (2002). *Zeitreihen – Statistische Modellierung, Schätzung und Prognose*. München: Vahlen.

Rubinstein, R.Y. (1981). *Simulation and the Monte Carlo method*. New York: Wiley.

SAS. (1990). *SAS Guide to Macro Processing, Version 6, Second edition*. Cary, NC: SAS Institute Inc.

SAS/ETS. (1999). *User's guide, Version 8, Volume 1 and 2*. Cary, North Carolina 27513: SAS Institute Inc.

SAS/ETS Software (2001). The Varimax Procedure. In *SAS/ETS Software: Changes and Enhancements, Release 8.2, Chapter 4*. Cary, North Carolina 27513: SAS Institute Inc. <http://support.sas.com/rnd/app/da/new/801ce/ets/a-index.htm>.

SAS/GRAPH. (1991). *Software: Usage, Version 6, First edition*. Cary, NC: SAS Institute Inc.

SAS/IML. (1990). *Usage and Reference, Version 6, First edition*. Cary, NC: SAS Institute Inc.

SAS/Language. (1990). *Reference, Version 6, First edition*. Cary, NC: SAS Institute Inc.

Schiepek, G. (2003). A dynamic systems approach to clinical case formulation. *European Journal of Psychological Assessment*, 19, 175–184.

Schlittgen, R. & Streitberg, B.H.J. (1994). *Zeitreihenanalyse* (5., völlig überarb. und erw. ed.). München: R. Oldenbourg Verlag.

Schmitz, B. (1989). *Einführung in die Zeitreihenanalyse: Modelle, Softwarebeschreibung, Anwendungen*. Bern: Huber.

Schmitz, B. & Wiese, B. (1999). Zeitreihenanalyse als Instrumentarium der Prozeßforschung. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 19, 215–218.

Schulz, C.A. (2003). Strategien zum komplexen Problemlösen. Eine experimentelle Analyse der Rolle von Analogien beim Wissenserwerb in komplexen, dynamischen Systemen. Diplomarbeit, Psychologisches Institut der Universität Heidelberg, Hauptstraße 47-51, 69117 Heidelberg.

Stadnytska, T. (2004). Stationaritätsverletzungen bei ARMA-Modellen: Eine Simulationsstudie. Diplomarbeit, Psychologisches Institut der Universität Heidelberg, Hauptstraße 47-51, 69117 Heidelberg.

Stapel, E. Purplemath - your algebra resource: the remainder theorem. (2004). <http://www.purplemath.com/modules/remaindr.htm>.

Stewart, J. & Gill, L. (1998). *Econometrics*. London: Prentice Hall.

Stier, W. (2001). *Methoden der Zeitreihenanalyse*. Berlin: Springer.

Stock, J.H. (1987). Asymptotic properties of least-squares estimators of cointegrating vectors. *Econometrica*, 55, 1935–1056.

Stock, J.H. & Watson, M.W. (1988). Testing for common trends. *Journal of the American Statistical Association*, 83, 1097–1107.

Stoll, R.R. (1952). *Linear algebra and matrix theory*. New York: McGraw-Hill Book Company, Inc.

Thome, H. (1997). Scheinregressionen, kointegrierte Prozesse und Fehlerkorrekturmodelle: Eine Strategie zur Modellierung langfristiger Gleichgewichtsbeziehungen und kurzfristiger Effekte. *Zeitschrift für Soziologie*, 26 (3), 202–221.

- Van Orden, G.C., Holden, J.G. & Turvey, M.T. (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology: General*, 132, 331–350.
- Van Orden, G.C., Moreno, M.A. & Holden, J.G. (2002). A proper metaphysics for cognitive performance. *Nonlinear Dynamics, Psychology, and Life Sciences*, 7, 47–58.
- Wagenmakers, E.-J., Farrell, S. & Ratcliff, R. Estimation and interpretation of $1/f^\alpha$ noise in human cognition. <http://rats.psy.ohio-state.edu/roger/pdf/fnoiseproof.pdf>. Psychonomic Bulletin and Review (in press).
- Ward, L.M. (2002). *Dynamical cognitive science*. Cambridge, MA: MIT Press.
- Werner, J. (1997). *Lineare Statistik - Allgemeines lineares Modell*. Weinheim: Beltz - Psychologie Verlags Union.
- Werner, J. (Hrsg.). (2005). *Zeitreihenanalysen mit Beispielen aus der Psychologie*. Berlin: Logos Verlag.
- Wiener, N. (1956). The theory of prediction. In Beckenbach, E.F. (Ed.), *Modern mathematics for engineers, series 1*, pp. 165–190. New York: McGraw-Hill.
- Witt, R. & Witte, A. (2000). Crime, prison, and female labor supply. *Journal of Quantitative Criminology*, 16(1), 69–85.
- Wolters, J. (2000). Dynamische Regressionsmodelle. Technical Report: FU Berlin.
- Yoshinaga, H., Miyazima, S. & Mitake, S. (2000). Fluctuation of biological rhythm in finger tapping. *Physica A*, 280, 582–586.
- Yule, G.U. (1926). Why do we sometimes get nonsense correlation between time series? *J. Roy. Stat. Soc.*, 89, 1–64.
- Zorn, C. Longitudinal data analysis. (2001). <http://www.emory.edu/POLS/zorn/-Classes/POLS571>. Course Information.

Anhang

Anhang A: Backward-Shift-Operator

Die Verwendung des Rückwärtsverschiebungs-Operators (oder Backward-Shift- bzw. Backshift-Operators) bzw. Lag-Operators B hat sich für die formale Behandlung stochastischer Prozesse sowohl in der univariaten als auch der multivariaten Zeitreihenanalyse bewährt und ist für den univariaten Fall folgendermaßen definiert:

$$B^j(x_t) := x_{t-j} \text{ für } j = 1, 2, \dots$$

wobei für eine Konstante c gilt (vgl. Stier, 2001, S. 44):

$$B^j(c) = c; j = 1, 2, \dots$$

Die Zeitreihe x_t wird dabei um j Zeitpunkte zurückverschoben.

Für die Rückwärtsverschiebung um nur einen Zeitpunkt ($j = 1$) gilt folglich:

$$B(x_t) = x_{t-1}$$

So kann beispielsweise im multivariaten Fall ein VAR(p)-Prozess

$$\mathbf{x}_t = \mathbf{c} + \Phi_1 \mathbf{x}_{t-1} + \Phi_2 \mathbf{x}_{t-2} + \dots + \Phi_p \mathbf{x}_{t-p} + \boldsymbol{\epsilon}_t$$

in Backshift-Operator-Schreibweise folgendermaßen dargestellt werden:

$$\Phi(B)\mathbf{x}_t = \mathbf{c} + \boldsymbol{\epsilon}_t.$$

Bei der Verschiebung einer Zeitreihe zunächst um j und dann zusätzlich um k Zeitpunkte ist mithilfe des Operators folgende Schreibweise möglich:

$$B^j B^k(x_t) = B^{j+k}(x_t) = x_{t-j-k}$$

Auch die Darstellung der Differenzenbildung kann durch den Operator erleichtert werden:

$$1. \text{ Differenz: } (1 - B)x_t = x_t - (B)x_t = x_t - x_{t-1}$$

$$2. \text{ Differenz: } (1 - B)^2 x_t = (1 - 2B + B^2)x_t = x_t - 2x_{t-1} + x_{t-2}$$

$$d\text{-te Differenz: } (1 - B)^d x_t$$

Anhang B: Granger-Kausalität

Das Kausalitätskonzept ist zwar „philosophisch vieldeutig (und auch umstritten)“ (Stier, 2001, S. 83), dennoch gilt als allgemein akzeptiert, dass sich die Korrelation zwischen zwei Variablen insofern von der Kausalität unterscheidet, als diese eine Ursache-Effekte-Beziehung mit einer Wirkrichtung darstellt. Die Vieldeutigkeit des Konzepts dagegen ist in der Uneinigkeit darüber begründet, wie Kausalität im Zusammenhang mit Variablen oder Datensätzen definiert werden soll, die nicht im Rahmen eines kontrollierten Experiments erhoben und analysiert werden. Hinter der für das Kointegrationskonzept relevanten Granger-Kausalität verbirgt sich die zunächst einfache Idee, dass die Ursache dem Effekt nicht zeitlich nachgelagert sein kann. Ist die Variable Y kausal für die Variable X , dann kann X unter Mitberücksichtigung der Variable Y genauer prognostiziert werden als wenn X nur auf Grundlage seiner eigenen Vergangenheit vorhergesagt werden würde. Nach Meinung von Stier (2001, S. 83) wird Kausalität in diesem Fall auf Prognostizierbarkeit reduziert.

Als Gütemaß für die Prognosegenauigkeit ist der mittlere quadratische Fehler (Mean Square Error (MSE)) zu betrachten, der bekannterweise zugleich Erwartungstreue und Varianz miteinschließt. Wenn der MSE des nicht-erwartungstreuen gegenüber dem des erwartungstreuen Schätzers kleiner ist, wird ein Bias durch eine geringere Varianz ausgeglichen (Werner, 1997, S. 24). Vor diesem Hintergrund ist Y dann Granger-kausal zu X , wenn gilt:

$$MSE_X(s|\Omega_t) < MSE_X(s|\Omega_t \setminus \{Y_h, h < t\})$$

Darin bezeichnen Ω_t die Menge, die alle möglichen bis zum (aber ausschließlich vom) Zeitpunkt t verfügbaren Informationen enthält; s den Prognosehorizont von mindestens $s = 1, 2, \dots$. Dabei wird postuliert, dass für X eine optimale Prognose - also mit minimalem MSE - vorliegt.

Folglich ist Y dann Granger-kausal zu X , wenn unter Berücksichtigung der Informationsmenge Ω_t der mittlere quadratische Prognosefehler von X für mindestens einen der Prognoseschritte $s = 1, 2, \dots$ kleiner ist, als wenn bei der Prognose von X die Historie von Y ($\{Y_h, h < t\}$) in der Informationsmenge Ω_t nicht enthalten ist. Granger-

Kausalität beschreibt also den Fall, in dem eine Berücksichtigung von Y zu einer genaueren Prognose von X führt, als wenn diese Variable ohne die Y -Information vorhergesagt würde. Speziell mit der Vorhersage in kointegrierten Systemen befasst sich eine Arbeit von Engle und Yoo (1987).

Granger selbst fasst in seiner „Nobel Lecture“ zusammen (Bank of Sweden / The Royal Swedish Academy of Sciences, 2003a):

The statement about causality has just two components:

1. The cause occurs before the effect; and
2. The cause contains information about the effect that is unique, and is in no other variable.

In vielen Quellen wird auch von Wiener-Granger-Kausalität gesprochen, da das Abhängigkeitskonzept sowohl auf Wiener (1956) als auch auf Granger (1969) zurückgeführt werden kann (vgl. Werner, 2005, S. 127f.).

Anhang C: Wiener Prozess

Ein Brownscher oder Wiener Prozess ist ein in kontinuierlicher Zeit ablaufender stochastischer Prozess W_t mit $W_t \sim N(0, t)$ und $W_0 = 0$, wobei seine Zuwächse ebenfalls normalverteilt sind. Ein Wiener Prozess ist nach Stier (2001, S. 288) als ein kontinuierlicher Random Walk oder als Grenzfall eines diskreten Random Walks zu begreifen, bei dem das Zeitintervall sukzessiver Prozesswerte gegen Null geht. In Stier (2001, S. 288) findet sich eine anschauliche Beschreibung dieser Prozesse, die Hinweise auf exponentielle/geometrische Wiener Prozesse und deren interessante mathematische Eigenschaften enthält.

Anhang D: SAS-Programme

Um es der Leserin bzw. dem Leser zu ermöglichen, die wesentlichen Programm- und Prozedurschritte zur Generierung der im Rahmen dieser Arbeit simulierten multivariaten Prozesse bei Bedarf selbst durchzuführen oder z.B. für eine andere Fragestellung zu modifizieren, seien an dieser Stelle einige kurze Erläuterungen zu

den Programmen gegeben, die in diesem Anhang - geordnet nach den verwendeten Makros und den zugehörigen Hauptprogrammen - aufgeführt sind. Dabei soll die Programmierung in SAS exemplarisch für einen VAR(1)-Prozess mit $\phi_{11} = 0.8$, $\phi_{12} = 0.2111111$, $\phi_{21} = 0.9$ und $\phi_{22} = 0.05$ beschrieben werden (vgl. z.B. Anhang D3, S. 147ff.):

```
*Phi-Matrix initialisieren;
%let phi11 = 0.8;
%let phi12 = 0.2111111;
%let phi21 = 0.9;
%let phi22 = 0.05;
```

Die Anzahl der Wiederholungen, die im Rahmen dieser Arbeit konstant bei 200 liegt, und die Anzahl der Messzeitpunkte, die in 50er-Schritten, in diesem Ausschnitt beginnend mit 50 Messzeitpunkten bis zu 350 (bei den Programmen zu den Testverfahren, vgl. Anhang D3, S. 147ff. bzw. Anhang D4, S. 153ff.) bzw. 500 Messzeitpunkten (bei den Programmen zur Schätzung von Kointegrationsvektor und Anpassungskoeffizient, vgl. Anhang D5, S. 157ff.) variiert werden, können vorab angegeben werden. Die Variation der Messzeitpunkte ist an die jeweilige Fragestellung angepasst. Der hier angeführte Ausschnitt bezieht sich auf einen Durchlauf mit 50 Messzeitpunkten. Die weiteren Variationen werden unten angehängt.

```
*Variierbare Anzahl der Meßzeitpunkte und der Wiederholungen;
*Anzahl der Messzeitpunkte variiert T=50 bis T=350 in 50er-Schritten;
*Anzahl der Wiederholungen konstant WD = 200;
##### T = 50 #####;
***** Anzahl der Meßzeitpunkte *****;
%let MZ=50;
***** Anzahl der Wiederholungen *****;
%let WD=200;
```

Außerdem kann eine Ausgabedatei (in diesem Fall: AD_2B_10) über einen Pfad spezifiziert werden, die deshalb permanent ist, weil sie in einer eigens angelegten Library (d_AD_2B) angelegt wird, damit in dieser Datei die relevanten Schlussergebnisse abgelegt und dauerhaft gespeichert werden können:

```
*Benennung der permanenten Ausgabedatei für die entsprechende Matrix;
%let ausgabedatei = d_AD_2B.AD_2B_10;
```

Darüber hinaus kann über die Normierung von $\beta_1 = 1$ das zur angegebenen Matrix zugehörige β_2 vorab errechnet werden, damit es bei der Simulation der Einzelreihen später als Gewichtungsfaktor für die Generierung der Linearkombination bzw. Differenzenreihe zur Verfügung steht.

```
*Berechnung von beta (Population);
*Beta1-Normierung;
%let beta1 = 1;
*Beta2;
%let beta2 = &phi12/(&phi11-1);
```

Die Startzahlen werden jeweils getrennt für die beiden Einzelreihen generiert, indem zunächst jeweils 200 Startzahlen in die temporären Dateien `rannor1` bzw. `rannor2` abgelegt werden.

```
*Erzeugung von Anfangswerten für die Funktion RANNOR;
*RANNOR1 für den Komponentenprozess y1t mit a1t;
data rannor1;
keep a1t3;
do ii=1 to &WD;
a1t1=ranuni(8643989);
a1t2=a1t1*264509213;
a1t3=round(a1t2);
output;
end;
run;
proc transpose data=rannor1 out=ran.rannor1;
run;
data ran.startwert1;
set ran.rannor1;
rename COL1-COL&WD = A1-A&WD _NAME_=rannor1;
run;
```

Diese Dateien werden danach in der Library `ran` unter demselben Namen permanent abgespeichert und dann in einer Datei (`startwerte`) in derselben Library zusammengefügt, wo sie für alle weiteren Prozedurschritte nun dauerhaft verfügbar sind.

```
data ran.startwerte;
merge ran.startwert1 ran.startwert2;
run;
```

Anschließend werden jeweils 200 Einzelreihen (`y1t` und `y2t`) und 200 Differenzenreihen (`ydt`) mit je 50 Messzeitpunkten erzeugt. Um eine sequentielle Generierung der Reihen zu umgehen, können die Variablen dabei über `arrays` definiert werden.

```

data ydiff;
set ran.startwerte;

keep y1t1-y1t&WD i ii;
keep y2t1-y2t&WD i ii;
keep ydt1-ydt&WD i ii;

array y1t{&WD} (&WD*0);
array y2t{&WD} (&WD*0);
array ydt{&WD} (&WD*0);

array y1t_1{&WD} (&WD*0);
array y2t_1{&WD} (&WD*0);

array a1t{&WD} (&WD*0);
array a2t{&WD} (&WD*0);

array A{&WD} (&WD*0);
array B{&WD} (&WD*0);

```

In diesem Zusammenhang haben die beiden `do`-Schleifen, die jeweils mit dem `end`-Befehl geschlossen werden, unterschiedliche Funktionen. Während die erste Schleife dem aktuellen Wert der Reihe y_t ³¹ seinem um einen Zeitpunkt zurückgesetzten Vorgängerwert y_{t-1} zuordnet und gleichzeitig dem aktuellen Zufallsschock a_t der Wert 0 zugewiesen wird, werden im Rahmen der zweiten Schleife alle Messzeitpunkte durchlaufen. In diesem Zusammenhang durchläuft die Reihe zunächst 100 Messzeitpunkte, die anschließend aber keine Berücksichtigung finden: es wird erst ab $i > 0$ gezählt. Über den Schritt `call rannor` wird für jede Zeitreihe aus der permanenten Datei `startwerte` ein neuer Startwert herausgegriffen.

```

do ii = 1 to &WD;
y1t(ii)=0; y2t(ii)=0; ydt(ii)=0; y1t_1(ii)=0; y2t_1(ii)=0; a1t(ii)=0; a2t(ii)=0;

do i = -100 to &MZ;

call rannor (A(ii), a1t(ii));
call rannor (B(ii), a2t(ii));

y1t(ii) = &phi11*y1t_1(ii) + &phi12*y2t_1(ii) + a1t(ii);
y2t(ii) = &phi21*y1t_1(ii) + &phi22*y2t_1(ii) + a2t(ii);
ydt(ii) = y1t(ii) + &beta2*y2t(ii);

y1t_1(ii) = y1t(ii);
y2t_1(ii) = y2t(ii);

```

³¹bzw. in der Notation des theoretischen Teils x_t


```

if i > 0 then output;
end;
end;
run;

```

Innerhalb der beiden `do`-Schleifen sind die zu simulierenden mathematischen Modelle der Komponentenprozesse und der Differenzenreihe aufgeführt, die aus dem einführenden theoretischen Teil bekannt sind (vgl. Abschnitt 1.1, S. 16ff.). So wird die durch den Kointegrationsvektor gewichtete Linearkombination `ydt` z.B. durch die Gleichung

$$ydt(ii) = y1t(ii) + \beta_2 y2t(ii)$$

spezifiziert. Es wird deutlich, dass bei der Simulation der Zeitreihen eine direkte Übertragung der mathematischen Modelle möglich ist.

Die bisher beschriebenen Programmier-Schritte sind in ihrer Grundstruktur allen Simulationsprogrammen gemeinsam. In der Regel werden in den Simulationsprogrammen anschließend die Daten zunächst geordnet und für die Auswertung sinnvoll in Dateien abgelegt. Das geschieht unter anderem über ein Makro (hier: `diffset_eig`) und durch ein auf die Fragestellung zugeschnittenes Datenmanagement. In weiteren Schritten ermöglichen einzelne Makros, dass die Auswertungsprozedur `proc varmax` für die Einzelreihen und die Differenzenreihe mit jeweils 200 Wiederholungen und den wie oben beschrieben variierten Messzeitpunkten durchgeführt werden kann. Die für die Hypothesen relevanten Auswertungsergebnisse werden wiederum zielgerichtet aus den temporären Dateien herausgegriffen und in permanenten Dateien abgelegt. Danach sind weitere Schritte möglich. So können z.B. arithmetisches Mittel oder Minimal- bzw. Maximalwerte über die Prozedur `proc univariate` (ebenfalls über ein Makro und damit über alle Wiederholungen) bestimmt oder die Ergebnisse durch die Prozedur `proc freq` ausgezählt werden. Die für jede Matrix generierten Ergebnisdateien werden abschließend zu einer Gesamtdatei zusammengefügt:

```

##### ZUSAMMENFÜGEN DER EINZELDATEIEN #####;
data d_AD_1A.d_COMBINE_AD_1A;
set d_AD_1A.AD_1A_1 d_AD_1A.AD_1A_2 d_AD_1A.AD_1A_3 d_AD_1A.AD_1A_4
    d_AD_1A.AD_1A_5 d_AD_1A.AD_1A_6 d_AD_1A.AD_1A_7 d_AD_1A.AD_1A_8
    d_AD_1A.AD_1A_9 d_AD_1A.AD_1A_10 d_AD_1A.AD_1A_11 d_AD_1A.AD_1A_12
    d_AD_1A.AD_1A_13 d_AD_1A.AD_1A_14 d_AD_1A.AD_1A_15;
run;

```

Anhang D1: SAS-Programm zur Generierung von Φ_1 -Matrizen mit einer Einheitswurzel

```

proc iml;
start phikoeff;
reset print;
print, "Gleichung zur Bestimmung der phi-Elemente für eigval(phi)=1";
*phi12 = .05;
*phi21 = .1;
*phi22 = .9;
*phi11 = 1 - (phi12*phi21)/(1-phi22);
*phi11 = .95 ;
*phi21 = .1 ;
*phi22 = .9 ;
*phi12 = (1 - phi11 - phi22 + phi11*phi22)/phi21;
*auflösen nach phi22;
phi11 = .6 ;
phi12 = -.4;
phi21 = -.55;
phi22 = (1 -phi11-phi12*phi21)/(1-phi11);
finish;
run phikoeff;
quit;
proc iml;
start rang;
reset noprint;
print, "Test auf vollen Rang - Determinante";
*phi = {phi11 phi12,
*      phi21 phi22};
phi = {.9 -.1,
      -.2 .8};
d = det(phi);
print d;
finish;
run rang;
proc iml;
start eigenwert;
reset noprint;
print, "Überprüfung, ob eigval(phi)=1";
*phi = {phi11 phi12,
*      phi21 phi22};
phi = {.1 -.9,
      -.7 .3};
val = EIGVAL(phi);
print val;
finish;
run eigenwert;
quit;

```

Anhang D2: SAS-Programm zur Bestimmung der Populationsparameter von Anpassungskoeffizient α und Kointegrationsvektor β über das Fehlerkorrekturmodell

```

Proc IML;
start alphabeta;
reset noprint;
phi11 = .84 ;
phi12 = .8;
phi21 = .1;
phi22 = .5;
beta1 = 1;
beta2 = phi12/(phi11-1);
print beta2;
alpha1 = phi11-1;
alpha2 = phi21;
print alpha1;
print alpha2;
print, "Probe => a=b?";
a = phi22-1;
b = phi21*phi12/(phi11-1);
print a;
print b;
finish;
run alphabeta;
quit;

```

Anhang D3: SAS-Simulationsprogramm zum Dickey-Fuller-Kointegrationstest

Verwendete Makros

```

##### MAKRO 1 #####;
%macro diffset_eig(datei);
%do i=1 %to &WD;
    data dat&i;
        set &datei;
        where ii=&i;
        keep y1t&i y2t&i ydt&i;
    %end ;

    data ydiff;
%do i=1 %to &WD;
    merge dat&i;
%end ;
run;
%mend diffset_eig;

##### MAKRO 2 #####;
%macro varmaxdf_eig(datei);

```

```

%do i=1 %to &WD;
proc varmax data=&datei;
model y1t&i y2t&i / p=1 noint DFTEST;
quit;
%end;

%mend varmaxdf_eig;

##### MAKRO 3 #####;
%macro varmaxdfydt_eig(datei);

%do i=1 %to &WD;
proc varmax data=&datei;
model ydt&i / p=1 noint DFTEST;
quit;
%end;

%mend varmaxdfydt_eig;

```

Hauptprogramm

```

*Dickey-Fuller Unit Root Test über PROC VARMAX;
*für die Differenzenreihe ydt = y1t + beta*y2t und y1t und y2t ;
*über proc varmax;
*Variierbare Anzahl der Meßzeitpunkte und der Wiederholungen;
*Anzahl der Messzeitpunkte variiert T=50 bis T=350 in 50er-Schritten;
*Anzahl der Wiederholungen konstant WD = 200;
##### T = 50 #####;
***** Anzahl der Meßzeitpunkte *****;
%let MZ=50;
***** Anzahl der Wiederholungen *****;
%let WD=200;
*Phi-Matrix initialisieren;
%let phi11 = 0.8;
%let phi12 = 0.2111111;
%let phi21 = 0.9;
%let phi22 = 0.05;
*Benennung der permanenten Ausgabedatei für die entsprechende Matrix;
%let ausgabedatei = d_AD_2B.AD_2B_10;
*Berechnung von beta (Population);
*Beta1-Normierung;
%let beta1 = 1;
*Beta2;
%let beta2 = &phi12/(&phi11-1);
*Erzeugung von Anfangswerten für die Funktion RANNOR;
*RANNOR1 für den Komponentenprozess y1t mit a1t;
data rannor1;
keep a1t3;
do ii=1 to &WD;
a1t1=ranuni(8643989);
a1t2=a1t1*264509213;
a1t3=round(a1t2);

```

```

output;
end;
run;
proc transpose data=rannor1 out=ran.rannor1;
run;
data ran.startwert1;
set ran.rannor1;
rename COL1-COL&WD = A1-A&WD _NAME_=rannor1;
run;
*RANNOR2 für den Komponentenprozess y2t mit a2t;
data rannor2;
keep a2t3;
do ii=1 to &WD;
a2t1=ranuni(1325372);
a2t2=a2t1*985476322;
a2t3=round(a2t2);
output;
end;
run;
proc transpose data=rannor2 out=ran.rannor2;
run;
data ran.startwert2;
set ran.rannor2;
rename COL1-COL&WD = B1-B&WD _NAME_=rannor2;
run;
data ran.startwerte;
merge ran.startwert1 ran.startwert2;
run;
#####;
data ydiff;
set ran.startwerte;

keep y1t1-y1t&WD i ii;
keep y2t1-y2t&WD i ii;
keep ydt1-ydt&WD i ii;

array y1t{&WD} (&WD*0);
array y2t{&WD} (&WD*0);
array ydt{&WD} (&WD*0);

array y1t_1{&WD} (&WD*0);
array y2t_1{&WD} (&WD*0);

array a1t{&WD} (&WD*0);
array a2t{&WD} (&WD*0);

array A{&WD} (&WD*0);
array B{&WD} (&WD*0);

do ii = 1 to &WD;
y1t(ii)=0; y2t(ii)=0; ydt(ii)=0; y1t_1(ii)=0; y2t_1(ii)=0; a1t(ii)=0; a2t(ii)=0;

do i = -100 to &MZ;

call rannor (A(ii), a1t(ii));

```

```

call rannor (B(ii), a2t(ii));

y1t(ii) = &phi11*y1t_1(ii) + &phi12*y2t_1(ii) + a1t(ii);
y2t(ii) = &phi21*y1t_1(ii) + &phi22*y2t_1(ii) + a2t(ii);
ydt(ii) = y1t(ii) + &beta2*y2t(ii);

y1t_1(ii) = y1t(ii);
y2t_1(ii) = y2t(ii);

if i > 0 then output;
end;
end;
run;
#####;
%diffset_eig(work.ydiff);
data ydiff;
set ydiff;
beta2 = &beta2;
t+1;
run;
filename routed 'output-filename';
proc printto print=routed new;
run;
*Makro für y1t und y2t;
options nonumber center nodate label pagesize=100 ;
%varmaxdf_eig(work.ydiff);

proc printto print=print;
run;
    data df_y1ty2t;
        infile routed;
        input  spalte1$ spalte2$ spalte3$ spalte4$ spalte5$ spalte6$ spalte7$
        spalte8$ spalte9$ spalte10$ spalte11$ spalte12$ @ ;
    *DF1;
    if spalte2 = 'Zero' and spalte3 = 'Mean' and spalte7 = '<.0001' then DF1 = .0001;
    if spalte2 = 'Zero' and spalte3 = 'Mean' and spalte7 ^= '<.0001' then DF1 = spalte7;
    *DF2;
    if spalte7 = 'Zero' and spalte8 = 'Mean' and spalte12 = '<.0001' then DF2 = .0001;
    if spalte7 = 'Zero' and spalte8 = 'Mean' and spalte12 ^= '<.0001' then DF2 = spalte12;
    keep DF1 DF2;
run;
data dickey1;
set df_y1ty2t;
keep DF1;
if DF1 = '' then delete;
run;
data dickey2;
set df_y1ty2t;
keep DF2;
if DF2 = '' then delete;
run;
data prob_df12;
merge dickey1 dickey2;
keep DF1 DF2;
run;

```

```

data sig_df12;
set prob_df12;
*Signifikanzen f. DF1(5%-Niveau);
if DF1 lt .05 then DF_y1t = 's';
if DF1 gt .05 then DF_y1t = 'n' ;
if DF1 = .05 then DF_y1t = 'n' ;
*und für DF2;
if DF2 lt .05 then DF_y2t = 's';
if DF2 gt .05 then DF_y2t = 'n' ;
if DF2 = .05 then DF_y2t = 'n' ;
run;

*Makro für Linearkombination ydt;
filename routed 'output-filename';
proc printto print=routed new;
run;

options nonumber center nodate label pagesize=100 ;
%varmaxdfydt_eig(work.ydiff);
proc printto print=print;
run;
    data df_ydiff;
        infile routed;
        input  spalte1$ spalte2$ spalte3$ spalte4$ spalte5$ spalte6$ spalte7$ spalte8$
            spalte9$ spalte10$ spalte11$ spalte12$ spalte13$ spalte14$ spalte15$ spalte16$ @ ;
        if spalte2 = 'Zero' and spalte3 = 'Mean' and spalte7 = '<.0001' then DFdiff = 0.0001;
        if spalte2 = 'Zero' and spalte3 = 'Mean' and spalte7 ^= '<.0001' then DFdiff = spalte7;
        keep DFdiff;
run;
data dickeydiff;
set df_ydiff;
keep DFdiff;
if DFdiff = '' then delete;
run;
data prob_diff;
set dickeydiff;
keep DFdiff;
run;
data sig_dfdiff;
set prob_diff;
*Signifikanzen f. DFdiff (5%-Niveau);
if DFdiff lt .05 then DF_ydt = 's';
if DFdiff gt .05 then DF_ydt = 'n' ;
if DFdiff = .05 then DF_ydt = 'n' ;
run;

data sig_all;
merge sig_df12 sig_dfdiff;
run;

proc freq data=sig_all;
*output out=frequencies;
tables DF_y1t/norow nocol nocum out=y1t;
tables DF_y2t/norow nocol nocum out=y2t;
tables DF_ydt/norow nocol nocum out=ydt;

```

```

run;

data y1t_r;
set y1t;
rename COUNT=COUNT_y1t PERCENT=PERCENT_y1t;
run;

data y2t_r;
set y2t;
rename COUNT=COUNT_y2t PERCENT=PERCENT_y2t;
run;

data ydt_r;
set ydt;
rename COUNT=COUNT_ydt PERCENT=PERCENT_ydt;
run;

data frequencies;
merge y1t_r y2t_r ydt_r;
run;

data cointegration_df;
##### Matrizen einfügen zum Nachvollziehen;
phi_11 = &phi11;
phi_12 = &phi12;
phi_21 = &phi21;
phi_22 = &phi22;
##### MZ und WD;
T = &MZ;
Wdh = &WD;
*globale Abhängigkeit über Nebendiagonale;
*Betrag von phi12;
k12 = &phi12;
n12 = (k12)**2;
nd12 = sqrt(n12);
*Betrag von phi21;
k21 = &phi21;
n21 = (k21)**2;
nd21 = sqrt(n21);
*globale Abh. als Summe der Beträge;
Abh_glob = nd12 + nd21;
set frequencies;
drop k12 n12 nd12 k21 n21 nd21;
run;
quit;
proc append base = &ausgabedatei data = work.cointegration_df force;
proc datasets library=work kill;
run;
quit;

##### ZUSAMMENFÜGEN DER EINZELDATEIEN #####;
data d_AD_1A.d_COMBINE_AD_1A;
set d_AD_1A.AD_1A_1 d_AD_1A.AD_1A_2 d_AD_1A.AD_1A_3 d_AD_1A.AD_1A_4
    d_AD_1A.AD_1A_5 d_AD_1A.AD_1A_6 d_AD_1A.AD_1A_7 d_AD_1A.AD_1A_8

```



```

d_AD_1A.AD_1A_9 d_AD_1A.AD_1A_10 d_AD_1A.AD_1A_11 d_AD_1A.AD_1A_12
d_AD_1A.AD_1A_13 d_AD_1A.AD_1A_14 d_AD_1A.AD_1A_15;
run;

```

Anhang D4: SAS-Simulationsprogramm zum Johansen-Kointegrations- test

Verwendete Makros

```

##### MAKRO 1 #####;
%macro johset_eig(datei);
  %do i=1 %to &WD;
    data dat&i;
      set &datei;
      where ii=&i;
      keep y1t&i y2t&i;
    %end ;

    data johansen;
    %do i=1 %to &WD;
      merge dat&i;
    %end ;
    run;
  %mend johset_eig;

##### MAKRO 2 #####;
%macro johansen_eig(datei);

  %do i=1 %to &WD;
  proc varmax data=&datei ;
  model y1t&i y2t&i / p=1 noint lagmax=0 cointttest=(johansen);
  quit;
  %end;

%mend johansen_eig;

```

Hauptprogramm

```

##### Johansen-Testverfahren;
*Variierbare Anzahl der Messzeitpunkte und Wiederholungen;
*Anzahl der Messzeitpunkte variiert T=50 bis T=500 in 50er-Schritten;
*Anzahl der Wiederholungen konstant WD = 200;
##### T = 50 #####;
##### Anzahl der Messzeitpunkte #####;
%let MZ=50;
##### Anzahl der Wiederholungen #####;
%let WD=200;
*Phi-Matrix initialisieren;
%let phi11= 0.8;
%let phi12= 0.2111111;
%let phi21= 0.9;

```

```

%let phi22= 0.05;
*Benennung der permanenten Ausgabedatei für die entsprechende Matrix;
%let ausgabedatei=j_AD_2B.AD_2B_10;
*Erzeugung von Anfangswerten für die Funktion RANNOR;
*RANNOR1 für den Komponentenprozess y1t mit a1t;
data rannor1;
keep a1t3;
do ii=1 to &WD;
a1t1=ranuni(8643989);
a1t2=a1t1*264509213;
a1t3=round(a1t2);
output;
end;
run;
proc transpose data=rannor1 out=ran.rannor1;
run;
data ran.startwert1;
set ran.rannor1;
rename COL1-COL&WD = A1-A&WD _NAME_=rannor1;
run;
*RANNOR2 für den Komponentenprozess y2t mit a2t;
data rannor2;
keep a2t3;
do ii=1 to &WD;
a2t1=ranuni(1325372);
a2t2=a2t1*985476322;
a2t3=round(a2t2);
output;
end;
run;
proc transpose data=rannor2 out=ran.rannor2;
run;
data ran.startwert2;
set ran.rannor2;
rename COL1-COL&WD = B1-B&WD _NAME_=rannor2;
run;
data ran.startwerte;
merge ran.startwert1 ran.startwert2;
run;
#####;
data johansen;
set ran.startwerte;

keep y1t1-y1t&WD i ii;
keep y2t1-y2t&WD i ii;

array y1t{&WD} (&WD*0);
array y2t{&WD} (&WD*0);

array y1t_1{&WD} (&WD*0);
array y2t_1{&WD} (&WD*0);

array a1t{&WD} (&WD*0);
array a2t{&WD} (&WD*0);

```

```

array  A{&WD} (&WD*0);
array  B{&WD} (&WD*0);

do ii = 1 to &WD;
y1t(ii)=0; y2t(ii)=0; y1t_1(ii)=0; y2t_1(ii)=0; a1t(ii)=0; a2t(ii)=0;

do i = -100 to &MZ;

call rannor (A(ii), a1t(ii));
call rannor (B(ii), a2t(ii));

y1t(ii) = &phi11*y1t_1(ii) + &phi12*y2t_1(ii) + a1t(ii);
y2t(ii) = &phi21*y1t_1(ii) + &phi22*y2t_1(ii) + a2t(ii);

y1t_1(ii) = y1t(ii);
y2t_1(ii) = y2t(ii);

if i > 0 then output;
end;
end;
run;
#####;
%johset_eig(work.johansen);
  data johansen;
    set johansen;
    t+1;
  run;

filename routed 'output-filename';
proc printto print=routed new;
run;

options nonumber center nodate label pagesize=100 ;

%johansen_eig(work.johansen);

  proc printto print=print;
run;
  data output;
    infile routed;
    input  spalte1$ spalte2$ spalte3$ spalte4$ spalte5$ spalte6$ spalte7$ spalte8$
    spalte9$ spalte10$ spalte11$ spalte12$ @ ;
    if spalte1 = '0' and spalte2 = '0' then trace0= spalte4;
    if spalte1 = '0' and spalte2 = '0' then critval0=spalte5;
  if spalte8 = '0' and spalte9 = '0' then trace0= spalte11;
    if spalte8 = '0' and spalte9 = '0' then critval0=spalte12;
    if spalte1 = '1' and spalte2 = '1' then trace1= spalte4;
    if spalte1 = '1' and spalte2 = '1' then critval1=spalte5;
  if spalte8 = '1' and spalte9 = '1' then trace1= spalte11;
    if spalte8 = '1' and spalte9 = '1' then critval1=spalte12;
    keep trace0 critval0 trace1 critval1;

run;

data zero;

```

```

set output;
keep trace0 critval0;
if trace0 ='' and critval0='' then delete;
run;
data one;
set output;
keep trace1 critval1;
if trace1 ='' and critval1='' then delete;
run;
data joh_test;
merge zero one;
tr0=input(trace0,10.);
krW0=input(critval0,10.);
tr1=input(trace1,10.);
krW1=input(critval1,10.);
keep tr0 krW0 tr1 krW1;
run;

data joh_sig;
set joh_test;
##### Signifikanz #####;
if tr0 gt krW0 then rang0 = 's' ;
if tr0 lt krW0 then rang0 = 'n' ;
if tr1 gt krW1 then rang1 = 's' ;
if tr1 lt krW1 then rang1 = 'n' ;
run;

proc freq data=joh_sig;
*output out=frequencies;
tables rang0/norow nocol nocum out=freq0;
tables rang1/norow nocol nocum out=freq1;
run;
data freq0_r;
set freq0;
rename COUNT=COUNT_0 PERCENT=PERCENT_0;
run;
data freq1_r;
set freq1;
rename COUNT=COUNT_1 PERCENT=PERCENT_1;
run;
data frequencies;
merge freq0_r freq1_r;
run;
data cointegration_joh;
##### Matrizen einfügen zum Nachvollziehen;
phi_11 = &phi11;
phi_12 = &phi12;
phi_21 = &phi21;
phi_22 = &phi22;
##### MZ und WD;
T = &MZ;
Wdh = &WD;
*globale Abhängigkeit über Nebendiagonale;
*Betrag von phi12;
k12 = &phi12;

```

```

n12 = (k12)**2;
nd12 = sqrt(n12);
*Betrag von phi12;
k21 = &phi21;
n21 = (k21)**2;
nd21 = sqrt(n21);
*globale Abh. als Summe der Beträge;
Abh_glob = nd12 + nd21;
set frequencies;
drop k12 n12 nd12 k21 n21 nd21;
run;
quit;
proc append base = &ausgabedatei data = work.cointegration_joh force;
proc datasets library=work kill;
run;
quit;

##### ZUSAMMENFÜGEN DER EINZELDATEIEN #####;
data j_AD_1A.j_COMBINE_AD_1A;
set j_AD_1A.AD_1A_1 j_AD_1A.AD_1A_2 j_AD_1A.AD_1A_3 j_AD_1A.AD_1A_4
    j_AD_1A.AD_1A_5 j_AD_1A.AD_1A_6 j_AD_1A.AD_1A_7 j_AD_1A.AD_1A_8
    j_AD_1A.AD_1A_9 j_AD_1A.AD_1A_10 j_AD_1A.AD_1A_11 j_AD_1A.AD_1A_12
    j_AD_1A.AD_1A_13 j_AD_1A.AD_1A_14 j_AD_1A.AD_1A_15;
run;

```

Anhang D5: SAS-Simulationsprogramm zur Schätzung von Anpassungs- koeffizient α und Kointegrationsvektor β

Verwendete Makros

```

##### MAKRO 1 #####;
%macro estset_eig(datei);
%do i=1 %to &WD;
    data dat&i;
        set &datei;
        where ii=&i;
        keep y1t&i y2t&i;
    %end;

    data estimate;
%do i=1 %to &WD;
    merge dat&i;
%end;
run;
%mend estset_eig;

##### MAKRO 2 #####;
%macro ecm_eig(datei);

```

```

%do i=1 %to &WD;
proc varmax data=&datei;
model y1t&i y2t&i / p=1 noint print=(iarr) ecm=(rank=1 normalize=y1t&i);
quit;
%end;

%mend ecm_eig;

##### MAKRO 3 #####;
%macro meansd_eig(datei);
proc univariate data=&datei noprint;
var B1 B2 A1 A2;
output out=goodness
MEAN=m_B1 m_B2 m_A1 m_A2
STD=s_B1 s_B2 s_A1 s_A2
MIN=min_B1 min_B2 min_A1 min_A2
MAX=max_B1 max_B2 max_A1 max_A2
Range=range_B1 range_B2 range_A1 range_A2;
run;
%mend meansd_eig;

```

Hauptprogramm

```

##### ALPHA und BETA: Güte der Schätzungen;
*Variierbare Anzahl der Meßzeitpunkte und der Wiederholungen;
*Anzahl der Messzeitpunkte variiert T=50 bis T=500 in 50er-Schritten;
*Anzahl der Wiederholungen konstant WD = 200;
##### T = 50 #####;
***** Anzahl der Meßzeitpunkte *****;
%let MZ=50;
***** Anzahl der Wiederholungen *****;
%let WD=200;

* Phi-Matrix initialisieren;
%let phi11= 0.8;
%let phi12= 0.2111111;
%let phi21= 0.9;
%let phi22= 0.05;
*Benennung der permanenten Ausgabedatei für die entsprechende Matrix;
%let ausgabedatei= e_AD_2B.AD_2B_10;
* Beta1-Normierung;
%let beta1 = 1;

*Erzeugung von Anfangswerten für die Funktion RANNOR;
*RANNOR1 für den Komponentenprozess y1t mit a1t;
data rannor1;
keep a1t3;
do ii=1 to &WD;
a1t1=ranuni(8643989);
a1t2=a1t1*264509213;
a1t3=round(a1t2);
output;
end;

```

```

run;
proc transpose data=rannor1 out=ran.rannor1;
run;
data ran.startwert1;
set ran.rannor1;
rename COL1-COL&WD = A1-A&WD _NAME_=rannor1;
run;
*RANNOR2 für den Komponentenprozess y2t mit a2t;
data rannor2;
keep a2t3;
do ii=1 to &WD;
a2t1=ranuni(1325372);
a2t2=a2t1*985476322;
a2t3=round(a2t2);
output;
end;
run;
proc transpose data=rannor2 out=ran.rannor2;
run;
data ran.startwert2;
set ran.rannor2;
rename COL1-COL&WD = B1-B&WD _NAME_=rannor2;
run;
data ran.startwerte;
merge ran.startwert1 ran.startwert2;
run;
#####;
data estimate;
set ran.startwerte;

keep y1t1-y1t&WD i ii;
keep y2t1-y2t&WD i ii;

array y1t{&WD} (&WD*0);
array y2t{&WD} (&WD*0);

array y1t_1{&WD} (&WD*0);
array y2t_1{&WD} (&WD*0);

array a1t{&WD} (&WD*0);
array a2t{&WD} (&WD*0);

array A{&WD} (&WD*0);
array B{&WD} (&WD*0);

do ii = 1 to &WD;
y1t(ii)=0; y2t(ii)=0; y1t_1(ii)=0; y2t_1(ii)=0; a1t(ii)=0; a2t(ii)=0;

do i = -100 to &MZ;

call rannor (A(ii), a1t(ii));
call rannor (B(ii), a2t(ii));

y1t(ii) = &phi11*y1t_1(ii) + &phi12*y2t_1(ii) + a1t(ii);
y2t(ii) = &phi21*y1t_1(ii) + &phi22*y2t_1(ii) + a2t(ii);

```

```

y1t_1(ii) = y1t(ii);
y2t_1(ii) = y2t(ii);

if i > 0 then output;
end;
end;
run;
#####;
%estset_eig(work.estimate);
data estimate;
set estimate;
t+1;
run;
filename routed 'output-filename';
proc printto print=routed new;
run;

options nonumber center nodate label pagesize=100 ;

%ecm_eig(work.estimate);
proc printto print=print;
run;
  data output;
    infile routed;
    input  spalte1$ spalte2$ spalte3$ spalte4$ spalte5$ spalte6$ spalte7$ spalte8$
          spalte9$ spalte10$ spalte11$ spalte12$ spalte13$ spalte14$ @ ;
    if spalte1 = 'Estimate' and spalte2 = 'given' then beta1 = spalte8;
    if spalte1 = 'Estimate' and spalte2 = 'given' then beta2 = spalte10;
    if spalte1 = 'given' and spalte2 = 'RANK=1' then alpha1 = spalte7;
    if spalte1 = 'given' and spalte2 = 'RANK=1' then alpha2 = spalte9;
    keep beta1 beta2 alpha1 alpha2;
run;
data beta;
set output;
keep beta1 beta2;
if beta1 = '' and beta2 = '' then delete;
run;
data alpha;
set output;
keep alpha1 alpha2;
if alpha1 = '' and alpha2 = '' then delete;
run;
data estall;
merge beta alpha;
*Umwandlung von Textvariable in numerische Variable;
B1=input(beta1,10.);
B2=input(beta2,10.);
A1=input(alpha1,10.);
A2=input(alpha2,10.);
keep B1 B2 A1 A2;
run;
*Berechnung der Kennwerte, v.a. Mittelwert und Streuung, mit proc univariate;
%meansd_eig(work.estall);
##### DATEIBEARBEITUNG #####;

```



```

data guete;
##### Matrizen einfügen zum Nachvollziehen;
phi_11 = &phi11;
phi_12 = &phi12;
phi_21 = &phi21;
phi_22 = &phi22;
##### globale Abhängigkeit über Nebendiagonale;
*Betrag von phi12;
k12 = &phi12;
n12 = (k12)**2;
nd12 = sqrt(n12);
*Betrag von phi21;
k21 = &phi21;
n21 = (k21)**2;
nd21 = sqrt(n21);
*globale Abhängigkeit als Summe der Beträge;
Abh_glob = nd12 + nd21;
##### Pop-Werte alpha und beta #####;
##### beta1_pop;
beta1_pop = &beta1;
##### beta2_pop;
beta2_pop = &phi12/(&phi11-1);
##### alpha1_pop;
alpha1_pop = &phi11-1;
##### alpha2_pop;
alpha2_pop = &phi21;
##### Probe muss 1 sein!;
probe = (&phi22-1)/(&phi21*&phi12/(&phi11-1));
##### Ergebnisse proc univariate;
set goodness;
##### MZ und WD;
T = &MZ;
Wdh = &WD;
##### MSE => s. Werner, 1997, S. 23f. #####;
##### MSE beta1;
MSE_B1 = (s_B1)**2 + (beta1_pop - m_B1)**2;
##### MSE beta2;
MSE_B2 = (s_B2)**2 + (beta2_pop - m_B2)**2;
##### MSE beta;
MSE_B = (MSE_B1 + MSE_B2)/2;
##### MSE alpha1;
MSE_A1 = (s_A1)**2 + (alpha1_pop - m_A1)**2;
##### MSE alpha2;
MSE_A2 = (s_A2)**2 + (alpha2_pop - m_A2)**2;
##### MSE alpha;
MSE_A = (MSE_A1 + MSE_A2)/2;
##### MSE alphabeta_ges;
MSE_AB = (MSE_A + MSE_B)/2;
drop k12 n12 nd12 k21 n21 nd21;
run;
quit;
proc append base = &ausgabedatei data = work.guete force;
proc datasets library=work kill;
run;
quit;

```

```

##### ZUSAMMENFÜGEN DER EINZELDATEIEN #####;
data e_AD_1A.COMBINE_AD_1A;
set e_AD_1A.AD_1A_1 e_AD_1A.AD_1A_2 e_AD_1A.AD_1A_3 e_AD_1A.AD_1A_4
    e_AD_1A.AD_1A_5 e_AD_1A.AD_1A_6 e_AD_1A.AD_1A_7 e_AD_1A.AD_1A_8
    e_AD_1A.AD_1A_9 e_AD_1A.AD_1A_10 e_AD_1A.AD_1A_11 e_AD_1A.AD_1A_12
    e_AD_1A.AD_1A_13 e_AD_1A.AD_1A_14 e_AD_1A.AD_1A_15;
run;

```

Anhang D6: SAS-Simulationsprogramm zur Replikation von Granger und Newbold (1974)

Verwendete Makros

```

%macro regress_eig(datei);

    %do i=1 %to 500;
proc reg data=&datei OUTEST=result&i;
model u&i=v&i;
quit;
%end;
%mend regress_eig;

```

Hauptprogramm

```

*REPLIKATION von GRANGER & NEWBOLD (1974): SCHEINREGRESSION;
*500 WIEDERHOLUNGEN;
*Anzahl der MZP variiert: unterschiedliche Auflösungen für Graphik;
##### MZ=20 #####;
%let MZ=20;
*Benennung der permanenten Ausgabedatei;
%let ausgabedatei = GN_1974.RW12_reg;
*1. Erzeugung von Anfangswerten für die Funktion Rannor (Prozess u);
data rannor1;
keep a3;
do i=1 to 500;
a1=ranuni(8643989);
a2=a1*264509213;
a3=round(a2);
output;
end;
run;
proc transpose data=rannor1 out=ran.startwert1;
run;
##### Random Walk 1: u #####;
data RW1;
set ran.startwert1;
keep u1-u500 t;
array u(500) u1-u500;
array au(500) au1-au500;
array ut1_(500) ut1_1-ut1_500;

```

```

array cols(500) col1-col500;
do i=1 to 500;
u(i)=0;au(i)=0;ut1_(i)=0;
end;
do t=1 to &MZ;
do i=1 to 500;
call rannor(cols(i), au(i));
ut1_(i) = u(i);
u(i) = u(i) + au(i);
end;
if t gt 0 then output;
end;
run;
#####;
*2. Erzeugung von Anfangswerten für die Funktion Rannor (Prozess v);
data rannor2;
keep a3;
do i=1 to 500;
a1=ranuni(1325372);
a2=a1*985476322;
a3=round(a2);
output;
end;
run;
proc transpose data=rannor2 out=ran.startwert2;
run;
##### Random Walk 2: v #####;
data RW2;
set ran.startwert2;
keep v1-v500 t;
array v(500) v1-v500;
array av(500) av1-av500;
array vt1_(500) vt1_1-vt1_500;
array cols(500) col1-col500;
do i=1 to 500;
v(i)=0;av(i)=0;vt1_(i)=0;
end;
do t=1 to &MZ;
do i=1 to 500;
call rannor(cols(i), av(i));
vt1_(i) = v(i);
v(i) = v(i) + av(i);
end;
if t gt 0 then output;
end;
run;
#####;
##### RW1 und RW2 #####;
*1 und 2*;
data RW1_reg_RW2;
merge RW1 RW2;
run;

filename routed 'output-filename';
proc printto print=routed new;

```

```
run;

options nonumber center nodate label pagesize=100 ;

%regress_eig(work.RW1_reg_RW2);
  proc printto print=print;
run;
  data reg_result;
  infile routed;
  input  spalte1$ spalte2$ spalte3$ spalte4$ spalte5$ spalte6$ spalte7$ spalte8$
  spalte9$ spalte10$ spalte11$ spalte12$ spalte13$ spalte14$ spalte15$ spalte16$
  spalte17$ spalte18$ @ ;
  if spalte1='Model' and spalte7='Error' and spalte6 ^= '<.0001' then _fprob_=spalte6;
  if spalte1='Model' and spalte7='Error' and spalte6 = '<.0001' then _fprob_=0.0001;
  keep _fprob_;
run;
data RW1_RW2_spurious;
set reg_result;
keep _fprob_;
if _fprob_ = '' then delete;
run;
data prob_RW12;
merge RW1_RW2_spurious;
F_PROB=input(_fprob_,10.);
keep F_PROB;
run;
data sig_RW12;
set prob_RW12;
if F_PROB lt .01 then FPROB = 'ss';
if F_PROB gt .01 lt .05 then FPROB = 's';
if F_PROB gt .05 then FPROB = 'n' ;
if F_PROB = .05 then FPROB = 'n' ;
run;
proc freq data=sig_RW12;
tables FPROB/norow nocol nocum out=FPROB;
run;
data FPROB;
set FPROB;
t = &MZ;
run;
proc append base = &ausgabedatei data = work.FPROB force;
*proc datasets library=work kill;
run;
quit;
```

Anhang D7: SAS-Simulationsprogramm zur Überprüfung der Faustregel von Granger und Newbold (1974)

Verwendete Makros

```

##### MAKRO 1 #####;
%macro spuriousset_eig(datei);
  %do i=1 %to &WD;
    data dat&i;
      set &datei;
      where ii=&i;
      keep y1t&i y2t&i;
  %end ;

  data spurious;
  %do i=1 %to &WD;
    merge dat&i;
  %end ;
  run;
%mend spuriousset_eig;

##### MAKRO 2 #####;
%macro regression1_eig(datei);

  %do i=1 %to &WD;
  proc reg data=&datei CORR;
  model y1t&i=y2t&i / noint DW;
  quit;
  %end;
%mend regression1_eig;

##### MAKRO 3 #####;
%macro regression2_eig(datei);

  %do i=1 %to &WD;
  proc reg data=&datei;
  model y1t&i=y2t&i / noint RSQUARE B;
  quit;
  %end;
%mend regression2_eig;

```

Hauptprogramm

```

*Indikation von Spurious Regression durch R2 und die DW-Statistik;

*Variierbare Anzahl der Meßzeitpunkte und der Wiederholungen;
***** Anzahl der Meßzeitpunkte *****;
%let MZ=20000;
***** Anzahl der Wiederholungen *****;
%let WD=20;

data spurious;

```

```

array y1t{&WD} (&WD*0);
array y2t{&WD} (&WD*0);
array y1t_1{&WD} (&WD*0);
array y2t_1{&WD} (&WD*0);

phi11 = 1;
phi12 = 0;
phi21 = 0;
phi22 = 1;

do ii = 1 to &WD;
do i = 1 to &MZ;

a1t = rannor(12345);
a2t = rannor(67890);
*y1t_1 = 0;
*y2t_1 = 0;
*t = 0;

y1t(ii) = phi11*y1t_1(ii) + phi12*y2t_1(ii) + a1t;
y2t(ii) = phi21*y1t_1(ii) + phi22*y2t_1(ii) + a2t;

y1t_1(ii) = y1t(ii);
y2t_1(ii) = y2t(ii);

put i ii y1t(ii) y2t(ii);
t = t+1;
output;
end;
end;
run;
%spuriousset_eig(work.spurious);
data spurious;
set spurious;
t+1;
run;

filename routed 'output-filename';
proc printto print=routed new;
run;

options nonumber center nodate label pagesize=100 ;
%regression1_eig(work.spurious);
proc printto print=print;
run;
data scheinregression1;
infile routed;
input spalte1$ spalte2$ spalte3$ spalte4$ spalte5$ spalte6$ spalte7$ spalte8$ @;
if spalte1 = 'Durbin-W' and spalte2 = 'D' then _DW_ = spalte3;
if spalte1 = 'Variable' and spalte5 = '1.0000' then _CORR_ = spalte6;
keep _DW_ _CORR_;
run;
data durbinwatson;
set scheinregression1;
keep _DW_;

```

```

if _DW_ = '' then delete;
run;
data correlation;
set scheinregression1;
keep _CORR_;
if _CORR_ = '' then delete;
run;

data statistics1;
merge durbinwatson correlation;
DW=input(_DW_,10.);
CORR=input(_CORR_,10.);
keep DW CORR;
run;
*2. Teil;
filename routed 'output-filename';
proc printto print=routed new;
run;

options nonumber center nodate label pagesize=100 ;
%regression2_eig(work.spurious);
proc printto print=print;
run;
data scheinregression2;
  infile routed;
  input  spalte1$ spalte2$ spalte3$ spalte4$ spalte5$ spalte6$ spalte7$ spalte8$
        spalte9$ spalte10$ spalte11$ spalte12$ spalte13$ spalte14$ spalte15$ spalte16$
        spalte17$ spalte18$ spalte19$ @;
if spalte12 = 'R-Square' and spalte18 = 'R-Sq' then _RSQ_ = spalte13;
if spalte1 = 'Coeff' and spalte5 = 'Estimate' then _REGCOEFF_ = spalte19;
keep _REGCOEFF_ _RSQ_;
run;
data rsquare;
set scheinregression2;
keep _RSQ_;
if _RSQ_ = '' then delete;
run;
data regcoeff;
set scheinregression2;
keep _REGCOEFF_;
if _REGCOEFF_ = '' then delete;
run;

data statistics2;
merge regcoeff rsquare;
RSQ=input(_RSQ_,10.);
REGCOEFF=input(_REGCOEFF_,10.);
keep RSQ REGCOEFF;
run;
data statistics;
set statistics1;
set statistics2;
run;

data scheinrg.Statistiken;

```

```

set statistics;
*DW;
if DW > 2.2 then autocorrelation = 'negative' ;
if DW < 1.8 then autocorrelation = 'positive' ;
if DW > 1.8 and DW < 2.2 then autocorrelation = 'uncorrelated' ;
*CORR;
if CORR > 0 and CORR > 0.25 and CORR < 0.75 then correlation =
'moderate/positive';
if CORR > 0 and CORR > 0.75 then correlation = 'high/positive';
if CORR > 0 and CORR < 0.25 then correlation = 'low/positive';
if CORR < 0 and CORR < -0.25 and CORR > -0.75 then correlation =
'moderate/negative';
if CORR < 0 and CORR < -0.75 then correlation = 'high/negative';
if CORR < 0 and CORR > -0.25 then correlation = 'low/negative';
*RSQ;
if RSQ > 0.25 and RSQ < 0.75 then rsquare = 'moderate';
if RSQ > 0.75 then rsquare = 'high';
if RSQ < 0.25 then rsquare = 'low';
*REGCOEFF;
if REGCOEFF > 0 and REGCOEFF > 0.25 and REGCOEFF < 0.75 then regcoefficient =
'moderate/positive';
if REGCOEFF > 0 and REGCOEFF > 0.75 then regcoefficient = 'high/positive';
if REGCOEFF > 0 and REGCOEFF < 0.25 then regcoefficient = 'low/positive';
if REGCOEFF < 0 and REGCOEFF < -0.25 and REGCOEFF > -0.75 then regcoefficient =
'moderate/negative';
if REGCOEFF < 0 and REGCOEFF < -0.75 then regcoefficient = 'high/negative';
if REGCOEFF < 0 and REGCOEFF > -0.25 then regcoefficient = 'low/negative';
run;
proc freq data=scheinrg.Statistiken;
tables autocorrelation/norow nocol nopercent nocum;
tables correlation/norow nocol nopercent nocum;
tables rsquare/norow nocol nopercent nocum;
tables regcoefficient/norow nocol nopercent nocum;
run;
proc corr data=scheinrg.statistiken;
run;
proc datasets library=work kill;
run;
quit;

```

Anhang D8: SAS-Programme zur Generierung von Diagrammen und Graphiken

```

##### DIAGRAMM ZUR REPLIKATION VON GRANGER UND NEWBOLD #####;
data GN_1974.SPREG;
set GN_1974.RW12_reg;
if FPROB='' then delete;
rename FPROB=regression;
run;

title1 'Replikation: Granger und Newbold (1974)';

```



```

title2 'Prozentsatz signifikanter Regressionen: 2 unabh. Zufallsprozesse';
title3 '500 Wiederholungen';
title4 'T=20 bis T=300 (variiert in 20er-Schritten)';
title5 'Signifikanzniveau: 0.05 (s), 0.01 (ss); nichtsignifikant: 0.05 (n) ';
footnote1 justify=right 'Diplomarbeit ESK';

axis1 order= (0 to 300 by 20) label=('Variierte Anzahl der Messzeitpunkte T');
axis2 order= (0 to 100 by 25) label=('Prozent');

symbol1 value=star interpol=join color=black height=3 width=2;
symbol2 v=diamond i=join color=black height=3;
symbol3 v=dot i=join color=black height=3;

proc gplot data=GN_1974.SPREG;
plot PERCENT*t=regression /haxis=axis1 vaxis=axis2 vref=75 lvref=2 frame;
run;
quit;

title1 'Replikation: Granger und Newbold (1974)';
title2 'Prozentsatz signifikanter Regressionen: 2 unabh. Zufallsprozesse';
title3 '500 Wiederholungen';
title4 'T=20 bis T=200 (variiert in 20er-Schritten)';
title5 'Signifikanzniveau: 0.05 (s), 0.01 (ss); nichtsignifikant: 0.05 (n) ';
footnote1 justify=right 'Diplomarbeit ESK';

axis1 order= (0 to 200 by 20) label=('Variierte Anzahl der Messzeitpunkte T');
axis2 order= (0 to 100 by 25) label=('Prozent');

symbol1 value=star interpol=join color=black height=3 width=2;
symbol2 v=diamond i=join color=black height=3;
symbol3 v=dot i=join color=black height=3;

proc gplot data=GN_1974.SPREG;
plot PERCENT*t=regression /haxis=axis1 vaxis=axis2 vref=75 lvref=2 frame;
run;
quit;

title1 'Replikation: Granger und Newbold (1974)';
title2 'Prozentsatz signifikanter Regressionen: 2 unabh. Zufallsprozesse';
title3 '500 Wiederholungen';
title4 'T=20 bis T=100 (variiert in 20er-Schritten)';
title5 'Signifikanzniveau: 0.05 (s), 0.01 (ss); nichtsignifikant: 0.05 (n) ';
footnote1 justify=right 'Diplomarbeit ESK';

axis1 order= (0 to 100 by 20) label=('Variierte Anzahl der Messzeitpunkte T');
axis2 order= (0 to 100 by 25) label=('Prozent');

symbol1 value=star interpol=join color=black height=3 width=2;
symbol2 v=diamond i=join color=black height=3;
symbol3 v=dot i=join color=black height=3;

proc gplot data=GN_1974.SPREG;
plot PERCENT*t=regression /haxis=axis1 vaxis=axis2 vref=75 lvref=2 frame;
run;
quit;

```

```

##### VERLAUFSDIAGRAMME DER MULTIVARIATEN PROZESSE #####;
proc gplot data=single;
  symbol1 v=none i=join l=1 c=black;
  symbol2 v=none i=join l=2 c=red;
  symbol3 v=none i=join l=3 c=green;

  plot y1t * t = 1
        y2t * t = 2
        ydiff * t = 3 / overlay ;
run;

##### EXEMPLARISCHE PROGRAMME FÜR DIE ERGEBNISSCHAUBILDER #####;

***** BEISPIEL 1 *****;
*DICKEY-FULLER-TEST --- nicht-auseinanderdriftende Zeitpfade ---;
*NAD_0;
data EMP_GRAP.EMP_DF_NAD_0;
set D_NAD_0.D_COMBINE_NAD_0;
keep Abh_glob T DF_ydt PERCENT_ydt;
if DF_ydt='n' then delete;
if PERCENT_ydt='.' then delete;
if T='250' then delete;
if T='300' then delete;
if T='350' then delete;
run;

***** BEISPIEL 2 *****;
##### H_1A ##### ;
*DF-Test;
*auseinanderdriftende Zeitpfade AD_1A;
*T=50;
data workout1;
set D_AD_1A.D_COMBINE_AD_1A;
keep Abh_glob T DF_y1t PERCENT_y1t DF_y2t PERCENT_y2t;
if DF_y1t='n' then delete;
if DF_y2t='n' then delete;
run;
data workout2;
set D_AD_1A.D_COMBINE_AD_1A;
keep Abh_glob T DF_ydt PERCENT_ydt;
if DF_ydt='n' then delete;
if PERCENT_ydt='.' then delete;
run;
data HYPO.H_1A_df;
merge workout1 workout2;
if T='100' then delete;
if T='150' then delete;
if T='200' then delete;
if T='250' then delete;
if T='300' then delete;
if T='350' then delete;
run;

goptions reset=global gunit=pct noborder cback=white

```

```

colors=(black) ctext=black ctitle=black
htext=6 ftext='Arial' ftitle='Arial'
devmap=winansi keymap=winansi;
title1 h=2 ;
title2 h=7 'HYPOTHESE 1A';
title3 h=6.5 'T=50';

proc gplot data=HYPO.H_1A_df;
symbol1 color=black interpol=L line=33 width=3 value=dot height=4;
symbol2 color=black interpol=join line=3 width=3 value=dot height=4;
symbol3 color=black interpol=spline line=1 width=3 value=dot height=4;

axis1 major=(height=2) minor=none label=('Abh_glob');
axis2 order=(0 to 100 by 25) label=('');

plot PERCENT_y1t*Abh_glob PERCENT_y2t*Abh_glob PERCENT_ydt*Abh_glob
/ overlay vaxis=axis2 haxis=axis1;

note c=black height=5
move=(10,26) '1. Komponentenprozess'
move=(10,43) '2. Komponentenprozess'
move=(70,62) 'Linearkombination';

run;
quit;
*****NAD_0*****;
goptions reset=global gunit=pct noborder cback=white
colors=(black) ctext=black ctitle=black
htext=6 ftext='Arial' ftitle='Arial'
devmap=winansi keymap=winansi;
title1 h=2 ;
title2 H=7 'nicht-auseinanderdriftende Zeitpfade';
title3 H=6.5 'Variation minimaler Abhängigkeiten';

proc gplot data=EMP_GRAP.EMP_DF_NAD_0;
symbol1 color=red interpol=spline line=1 width=3 height=4;
symbol2 color=green interpol=spline line=1 width=3 height=4;
symbol3 color=blue interpol=spline line=1 width=3 height=4;
symbol4 color=orange interpol=spline line=1 width=3 height=4;
symbol5 color=yellow interpol=spline line=1 width=3 height=4;
symbol6 color=pink interpol=spline line=1 width=3 height=4;
symbol7 color=brown interpol=spline line=1 width=3 height=4;
*ACHSEN;
axis1 order=(0.105 to 0.125 by 0.005) label=('Abh_glob');
axis2 label=('');

plot PERCENT_ydt*Abh_glob=t /haxis=axis1 vaxis=axis2 lvref=2 frame;
run;
quit;

```

Anhang E: Detail-Herleitung zum Engle-Granger-Repräsentationstheorem

Die Herleitung von Gleichung 3.3, S. 52, aus Gleichung 3.2, S. 52, erfolgt laut mathematischer und ökonometrischer Lehrbücher (vgl. Hartley & Hawkes, 1970, S. 41; Stewart & Gill, 1998, S. 175ff.; Stoll, 1952, S. 160) gemäß des sogenannten **Remainder Theorems** für Polynome, das in der Internetressource *Purplemath* (vgl. Stapel, 2004) genauer beschrieben wird. Stewart und Gill (1998, S. 176) geben außerdem an, dass sich Lag-Polynome wie *Polynome komplexer Zahlen* verhalten.

◇ ◇ ◇ ◇ ◇ ◇ ◇

Engle und Granger (1987) erbringen die Herleitung von Gl. 3.2, S. 52, zu Gl. 3.3, S. 52:

Gl. (3.3) in Engle und Granger (1987, S. 255):

$$\mathbf{A}(B)\mathbf{x}_t = d(B)\boldsymbol{\epsilon}_t$$

darin sind

- $d(B) = 1$
- $\mathbf{A}(B)$ entspricht (u.a. gemäß der Definition auf S. 254) $\Phi(B)$
- ⇒ entspricht Gl. 3.2: $\Phi(B)\mathbf{x}_t = \boldsymbol{\epsilon}_t$

Gl. (3.4) in Engle und Granger (1987, S. 256):

$$\mathbf{A}^*(B)(1 - B)\mathbf{x}_t = -\gamma\mathbf{z}_{t-1} + d(B)\boldsymbol{\epsilon}_t$$

darin sind

- S. 253 (Def. von Kointegration): Linearkombination $\mathbf{z}_t = \boldsymbol{\alpha}'\mathbf{x}_t$
 → $\mathbf{z}_{t-1} = \boldsymbol{\alpha}'\mathbf{x}_{t-1}$
 → Def. von Kointegration (S. 253): $\boldsymbol{\alpha}$ entspricht „unserem“ Kointegrationsvektor $\boldsymbol{\beta}'$
- S. 256, (3): $\mathbf{A}(1) = \gamma\boldsymbol{\alpha}'$
 → Fehlerkorrekturmodell (S. 254): γ entspricht „unserem“ Anpassungskoeffizienten $-\alpha$
 ⇒ dadurch ist nebenbei der Ausdruck $-\Phi(1) = \boldsymbol{\alpha}\boldsymbol{\beta}'$ aus Gl. 3.5 nachgewiesen

Nach Einsetzen und Anpassung der Notation lässt sich ableiten:

$$\begin{array}{lll} \Phi^*(B)(1 - B)\mathbf{x}_t & = & -\Phi(1)\mathbf{x}_{t-1} + \boldsymbol{\epsilon}_t \quad | +\Phi(1)\mathbf{x}_{t-1} \\ \Phi^*(B)(1 - B)\mathbf{x}_t + \Phi(1)\mathbf{x}_{t-1} & = & \boldsymbol{\epsilon}_t \quad | B(\mathbf{x}_t) = \mathbf{x}_{t-1} \text{ (Backshift-Op.)} \\ \Phi^*(B)(1 - B)\mathbf{x}_t + \Phi(1)B\mathbf{x}_t & = & \boldsymbol{\epsilon}_t \quad | \mathbf{x}_t \text{ ausklammern} \\ [\Phi^*(B)(1 - B) + \Phi(1)B]\mathbf{x}_t & = & \boldsymbol{\epsilon}_t \end{array}$$

Gemäß Gl. 3.3 gilt:

$$\Phi^*(B)(1 - B) + \Phi(1)B = \Phi(B) \quad \text{Einsetzen!}$$

Daraus ergibt sich direkt Gl. 3.2:

$$\Phi(B)\mathbf{x}_t = \boldsymbol{\epsilon}_t$$

Die Definitionen aus Gl. 3.4 ergeben sich aus der Definition der Eigenschaften von Φ und Φ^* , die für Φ durch das bekannte Matrix-Lag-Polynom beschrieben sind:

$$\Phi(B) := I_m - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p$$

Φ^* wiederum hat die Eigenschaften der *Adjungierten* von Φ .

◇ ◇ ◇ ◇ ◇ ◇ ◇

Wie bereits beschrieben und in Anhang A, S. 139, erläutert, steht B auch in dieser Herleitung für den Backward-Shift-Operator, der nach Stewart und Gill (1998, S. 176) als eine *komplexe Variable* aufgefasst werden kann.

Tabellenverzeichnis

3.1	Grundidee des Kointegrationskonzepts: VAR(1)-Prozesse; Modellmatrix Φ_1 , zugehörige Eigenwerte (EW) und Dickey-Fuller-Tests (DF) der VARMAX-Prozedur über Komponentenprozesse (black, red) und Linearkombination (green) für T=1000.	36
3.2	Für die folgenden VAR(1)-Prozesse ermittelte R^2 - und DW-Statistiken. $T = 20000$, 20 Wiederholungen.	48
3.3	Funktion des Anpassungskoeffizienten: VAR(1)-Prozesse; Modellmatrix Φ_1 , zugehörige Eigenwerte (EW) und Populationsparameter α_1 und α_2 . Dickey-Fuller-Tests (DF) der VARMAX-Prozedur über Komponentenprozesse (black, red) und Linearkombination (green) für T=1000. Dabei $\beta_1 = 1$ und $\beta_2 = -1$. Veranschaulichung in Abbildung 3.2, S. 55.	56
3.4	Funktion des Kointegrationsvektors: VAR(1)-Prozesse; Modellmatrix Φ_1 , zugehörige Eigenwerte (EW) und Populationsparameter β_1 und β_2 . Dickey-Fuller-Tests (DF) der VARMAX-Prozedur über Komponentenprozesse (black, red) und Linearkombination (green) für T=1000. Veranschaulichung in Abbildung 3.3, S. 57.	58
4.1	Simulierte VAR(1)-Prozesse mit nicht-auseinanderdriftenden Zeitpfaden der Komponentenprozesse.	88
4.2	Simulierte VAR(1)-Prozesse mit auseinanderdriftenden Zeitpfaden der Komponentenprozesse.	89
5.1	Dickey-Fuller-Kointegrationstest für VAR(1)-Prozesse mit nicht-auseinanderdriftenden Zeitpfaden der Komponentenprozesse.	94
5.2	Dickey-Fuller-Kointegrationstest für VAR(1)-Prozesse mit auseinanderdriftenden Zeitpfaden der Komponentenprozesse.	95
5.3	Johansen-Testung auf eine Kointegrationsbeziehung ($H_0 : r = 0$ vs. $H_1 : r > 0$) der VAR(1)-Prozesse mit nicht-auseinanderdriftenden Zeitpfaden der Komponentenprozesse. \Rightarrow Fortsetzung auf der folgenden Seite \Rightarrow	97
5.4	Johansen-Testung auf eine Kointegrationsbeziehung ($H_0 : r = 0$ vs. $H_1 : r > 0$) der VAR(1)-Prozesse mit auseinanderdriftenden Zeitpfaden der Komponentenprozesse.	99

5.5 MSEs für α und β ; VAR(1)-Prozesse mit **nicht-auseinanderdriftenden** Zeitpfaden der Komponentenprozesse. \Rightarrow Fortsetzung auf der nächsten Seite \Rightarrow . . . 108

5.6 MSEs für α und β ; VAR(1)-Prozesse mit **auseinanderdriftenden** Zeitpfaden der Komponentenprozesse ($\phi_{11} + \phi_{12} < 1$). \Rightarrow Fortsetzung auf der nächsten Seite \Rightarrow 110

5.7 MSEs für α und β ; VAR(1)-Prozesse mit **auseinanderdriftenden** Zeitpfaden der Komponentenprozesse ($\phi_{11} + \phi_{12} > 1$). 112

Abbildungsverzeichnis

1.1	Bivariater Prozess mit 2 (unabhängigen) Komponentenprozessen; $T=1000$	17
2.1	Stationärer Prozess (grün), instationärer Prozess (rot); jeweils für $T = 1000$ und $T = 50000$ zur Veranschaulichung des langfristigen Verlaufs stationärer und instationärer Prozesse.	21
3.1	Grundidee der Kointegrationsmethodologie: demonstriert durch multivariate Prozesse mit zunehmender Abhängigkeit der Einzelreihen, beginnend mit einem Modell mit völlig unabhängigen Komponentenprozessen (links oben). Modelleigenschaften in Tabelle 3.1, S. 36.	37
3.2	Graphische Veranschaulichung der Funktion des Anpassungskoeffizienten. Modelleigenschaften in Tabelle 3.3, S. 56. $ \alpha_1 $ von links oben nach rechts unten größer werdend.	55
3.3	Graphische Veranschaulichung der Funktion des Kointegrationsvektors. Modelleigenschaften in Tabelle 3.4, S. 58. Zunehmende Annäherung von $\beta' = (1 \quad -8)$ an $\beta' = (1 \quad -1)$ von links oben nach rechts unten.	57
5.1	Ergebnisse des Dickey-Fuller-Kointegrationstests : prozentualer Anteil stationärer Linearkombinationen.	96
5.2	Ergebnisse des Johansen-Kointegrationstests : prozentualer Anteil einer signifikanten Kointegrationsbeziehung.	100
5.3	HYPOTHESE 0 (Variation minimaler Abhängigkeiten für nicht-auseinanderdriftende Zeitpfade) mittels Dickey-Fuller-Kointegrationstest (links: prozentualer Anteil stationärer Linearkombinationen) bzw. Johansen-Kointegrationstest (rechts: prozentualer Anteil einer signifikanten Kointegrationsbeziehung).	102
5.4	HYPOTHESE 1A (prozentualer Anteil stationärer Komponentenprozesse und stationärer Linearkombinationen) mittels Dickey-Fuller-Kointegrationstest für auseinanderdriftende Zeitpfade (Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$).	104
5.5	HYPOTHESE 1B: Vergleich des Verhaltens des Johansen-Testverfahrens (prozentualer Anteil einer signifikanten Kointegrationsbeziehung) für Modelle mit nicht-auseinanderdriftenden vs. auseinanderdriftenden Zeitpfaden (Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$).	105

5.6	HYPOTHESE 1C: Vergleich des Verhaltens des Dickey-Fuller- vs. Johansen-Kointegrationstests für Modelle mit <i>nicht</i> -auseinanderdriftenden Zeitpfaden (links) und Modelle mit auseinanderdriftenden Zeitpfaden (rechts; Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$); prozentualer Anteil stationärer Linearkombinationen (DF-Test) bzw. einer signifikanten Kointegrationsbeziehung (Johansen-Test).	106
5.7	Ergebnisse der Parameterschätzungen von Kointegrationsvektor und Anpassungskoeffizient für <i>nicht</i> -auseinanderdriftende Zeitpfade.	113
5.8	Ergebnisse der Parameterschätzungen von Kointegrationsvektor und Anpassungskoeffizient für auseinanderdriftende Zeitpfade ($\phi_{11} + \phi_{12} < 1$).	114
5.9	HYPOTHESE 2A zur Schätzung des Kointegrationsvektors für VAR(1)-Modelle mit <i>nicht</i> -auseinanderdriftenden (links) und auseinanderdriftenden Zeitpfaden (rechts; Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$).	116
5.10	HYPOTHESE 2B: übergeordneter Vergleich der Schätzgüte von Kointegrationsvektor vs. Anpassungskoeffizient für Modelle mit <i>nicht</i> -auseinanderdriftenden Zeitpfaden.	117
5.11	HYPOTHESE 2B: übergeordneter Vergleich der Schätzgüte von Kointegrationsvektor vs. Anpassungskoeffizient für Modelle mit auseinanderdriftenden Zeitpfaden (Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$).	118
5.12	HYPOTHESE 2C: übergeordneter Vergleich der Schätzgüte von Kointegrationsvektor (links) vs. Anpassungskoeffizient (rechts) für Modelle mit nicht-auseinanderdriftenden (grün) vs. auseinanderdriftenden (rot) Zeitpfaden (Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$).	119
5.13	HYPOTHESE 2C: Einzelvergleiche der Schätzgüte von α_1 (links) vs. α_2 (rechts) für Modelle mit nicht-auseinanderdriftenden (grün) vs. auseinanderdriftenden (rot) Zeitpfaden (Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$).	120
5.14	HYPOTHESE 2D zur Schätzung von Kointegrationsvektor (links) und Anpassungskoeffizient (rechts) für Modelle mit <i>nicht</i> -auseinanderdriftenden Zeitpfaden; unterschieden nach der variierten Abhängigkeit der Komponentenprozesse.	121

5.15 HYPOTHESE 2D zur Schätzung von Kointegrationsvektor (links) und Anpassungskoeffizient (rechts) für Modelle mit auseinanderdriftenden Zeitpfaden (Modellmatrizen mit $\phi_{11} + \phi_{12} < 1$); unterschieden nach der variierten Abhängigkeit der Komponentenprozesse. 122

Erklärung

Hiermit versichere ich, dass ich die vorliegende Diplomarbeit selbständig und ohne Nutzung von anderen als den angegebenen Quellen und Hilfsmitteln angefertigt habe. Alle den Quellen wörtlich oder inhaltlich entnommenen Informationen wurden als solche kenntlich gemacht.

Heidelberg, Juli 2005

.....

Esther Stroe-Kunold

Freiwillige Erklärung

Ich stimme ausdrücklich zu, dass meine durch

Prof. Dr. J. Werner und *Prof. Dr. J. Funke*

betreute Diplomarbeit mit dem Titel

*Multivariate Analyse instationärer Zeitreihen:
Integration und Kointegration in Theorie und Simulation*

nach Beendigung der Diplomprüfung wissenschaftlichen Zwecken zugänglich gemacht und in der Institutsbibliothek aufgestellt wird (Veröffentlichungen nach § 6 Abs. 1 UrhG), sowie hieraus im Rahmen des § 51 UrhG zitiert werden kann.

Sämtliche Verwertungsrechte nach § 15 UrhG verbleiben bei der Verfasserin der Diplomarbeit.

Heidelberg, Juli 2005

.....

Esther Stroe-Kunold

Stroe-Kunold, Esther

Brunnengasse 10

69117 Heidelberg