

Internationale Aktienmärkte der G5-Länder: eine ökonometrische Modellierung

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum politicarum

(Doktor der Wirtschaftswissenschaft)

eingereicht an der

Wirtschaftswissenschaftlichen Fakultät

der Universität-Paderborn

von

Frau Dipl.Math. Elena Schneider

Dekan der Fakultät für Wirtschaftswissenschaften:

Prof. Dr. Peter F. E. Sloane

Gutachter:

1. Prof. Dr. Th. Gries

2. Prof. Dr. M. Kraft

Paderborn, im Oktober 2005

Danksagungen

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftliche Mitarbeiterin am Lehrstuhl für Wachstums- und Konjunkturtheorie der Universität Paderborn. Ich möchte mich an dieser Stelle bei allen bedanken, die zum Gelingen dieser Arbeit beigetragen haben.

Zuerst gilt mein besonderer Dank Prof. Dr. Thomas Gries, der mir mit seiner Betreuung stets zur Seite stand und diese Arbeit mit vielen interessanten Ideen und zahlreichen Diskussionen unterstützt hat. Für die fachliche Unterstützung bei der Auswahl empirischer Methoden und die fruchtbare Zusammenarbeit während der Promotion bedanke ich mich herzlich bei Prof. Dr. Manfred Kraft.

Ich möchte mich bei allen Kollegen am Lehrstuhl für die gute Zusammenarbeit und die vielen interessanten Diskussionen und Gespräche bedanken. Ausdrücklich nennen möchte ich die Kollegen, mit denen ich eng wissenschaftlich zusammengearbeitet habe: Dr. Stefan Jungblut, Axel Wieneke und Björn Angermayer danke ich für die hilfreichen inhaltliche Ratschläge.

Bei Irina Harder, Handan Elitok, Stephan Eisfelder, Andreas Schneider, Ella Schneider und Michael Lubomirski bedanke ich mich zusätzlich für das Korrekturlesen des Manuskripts.

Schließlich möchte ich mich bei Andrei, Annemarie, meinen Eltern, Schwiegereltern und Freunden für ihren Rückhalt, ihre Unterstützung und Geduld in den letzten Jahren bedanken.

Vielen herzlichen Dank!

Paderborn, im Oktober 2005

Elena Schneider

Inhaltsverzeichnis

1	Aktienmärkte: Einleitung und Motivation	1
1.1	Motivation	1
1.2	Problemstellung und Gang der Untersuchung	7
2	Ökonometrische Methoden	9
2.1	Stochastische Prozesse und ihre Eigenschaften	10
2.1.1	Einführung. VAR-Prozesse	10
2.1.2	Schätzen von VAR-Modellen. Strukturanalyse	17
2.1.3	Probleme	27
2.2	Trends und integrierte Prozesse	29
2.2.1	Deterministischer oder stochastischer Trend?	29
2.2.2	Integrierte Prozesse. Tests	33
2.2.3	Probleme	43
2.3	Kointegrationskonzept	48
2.3.1	Kointegration. Granger-Repräsentationstheorem	48
2.3.2	Schätzverfahren. Kointegrationstests	52
2.3.3	Probleme	61
3	Empirische Analyse und Folgerungen	64

<i>INHALTSVERZEICHNIS</i>	II
3.1 Datenbeschreibung	64
3.1.1 Explorative Statistiken	69
3.1.2 Korrelationsanalyse	73
3.2 Kointegrationsanalyse der Aktienmärkte	76
3.2.1 Einheitswurzel-Tests	76
3.2.2 Paarweise Kointegration versus Granger-Kausalität	81
3.2.3 Multivariate Kointegration	84
3.2.4 Dynamische Analyse	90
4 Ökonomische Konzepte	94
4.1 Die Markteffizienz-Hypothese	94
4.1.1 Die MEH und der Aktienmarkt	94
4.1.2 Die Random-Walk-Hypothese	97
4.1.3 Die MEH und empirische Nachweise	100
4.2 Portfolio-Theorie	107
4.2.1 Die internationale Diversifikation	109
4.2.2 Das Capital-Asset-Pricing-Model	112
4.2.3 Das CAPM und Empirische Nachweise	119
4.3 Relevanz der Kointegration für Aktienpreise	123
5 Zusammenfassung der Arbeit und Ausblick	126
Literaturverzeichnis	133

Abbildungsverzeichnis

1.1	Verlauf der nationalen Aktienindizes der G5-Länder. Die nominalen Indizes sind in DM/Euro konvertiert; Basisjahr 2000.	5
2.1	Der nominale Aktienindex der USA in der Zeitperiode von 06.1964 bis 06.2000	29
3.1	Vergleich des Handelsvolumens der internationalen Aktienmärkte der G5-Länder im Jahr 2000	65
3.2	IAF für den deutschen Aktienindex	92

Tabellenverzeichnis

1.1 Renditen der internationalen Aktienmärkte	3
1.2 Risiken der internationalen Aktienmärkte	4
2.1 Regression der integrierten Variablen	35
2.2 Asymptotische Verteilung für AR(1)-Prozess	38
3.1 Statistiken der Aktienmärkte. (DM/Euro)	69
3.2 Statistiken der Aktienmärkte. EW	70
3.3 Statistiken der Renditen der Aktienmärkte in (DM/Euro)	70
3.4 Statistiken der realen Renditen der Aktienmärkte in EW	71
3.5 Korrelationsmatrix der Aktienindizes in (DM/Euro)	73
3.6 Korrelationsmatrix der Aktienindizes in eigener Währung	74
3.7 Korrelationsmatrix der Renditen in (DM/Euro)	74
3.8 Korrelationsmatrix der Renditen in eigener Währung	75
3.9 Ng-Perron-Test. Modell mit einem Achsenabschnitt	77
3.10 Ng-Perron-Test. Modell mit einem Trend	78
3.11 DF-GLS-Test. Modell mit einem Achsenabschnitt	78
3.12 DF-GLS-Test. Modell mit einem Trend	79
3.13 Modifizierter ADF-Test für Aktienindizes und Renditen.	79

3.14	Paarweise Kointegration.	81
3.15	Granger-Kausalität-Test.	83
3.16	Johansens Kointegrationstest. Modell 1	86
3.17	Johansens Kointegrationstest. Modell 2	87
3.18	Johansens Kointegrationstest. Modell 3	87
3.19	Der normalisierte Kointegrationsvektor. Modell 1	88
3.20	Der normalisierte Kointegrationsvektor. Modell 2	89
3.21	Die Matrix der kurzfristigen Anpassungen. Modell 1	90
3.22	Die Matrix der kurzfristigen Anpassungen. Modell 2	91
3.23	Varianz-Dekomposition. Modell 1	93
3.24	Varianz-Dekomposition. Modell 2	93

Kapitel 1

Aktienmärkte: Einleitung und Motivation

1.1 Motivation

Nach der Liberalisierung der Kapitalbewegungen im letzten Drittel des zwanzigsten Jahrhunderts wird von einem hohen Integrationsgrad der internationalen Kapitalmärkte ausgegangen. Potenziell können die internationalen Aktienmärkte, beispielsweise neu entstandene Märkte in Osteuropa, Russland und China, höhere Gewinnchancen für Investoren versprechen. Gleichzeitig könnten diese Investitionen mit einem hohen Risiko verbunden sein, das auf die wirtschaftliche und politische Situation des Landes zurückzuführen ist. Die Existenz stabiler langfristiger struktureller Gleichgewichtsbeziehungen zwischen internationalen Aktienmärkten könnte die Risikoposition der Investoren allerdings theoretisch verbessern. Für Anleger ist deswegen die tatsächliche internationale Integration der Kapitalmärkte in Bezug auf internationale Diversifikationsstrategien von besonderem Interesse.

Weitere Fragen, die mit der Existenz der langfristigen Gleichgewichtsbeziehungen verbunden sind, beziehen sich auf kurzfristige Anpassungen der Aktienmärkte zum Gleichgewicht nach verschiedenen aufgetretenen Schocks. Wie stark ist die Abweichung eines Marktes vom existierenden Gleichgewicht nach Auftreten unerwarteter Neuigkeiten auf dem eigenen bzw. auf einem anderen Markt? Wie lange braucht das Gesamtsystem, um zu dem stabilen Gleichgewicht zurückzukehren? Welcher Markt bzw. welche Märkte beeinflussen den Rest des Systems am stärksten?

Langfristige Anlagestrategien gewinnen für private Investoren zusätzlich an Bedeutung in Verbindung mit der aktuellen politischen Diskussion über die Altersversorgung. Private Rentenversicherungen und Immobilien werden immer häufiger als zusätzliche

Vorsorge zur staatlichen Rente gesehen. Dementsprechend entsteht die Frage: Kann eine diversifizierte langfristige Aktienanlage (über 30 oder 40 Jahren) als eine Alternative zur traditionellen zusätzlichen Altersabsicherung gesehen werden?

Langfristige Beziehungen zwischen internationalen Aktienmärkten wurden bereits in mehreren empirischen Untersuchungen festgestellt¹. Mehrere von ihnen liefern jedoch widersprüchliche Ergebnisse². Ein Grund für solche Diskrepanzen könnte in der Vielfalt der relevanten Parameter und Freiheitsgrade bei der Auswahl eines geeigneten ökonometrischen Modells liegen. So werden mehrere Parameter immer wieder alternativ berücksichtigt: Zeithorizont, Anzahl der Verzögerungen in dem Vektorkorrekturmodell, Datenfrequenz, Trendeinbeziehung, Strukturbrüche und Volatilität des Marktes sowie Saisoneffekte. Diese Parameter werden nie gemeinsam in die Analyse eines Datensatzes einbezogen. Vor diesem Hintergrund wird in der vorliegenden Arbeit eine ausführliche Datenanalyse der Aktienmärkte durchgeführt.

Wenn die Gleichgewichtsbeziehungen der internationalen Aktienmärkte festgestellt werden, bleibt offen, ob diese Beziehungen als ökonomisches Gleichgewicht interpretierbar sind. Das Gleichgewichtsmodell CAPM (internationales Capital-Asset-Pricing-Model)³ könnte zur möglichen Erklärung eines gemeinsamen Fundamentalprozesses auf den Aktienmärkten herangezogen werden. Weitere Theorien, wie die Markteffizienz-Hypothese (MEH)⁴ oder Behavioral Finance⁵, können ebenfalls für die theoretische Interpretation verwendet werden. Das CAPM beruht zwar auf einer idealisierten Welt, wie auch die MEH, spiegelt aber die wichtigsten Ideen der Gleichgewichtskonzepte wider. Die *'behavioral'*-Sicht hilft einerseits, das Verhalten der Märkte besser zu verstehen, andererseits bereitet die Einbindung dieser Hypothese jedoch nicht weniger Schwierigkeiten als die der idealisierten MEH und CAPM.

Bei einer Entscheidung für CAPM als eine mögliche Interpretation der empirisch gefundenen Gleichgewichtsbeziehungen der Aktienmärkte wird an erster Stelle an das Risiko/Rendite-Verhältnis einer Aktie gedacht. Das Rendite/Risiko-Verhältnis der Aktien ist ebenfalls der erste Baustein bei einer Portfolio-Analyse. Das internationale CAPM ist unmittelbar mit der internationalen Diversifikation eines Portfolios verbun-

¹Vgl. Taylor and Tonks (1989), Kasa (1992), Ashanapalli and Doukas (1993), Corhay und andere (1993), Leachman und Francis (1995), Dickinson (2000).

²Vgl. Richards (1995), Godbout und van Norden (1997) und Ahlgren und Antell (2002).

³CAPM beschreibt einen Zusammenhang zwischen dem systematischen Risiko und der erwarteten Rendite einer Aktie.

⁴An erster Stelle verbindet MEH mit Informationseffizienz und mit ihrer mathematischen Formalisierung als Random-Walk-Hypothese (RWH). Entsprechend der RWH folgen die Aktienpreise einem zufälligen Random-Walk-Prozess. Die MEH stellt eine Restriktion für die erwartete Rendite dar, berücksichtigt aber kein Marktrisiko. Eine Bewertung des Risiko/Rendite-Ausgleichs gehört mittlerweile zu den wichtigen Problemen der modernen Finanzökonomie.

⁵Behavioral Finance ist die Theorie der Finanzmärkte, die das psychologische Verhalten der Investoren und den soziologischen Aspekt bei der Preisbildung auf Aktienmärkten mit einbezieht. Nach der Meinung der Behavioralisten überschätzen beispielsweise die Investoren die historische Performance des Marktes, vgl. Shiller (2003).

den. Deswegen steht eine Abschätzung der Rendite/Risiko-Situation für die international agierenden Anleger bei einer Portfolio-Entscheidung an erster Stelle. Eine vorläufige empirische Analyse der Rendite/Risiko-Situation auf den internationalen Aktienmärkten der G5-Länder ist in den Tabellen (1.1) und (1.2) dargestellt.

Tabelle 1.1: Durchschnittliche Erträge der internationalen Aktienmärkte der G5-Länder, 1964-2004 (nominal, Prozent pro Jahr)

Aktienmarkt	Eigenes Land	Wechselkurs- gewinn	Für den deutschen Investor
USA	6,26	-2,20	4,06
Japan	6,03	0,79	6,83
France	6,67	-2,65	4,02
Germany	4,05	0,00	4,05
UK	7,57	-3,60	3,96
Gleichgewichteter Index	6,12	-1,53	4,59
Volumengewichteter Index	6,20	-1,72	4,48
Gleichgewichteter Index (ohne Deut.)	6,63	-1,91	4,72

Notiz: Die Gewichte für den volumengewichteten Index stammen aus Morgan Stanley Capital International Perspectives, Juni 2000.

Aus der Tabelle (1.1) ist ersichtlich, dass aus langfristiger Sicht⁶ deutsche Investoren kaum eine Gelegenheit hatten, aus den anderen europäischen Aktien eine Überschussrendite zu erwirtschaften. An erster Stelle vermindern die Wechselkursrisiken die Gewinne aus den internationalen Anlagen (bis auf das Niveau der deutschen Renditen von 4,05% pro Jahr). Die international diversifizierte Portfolios (japanische und amerikanische Aktien eingeschlossen), die durch einen gleichgewichteten sowie durch einen volumengewichteten Aktienindex repräsentiert sind, konnten den deutschen Aktienmarkt überbieten. Der Grund dafür ist die hohe Überschussrendite des japanischen Aktienmarktes. Obwohl sich die anhaltende Stagnation des wirtschaftlichen Wachstums in Japan in den Verlusten der japanischen Börse widerspiegelt, überwiegt der Ertrag aus dem Zeitraum der enormen Wachstumsphase immer noch.

Die Entscheidung bezüglich einer international ausgerichteten Investition, die allein nach der ex-post-Betrachtung der Erträge getroffen wurde, kann nicht unbedingt zu einer erfolgreichen Anlagestrategie führen. Ein mitentscheidender Faktor dabei ist das Risiko, das mit den ausländischen Märkten verbunden ist. Die Risiken werden

⁶Die Renditen der Aktienmärkte wurden über eine Periode von vierzig Jahren gerechnet. Die kurzfristigen Gewinne können beispielsweise in den Wachstumsphasen der deutschen Börse erwirtschaftet werden.

Tabelle 1.2: Risiken der internationalen Aktienmärkte der G5-Länder, 1964-2004

Aktienmarkt	Wechselkurs-		
	Inlandsrisiko	risiko	Gesamtrisiko
USA	12,24	39,79	17,57
Japan	17,57	35,51	22,84
Frankreich	17,79	57,83	18,36
Germany	17,04	0,00	17,04
UK	16,51	130,35	17,92
Gleichgewichteter Index	16,23	52,69	18,75
Volumengewichteter Index	14,27	43,36	18,46
Gleichgewichteter Index (ohne Deut.)	16,03	65,87	19,17

Notiz: Die Standardabweichung des Preises von seinem Mittelwert wurde über die Gesamtperiode gerechnet. Die Gewichte für den volumengewichteten Index sind aus Morgan Stanley Capital International Perspectives, Juni 2000.

üblicherweise durch die Standardabweichung des Aktienpreises von seinem Mittelwert gemessen. Aus der Sicht der deutschen Investoren ist der deutsche Markt mit dem kleinsten Risiko verbunden (vgl. Tabelle (1.2)). Das Wechselkursrisiko, das beispielsweise bei dem britischen Pfund durchschnittlich bis 130% steigt, kann eine Entscheidung bezüglich Erwerben britischer Aktien erschweren.

Die relativ kleine durchschnittliche Rendite des deutschen Aktienmarktes ist mit dem ebenfalls kleinsten Risiko verbunden. Dagegen ist die höchste Rendite des japanischen Marktes durch das höchste Risiko ausgezeichnet. Die gegenläufigen Interessen der Anleger liegen im Erwirtschaften eines angemessenen Ertrags aus internationaler Anlage bei einem möglichst minimalen Risiko. Eine theoretische Lösung des Optimierungsproblems bietet den Investoren die Portfolio-Theorie an, die letztendlich zum internationalen CAPM als ein Gleichgewichtsmodell führt.

Langfristige Beziehungen der Aktienmärkte können auf zwei Wegen festgestellt werden. Zum einen kann von einer ökonomischen Theorie ein Modell konzipiert werden. Nach einer empirischen Schätzung des Modells wird es validiert, d.h. das Verhalten des Modells wird mit dem Realitätsverhalten verglichen. Die zweite Vorgehensweise beginnt mit einer rein empirischen Datenanalyse. Bei der Datenanalyse könnten allerdings scheinbare Verbindungen identifiziert werden, die kaum eine ökonomische Interpretation finden. Deswegen ist eine Vorstellung über den ökonomischen Hintergrund bei der empirischen Datenanalyse notwendig.

Wenn für eine Datenanalyse bei der empirischen Forschung entschieden wird, stellt

man sich die Frage nach der möglichen Beschreibung der stattgefundenen Prozesse an erster Stelle. Welche statistischen Modelle können am besten die vorliegenden Daten beschreiben?

Die Modellierung der Phänomene kann beispielsweise mit der Betrachtung des zeitlichen Verlaufs der interessierenden Größen begonnen werden. Für eine Darstellung des Verhaltens der internationalen Aktienindizes im Zeitablauf werden die Zeitreihenmodelle angewendet. Das klassische Komponentenmodell der Zeitreihenanalyse zerlegt eine Zeitreihe in einen Trend, eine zyklische Komponente und eine Restgröße. Durch die Trendkomponente wird die Instationarität des Prozesses berücksichtigt. Die zyklische oder kurzfristige Komponente beschreibt die Anpassungen des Prozesses zu diesem Trend. Der Restkomponente werden nicht vorhersehbare Änderungen zugewiesen.

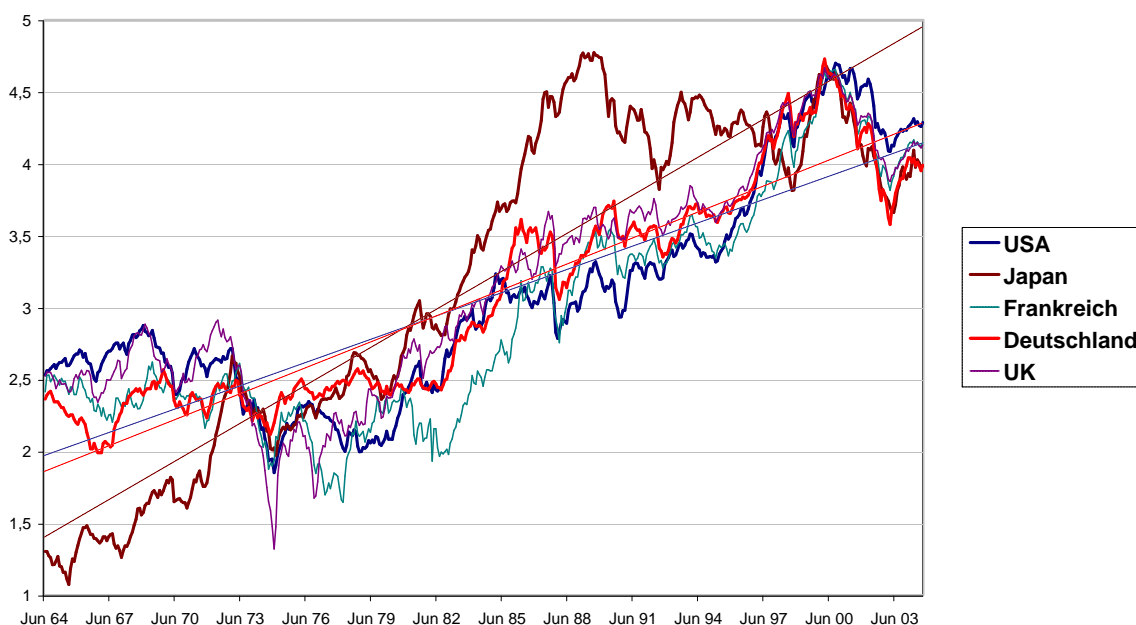


Abbildung 1.1: Verlauf der nationalen Aktienindizes der G5-Länder. Die nominalen Indizes sind in DM/Euro konvertiert; Basisjahr 2000.

Bei der Analyse der Grafik (1.1) für die Aktienpreisentwicklung über die letzten vierzig Jahre könnten einige Gemeinsamkeiten in der Bewegung der Märkte festgestellt werden. Erstens wird ein Wachstumsverhalten im Verlauf aller Zeitreihen beobachtet. Dies bewirkt die sogenannte Instationarität eines Prozesses, der die Bewegung der Aktienpreise generiert. Zweitens weichen in bestimmten Perioden die Preise der unterschiedlichen nationalen Aktienindizes von einem eventuellen Trend ab, allerdings laufen sie nicht weit auseinander.

In den 80er Jahren begann eine Diskussion, ob die Hypothese eines deterministischen Trends wirklich auf alle ökonomischen Variablen zutrifft, oder ob der beobachtete

Trend selbst stochastischer Natur ist⁷. Seitdem wurde die Frage nach einem geeigneten Test, der zur die Feststellung der Natur der Trends beitragen kann, in der Ökonometrie stark verfolgt. Das führte beispielsweise zur Entwicklung mehrerer alternativer Einheitswurzel-Tests wie des erweiterten Dickey-Fuller-GLS-Tests oder des Kwiatkowski, Phillips, Schmidt, und Shin Tests (KPSS)⁸.

Die zweite Beobachtung aus der Grafik (1.1), nämlich dass die Bewegung der Aktienindizes in einer bestimmten Bandbreite liegt, stellt die Frage nach stabilen Zusammenhängen zwischen den nationalen Aktienmärkten. Die Korrelationsanalyse stellt die erste Möglichkeit, um die linearen Verbindungen zwischen verschiedenen Zeitreihen zu untersuchen. Die Instationarität der Daten führt allerdings zur Scheinkorrelation, die die Ergebnisse unbrauchbar macht. Zur weiteren Analyse von Zusammenhängen zwischen Zeitreihen werden häufig die Regressionsmodelle angewendet. Die Parameter der entsprechenden Regressionsmodelle werden hauptsächlich mittels Kleinste-Quadrate-Methode geschätzt. Sowohl die Güteeigenschaften der Parameterschätzung als auch die der Anpassung des Modells insgesamt hängen aber von einer Reihe der Annahmen ab. Insbesondere müssen die zugrunde liegenden Zeitreihen, wieder um Scheinkorrelation auszuschließen, (schwach) stationär sein. Bei trendstationären Zeitreihen wird die Zeitreihe um den Trend bereinigt, bevor man ökonometrische Untersuchungen durchführt. Die Methode ist ebenfalls mit den Schwierigkeiten der Trendschätzung verbunden. Viele ökonomische Zeitreihen (nicht nur solche mit stochastischem Trend) sind jedoch differenzenstationär. Die Trendbereinigung erfolgt mittels Differenzenbildung. Dabei darf man aber nicht übersehen, dass diese Transformationen jeweils mit einem Informationsverlust über die interessierenden Niveauezusammenhänge verbunden sind und damit wichtige Langzeiteigenschaften für weitere Analysen verloren gehen.

Kointegrationsmodelle liefern geeignete Verfahren für Zusammenhangsanalysen differenzenstationärer Zeitreihen. Sie setzen nun genau an der besprochenen Eigenschaft der Instationarität der ökonomischen Daten an. In einem Kointegrationsmodell sollen nur Zeitreihen mit jeweils identischer Integrationsordnung verwendet werden. Kointegration der Variablen bedeutet, dass diese Variablen eine Beziehung erfüllen, so dass die Abweichungen von dieser Beziehung einen stationären Prozess darstellen. Die Beziehung wird bis auf Irregularitäten, die keine bis in die Zukunft reichenden Auswirkungen haben, innerhalb jeder Beobachtungsperiode eingehalten.

Zugleich ist die Kointegration durch das Granger-Repräsentationstheorem (1987) mit den Fehlerkorrektur-Modellen verbunden. Die Fehlerkorrektur-Modelle kombinieren langfristige Gleichgewichtsbeziehungen, welche auf der ökonomischen Theorie basieren, mit der kurzfristigen Dynamik in den Modellen. Die kurzfristigen Anpassungsprozesse nach den Schocks zu dem gemeinsamen fundamentalen Gleichgewicht ist aus ökonomischer Sicht ebenfalls von großem Interesse.

⁷Vgl. Nelson und Plosser (1982), De Jong und andere (1992).

⁸vgl. Maddala (1998), Kapitel 4.

1.2 Problemstellung und Gang der Untersuchung

Vor dem Hintergrund der vorgeführten Überlegungen lässt sich die Zielsetzung dieses Beitrags wie folgt beschreiben:

- Im ersten Schritt wird nach langfristigen strukturellen Gleichgewichtsbeziehungen zwischen den internationalen Aktienmärkten mittels empirischer Datenanalyse gesucht. Eine passende Methode und verschiedene Parameter bei der Modellierung sollen dabei berücksichtigt werden.
- Wenn auf den Aktienmärkten ein gemeinsamer Fundamentalprozess festgestellt wird, wird die Frage bezüglich eines kurzfristigen Anpassungsprozesses zu diesem Gleichgewicht nach verschiedenen Schocks gestellt.
- Die Frage, ob sich der empirisch festgestellte Fundamentalprozess als ein ökonomisches Gleichgewicht interpretieren lässt, wird ebenfalls verfolgt.
- Aus theoretischer Sicht wird nach ökonomischen Theorien gesucht, die mit solchen Fundamentalgleichgewichten kompatibel sind und als Erklärungshilfe der empirischen Ergebnisse herangezogen werden.
- Die letzte Frage bezüglich praktischer Implikation für die international agierenden Anleger, die aus den empirischen Ergebnissen und deren ökonomischen Interpretationsmöglichkeiten hervor kommt, wird versucht zu beantworten.

Die vorliegende Dissertationsarbeit ist wie folgt gegliedert:

Im **Kapitel 2** sind die grundlegenden ökonometrischen Konzepte zunächst zusammengefasst. Ein ausführlicher Ausblick über die weiterführenden Tests und Methoden, die manchmal nur in neuesten Forschungsartikeln zu finden sind, bringt den Leser auf den notwendigen Wissensstand der im Weiteren angewendeten ökonometrischen Methodik. Die Diskussion über Schwächen und Stärken der Methoden ist ebenfalls Bestandteil des Kapitels.

Kapitel 3 beinhaltet Details der statistischen Untersuchung und stellt die Ergebnisse der durchgeführten Testläufe vor. Die Präsentation der statistischen Ergebnisse und ihre ökonomische Interpretation stellt den zentralen Teil des Beitrags dar.

Kapitel 4 setzt sich mit den ökonomischen Ideen, die zur Interpretation der empirischen Ergebnissen herangezogen werden können, auseinander. Es werden MEH, Portfolio-Theorie und damit verbundene internationale CAPM dargestellt. Weiterhin sind die mathematische Formalisierung der Theorien, die empirischen Überprüfungen sowie kritische Bemerkungen zu diesen Modellen Bestandteile des Kapitels. Eine Diskussion über Kompatibilität theoretischer Überlegungen mit empirischen Ergebnissen ist ebenfalls in diesem Kapitel zusammengefasst.

Eine kurze Zusammenfassung und eine kritische Würdigung, die im **Kapitel 5** dargestellt sind, schließen die Dissertationsarbeit ab.

Kapitel 2

Ökonometrische Methoden

Empirische Überprüfung theoretischer Modelle ist häufig mit Schwierigkeiten verbunden. Unrealistische Annahmen und Vereinfachungen durch ein mathematisches Modell, das ein ökonomisches Verhalten darstellen soll, können zu inkonsistenten ökonomischen Ergebnissen führen. Eine andere Vorgehensweise, die von Daten ausgehende Untersuchung (Datenanalyse), fordert vorerst kein ökonomisches Modell, kann allerdings scheinbare Verbindungen identifizieren, die kaum eine ökonomische Interpretation finden. In dieser Arbeit wird vorerst der zeitliche Verlauf der Aktienpreise und die Verbindungen zwischen den Aktienmärkten aus einem rein mathematisch-statistischen Konzept modelliert werden. Die Kompatibilität der empirischen Ergebnisse mit ökonomischen Theorien wird im zweiten Schritt überprüft.

Die ökonometrische Modellierung der Aktienmärkte geht von den Beobachtungen des zeitlichen Verlaufs der Aktienindizes aus. Eine geeignete Beschreibung des Wachstumsverhaltens der Aktienpreise steht an erster Stelle der Überlegungen. Die Ideen sowohl eines deterministischen als auch eines stochastischen Trends können angewendet werden. Die zweite Entscheidung bei der Modellierung beinhaltet die Auswahl einer geeigneten Methode zur Identifizierung der Verbindungen zwischen den Aktienmärkten. Die Kointegrationsmethode passt aus mehreren Gründen zu dieser Fragestellung. Die Diskussion der Vorteile und der Nachteile der ausgewählten Methodik, die in diesem Kapitel ausführlich dargestellt wird, gehört ebenfalls zu den Modellentscheidungen.

2.1 Stochastische Prozesse und ihre Eigenschaften

2.1.1 Einführung. VAR-Prozesse

Die Darstellung einer Zeitreihe als eine Realisierung eines stochastischen Prozesses ist ein Baustein der Zeitreihentheorie. Für eine deutliche Darstellung des Konzeptes wird mit der Definition des stochastischen Prozesses begonnen.

Definition 2.1 Sei $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum, (E, \mathfrak{E}) ein messbarer Raum und T eine beliebige Indexmenge. Ein stochastischer Prozess $X_t(\omega)$ ist eine Abbildung $X_t(\omega) : \Omega \times T \rightarrow E$, so dass gilt

i) $\forall t \in T$ fest

$$X_t(\cdot) : \Omega \ni \omega \rightarrow X_t(\omega) \in E$$

ist eine $\mathfrak{A} - \mathfrak{E}$ -messbare Zufallsvariable;

ii) $\forall \omega \in \Omega$ fest

$$X(\omega) : T \ni t \rightarrow X_t(\omega) \in E^T.$$

$E^T = \{f : T \rightarrow E\}$ ist der Raum aller Abbildungen von T nach E . $X(\omega)$ heißt Trajektorie oder Pfad des Prozesses, $X_t(\cdot)$ - Zustand des Prozesses zum Zeitpunkt t , (E, \mathfrak{E}) Zustandsraum. Für $T = \mathbb{Z}$ ist (X_t) ein Prozess mit einem diskreten Zeitparameter. Wenn $T = \mathbb{R}$ ist, dann hat man einen kontinuierlichen stochastischen Prozess.

Ein stochastischer Prozess $(X_t)_{t \in T}$ kann deswegen auf zwei verschiedene Arten interpretiert werden. Die erste Interpretation ist eine Darstellung des stochastischen Prozesses als eine Menge unterschiedlicher Zeitpfade. Ein Zufallsvorgang wählt eine bestimmte Realisation von unendlich vielen Pfaden aus. Die zweite Interpretation stellt den stochastischen Prozess als eine Folge der Zufallsvariablen dar, wobei jedem Zeitpunkt t eine Zufallsvariable zugeordnet ist. In den meisten zeitreihenanalytischen Anwendungen sollte es sich als schwierig erweisen, den zugrundeliegenden Zufallsprozess substantiell zu spezifizieren. Ein stochastischer Prozess dient dann als eine Modellvorstellung für den Mechanismus, der die beobachtete Zeitreihe produziert hat.

Zu den wichtigsten Beispielen der stochastischen Prozesse gehören White-Noise-, Random-Walk- und Markov(1)-Prozess.

1. Ein White-Noise-Prozess oder ein reiner Zerfallsprozess stellt eine Folge $(\varepsilon_t)_{t \in T}$ von identisch verteilten und unabhängigen Zufallsvariablen ε_t dar. Der Prozess

hat einen diskreten Zeitparameter, $T = \mathbb{N}$ oder $T = \mathbb{Z}$. Wenn $\varepsilon_t \sim N(\mu, \sigma^2)$ -verteilt ist, dann nennt man $(\varepsilon_t)_{t \in T}$ einen Gauss-White-Noise-Prozess. White-Noise-Prozesse sind als solche uninteressant, dienen jedoch häufig als Grundbausteine für komplexere stochastische Prozesse.

2. Ein Random-Walk-Prozess (X_t) wird mit Hilfe eines White-Noise-Prozesses $(\varepsilon_t)_{t \in T}$ durch folgende Rekursion definiert

$$X_t = \begin{cases} \varepsilon_t & : t = 0 \\ X_{t-1} + \varepsilon_t & : t = 1, 2, \dots \end{cases}$$

In der Ökonomie finden Random-Walk-Prozesse viele Anwendungen, unter anderem bei der Modellierung von Preisbewegungen auf spekulativen Märkten.

3. Ein Markov(1)-Prozess wird ähnlich zu einem Random-Walk geschrieben

$$X_t = \begin{cases} \varepsilon_t & : t = 0 \\ aX_{t-1} + \varepsilon_t & : t = 1, 2, \dots, \end{cases}$$

mit $|a| < 1$. (ε_t) stellt einen White-Noise-Prozess dar. Sowohl für Random-Walk als auch für Markov(1)-Prozess hat (ε_t) einen Erwartungswert von Null und eine in allen Zeitpunkten gleichen Varianz.

Ein stochastischer Prozess ist bereits eindeutig definiert, wenn das System seiner endlich-dimensionalen Verteilungsfunktionen bekannt ist. In der Zeitreihenanalyse liegt das Interesse der Anwender allerdings weniger bei dem gesamten System der Verteilungsfunktionen, sondern bei bestimmten Kerngrößen. Zu den Kerngrößen eines stochastischen Prozesses gehören sein Erwartungswert, die Varianz, die Kovarianzen und die Korrelation. Sie stellen die Momente erster und zweiter Ordnung des Prozesses dar.

Definition 2.2 *Sofern $\forall t \in T$ der Erwartungswert bzw. die Varianz der Zufallsvariablen X_t existiert, ist die Mittelwertfunktion $\mu(t)$ des stochastischen Prozesses (X_t) gegeben durch*

$$\mu(t) = \mu_X(t) = E[X_t].$$

Die Varianzfunktion $\sigma^2(t)$ von (X_t) ist definiert durch

$$\sigma^2(t) = \sigma_X^2(t) = \text{Var}[X_t] = E[(X_t - E[X_t])^2].$$

Die Kovarianzfunktion $\gamma(s, t)$ ordnet jedem Paar von Zeitpunkten $s, t \in T$ die Kovarianz der Zufallsvariablen X_s und X_t zu:

$$\gamma(s, t) = \gamma_X(s, t) = \text{Cov}[X_s, X_t] = E[(X_s - \mu(s))(X_t - \mu(t))].$$

Die Korrelationsfunktion $\varrho(s, t)$ des stochastischen Prozesses ordnet jedem Paar von Zeitpunkten $s, t \in T$ die Korrelation

$$\varrho(s, t) = \varrho_X(s, t) = \text{Corr}[X_s, X_t] = \frac{\text{Cov}[(X_s, X_t)]}{\sqrt{\sigma_X^2(s)\sigma_X^2(t)}}$$

der entsprechenden Zufallsvariablen zu.

Die Momente der ersten und zweiten Ordnung für einen White-Noise-, Random-Walk- und Markov(1)-Prozess sehen dann wie folgt aus.

1. Alle Zufallsvariablen ε_t eines White-Noise-Prozesses sind identisch verteilt, besitzen folglich alle den gleichen Erwartungswert $\mu = E[\varepsilon_t]$ und die gleiche Varianz $\sigma^2 = \text{Var}[\varepsilon_t]$. Die Mittelwertfunktion und Varianzfunktion sind in diesem Fall konstant:

$$\mu(t) = \mu, \sigma^2(t) = \sigma^2 \text{ für alle } t.$$

Für $s \neq t$ sind die Zufallsvariablen ε_s und ε_t stochastisch unabhängig. Damit gilt für den Erwartungswert $E[\varepsilon_s \varepsilon_t] = E[\varepsilon_s]E[\varepsilon_t]$, woraus folgt für die Kovarianz $\text{Cov}[\varepsilon_s, \varepsilon_t] = 0$. Die Kovarianz- und Korrelationsfunktion sind wie folgt beschrieben:

$$\gamma(s, t) = \begin{cases} 0 & : s \neq t, \\ \sigma^2 & : s = t \end{cases}$$

$$\varrho(s, t) = \begin{cases} 0 & : s \neq t, \\ 1 & : s = t. \end{cases}$$

2. Für die Mittelwertfunktion eines Random-Walk-Prozesses gilt:

$$\mu_X(t) = t\mu.$$

Aus der Unabhängigkeit der ε_t folgt weiter:

$$\sigma_X^2(t) = t\sigma^2.$$

Nur für $\mu = 0$ ist die Mittelwertfunktion eines Random-Walk-Prozesses konstant, sonst ist sie linear steigend oder fallend ($\mu > 0$, $\mu < 0$). Für $\mu \neq 0$ stellt (X_t) einen Random-Walk mit Drift dar. Die Varianzfunktion ist für $\sigma^2 \neq 0$ auch eine wachsende Funktion der Zeit. Dementsprechend gilt für die Kovarianzfunktion und die Korrelationsfunktion:

$$\gamma(s, t) = \min(s, t)\sigma^2.$$

$$\varrho(s, t) = \frac{\min(s, t)}{\sqrt{st}}.$$

3. Die Darstellung $X_t = \sum_{i=0}^t a^i \varepsilon_{t-i}$ eines Markov(1)-Prozesses wird durch eine Rekursion erreicht. Wegen der Linearität des Erwartungswertes und $E[\varepsilon_t] = 0, \forall t$ ergibt sich:

$$\mu_X(t) = 0.$$

Unter Berücksichtigung der angenommenen konstanten Varianz des White-Noise-Prozesses (ε_t), ($Var[\varepsilon_t] = \sigma^2$) und der Unabhängigkeit der Variablen ε_t für alle t , gilt für die Varianzfunktion mit $t \rightarrow \infty$:

$$\sigma_X^2(t) \stackrel{t \rightarrow \infty}{=} \frac{\sigma^2}{1 - a^2}.$$

Für $t = s + \tau$ ($\tau > 0$) und wegen $Cov[\varepsilon_i, \varepsilon_j] = 0$ für $i \neq j$ folgt für die Kovarianzfunktion:

$$\gamma(s, t) \stackrel{s \rightarrow \infty}{=} \frac{\sigma^2 a^\tau}{1 - a^2} = \gamma_X(\tau).$$

Entsprechend gilt für die Korrelationsfunktion:

$$\varrho(s, t) = \frac{Cov[X_s, X_{s+\tau}]}{\sqrt{\sigma^2(s)\sigma^2(s+\tau)}} = \varrho(\tau).$$

Der entscheidende Unterschied des Markov(1)-Prozesses gegenüber dem Random-Walk-Prozess ist seine Invarianz gegenüber der Zeit. Er hat eine konstante Mittelwert- und Varianzfunktion. Die Varianz- und Kovarianzfunktion des Markov(1)-Prozesses hängen nur von der Zeitdifferenz (Lag τ) und nicht von der Zeit ab. Dieser Unterschied führt zu einer weiteren Klassifikation der stochastischen Prozesse, die auf wichtigsten Kennzahlen (Momente) des Prozesses basiert. Entsprechend der Zeitunabhängigkeit der Momente wird ein stochastischer Prozess als stationär oder schwach stationär bezeichnet.

Definition 2.3 Ein stochastischer Prozess (X_t) heißt:

- i) *mittelwertstationär*, wenn die Erwartungswertfunktion konstant ist:

$$\mu_X(t) = \mu \text{ für alle } t \in T$$

- ii) *varianzstationär*, wenn die Varianzfunktion konstant ist:

$$\sigma_X^2(t) = \sigma^2 \text{ für alle } t \in T$$

- iii) *kovarianzstationär*, wenn die Kovarianzfunktion $\gamma(s, t)$ des Prozesses nur von dem Lag $\tau = t - s$ abhängt:

$$\gamma(s, t) = \gamma(\tau) \text{ für alle } s, t \in T$$

iv) schwach stationär, wenn er mittelwert- und kovarianzstationär ist.

Damit bedeutet schwache Stationarität, dass der Erwartungswert, die Varianz und die Kovarianz invariant gegenüber den Zeitverschiebungen bleiben. Wie schon oben gezeigt wurde, sind White-Noise-Prozess und Markov(1)-Prozess schwach stationär, dagegen ist ein Random-Walk kein stationärer Prozess. Die drei oben genannten Beispiele der stochastischen Prozesse werden im Weiteren bei der Modellbildung und Testkonstruktion verwendet.

Für eine Modellierung realer Phänomene sind an erster Stelle solche linearen Prozesse von Interesse, die von den eigenen vergangenen Werten und (oder) von den Zufallsschocks der Vergangenheit und Gegenwart abhängig sind. Die zwei oft verwendeten Arten der linearen Prozesse sind ein Moving-Average-Prozess der Ordnung q ($MA[q]$) und ein Autoregressiver-Prozess der Ordnung p ($AR[p]$)¹.

Ein $MA[q]$ lässt sich durch eine Filtration von White-Noise-Prozess (ε_t) erhalten:

$$X_t = \Psi(L)\varepsilon_t \quad (2.1)$$

mit $\Psi(L) = \psi_0 - \psi_1L - \dots - \psi_qL^q$, wobei L einen Shift Operator ist (d.h. $L\varepsilon_t = \varepsilon_{t-1}$).

Ein $AR[p]$ lässt sich mit Hilfe des Shift Operators wie folgt schreiben:

$$X_t = \Phi(L)(X_t) + \varepsilon_t \quad (2.2)$$

mit $\Phi(L) = \varphi_1L + \varphi_2L^2 + \dots + \varphi_pL^p$.

Unter bestimmten Voraussetzungen besitzen der $MA[q]$ -Prozess und der $AR[p]$ -Prozess die im ersten Teil angesprochene Stationaritätseigenschaft. Ein endlicher $MA[q]$ -Prozess ist stets stationär. Für einen unendlichen Prozess soll die Koeffizientenfolge $(\psi_t)_{t=0}^{\infty}$ absolut summierbar sein. Ein endlicher $AR[p]$ -Prozess lässt sich genau dann als ein stationärer $MA[\infty]$ -Prozess darstellen, wenn alle Lösungen der charakteristischen Gleichung $1 - \phi_1z - \phi_2z^2 - \dots - \phi_pz^p = 0$ außerhalb des Einheitskreises liegen. Dann ist $AR[p]$ -Prozess stationär.

Beide oben genannten Prozesse lassen sich miteinander zu einem Auto-regressiven-Moving-Average-Prozess ($ARMA[p, q]$ -Prozess) kombinieren:

$$(1 - \Phi(L))(X_t) = \Psi(L)\varepsilon_t. \quad (2.3)$$

Ein $ARMA[p, q]$ Prozess wird zu einem stationären Prozess, wenn der autoregressive Teil sich in einen stationären $MA[\infty]$ -Prozess konvertieren lässt.

Bei den allgemeinen ökonomischen Modellen, die die Abhängigkeit zwischen vielen Größen darstellen, kommen mehrere Zeitreihen zusammen. Aus diesem Grund ist es

¹Eine ausführliche Darstellung der eindimensionalen ARMA-Prozesse ist in Schlittgen und Streiberg (1999), Kapitel 2 zu finden.

sinnvoll, einen Vektorprozess $(\vec{X}_t)_{t \in T}$, mit $\vec{X}_t \in \mathbb{R}^n$ zu betrachten. Im Weiteren wird für einen Vektorprozess statt (\vec{X}_t) einfach (X_t) geschrieben.²

Definition 2.4 Ein stochastischer Prozess (X_t) heißt ein Vektor-Moving-Average-Prozess der Ordnung q (VMA[q]), wenn er sich in der Form

$$X_t = \Psi(L)\varepsilon_t \quad (2.4)$$

darstellen lässt. Dabei sind Ψ_i die Matrizen der Koeffizienten aus $\mathbb{R}^{n \times n}$ für jedes $i = 1, 2, \dots, q$. (ε_t) ist ein Vektor-White-Noise-Prozess mit dem Mittelwert:

$$E[\varepsilon_t] = \mathbf{0}, \text{ für } \mathbf{0} \in \mathbb{R}^n.$$

Für die Kovarianzfunktion von Vektor-White-Noise-Prozess gilt:

$$E[\varepsilon_s \varepsilon_t'] = \begin{cases} \mathbf{\Lambda} & : \text{ falls } s = t, \\ \mathbf{0} & : \text{ sonst ;} \end{cases}$$

mit $\mathbf{\Lambda}$ einer $(n \times n)$ symmetrischen positiv definiten Matrix.

Analog zum eindimensionalen AR[p]-Prozess wird ein Vektor-Autoregressiver-Prozess (VAR[p]-Prozess) definiert.

Definition 2.5 Unter einem Vektor-Autoregressiven-Prozess der Ordnung p (VAR[p]) wird verstanden:

$$X_t = \Phi(L)X_t + \varepsilon_t \quad (2.5)$$

mit Φ_i der reellwertigen Matrizen $(n \times n)$ für jedes $i = 1, 2, \dots, p$ und einem (ε_t) Vektor-White-Noise-Prozess wie oben.

Eine Verallgemeinerung der beiden Vektorprozesse führt zu einem Vektor-ARMA[p, q]-Prozess:

$$(1 - \Phi(L))X_t = \Psi(L)\varepsilon_t \quad (2.6)$$

mit

$$\begin{aligned} \Psi(L) &= I_n - \Psi_1 L - \Psi_2 L^2 - \dots - \Psi_q L^q, \\ \Phi(L) &= \Phi_1 L + \Phi_2 L^2 + \dots + \Phi_p L^p. \end{aligned}$$

$(\Psi_t)_{t=0}^\infty$ stellt eine Folge der $(n \times n)$ Matrizen dar. Wenn jede von n^2 Folgen aus Elementen der Matrizen $(\psi_{ij,t})_{t=0}^\infty$ absolut summierbar ist, dann ist die Folge der Matrizen $(\Psi_t)_{t=0}^\infty$ absolut summierbar.

Gemäß dem eindimensionalen Fall werden Stationaritäts- und Invertierbarkeitsbedingungen für den mehrdimensionalen Fall in folgendem Satz zusammengefasst.

²Eine ausführliche Darstellung der Vektorprozesse ist bei Lütkepohl (1991), Kapitel 2 zu finden.

Satz 2.6 (i) Sei $\Psi_0 = I_n$. Wenn die Matrixfolge $(\Psi_t)_{t=0}^\infty$ absolut summierbar ist, dann ist ein $VMA[\infty]$ -Prozess schwach stationär.

(ii) Ein endlicher $VAR[p]$ -Prozess lässt sich als ein unendlicher $VMA[\infty]$ -Prozess mit absolut summierbarer Koeffizientenfolge $(\Psi_t)_{t=0}^\infty$ darstellen, wenn alle Lösungen der Determinanten Gleichung für $z \in \mathbb{C}^n$

$$\det(I_n - \Phi_1 z - \Phi_2 z^2 - \dots - \Phi_p z^p) = 0$$

außerhalb des Einheitskreises liegen. Insbesondere ist dann (X_t) schwach stationär.

Manchmal ist es hilfreich, einen $VAR[p]$ -Prozess als einen $VAR[1]$ -Prozess darzustellen. Die Vektoren ξ_t und v_t aus \mathbb{R}^{np} , genauso wie Matrix $F \in \mathbb{R}^{np \times np}$ werden folgendermaßen definiert :

$$\xi_t := \begin{pmatrix} X_t \\ X_{t-1} \\ \vdots \\ X_{t-p+1} \end{pmatrix} \quad v_t := \begin{pmatrix} \varepsilon_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} \quad F := \begin{pmatrix} \Phi_1 & \Phi_2 & \Phi_3 & \dots & \Phi_{p-1} & \Phi_p \\ I_n & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_n & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & I_n & \mathbf{0} \end{pmatrix}.$$

Der $VAR[p]$ -Prozess aus (2.5) kann umgeschrieben werden als:

$$\xi_t = F\xi_{t-1} + v_t \tag{2.7}$$

mit

$$E[v_s v_t'] = \begin{cases} \mathbf{Q} & : \text{ falls } s = t, \\ \mathbf{0} & : \text{ sonst;} \end{cases}$$

und

$$\mathbf{Q} := \begin{pmatrix} \Lambda & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{pmatrix}.$$

Aufgrund der Darstellung (2.7) kann die Stationaritätsbedingung für einen $VAR[p]$ -Prozess aus dem Satz (2.6) folgendermaßen formuliert werden:

Bemerkung 2.7 Wenn alle Eigenwerte der Matrix F im Betrag kleiner als 1 sind, dann ist der $VAR[p]$ -Prozess schwach stationär. Äquivalente Formulierung ist, dass ein $VAR[p]$ -Prozess schwach stationär ist, wenn alle Lösungen der charakteristischen Gleichung $\det(I_n - \Phi_1 z - \Phi_2 z^2 - \dots - \Phi_p z^p) = 0$ außerhalb des Einheitskreises liegen.

2.1.2 Schätzen von VAR-Modellen. Strukturanalyse

Ein stationärer $AR[p]$ -Prozess $y_t = \Phi(L)y_t + \varepsilon_t$ kann mit Hilfe eines Standard Regressionsmodells $y_t = \phi'x_t + \varepsilon_t$ mit $\varepsilon_t \sim iid(0, \sigma^2)$ und $x_t' = (1 \ y_{t-1} \ \dots \ y_{t-p})$ geschätzt werden. Allerdings kann die Annahme der Unabhängigkeit der Störterme von der erklärenden Variable x_t' nicht erfüllt werden. Folglich führt die Kleinste-Quadrat-Methode zum verfälschten Schätzer $\hat{\phi}$. Andererseits, wenn die Störterme ε_t einem Standard Gauss-White-Noise-Prozess folgen, kann gezeigt werden, dass der Kleinste-Quadrate-Schätzer $\hat{\phi}$ ein konsistenter Schätzer ist und einer asymptotischen Normalverteilung folgt³. t- und F-Statistiken können in diesem Fall nur asymptotisch angewendet werden. Der Kleinste-Quadrate-Schätzer für einen n-dimensionalen VAR[p]-Prozess und seine Eigenschaften werden im weiteren Abschnitt angesprochen.

Für einen n-dimensionalen stochastischen Prozess (X_t) wird angenommen, dass die Beobachtungen des Prozesses für das Zeitintervall $[-p, T]$ vorliegen. Für das VAR[p]-Modell aus (2.5) werden folgende Bezeichnungen eingeführt. Mit Y wird die Matrix aus Vektoren (X_t) mit $t = 1, \dots, T$ bezeichnet:

$$Y := (X_1 \ X_2 \ \dots \ X_T) \in \mathbb{R}^{n \times T}.$$

Die Matrix Φ ist die Matrix der Koeffizienten $\Phi_i \in \mathbb{R}^{n \times n}$:

$$\Phi := (\Phi_1 \ \Phi_2 \ \dots \ \Phi_p) \in \mathbb{R}^{n \times np}.$$

Wenn das Modell einen Achsenabschnitt $\alpha \in \mathbb{R}^n$ zulässt $X_t = \alpha + \Phi(L)X_t + \varepsilon_t$, dann ist die Matrix Φ wie folgt definiert:

$$\Phi := (\alpha \ \Phi_1 \ \Phi_2 \ \dots \ \Phi_p) \in \mathbb{R}^{n \times (np+1)}.$$

Die erklärenden Variablen in einem $VAR[p]$ -Modell sind die verzögerten Werte des Prozesses (X_t) . Sie werden als Matrix Z bezeichnet:

$$Z_t := \begin{pmatrix} X_{t-1} \\ \vdots \\ X_{t-p} \end{pmatrix} \in \mathbb{R}^{np \times 1}.$$

Entsprechend gilt für ein Modell mit einem Achsenabschnitt α :

$$Z_t := \begin{pmatrix} 1 \\ X_{t-1} \\ \vdots \\ X_{t-p} \end{pmatrix} \in \mathbb{R}^{(np+1) \times 1}.$$

Für die T Beobachtungen aus dem Zeitintervall $[0, T-1]$ ist die Matrix $Z = (Z_0 \ \dots \ Z_{T-1}) \in \mathbb{R}^{np \times T}$. Die Matrix der Störterme lässt sich wie folgt schreiben:

$$\varepsilon := (\varepsilon_1 \ \dots \ \varepsilon_T) \in \mathbb{R}^{n \times T}.$$

³Vgl. Hamilton (1994).

Das $VAR[p]$ -Prozess aus (2.5) lässt sich in einem Intervall $[-p, T]$ wie folgt umschreiben:

$$Y = \Phi Z + \varepsilon. \quad (2.8)$$

Diese Schreibweise gilt auch für ein Modell mit dem Achsenabschnitt α .

Der Kleinste-Quadrate-Schätzer für die unbekannte Matrix der Koeffizienten Φ lässt sich in einer Matrixform darstellen.

$$\hat{\Phi} = YZ'(ZZ')^{-1}. \quad (2.9)$$

Der Schätzfehler ist durch $\hat{E} = \hat{Y} - Y = \hat{\Phi}Z - Y$ definiert.

Satz 2.8 (Asymptotische Eigenschaften des Schätzers) *Ist ein n -di-mensionaler $VAR[p]$ -Prozess (X_t) schwach stationär und (ε_t) ein Standard Vektor-White-Noise-Prozess, dann ist der Kleinste-Quadrate-Schätzer für die Koeffizientenmatrix $\hat{\Phi}$ aus (2.9) ein konsistenter Schätzer, d.h. es gilt*

$$\lim_{n \rightarrow \infty} P(|\hat{\Phi}_n - \Phi| < \varepsilon) = 1, \forall \varepsilon > 0.$$

Eine Prognose ist eine Vorhersage der zukünftigen Ereignisse mit Hilfe der Vergangenheitsinformation. Aus ökonomischer Sicht unterscheidet sich die Prognose von der Wahrsagerei „durch die explizite Unterstellung eines Modells für den beobachteten Prozess.“⁴ Die Prognose ist dann ungültig, wenn das angenommene Modell nicht zutrifft oder sich ändert. Es ist leicht nachvollziehbar, dass die gesammelten Daten für die Vorhersagen über eine Variable benutzen werden sollen, wenn die Beobachtungswerte für die Variable zugänglich sind und Informationen bezüglich zukünftiger Entwicklung der Variable beinhalten. Formal kann diese Idee folgendermaßen ausgedrückt werden: Für den Wert der Variable y_t zu dem zukünftigen Zeitpunkt $T + h$ wird eine Prognose als die Funktion der vergangenen Werte y_T, y_{T-1}, \dots definiert

$$\hat{y}_{T+h} = f(y_T, y_{T-1}, \dots).$$

In vielen Anwendungen wird die Prognosefunktion $f(\cdot)$ als eine lineare Funktion angenommen.

$$\hat{y}_{T+h} = a + b_1 y_T + b_2 y_{T-1} + \dots$$

Bei der Handlung mit ökonomischen Variablen kann der zukünftige Wert einer Variable $y_{i,T+h}, i = 1, \dots, n$ von den anderen Variablen und ihren Vergangenheitswerten $y_{1,T}, y_{2,T}, \dots, y_{n,T}, y_{1,T-1}, \dots, y_{n,T-1}, \dots$ abhängig sein,

$$\hat{y}_{1,T+h} = f_1(y_{1,T}, y_{2,T}, \dots, y_{n,T}, y_{1,T-1}, \dots, y_{n,T-1}, \dots).$$

⁴Schlittgen und Streitberg (1999), Kapitel 4.

In der Realität steht eine endliche Anzahl der Beobachtungen eines Prozesses zur Verfügung. Häufig werden die Prognosen nur für die nächste Periode gemacht. Der prognostizierte Wert der i -ten Variable y_i lässt sich dann für eine Periode ($h = 1$) wie folgt schreiben:

$$\hat{y}_{i,T+1} = a_i + \sum_{i=1}^n b_{i1}y_{i,T} + \dots + \sum_{i=1}^n b_{ip}y_{i,T-p+1}.$$

Die Prognose geht von p Vergangenheitswerten aus, dabei wird auch eine lineare Form der Schätzfunktion unterstellt. Die Prognose kann entsprechend in einer Matrixform dargestellt werden:

$$\hat{Y}_{T+1} = a + B_1Y_T + B_2Y_{T-1} + \dots + B_pY_{T-p+1} = a + BY.$$

Die Differenz zwischen dem tatsächlich eingetretenen Wert Y_{T+1} und dem geschätzten Wert \hat{Y}_{T+1} wird mit ε_{T+1} bezeichnet ($\varepsilon_{T+1} = Y_{T+1} - \hat{Y}_{T+1}$). Sie stellt den Prognosefehler dar, der ein zentrales Element des Verfahrens zur Beurteilung der Prognosen ist. Im Weiteren wird ausschließlich von einem Prognoseverfahren gesprochen, das den mittleren quadratischen Prognosefehler minimiert und in dieser Weise als optimale Methode verstanden wird. Lütkepohl (1991) zeigt, dass die Bildung des bedingten mathematischen Erwartungswertes der Variable zum Zeitpunkt t für einen Vorhersagehorizont h eine optimale Prognosemethode ist,

$$\hat{y}_{t+h} \equiv E(y_{t+h}|\Omega_t) = E(y_{t+h}|y_s|s \leq t).$$

Die Informationsmenge bei der Bildung des Erwartungswertes zum Zeitpunkt t , Ω_t , beinhaltet nur die Vergangenheitswerte der Variable y_t .

Ausgehend von der oben beschriebenen Idee für eine Prognose bieten die n -dimensionalen endlichen VAR[p]-Modelle eine gute Möglichkeit für die Schätzung der Zukunftswerte der ökonomischen Variablen. Für einen n -dimensionalen VAR[p]-Prozess $X_t = \Phi(L)X_t + \varepsilon_t$ impliziert die Optimalitätsbedingung der bedingten Erwartungen

$$E_t(X_{t+h}) = \sum_{i=1}^p \Phi_i E_t(X_{t+h-i}), \tag{2.10}$$

dass $E_t(X_{t+h})$ eine optimale h -Schritt Prognose ist, vorausgesetzt, dass der Fehlertermprozess (ε_t) einen Standard White-Noise-Prozess präsentiert. Die Formel (2.10) kann als ein rekursiver Algorithmus für die Berechnung der Vorhersagen benutzt werden. Es wird mit einem $h = 1$ gestartet.

$$\begin{aligned} E_t(X_{t+1}) &= \Phi_1 X_t + \dots + \Phi_p X_{t-p+1} \\ E_t(X_{t+2}) &= \Phi_1 E_t(X_{t+1}) + \Phi_2 X_t + \dots + \Phi_p X_{t-p+2} \\ &\vdots \end{aligned}$$

Für einen n-dimensionalen VAR[1]-Prozess $X_t = \Phi X_{t-1} + \varepsilon_t$, mit dem Mittelwert 0, lässt sich mit der Rekursion der geschätzte Wert des Prozesses zum Zeitpunkt $t + h$ wie folgt berechnen:

$$\widehat{X}_{t+h} \equiv E(X_{t+h}) = \Phi^h X_t = \Phi \widehat{X}_{t+h-1}. \quad (2.11)$$

Der Prognosefehler ist dann $X_{t+h} - E(X_{t+h}) = \sum_{i=0}^{h-1} \Phi^i \varepsilon_{t+h-i}$. Die Prognose ist unverfälscht, weil $E_t(X_{t+h} - E_t(X_{t+h})) = 0$ gilt.

Für einen n-dimensionalen VAR[p]-Prozess aus (2.5) existiert eine äquivalente Darstellung als ein n-dimensionaler VAR[1]-Prozess (2.7). Der Mittelwert des Prozesses wird vorerst als 0 angenommen. Eine optimale Prognose (2.11) lautet entsprechend

$$\widehat{\xi}_t(h) = F^h \xi_t = F \widehat{\xi}_{t+h-1}. \quad (2.12)$$

Wie leicht zu erkennen ist, gilt für die optimale h-Schritt Prognose $\widehat{\xi}_{t+h}$

$$\widehat{\xi}_{t+h} = \begin{pmatrix} \widehat{X}_{t+h} \\ \widehat{X}_{t+h-1} \\ \vdots \\ \widehat{X}_{t+h-p+1} \end{pmatrix}.$$

Um die optimale Prognose für den VAR[p]-Prozess in Termen von Φ_i zu bekommen, wird die Matrix $\mathbf{J} := [\mathbf{I}_n \mathbf{0} \cdots \mathbf{0}] \in \mathbb{R}^{np \times n}$ definiert. Die optimale Prognose für den Prozess (X_t) lautet dann

$$\begin{aligned} \widehat{X}_{t+h} &= \mathbf{J} F \widehat{\xi}_{t+h-1} \\ &= [\Phi_1 \cdots \Phi_p] \widehat{\xi}_{t+h-1} \\ &= \Phi_1 \widehat{X}_{t+h-1} + \cdots + \Phi_p \widehat{X}_{t+h-p}. \end{aligned} \quad (2.13)$$

Für einen n-dimensionalen VAR[p]-Prozess mit einem konstanten Mittelwert $\mu \neq 0$ und einem Achsenabschnittterm α lässt sich zeigen, dass die optimale h-Schritt Prognose folgende Form hat:

$$\widehat{X}_{t+h} = \alpha + \Phi_1 \widehat{X}_{t+h-1} + \cdots + \Phi_p \widehat{X}_{t+h-p}.$$

Der Prognosefehler wird mit Hilfe der MA[∞]-Darstellung des VAR[p]-Prozesses beschrieben. Für einen stationären Prozess mit der VAR[1]-Darstellung aus (2.7) und mit dem Mittelwert Null, gilt

$$\begin{aligned} X_{t+h} - \widehat{X}_{t+h} &= \mathbf{J}(\xi_{t+h} - \widehat{\xi}_{t+h}) = \mathbf{J} \left[\sum_{i=0}^{h-1} F^i v_{t+h-i} \right] \\ &= \sum_{i=0}^{h-1} \mathbf{J} F^i \mathbf{J}' v_{t+h-i} = \sum_{i=0}^{h-1} \Psi_i \varepsilon_{t+h-i}. \end{aligned} \quad (2.14)$$

Die Matrizen Ψ_i sind die Koeffizientenmatrizen der $MA[\infty]$ -Darstellung aus Satz (2.6).

Die Koeffizientenmatrix Φ eines n -dimensionalen $VAR[p]$ -Prozesses mit einem Achsenabschnitt und einem unbekanntem Mittelwert, kann mit Hilfe der Kleinste-Quadrate-Methode durch $\widehat{\Phi} = (\widehat{\alpha} \ \widehat{\Phi}_1 \ \dots \ \widehat{\Phi}_p)$ geschätzt werden. Eine optimale Prognose für einen geschätzten $VAR[p]$ -Prozess lautet aus diesem Grund:

$$\widehat{X}_{t+h} = \widehat{\alpha} + \widehat{\Phi}_1 \widehat{X}_{t+h-1} + \dots + \widehat{\Phi}_p \widehat{X}_{t+h-p}.$$

Entsprechend sieht der Vorhersagefehler folgendermaßen aus:

$$\begin{aligned} X_{t+h} - \widehat{X}_{t+h} &= (X_{t+h} - \widehat{X}_{t+h}) + (\widehat{X}_{t+h} - \widehat{\widehat{X}}_{t+h}) \\ &= \sum_{i=0}^{h-1} \Psi_i \varepsilon_{t+h-i} + (X_{t+h} - \widehat{\widehat{X}}_{t+h}). \end{aligned}$$

Unter der Stationaritätsannahme für den Prozess (X_t) und den Gauss-White-Noise-Prozess für den Störterm ε_t kann gezeigt werden, dass diese Prognose eine unverfälschte Prognose ist (d.h. $E_t(X_{t+h} - \widehat{\widehat{X}}_{t+h}) = 0$), wenn der $VAR[p]$ -Prozess geschätzt wird⁵.

VAR -Modelle werden oft zur Analyse der Abhängigkeit zwischen den Variablen angewendet, weil sie die Korrelation in einem Variablensatz repräsentieren. Es muss allerdings beachtet werden, dass dabei im Allgemeinen lediglich die lineare Abhängigkeit der Variablen untersucht werden kann. Im folgenden Kapitel werden die drei möglichen theoretischen Ansätze zur Analyse und Interpretation der VAR -Modelle diskutiert, nämlich die Granger-Kausalität, die Impuls-Antwort-Funktionen und die Varianz-Dekomposition.

Granger-Kausalität

Für Zeitreihen können allgemeine Aussagen über die zeitliche Kausalität der Ereignisse getroffen werden. Angenommen, ein Ereignis A findet vor einem Ereignis B statt, dann ist es möglich, dass A Ursache von B ist, aber es ist unmöglich, dass B Ursache von A ist. Das bedeutet, dass nur Geschehen aus der Vergangenheit die heutigen Ereignisse verursachen kann, umgekehrt hat die Zukunft keinen unmittelbaren Einfluss auf Heute. Diese Grundidee entspricht der Granger-Kausalität (1969), die wie folgt formuliert werden kann: Die Variable X beeinflusst nach Granger die Variable Y , wenn die Vergangenheitswerte von X für die Vorhersage von Y relevant sind. Bei dem Granger-Kausalitätstest wird angenommen, dass alle für die Vorhersage X und Y relevanten Informationen nur in den entsprechenden Zeitreihen (X_t) und (Y_t) enthalten sind. Der Granger-Kausalitätstest beinhaltet eine Schätzung der folgenden Regressionen:

$$Y_t = \sum_{i=1}^n \alpha_i X_{t-i} + \sum_{j=1}^n \beta_j Y_{t-j} + u_{1t} \tag{2.15}$$

⁵Vgl. Lütkepohl (1991), Kapitel 2.

$$X_t = \sum_{i=1}^m \lambda_i X_{t-i} + \sum_{j=1}^m \delta_j Y_{t-j} + u_{2t} \quad (2.16)$$

Es werden vier Fälle der Kausalität unterschieden:

1. Einseitige Kausalität von Y durch X ($X \rightarrow Y$) liegt dann vor, wenn die geschätzten Koeffizienten vor X in der Gleichung (2.15) sich statistisch von Null unterscheiden, d.h. $\sum_{i=1}^n \alpha_i \neq 0$, und die geschätzten Koeffizienten vor Y in der Gleichung (2.16) sich statistisch von Null nicht unterscheiden, d.h. $\sum_{j=1}^m \delta_j = 0$.
2. Umgekehrt liegt die einseitige Kausalität von X durch Y ($Y \rightarrow X$) vor, wenn die geschätzten Koeffizienten vor X in der Gleichung (2.15) sich statistisch von Null nicht unterscheiden, d.h. $\sum_{i=1}^n \alpha_i = 0$, und die geschätzten Koeffizienten vor Y in der Gleichung (2.16) sich statistisch von Null unterscheiden, d.h. $\sum_{j=1}^m \delta_j \neq 0$.
3. Beidseitige Kausalität liegt vor, wenn die geschätzten Koeffizienten in beiden Gleichungen statistisch signifikant von Null verschieden sind.
4. Unabhängigkeit der beiden Variablen liegt vor bzw. es besteht keine Granger-Kausalität zwischen den Variablen, wenn die Y_t nur von eigenen Vergangenheitswerten Y_{t-j} abhängig ist, und nicht von X_{t-i} , bzw. X_t ist abhängig nur von X_{t-i} , und nicht von Y_{t-j} .

Allgemein besteht der Test auf Granger-Kausalität aus einer Schätzung der entsprechenden Regressionsgleichungen und aus dem Vergleich der entsprechenden quadrierten Fehlern der Schätzung. Bei dem Test wird wie folgt vorgegangen. Als erster Schritt wird die beschränkte Regressionsgleichung der Variable Y geschätzt,

$$Y_t = \sum_{j=1}^n \beta_j Y_{t-j} + e_t.$$

Die Summe der quadrierten Schätzfehler ist $RSS_R = \sum_{t=1}^T \hat{e}_t^2$, dabei ist T die Anzahl der Beobachtungen. Im nächsten Schritt wird die unbeschränkte Regressionsgleichung der Variable Y geschätzt.

$$Y_t = \sum_{i=1}^n \alpha_i X_{t-i} + \sum_{j=1}^m \beta_j Y_{t-j} + u_t.$$

Die Summe der quadrierten Residuen aus der unbeschränkten Gleichung ist $RSS_{UR} = \sum_{t=1}^T \hat{u}_t^2$. Die zu testende Null-Hypothese lautet: $H_0 : \sum_{i=1}^n \alpha_i = 0$, d.h. X_{t-i} beeinflussen den Wert von Y_t nicht. Die Teststatistik F :

$$F = \frac{(RSS_R - RSS_{UR})/n}{RSS_{UR}/(T - k)}$$

folgt der F-Verteilung mit n und $T - k$ Freiheitsgrad. Hier ist $k = n + m$ die Anzahl der geschätzten Parameter in der unbeschränkten Regression von Y und n die Anzahl der vor X_{t-i} stehenden Parameter. Wenn der berechnete Wert von F größer als der kritische Wert F_α mit dem gegebenen Signifikanzniveau α ist, dann wird die Null-Hypothese nicht angenommen, d.h. die Vergangenheitswerte X_{t-i} beeinflussen den heutigen Wert Y_t . Die Schritte werden für die Gleichung (2.16) wiederholt.

Der Granger-Kausalitätstest reagiert sehr empfindlich auf eine Änderung der Anzahl der geschätzten Parameter in den Regressionen. In vielen Untersuchungen wird beim Granger-Kausalitätstest eher eine große Anzahl der Zeitverzögerungen (Lags) verwendet.

In einem mehrdimensionalen Fall wird die Idee der Granger-Kausalität mit Hilfe einer optimalen Prognose formalisiert. Angenommen, dass $\hat{X}_{t+h} = E(X_{t+h}|\Omega_t)$ eine optimale (d.h. die Prognose minimiert den mittleren quadrierten Prognosefehler) h -Schritt Prognose für einen n -dimensionalen $VAR[p]$ -Prozess (X_t) ist. Bezeichne die Kovarianzmatrix der Schätzfehler mit $\Sigma_X(h)$

$$\Sigma_X(h) := E \left(\sum_{i=0}^{h-1} \Phi_i \varepsilon_{t+h-i} \left(\sum_{i=0}^{h-1} \Phi_i \varepsilon_{t+h-i} \right)' \right) = \sum_{i=0}^{h-1} \Phi_i \Sigma_\varepsilon (\Phi_i)' \quad (2.17)$$

Ein n -dimensionaler Prozess (Z_t) ist die Ursache für einen anderen n -dimensionalen Prozess (X_t) , wenn

$$\Sigma_X(h|\Omega_t) < \Sigma_X(h|\Omega_t \setminus \{Z_s | s \leq t\}) \quad (2.18)$$

für mindestens ein $h = 1, 2, \dots$. Der Prozess (Z_t) ist Granger-kausal für den Prozess (X_t) .

Eine Korrektur der Definition ist unumgänglich, wenn die Prozesse unterschiedliche Dimensionen haben. In diesem Fall ist ein Prozess (Z_t) für einen anderen Prozess (X_t) Granger-kausal, wenn gilt:

$$\Sigma_X(h|\Omega_t) \neq \Sigma_X(h|\Omega_t \setminus \{Z_s | s \leq t\}) \quad (2.19)$$

Ein Problem für die praktische Überprüfung der Granger-Kausalität ist die Auswahl der Informationsmenge Ω_t . Im Normalfall sind nicht alle relevanten Informationen weltweit für die Vorhersagen erhältlich. Dies schließt aus, dass eine optimale Prognose mit gegebenem Ω_t bestimmt werden kann. Auf Grund dessen wird in der Praxis eine weniger anspruchsvolle Definition der Kausalität benutzt. Statt aller relevanter Informationen werden nur die vergangenen und derzeitigen Werte der Prozesse berücksichtigt, damit ist $\Omega_t = \{X_s, Z_s | s \leq t\}$.

Weitere Probleme bei der Anwendung der Granger-Kausalität-Analyse für mehrdimensionale VAR -Modelle bringen die Datenfrequenz, die Anzahl der Variablen in der Informationsmenge oder eine saisonale Anpassung mit sich. Es könnte sein, dass

beispielsweise der Zinssatz in einem monatlichen Datensystem Granger-kausal zur Inflationsrate ist. Allerdings wird keine Gran-ger-Kausalität der beiden Variablen in den Quartalsdaten bestätigt. Es ist genauso möglich, dass in den saisonbereinigten Daten keine Granger-Kausalität zwischen Zinssatz und Inflation festgestellt wird. Eventuell kann eine weitere Variable, wie die Geldmenge, die Granger-Kausalität zwischen dem Zinssatz und der Inflation verursachen. Wird diese Variable aus der Informationsmenge ausgelassen, kann die Kausalität nicht festgestellt werden.

Der Hauptkritikpunkt der Granger-Kausalität-Analyse ist, dass mit ihrer Hilfe keine funktionale Kausalität festgestellt werden kann. Die Zeitreihen (zum Beispiel die Aktienpreise oder die Zinssätze), die ein zukunftsorientiertes Verhalten beschreiben, können gute vorläufige Indikatoren für mehrere ökonomische Größen darstellen. Allerdings können sie nicht die Bewegung dieser ökonomischen Größen verursachen. Jedoch lässt sich dadurch die Fundamentalebewertung der Aktienpreise begründen, wenn die Fundamentaldaten Granger-kausal zu den Aktienpreisen sind. Es ist immer zu beachten, dass die Granger-Kausalität als eine zeitliche, aber nicht als eine funktionale Kausalität interpretiert werden darf. Der Granger-Kausalitätstest kann für zukunftsorientierte Zeitreihen zur Schätzung der Markteffizienz oder zur zeitlichen Prognose verwendet werden.

Impuls-Antwort-Analyse

Im vorherigen Abschnitt wurde die Abhängigkeit eines Variablensystems $X_t \in \mathbb{R}^n$ von einem anderen System $Y_t \in \mathbb{R}^m$ untersucht. Die Definition der Granger-Kausalität beinhaltet alle Variablen beider Systeme. In der Praxis ist es häufig von besonderem Interesse, wie das Gesamtsystem auf eine Änderung nur einer Variable reagiert. Im zweidimensionalen Fall kann die Reaktion einer Variable auf die Änderung der anderen Variable als eine Kausalität interpretiert werden. Im mehrdimensionalen Kontext ist dies nicht mehr der Fall. Im folgenden Abschnitt wird die Reaktion des ganzen Variablensystems als eine Antwort auf den Impuls einer Variable dargestellt. Diese Art der Impuls-Antwort-Analyse wird oft auch Multiplikator-Analyse genannt. Als Beispiel kann ein System aus ökonomischen Variablen ausgewählt werden, das den Zinssatz als eine erklärende Variable beinhaltet. Von Interesse ist die Reaktion dieses Systems auf eine unerwartete Änderung des Zinssatzes, die als ein exogener Schock bezeichnet wird. Ein Anstieg oder eine Senkung des Zinssatzes wird von der Zentralbank administrativ und ohne vorläufige Ankündigung durchgeführt. Die Schocks im Preisniveau können beispielsweise auch exogen sein, wie im Falle der Ölkrise von 1973/74, als die OPEC eine Aktion zur Erhöhung der Preise beschlossen hatte.

Nehmen wir an, dass der Effekt einer Investitionserhöhung auf das System aus Investition, Einkommen und Konsum untersucht wird. Soll das Problem allgemeiner aufgefasst werden, ist in einem n -dimensionalen $VAR[1]$ -Modell die Reaktion des Systems auf einen Prognosefehler in einer Variable X_j zum Zeitpunkt $t = 0$ zu untersuchen. Alle Variablen $(X_{i,0})_{i=1}^n$ werden zum Zeitpunkt $t = 0$ auf Null (oder gleich dem

Mittelwert μ_i) gesetzt. Die Änderung in der Variable $X_{j,0}$ wird durch den Fehlerterm $\varepsilon_{j,0} = 1$ simuliert; für alle $i \neq j$ sind die Störterme ebenfalls Null. Verfolgen wir einen Einheitsschock in der ersten Variable zum Zeitpunkt $t = 0$.

$$X_0 = \begin{pmatrix} x_{1,0} \\ x_{2,0} \\ \vdots \\ x_{n,0} \end{pmatrix} = \begin{pmatrix} \varepsilon_{1,0} \\ \varepsilon_{2,0} \\ \vdots \\ \varepsilon_{n,0} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Zum Zeitpunkt $t = 1$ sieht das System wie folgt aus:

$$X_1 = \begin{pmatrix} x_{1,1} \\ x_{2,1} \\ \vdots \\ x_{n,1} \end{pmatrix} = \Phi X_0 = \begin{pmatrix} \phi_{11} \\ \phi_{12} \\ \vdots \\ \phi_{n1} \end{pmatrix}$$

. Zum Zeitpunkt $t = 2$ entsprechend:

$$X_2 = \begin{pmatrix} x_{1,2} \\ x_{2,2} \\ \vdots \\ x_{n,2} \end{pmatrix} = \Phi X_1 = \Phi^2 X_0$$

. Nach dem k -ten Iterationsschritt bekommt man

$$X_k = \begin{pmatrix} x_{1,k} \\ x_{2,k} \\ \vdots \\ x_{n,k} \end{pmatrix} = \Phi X_{k-1} = \Phi^k X_0,$$

die Reaktion des Systems auf einen Schock in der ersten Variable. Sie steht in der ersten Spalte der Matrix Φ^k . Analog hat man die Systemreaktion auf eine Änderung der j -ten Variable in der Periode 0 nach k Perioden in der j -ten Spalte der Matrix Φ^k . Auf diese Weise repräsentieren die Elemente der Matrix Φ^k die Effekte der Einheitsschocks in Variablen auf das System nach der k -ten Periode. Die Spalten der Matrix Φ^k werden Impuls-Antwort- oder dynamische Multiplikatoren genannt.

Eine Verbindung der Granger-Kausalität und der Impuls-Antwort-Analyse liefert folgende Bemerkung:

Bemerkung 2.9 *Eine Innovation in der k -ten Variable hat keinen Effekt auf die anderen Variablen, wenn die Variable nicht Granger-kausal zu diesen Variablen ist.*

Analog zur Granger-Kausalität ist die potentielle Unvollständigkeit eines VAR-Modells die größte Beschränkung für die Untersuchungen des wirtschaftlichen Variablen-systems. In der Realität sind beinahe alle Variablen untereinander abhängig, daher

werden in der empirischen Forschung VAR-Modelle benutzt. Alle Effekte der weggelassenen Variablen werden einem Störterm, also den Innovationen zugeordnet. Zu erheblichen Verzerrungen in der Impuls-Antwort-Analyse kann eine wichtige, aber aus dem System weggelassene Variable führen. Solche Analysen sind unbrauchbar für strukturelle Interpretationen. Damit kann das System nur begrenzt für eine Vorhersage verwendet werden.

Weitere Probleme bei der Interpretation einer Impuls-Antwort-Analyse stellen die Fehlermessmethoden, die Benutzung der saisonal bereinigten Daten oder gleichzeitig aggregierten Variablen dar⁶.

Varianz-Dekomposition

Nach dem Satz (2.6) existiert für einen stationären n-dimensionalen $VAR[p]$ -Prozess $X_t = \alpha + \sum_{i=1}^p \Phi_i X_{t-i} + \varepsilon_t$ eine $MA[\infty]$ -Darstellung, welche eine weitere Möglichkeit zur Interpretation des $VAR[p]$ -Modells gibt. Eine Orthogonalisierung des White-Noise-Prozesses (ε_t) führt zu einer weiteren äquivalenten Darstellung des Prozesses (X_t)⁷:

$$X_t = \alpha + \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i} = \alpha + \sum_{i=0}^{\infty} \Psi_i P P^{-1} \varepsilon_{t-i} = \alpha + \sum_{i=0}^{\infty} \Theta_i w_{t-i}.$$

Dabei ist $w_t := P^{-1} \varepsilon_t$ ein White-Noise-Prozess und $\Theta_i := \Phi_i P$. Die Matrix P erhält man aus einer Zerlegung der positiv definiten symmetrischen Kovarianzmatrix der Störterme ε_t , Σ_ε . Der Prozess (w_t) ist ein orthogonaler White-Noise-Prozess mit einer Kovarianzmatrix $\Sigma_w = \mathbf{I}_n$. Der Prognosefehler ist bei der optimalen h-Schritt Prognose \hat{X}_{t+h} mit (2.14) wie folgt berechnet:

$$X_{t+h} - \hat{X}_{t+h} = \sum_{i=0}^{h-1} \Psi_i \varepsilon_{t+h-i} = \sum_{i=0}^{h-1} \Theta_i w_{t+h-i}.$$

Die j-te Komponente des Vorhersagefehlers ist dann

$$\begin{aligned} x_{j,t+h} - \hat{x}_{j,t+h} &= \sum_{i=0}^{h-1} (\theta_{j1,i} w_{1,t+h-i} + \dots + \theta_{jn,i} w_{n,t+h-i}) \\ &= \sum_{k=1}^n (\theta_{jk,0} w_{k,t+h} + \dots + \theta_{jk,h-1} w_{k,t+1}). \end{aligned} \tag{2.20}$$

Auf diese Weise ist die j-te Komponente des Prognosefehlers als die Summe der Innovationen in allen anderen Komponenten des Prozesses (X_t) dargestellt. Die Vektoren (w_t) mit $t = 1, 2, \dots$ sind unkorreliert und haben eine Varianz gleich 1. Damit sieht

⁶Vgl. Lütkepohl (1987).

⁷Vgl. Lütkepohl (1991), Kapitel 2

der mittlere quadrierte Prognosefehler (MSE) für die Vorhersagekomponente $\hat{x}_{j,t+h}$ wie folgt aus:

$$MSE(\hat{x}_{j,t+h}) = E[(x_{j,t+h} - \hat{x}_{j,t+h})^2] = \sum_{k=1}^n (\theta_{jk,0}^2 + \dots + \theta_{jk,h-1}^2).$$

Folglich gilt für e_k , k -te Spalte der Matrix \mathbf{I}_n ,

$$\theta_{jk,0}^2 + \dots + \theta_{jk,h-1}^2 = \sum_{i=0}^{h-1} (e_j' \Theta_i e_k)^2.$$

Werden die beiden Seiten dieser Gleichung durch $MSE(\hat{x}_{j,t+h}) = \sum_{i=0}^{h-1} \sum_{k=1}^n \theta_{jk,i}^2$ dividiert, folgt

$$\frac{\theta_{jk,0}^2 + \dots + \theta_{jk,h-1}^2}{MSE(\hat{x}_{j,t+h})} = \frac{\sum_{i=0}^{h-1} (e_j' \Theta_i e_k)^2}{MSE(\hat{x}_{j,t+h})}$$

$$\omega_{jk,h} := \frac{\sum_{i=0}^{h-1} (e_j' \Theta_i e_k)^2}{MSE(\hat{x}_{j,t+h})}. \quad (2.21)$$

$\omega_{jk,h}$ stellt den Anteil der Varianz des Prognosefehlers von Variable j bei Innovationen in der Variable k an der Gesamtvarianz der j -ten Komponente dar. Auf diesem Weg wurde die Varianz des Prognosefehlers in die Komponenten zerlegt, die durch Innovationen in verschiedenen Variablen des Systems bestimmt sind. In Matrixform erhält man:

$$\Sigma_X(h) = MSE(\hat{X}_{t+h}) = \sum_{i=0}^{h-1} \Theta_i \Theta_i' = \sum_{i=0}^{h-1} \Psi_i \Sigma_\varepsilon \Psi_i'.$$

Die Diagonalelemente dieser Matrix sind die einzelnen mittleren quadrierten Prognosefehler $MSE(\hat{x}_{j,t+h})$, die in der Gleichung (2.21) verwendet werden können.

Die bei der Granger-Kausalität und der Impuls-Antwort-Analyse genannte Kritikpunkte können auch für die Varianz-Dekomposition-Methode angeführt werden. Eine ökonomische Interpretation dieser ist aufgrund einer hohen Sensibilität des Systems gegenüber der Anzahl der Variablen, der saisonal bereinigten Daten oder der aggregierten Zeitreihen sehr schwierig.

2.1.3 Probleme

Zu den Vorteilen der VAR-Methode gehört beispielsweise die Einfachheit des Modells. Es wird nicht zwischen endogenen und exogenen Variablen unterschieden, weil alle Variablen in einem VAR-Modell endogen sind. Die Schätzung der Parameter ist einfach, weil der gewöhnliche Kleinstquadratschätzer sogar für einzelne Gleichungen benutzt werden kann. Die Prognosen, die mit Hilfe der VAR-Modelle gemacht wurden, sind in

vielen Fällen besser als Prognosen, die mit komplexeren simultanen Gleichungssystemen errechnet wurden.⁸

Allerdings hat die Methodik mehrere Nachteile. Einer davon ist, dass ein VAR-Modell a-theoretisch ist. Ein VAR-Modell schließt viele Informationen aus. Das Einbeziehen oder das Ausschließen bestimmter Variablen in ein simultanes Gleichungssystem spielt jedoch eine kritische Rolle bei der Identifikation des Modells.

Ein großes praktisches Hindernis bei der VAR-Modellierung ist die Auswahl der Laglänge p für die Modellvariablen. Bei einem n -dimensionalen VAR[p]-Modell sind $n(n + p)$ Parameter zu schätzen. Die Anzahl der Parameter führt zur Beschränkung des Freiheitsgrades mit allen damit verbundenen Problemen.

Die Annahme der Stationarität der Variablen führt zu weiteren Einschränkungen des VAR-Modells. Wenn alle Prozesse im Modell nicht stationär sind, so müssen die Datenreihen transformiert werden, zum Beispiel Differenzen gebildet werden oder eine Trendbereinigung stattfinden. Dies führt einerseits zu einem Verlust der Information und andererseits zu bestimmten Effekten auf Verteilungen der Schätzer. Wenn gemischte Prozesse im Modell vorhanden sind (stationäre und nicht stationäre), dann ist die Transformation der Datenreihen besonders problematisch.

Die geschätzten Koeffizienten eines VAR-Modells sind oft schwierig zu interpretieren. Entsprechend werden die ökonomischen Systeme einer zusätzlichen Analyse wie Überprüfung auf Granger-Kausalität, Impuls-Antwort-Analyse oder Varianz-Dekomposition unterzogen. Der wichtigste Kritikpunkt der Granger-Kausalität-Analyse ist seine zeitliche und nicht funktionale Orientierung. Die Ergebnisse des Tests dürfen nicht ohne weiteres als die Identifikation bestimmter kausaler Verbindungen interpretiert werden.

Die größte Einschränkung, sowohl für die Granger-Kausalität als auch für die Impuls-Antwort-Analyse, stellt die potentielle Unvollständigkeit eines VAR-Modells für die zu untersuchenden Systeme der wirtschaftlichen Variablen. Ein Problem für die praktische Überprüfung der Granger-Kausalität oder bei der Impuls-Antwort-Analyse ist die Auswahl der Informationsmenge Ω_t . Die Sensibilität der Tests bezüglich der Anzahl der Verzögerungen im VAR[p]-Modell stellt ein zusätzliches Hindernis für die Analyse dar.

Gleiche Kritikpunkte wie bei der Granger-Kausalität oder der Impuls-Antwort-Analyse können für die Varianz-Dekomposition-Methode angeführt werden.

Eine ökonomische Interpretation wird zusätzlich durch die verschiedenen Datenfrequenzen, die saisonal bereinigten Datenreihen oder die aggregierten Zeitreihen erschwert. Ein weiterer Nachteil ist, dass die Prognosen aus den VAR-Modellen für eine politische Handlungsempfehlung wenig geeignet sind.

⁸Vgl. Lütkepohl (1997)

2.2 Trends und integrierte Prozesse

2.2.1 Deterministischer oder stochastischer Trend?

In den vorherigen Abschnitten wurde ein eindimensionales Zeitreihenmodell in Form eines stabilen $MA[\infty]$ -Prozesses dargestellt, $X_t = \mu + \Psi(L)\varepsilon_t$. Der (ε_t) ist ein White-Noise-Prozess mit einem Mittelwert gleich 0 und konstanter Varianz σ^2 . Erstens ist zu erwähnen, dass der Erwartungswert des Prozesses konstant und unabhängig von der Anzahl der Beobachtungen ist, $E(X_t) = \mu$. Zweitens, ist die h -Schritt-Prognose \hat{X}_{t+h} optimal, d.h. sie konvergiert zu dem Mittelwert: $\lim_{h \rightarrow \infty} \hat{X}_{t+h} = \mu$.

Die Annahme eines konstanten Mittelwertes bei der Modellierung kann für mehrere ökonomische und Finanzzeitreihen nicht akzeptiert werden. Bei vielen Zeitreihen wie dem BNP oder einem Aktienindex ist ein Trend in ihren Verläufen feststellbar. Beispielsweise zeigt die folgende Graphik des nominalen Aktienindex für die USA (2.1) einen exponentiellen Verlauf. Dieser Trend soll in allen zukünftigen Prognosen aufgenommen werden.

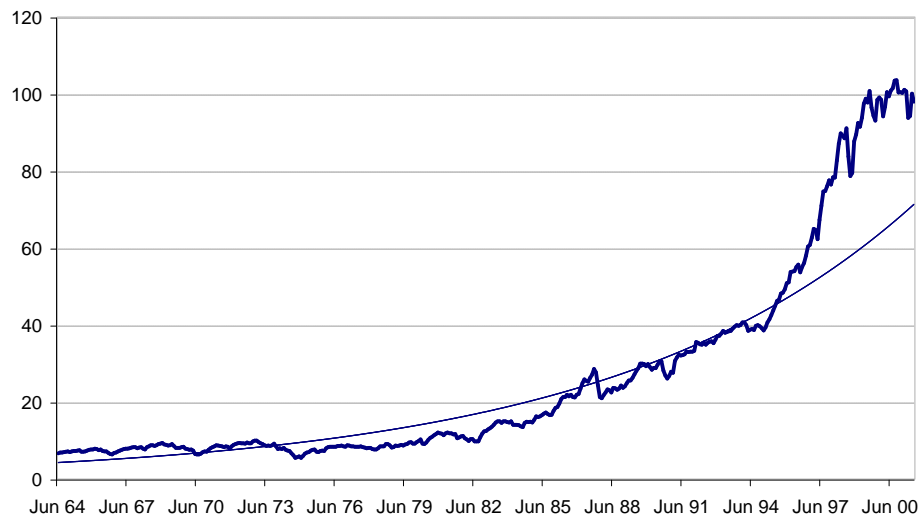


Abbildung 2.1: Der nominale Aktienindex der USA in der Zeitperiode von 06.1964 bis 06.2000

Zwei Approximationen eines Trends sind in der empirischen Wirtschaftsforschung verbreitet. Die erste Beschreibungsmöglichkeit liefert ein deterministischer Trend, der mit Hilfe eines $MA[\infty]$ -Prozesses wie folgt geschrieben wird:

$$X_t = \alpha + \delta t + \Psi(L)\varepsilon_t. \quad (2.22)$$

Damit ist gemeint, dass der Mittelwert des Prozesses X_t eine lineare Funktion der Zeit darstellt, $\mu = \alpha + \delta t$. Solche Prozesse werden als trend-stationäre Prozesse bezeichnet,

weil nach der Subtraktion des Trends $\alpha + \delta t$ aus der Gleichung (2.22) X_t zu einem stationären Prozess wird.

Die zweite Spezifikation eines Trends liefert der sogenannte Einheitswurzel-Prozess, der einen Trend stochastischer Natur beinhaltet:

$$(1 - L)X_t = \delta + \Psi(L)\varepsilon_t. \quad (2.23)$$

Diese Gleichung repräsentiert einen Random-Walk-Prozess mit Drift. Wie schon gezeigt wurde, ist der Erwartungswert des Prozesses $E(X_t) = \alpha + t\mu$ und die Varianz $Var(X_t) = t\sigma^2$. Der erste und der zweite Moment des Prozesses bleiben nicht invariant gegenüber der Zeit. Der Prozess wird in diesem Fall nur durch die Differenzbildung stationär, $\Delta X_t = \alpha + \Psi(L)\varepsilon_t$, deswegen werden solche Prozesse auch als differenzstationäre Prozesse bezeichnet.

In den 80er Jahren begann eine Diskussion, ob die Hypothese eines deterministischen Trends wirklich auf alle ökonomischen Variablen zutrifft, oder ob der beobachtete Trend selbst stochastischer Natur ist. Nelson und Plosser (1982) haben vierzehn amerikanische makroökonomische Zeitreihen auf Stationarität getestet. Sie fanden heraus, dass in dreizehn von vierzehn Fällen die Hypothese eines stochastischen Trends nicht abgelehnt werden konnte. Allerdings zeigten De Jong und andere (1992) mit Hilfe von Monte Carlo Simulationen, dass die Einheitswurzel-Tests eine niedrige Mächtigkeit haben. Seitdem wurde die Frage nach einem geeigneten Einheitswurzel-Test in der Ökonometrie stark verfolgt. Das führte zur Entwicklung mehrerer alternativer Tests wie des erweiterten Dickey-Fuller-GLS-Tests oder des Kwiatkowski, Phillips, Schmidt, und Shin Tests (KPSS)⁹.

Die Entscheidung für eine der beiden konkurrierenden Darstellungen eines Trends hat weitreichende ökonomische und ökonometrische Konsequenzen. Die Zeitreihen mit einem stochastischen Trend unterscheiden sich von den Zeitreihen mit einem deterministischen Trend nicht nur im Hinblick auf ihre ökonometrischen Eigenschaften, sie müssen auch völlig anders interpretiert werden. Wenn ein deterministischer Trend dem zu Grunde liegenden stochastischen Prozess unterstellt wird, dann wird von den stationären Fluktuationen um eine deterministische Funktion der Zeit ausgegangen. Die Wirkungen der Schocks sind nur vorübergehender Natur, und die Variable bewegt sich automatisch wieder auf den immanenten Wachstumspfad zu. Ein Prozess mit einem stochastischen Trend hat dagegen ein unendliches Gedächtnis, und alle exogenen Schocks haben einen permanenten Charakter. Im Falle eines stochastischen Trends ist die Wachstumskomponente ein nicht stationärer stochastischer Prozess, und die Fluktuationen der beobachteten Variablen werden nicht nur der Restkomponente ε_t zugeordnet, sondern den beiden Komponenten. Ein stochastischer Trend beinhaltet also auch Unsicherheit in Bezug auf die zukünftigen Entwicklung der langfristigen Komponente einer Zeitreihe.

⁹vgl. Maddala (1998), Kapitel 4.

Eine wichtige Frage für Ökonomen ist: Sind als Basis für Prognosen der meisten ökonomischen Zeitreihen eher trendstationäre oder differenzstationäre Modelle besser geeignet? Eine weitverbreitete Vermutung ist, dass die Anwendung der Einheitswurzel-Modelle für die Prognose bessere Ergebnisse liefert. Das Hauptargument ist, dass diese Modelle die mit den Prognosen verbundene Unsicherheit besser charakterisieren¹⁰. Einen Vergleich der Vorhersagemethoden führten Meese und Geweke (1984) durch, indem sie die Präzision der Vorhersagen von 150 quartals- und monatlichen makroökonomischen Zeitreihen untersuchten. Sie stellten fest, dass die auf den trendbereinigten Daten basierten Prognosen im Durchschnitt korrekter waren als diejenigen, die auf den Differenzen der Daten basierten.

Eine andere plausible Überlegung spricht gegen einen deterministischen Trend in den Zeitreihen. Die Trendspezifikation aus (2.22) kann aufgrund der realen Beobachtungen nicht vertretbar sein. Es stellt sich die Frage, wieso ein Trend als lineare Funktion ($\alpha + \delta t$) und nicht als quadratische ($\delta t + \gamma t^2$) oder exponentielle Funktion ($e^{\delta t}$) definiert wird? Beispielsweise scheint der Aktienindex der USA durch einen exponentiellen Trend besser charakterisiert zu sein als durch einen linearen Trend. Ein exponentieller Trend unterstellt ein konstantes Wachstum des Prozesses, denn $(X_t)'_t = \delta e^{\delta t} = \delta X_t$. Die konstante Wachstumsrate wird in mehreren ökonomischen Variablen beobachtet und ist bei der Lösung vieler Modelle notwendig. Beispielsweise kann die konstante Wachstumsrate der Preise (entspricht der konstanten Inflation) durch monetäre Politik der Zentralbank erreicht werden.

Ein formaler Vergleich der beiden alternativen Darstellungen des Trends werden im Weiteren präsentiert.

An erster Stelle ist ein Vergleich der beiden Prozesse im Bezug auf die Prognose interessant. Für einen Prozess mit einem deterministischen Trend (2.22) lautet die optimale h-schritt Prognose

$$\widehat{X}_{t+h} = \alpha + \delta(t+h) + \sum_{j=h}^{\infty} \psi_j \varepsilon_{t-(h+j)} \quad (2.24)$$

Die absolute Summierbarkeit der Koeffizienten ψ_j der $MA[\infty]$ -Darstellung des Prozesses impliziert, dass die Prognose zu dem Zeittrend $\alpha + \delta(t+h)$ erwartungsgemäß konvergiert, $E(\widehat{X}_{t+h} - \alpha + \delta(t+h))^2 \rightarrow 0$, für $h \rightarrow \infty$.

Für einen Einheitswurzel-Prozess aus (2.23) lautet die optimale h-schritt Prognose

$$\widehat{X}_{t+h} = h\delta + X_t + \left(\sum_{j=1}^h \psi_j\right)\varepsilon_t + \left(\sum_{j=2}^{h+1} \psi_j\right)\varepsilon_{t-1} + \dots \quad (2.25)$$

Für einen Random-Walk mit Drift ($\psi_1 = \psi_2 = \dots = 0$) lautet die optimale Prognose $\widehat{X}_{t+h} = h\alpha + X_t$. Für einen $ARMA(1,1)$ -Prozess ($\psi_1 \neq 0, \psi_2 = \dots = 0$) ist die

¹⁰Vgl. Phillips (1991b)

Prognose $\widehat{X}_{t+h} = h\delta + X_t + \psi_1\varepsilon_t$. Für $\delta = 0$ und $h = 1$ lässt sich die Optimale Prognose umschreiben $\widehat{X}_{t+1} = (1 + \psi_1) + X_t - \psi_1\widehat{X}_t$. Die Gleichung beschreibt die so genannten adaptiven Erwartungen. Der zukünftige Wert des Prozesses wird durch den jetzigen Wert und eine Korrektur der Prognose aus der Vorperiode dargestellt.

An der Darstellung der optimalen Prognose für beide Prozesse ist sofort der grundlegende Unterschied zwischen den Prozessen zu sehen. Der Parameter δ bei dem Einheitswurzel-Prozess spielt die gleiche Rolle wie bei dem Prozess mit einem deterministischen Trend. Die beiden Vorhersagen (2.24) und (2.25) konvergieren zu einer linearen Vorhersagefunktion mit der Steigung δ . Der Unterschied liegt am Achsenabschnitt dieser Linie. Während die Prognose für den trend-stationären Prozess zu einer Linie mit gleichem Achsenabschnitt wie bei dem Prozess X_t konvergiert, ändert sich der Achsenabschnitt der Prognosefunktion für einen Einheitswurzel-Prozess kontinuierlich mit jeder Beobachtung von X .

An zweiter Stelle ist der Vergleich des Prognosefehlers der beiden Prozessen vom Interesse. Die Spezifikation des Prozesses führt zu Unterschieden in der Varianz des Prognosefehler. Für einen trend-stationären Prozess aus (2.22) ist der Prognosefehler:

$$X_t - \widehat{X}_{t+h} = \varepsilon_{t+h} + \psi_1\varepsilon_{t+h-1} + \dots + \psi_{h-1}\varepsilon_{t+1}.$$

Der mittlere quadrierte Fehler ist $E[X_t - \widehat{X}_{t+h}]^2 = \{1 + \psi_1 + \dots + \psi_{h-1}\}^2\sigma^2$. Mit einem steigenden Vorhersagehorizont $h \rightarrow \infty$ ist der mittlere quadrierte Fehler beschränkt. Es ist zu bemerken: der Grenzwert des Erwartungswertes der quadrierten Fehler ist die Varianz der stationären Komponente $\Psi(L)\varepsilon_t$, $\lim_{h \rightarrow \infty} (E[X_t - \widehat{X}_{t+h}]^2) = \{1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{s-1}^2\}\sigma^2$.

Im Gegensatz zum trendstationären Prozess sieht der Fehler für einen Einheitswurzel-Prozess folgendermaßen aus:

$$X_t - \widehat{X}_{t+h} = \varepsilon_{t+h} + (1 + \psi_1)\varepsilon_{t+h-1} + \dots + (1 + \psi_1 + \dots + \psi_{s-1})\varepsilon_{t+1}$$

mit dem mittleren quadrierten Fehler $E[X_t - \widehat{X}_{t+h}]^2 = \{1 + (1 + \psi_1)^2 + \dots + (1 + \psi_1 + \dots + \psi_{s-1})^2\}\sigma^2$. Für einen $ARIMA(0, 1, 1)$ -Prozess folgt $E[X_t - \widehat{X}_{t+h}]^2 = \{1 + (h - 1)(1 + \psi_1)^2\}\sigma^2$. Damit sieht man sofort, dass mit dem steigenden Vorhersagehorizont auch der mittlere quadrierte Prognosefehler für einen Einheitswurzel-Prozess steigt.

An dritter Stelle unterscheiden sich die beiden Prozesse in der Reaktion auf Innovationschocks bei Impuls-Antwort-Analyse. Impuls-Antwort- oder dynamische Multiplikatoren für einen trend-stationären Prozess sind gegeben durch¹¹

$$\frac{\partial \widehat{X}_{t+h}}{\partial \varepsilon} = \psi_h. \tag{2.26}$$

¹¹Hier ist eine stetige Darstellung der Impuls-Antwort-Funktionen entsprechend Hamilton (1994), Kapitel 11 gegeben.

Der Effekt eines stochastischen Innovationsschocks ε_t zum Zeitpunkt t verschwindet mit dem wachsenden Vorhersagehorizont

$$\lim_{h \rightarrow \infty} \left(\frac{\partial X_t(h)}{\partial \varepsilon} \right) = \lim_{h \rightarrow \infty} (\psi_h) = 0.$$

Der dynamische Multiplikator für einen differenz-stationären Prozess ist gegeben durch

$$\frac{\partial \hat{X}_{t+h}}{\partial \varepsilon} = 1 + \psi_1 + \dots + \psi_h. \quad (2.27)$$

Bei dem Einheitswurzel-Prozess bleibt der permanente Effekt in dem Niveau des (X_t) nach dem Schock erhalten.

$$\lim_{h \rightarrow \infty} \left(\frac{\partial X_t(h)}{\partial \varepsilon} \right) = \lim_{h \rightarrow \infty} (1 + \psi_1 + \dots + \psi_h) = \Psi(1).$$

2.2.2 Integrierte Prozesse. Tests

Wie in Kapitel 2.2.1 festgestellt wurde, lässt sich eine einzelne stationäre Zeitreihe, die nicht direkt von der Zeit t abhängt, als ein endlicher $ARMA[p, q]$ -Prozess approximieren. In der Mehrzahl der Fälle weisen die beobachteten ökonomische Zeitreihen keine Stationarität auf, aber durch einfache oder mehrfache Differenzenbildung können sie in stationären Reihen transformiert werden. Es ist auch darauf hinzuweisen, dass die Differenzenbildung in der Tat mit einem nennenswerten Informationsverlust über die Niveauezusammenhänge verbunden ist. Kenntnisse über die fundamentalen Eigenschaften eines instationären integrierten Prozesses sind auch für das Verständnis der Tests auf Stationarität und der Existenz der langfristigen Gleichgewichtsbeziehungen wichtig¹².

Definition 2.10 *Ein stochastischer Prozess (X_t) ohne eine einzige deterministische Komponente, der sich nach einer d -fachen Differenzenbildung als ein stationärer invertierbarer $ARMA[p, q]$ -Prozess darstellen lässt, nach einer $(d-1)$ -fachen aber noch nicht stationär ist, heißt integriert vom Grad d , $X_t \sim I(d)$. Für die Vektor-Zeitreihen $(X_t) \in \mathbb{R}^n$ sagt man $X_t \sim I(d)$, wenn d der Integrationsordnung jeder Komponentenreihe $(X_{i,t}), i = 1, 2, \dots, n$ ist.*

Es wurde gezeigt, dass ein Random-Walk-Prozess im Gegensatz zum White-Noise-Prozess ein nicht stationäres Verhalten aufweist. Aber für die erste Differenz gilt:

$$\Delta X_t = X_t - X_{t-1} = \varepsilon_t, \forall t \in T,$$

¹²Vgl. Banerjee (1997)

damit ist Random-Walk ein von der Ordnung 1 integrierter Prozess ($X_t \sim I(1)$). Für einen stationären Prozess schreibt man oft $X_t \sim I(0)$. Für einen $ARMA[p, q]$ -Prozess, der von Integrationsordnung d ist, schreibt man auch $ARIMA[p, d, q]$ -Prozess.

Ein grosses Problem der trendbehafteten Zeitreihen ist das Phänomen der so genannten Scheinregression. Signifikante Zusammenhänge werden zwischen Variablen mittels der Regressionsanalyse festgestellt, obwohl sie tatsächlich nicht existieren. Das Phänomen wurde noch im Jahr 1926 von Yule entdeckt und während der nächsten Jahren von mehreren Wissenschaftlern wie Granger, Newbold, Phillips, Entorf für die unterschiedlichen Prozesse untersucht¹³.

Die scheinbare Regression entsteht unter anderem zwischen zwei unkorrelierten Random-Walk Prozessen

$$y_t = y_{t-1} + u_t, \quad u_t \sim iin(0, \sigma_u^2)$$

$$x_t = x_{t-1} + v_t, \quad v_t \sim iin(0, \sigma_v^2),$$

dabei wird angenommen, dass die Störterme u_t und v_t weder seriell untereinander noch mit einander korreliert sind. Es wird folgende Regression betrachtet:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t.$$

Aufgrund der Tatsache, dass y_t und x_t die unkorrelierten Random-Walk-Prozesse sind, wird erwartet, dass die Güte der Regression R^2 zu Null tendiert. Allerdings ist dies nicht der Fall. Yule (1926) entdeckte dieses Phänomen für lange nicht stationäre Zeitreihen. Granger und Newbold (1974) haben die scheinbare Regression mit den generierten Daten untersucht. Die Regression gab zwar einen hohen R^2 Koeffizient, allerdings eine niedrige Durbin-Watson-Statistik (DW-Statistik). Wurde eine Regression für die ersten Differenzen durchgeführt, so tendierte R^2 gegen Null und DW-Statistik zu zwei. Das demonstriert, dass keine Verbindung zwischen y_t und x_t existiert und ein hohes R^2 von der ursprünglichen Regression scheinbar ist. Entorf (1992) erhielt gleiche Ergebnisse für Random Walk Prozesse mit Drift.

Die folgende Tabelle¹⁴ zeigt, wann die Regression zwischen y_t und x_t gültig ist.

Die entscheidende Frage bei der Modellierung einer Zeitreihe ist, ob sie die Realisation eines differenz-stationären oder eines trend-stationären Prozesses ist. Der Einheitswurzel-Test ist ein Testverfahren, mit dem diese Frage beantwortet werden kann. In dem folgenden Abschnitt werden Tests auf den Integrationsgrad des Prozesses beschrieben sowie ihre Stärken und Schwächen diskutiert.

Dickey-Fuller-Test (DF-Test)

¹³Maddalla (1998), Kapitel 1.

¹⁴Vgl. Maddalla (1998), Kapitel 1.

Tabelle 2.1: Regression der integrierten Variablen

x_t / y_t	Deterministischer Trend	Stochastischer Trend
Deterministischer Trend	gültige Regression	Scheinbare Regression
Stochastischer Trend	Scheinbare Regression	Scheinbare Regression ausser x_t und y_t kointegriert

Der Test auf Integrationsgrad, der für AR[1]-Prozesse gedacht ist, wurde erstmal von Dickey und Fuller (1979) dargestellt. Daher wird der Test auch als Dickey-Fuller Test oder kurz DF-Test bezeichnet. Tatsächlich handelt es sich um ein Klasse von Tests. Bei dem DF-Test werden folgende drei Fälle untersucht:

- (i) Als erstes wird ein AR[1]-Modell ohne deterministischen Trend betrachtet:

$$X_t = \rho X_{t-1} + \varepsilon_t, \quad (2.28)$$

mit einem Gauss-White-Noise-Prozess (ε_t). Die Annahme bezüglich des datengenerierenden Prozesses ist ein Random-Walk. Dementsprechend wird die Nullhypothese $H_0 : \rho = 1$ gegen die Alternativhypothese $H_1 : |\rho| < 1$ getestet.

- (ii) Ein zweites Modell ist ein AR[1]-Modell mit Achsenabschnitt:

$$X_t = \alpha + \rho X_{t-1} + \varepsilon_t, \quad (2.29)$$

mit einem Gauss-White-Noise-Prozess (ε_t). In diesem Fall werden zwei verschiedene Nullhypothesen gestellt. Einmal lautet die Nullhypothese $H_0 : \alpha = 0, \rho = 1$, damit wird als datengenerierender Prozess ein Random-Walk unterstellt. Die andere Nullhypothese ist $H_0 : \alpha \neq 0, \rho = 1$, hier ist der wahre Prozess ein Random-Walk mit Drift. Die beiden Hypothesen werden gegen die Alternativhypothese $H_1 : |\rho| < 1$ getestet.

- (iii) Das dritte Modell beinhaltet einen deterministischen Trend.

$$X_t = \alpha + \rho X_{t-1} + \delta t + \varepsilon_t, \quad (2.30)$$

mit einem Gauss-White-Noise-Prozess (ε_t) und für ein beliebiges α . Jetzt ist der wahre Prozess wieder ein Random-Walk mit Drift. Die Nullhypothese $H_0 : \rho = 1, \delta = 0$ wird gegen die Alternativhypothese $H_1 : |\rho| < 1, \delta \neq 0$ getestet.

Im Fall (2.28) ist unter Annahme $\rho = 1$ der Kleinstquadratschätzer für ρ :

$$\hat{\rho}_N = \frac{\sum_{t=1}^N X_{t-1} X_t}{\sum_{t=1}^N X_{t-1}^2}.$$

Für die Teststatistik, die auf dem Schätzfehler basiert (F-Statistik), gilt hier:

$$N(\widehat{\rho}_N - 1) = \frac{N^{-1} \sum_{t=1}^N X_{t-1} \varepsilon_t}{N^{-2} \sum_{t=1}^N X_{t-1}^2}.$$

Die Verteilung der F-Statistik ist als Dickey-Fuller Verteilung bekannt. Dickey and Fuller (1979) haben keine Grenzverteilung mit Hilfe von Funktionalen auf Wiener-Prozess formuliert, allerdings diese Verteilung numerisch berechnet und entsprechende Tabellen gemacht. Die funktionale Form des Wiener-Prozesses hat danach Phillips (1987) angewendet.

Ausser der F-Statistik kann die t-Statistik für den Test der Nullhypothese $H_0 : \rho = 1$ benutzt werden:

$$t_{\widehat{\rho}_N} = \frac{\widehat{\rho}_N - 1}{\widehat{\sigma}_{\widehat{\rho}_N}}$$

mit

$$\widehat{\sigma}_{\widehat{\rho}_N} = \left\{ \frac{\hat{s}_N^2}{\sum_{t=1}^N X_{t-1}^2} \right\}^{1/2},$$

der Standardabweichung des geschätzten Koeffizientes $\widehat{\rho}$. \hat{s}_N^2 ist die geschätzte Varianz der stochastischen Fehlerterme ε_t ,

$$\hat{s}_N^2 = \frac{1}{N-1} \sum_{t=1}^N (X_t - \widehat{\rho}_N X_{t-1})^2.$$

Im Fall (2.29) muss unterschieden werden, ob der wahre Wert des Achsenabschnitts $\alpha = 0$ oder $\alpha \neq 0$ ist. Das hat einen starken Einfluss auf die asymptotische Verteilung der Schätzer $\widehat{\alpha}_N$ und $\widehat{\rho}_N$. Die Kleinstquadratschätzer $\widehat{\alpha}_N$ und $\widehat{\rho}_N$ werden als Lösungen von Normalgleichungen der 2-Variablen Regression erhalten. In der Matrixform sehen die Schätzer unter der Nullhypothese $H_0 : \alpha = 0, \rho = 1$ wie folgt aus:

$$\begin{pmatrix} \widehat{\alpha}_N \\ \widehat{\rho}_N \end{pmatrix} = \begin{pmatrix} N & \sum X_{t-1} \\ \sum X_{t-1} & \sum X_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum X_t \\ \sum X_{t-1} X_t \end{pmatrix}.$$

Die F-Statistik ist dann

$$N \begin{pmatrix} \widehat{\alpha}_N - 0 \\ \widehat{\rho}_N - 1 \end{pmatrix} = \begin{pmatrix} N & \sum X_{t-1} \\ \sum X_{t-1} & \sum X_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum \varepsilon_t \\ \sum X_{t-1} \varepsilon_t \end{pmatrix}.$$

Für den t-Test hat man die Statistik für $\widehat{\rho}_N$:

$$t_{\widehat{\rho}_N} = \frac{\widehat{\rho}_N - 1}{\widehat{\sigma}_{\widehat{\rho}_N}}$$

mit

$$\widehat{\sigma}_{\widehat{\rho}_N}^2 = \hat{s}_N^2 \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} N & \sum X_{t-1} \\ \sum X_{t-1} & \sum X_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

als Schätzer für die Varianz des geschätzten Koeffizientes ρ und

$$\hat{s}_N^2 = (N - 1)^{-1} \sum_{t=1}^N (X_t - \hat{\alpha}_N - \hat{\rho}_N X_{t-1})^2$$

als Schätzer für die Varianz der stochastischen Fehlerterme ε_t .

Die Grenzverteilung der beiden Teststatistiken ist keine Standardverteilung. Die Tabellen für die kritischen Werte für beide Teststatistiken findet man bei Hamilton (1994).

Für den Fall $\alpha \neq 0$, hat man unter Annahme der Nullhypothese als datengenerierenden Prozess ein Random-Walk mit Drift, $X_t = \alpha + X_{t-1} + \varepsilon_t$. Nach der Rekursion des Prozesses erhält man eine andere Schreibweise für den Prozess:

$$X_t = X_0 + \alpha t + \sum_{i=1}^t \varepsilon_i.$$

Hier wird der Random-Walk-Prozess mit Drift durch einen deterministischen Trend asymptotisch dominiert. Als Resultat erhält man die asymptotische Normalverteilung für die Kleinste-Quadrate-Schätzer der Koeffizienten α und ρ und damit folgen t- und F-Statistik Standard t- und F-Verteilungen. Den Beweis für die asymptotische Normalität der Test-Statistiken hat West (1988) präsentiert.

In dem dritten Regressionsmodell (2.30) mit einem deterministischen Trend, $X_t = \alpha + \rho X_{t-1} + \delta t + \varepsilon_t$, ist (X_{t-1}) asymptotisch equivalent zu einem deterministischen Zeitrend, wenn $\alpha \neq 0$ ist¹⁵. Weil ein deterministischer Trend δt als ein separater Regressor in der Regressionsgleichung eingeführt ist, fällt der Effekt des Drift-Terms α im datengenerierenden Prozess auf asymptotische Verteilungen aus, d.h. die asymptotische Verteilung des Kleinstquadratschätzers für ρ ist invariant zu dem Wert von α . Dadurch hat man als datengenerierenden Prozess entweder einen Random-Walk mit Drift oder einen Random-Walk mit einem linearen Trend.

Das Regressionsmodell aus (2.30) kann äquivalent umgeschrieben werden:

$$\begin{aligned} X_t &= (1 - \rho)\alpha + \rho(X_{t-1} - \alpha(t - 1)) + (\delta + \rho\alpha)t + \varepsilon_t \\ &\equiv \alpha^* + \rho^* Y_{t-1} + \delta^* t + \varepsilon_t \end{aligned}$$

mit $\alpha^* = (1 - \rho)\alpha$, $\rho^* = \rho$, $\delta^* = (\delta + \rho\alpha)$ und $Y_t = X_t - \alpha t$.

Unter Annahme der Nullhypothese $H_0 : \rho = 1, \delta = 0$, hat man dann $\alpha^* = 0, \rho^* = 1, \delta^* = \alpha$ und mit $X_t = \alpha + X_{t-1} + \varepsilon_t$ gilt:

$$Y_t = X_t - \alpha t = X_0 + \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t,$$

¹⁵Analog zu dem Fall Random-Walk mit Drift.

damit ist (Y_t) ein Random-Walk-Prozess. Die Schätzer sind dann in folgender Matrixform gegeben:

$$\begin{pmatrix} \widehat{\alpha}_N^* \\ \widehat{\rho}_N^* \\ \widehat{\delta}_N^* \end{pmatrix} = \begin{pmatrix} N & \sum Y_{t-1} & \sum t \\ \sum Y_{t-1} & \sum Y_{t-1}^2 & \sum Y_{t-1}t \\ \sum t & \sum Y_{t-1}t & \sum t^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum X_t \\ \sum Y_{t-1}X_t \\ \sum tX_t \end{pmatrix}.$$

Für die F-Statistik gilt sofort:

$$N \begin{pmatrix} \widehat{\alpha}_N^* - \alpha \\ \widehat{\rho}_N^* - 1 \\ \widehat{\delta}_N^* - \alpha \end{pmatrix} = \begin{pmatrix} N & \sum Y_{t-1} & \sum t \\ \sum Y_{t-1} & \sum Y_{t-1}^2 & \sum Y_{t-1}t \\ \sum t & \sum Y_{t-1}t & \sum t^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum u_t \\ \sum Y_{t-1}u_t \\ \sum tu_t \end{pmatrix}.$$

Es ist zu bemerken: Die Schätzer $\widehat{\rho}_N$ und $\widehat{\rho}_N^*$ für ρ sind identisch. Die kritischen Werte für die nicht standardverteilten Statistiken können wieder aus Hamilton (1994) entnommen werden. Wie schon früher gezeigt wurde: Ist der datengenerierende Prozess ein Random-Walk mit einem linearen Trend, d.h. $\delta \neq 0$, dann ist die F-Teststatistik für $\widehat{\alpha}_N, \widehat{\rho}_N$ und $\widehat{\delta}_N$ asymptotisch normal verteilt.

Tabelle 2.2: Asymptotische Verteilungen der Teststatistik für unterschiedliche datengenerierende Prozesse

Regression	DGP	asymptotische Verteilung
kein determ.Trend ($\alpha = \delta = 0$)	$\alpha = \delta = 0$	Funktional von Wiener Prozess
mit Drift ($\alpha \neq 0, \delta = 0$)	$\alpha = \delta = 0$ $\alpha \neq 0, \delta = 0$	Funktional von Wiener Prozess Normal Verteilung
mit linearem Trend ($\alpha \neq 0, \delta \neq 0$)	$\alpha = \delta = 0$ $\alpha \neq 0, \delta \neq 0$	Funktional von Wiener Prozess Normal Verteilung

In der Tabelle (2.2) sind die Ergebnisse für alle drei Fälle zusammengefasst¹⁶.

Erweiterter Dickey-Fuller-Test (ADF-Test)

Im Gegensatz zum einfachen DF-Test, der für $AR[1]$ -Prozesse gedacht ist, prüft der erweiterte Dickey-Fuller-Test (ADF-Test) einen $AR[p]$ -Prozess auf Stationarität. Bei dem Kleinste-Quadrate-Schätzer wird die Autokorrelation der höheren autoregressiven Terme berücksichtigt.

Die Herleitung der Grenzverteilungen der Teststatistik im ADF-Test ist analog der Vorgehensweise bei dem einfachen DF-Test. Dafür wird der $AR[p]$ -Prozess $(1 -$

¹⁶Vgl. Maddala (1998), Kapitel 3.

$\varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p) X_t = \varepsilon_t$ in einer äquivalenten Form mit Hilfe der Differenzen dargestellt. Definiert man zunächst $\rho = \varphi_1 + \varphi_2 + \dots + \varphi_p$, und $\xi_j = -(\varphi_{j+1} + \varphi_{j+2} + \dots + \varphi_p)$, für $j = 1, 2, \dots, p-1$, dann lässt sich der $AR[p]$ -Prozess wie folgt umschreiben¹⁷:

$$\{(1 - \rho L) - (\xi_1 L + \xi_2 L^2 + \dots + \xi_{p-1} L^{p-1})(1 - L)\} X_t = \varepsilon_t. \quad (2.31)$$

Die Existenz einer Einheitswurzel bei einem $AR[p]$ -Prozess (X_t) bedeutet, dass mindestens eine Lösung der charakteristischen Gleichung $(1 - \varphi_1 z - \varphi_2 z^2 - \dots - \varphi_p z^p) = 0$ gleich 1 ist. Aus dieser Annahme für $z = 1$ folgt, dass $(1 - \varphi_1 - \varphi_2 - \dots - \varphi_p) = 0$ gilt, und entsprechend der oben stehenden Definitionsgleichung von ρ ist $\rho = 1$. Alle anderen $(p-1)$ Lösungen der charakteristischsten Gleichung $1 - \xi_1 z - \xi_2 z^2 + \dots + \xi_{p-1} z^{p-1} = 0$ liegen außerhalb des Einheitskreises. Unter der Annahme der Nullhypothese einer Einheitswurzel für (X_t), kann der Ausdruck aus 2.31 in folgender Form geschrieben werden:

$$(1 - \xi_1 L - \xi_2 L^2 - \dots - \xi_{p-1} L^{p-1})(1 - L) X_t = \varepsilon_t.$$

Auf Grund der Annahmen bezüglich der Wurzel der charakteristischen Gleichung ist der Prozess

$$(1 - L) X_t = (1 - \xi_1 L - \xi_2 L^2 + \dots + \xi_{p-1} L^{p-1})^{-1} \varepsilon_t$$

invertierbar und dadurch auch stationär.

Ähnlich wie im zweiten Fall (datengenerierender Prozess ist ein Random-Walk-Prozess ohne Drift) des einfachen DF-Tests lautet die Nullhypothese $H_0 : \alpha = 0, \rho = 1$. Es wird H_0 gegen die Alternativhypothese $H_1 : |\rho| < 1$ getestet. Für die Herleitung der Grenzverteilungen der Kleinste-Quadrate-Schätzer der Koeffizienten wird die so genannte erweiterte Dickey-Fuller Regression benutzt:

$$\begin{aligned} X_t &= \xi_1 \Delta X_{t-1} - \xi_2 \Delta X_{t-2} + \dots + \xi_{p-1} \Delta X_{t-p+1} + \alpha + \rho X_{t-1} + \varepsilon_t \\ &\equiv Z_t' \beta + \varepsilon_t, \end{aligned} \quad (2.32)$$

mit $\beta = (\xi_1, \xi_2, \dots, \alpha, \rho)'$ und $Z_t = (\Delta X_{t-1}, \Delta X_{t-2}, \dots, \Delta X_{t-(p-1)}, 1, X_{t-1})'$. Der Schätzfehler, der für F-Statistik verwendet wird, ist durch

$$\hat{\beta} - \beta = \left[\sum_{t=1}^N Z_t Z_t' \right]^{-1} \left[\sum_{t=1}^N Z_t \varepsilon_t \right]$$

¹⁷Die Darstellungsform wird von Sims, Stock und Watson kanonische Form genannt.

gegeben. Dabei ist

$$\sum_{t=1}^N Z_t Z_t' = \begin{pmatrix} \sum u_{t-1}^2 & \sum u_{t-1}u_{t-2} & \cdots & \sum u_{t-1}u_{t-(p-1)} & \sum u_{t-1} & \sum u_{t-1}x_{t-1} \\ \sum u_{t-2}u_{t-1} & \sum u_{t-2}^2 & \cdots & \sum u_{t-2}u_{t-(p-1)} & \sum u_{t-2} & \sum u_{t-2}x_{t-1} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ \sum u_{t-(p-1)}u_{t-1} & \sum u_{t-(p-1)}u_{t-2} & \cdots & \sum u_{t-(p-1)}^2 & \sum u_{t-(p-1)} & \sum u_{t-(p-1)}x_{t-1} \\ \sum u_{t-1} & \sum u_{t-2} & \cdots & \sum u_{t-(p-1)} & N & \sum x_{t-1} \\ \sum x_{t-1}u_{t-1} & \sum x_{t-1}u_{t-2} & \cdots & \sum x_{t-1}u_{t-(p-1)} & \sum x_{t-1} & \sum x_{t-1}^2 \end{pmatrix}$$

und

$$\sum_{t=1}^N Z_t \varepsilon_t = \begin{pmatrix} \sum u_{t-1} \varepsilon_t \\ \sum u_{t-2} \varepsilon_t \\ \vdots \\ \sum u_{t-(p-1)} \varepsilon_t \\ \sum \varepsilon_t \\ \sum x_{t-1} \varepsilon_t \end{pmatrix}$$

mit der Summierung für $t = 1$ bis N und $u_t = \Delta X_t$ ein stationärer Prozess.

Dementsprechend ist die t-Statistik unter Gültigkeit der Nullhypothese:

$$t_{\hat{\rho}_N} = \frac{\hat{\rho}_N - 1}{[\hat{s}_N^2 e_{p+1}' (\sum Z_t Z_t')^{-1} e_{p+1}]^{-1/2}},$$

dabei ist $e_{p+1} \in \mathbb{R}^{p+1}$ ein Vektor mit Eins auf der letzten Position und ansonsten Null. \hat{s}_N^2 ist die geschätzte Varianz des Fehlers:

$$\hat{s}_N^2 = (N - (p + 1))^{-1} \sum_{t=1}^N (X_t - \hat{\xi}_{1,N} \Delta X_{t-1} - \hat{\xi}_{2,N} \Delta X_{t-2} - \cdots - \hat{\xi}_{p-1,N} \Delta X_{t-p+1} - \hat{\alpha}_N - \hat{\rho}_N X_{t-1})^2.$$

Die Verteilungen der Statistiken folgen den gleichen DF-Verteilungen wie beim einfachen DF-Test. Entsprechend kann der erweiterte DF-Test für einen $AR[p]$ -Prozess mit einem linearen Trend formuliert werden.

Phillips-Perron-Test

Die Idee des Phillips-Perron-Tests wurde zum ersten Mal von Phillips (1987) präsentiert und von Phillips und Perron (1988) erweitert. Der einfache Dickey-Fuller-Test wird dabei unter der Annahme der Autokorrelation von Störtermen (ε_t) modifiziert. Die entsprechende Anpassung ist bei den Teststatistiken vorzunehmen.

Die zwei Grössen σ^2 und σ_ε^2 werden beim Test berücksichtigt. Die erste Grösse stellt die Varianz der Summe der Fehlerterme ε_t dar:

$$\sigma^2 = \lim_{N \rightarrow \infty} \frac{1}{N} E\left(\left(\sum_{j=1}^N \varepsilon_j\right)^2\right).$$

Die zweite Grösse ist die Varianz der Fehler ε_t

$$\sigma_\varepsilon^2 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N E(\varepsilon_j^2).$$

Phillips (1987) und Phillips und Perron (1988) zeigten, dass folgende Schätzer für die Summe der Fehler σ^2 und für die Varianz der Fehler σ_ε^2 konsistent sind.

$$s^2 = \frac{1}{N} \sum_{t=1}^N \varepsilon_t^2 + \frac{2}{N} \sum_{\tau=1}^l \sum_{t=\tau+1}^N \varepsilon_t \varepsilon_{t-\tau}$$

$$s_\varepsilon^2 = \frac{1}{N} \sum_{t=1}^N \varepsilon_t^2.$$

Entsprechend ändern sich die Teststatistiken für ein $AR[1]$ -Prozess.

(i) $AR(1)$ -Modell ohne Drift, $\rho = 0$:

$$Z_{\hat{\rho}_N} = N(\hat{\rho}_N - 1) - \frac{1}{2} \frac{(s^2 - s_\varepsilon^2)}{N^{-2} \sum_1^N X_{t-1}^2}$$

$$Z_t = \frac{s_\varepsilon}{s} t \hat{\rho}_N - \frac{1}{2} \frac{(s^2 - s_\varepsilon^2)}{s(N^{-2} \sum_1^N X_{t-1}^2)^{1/2}}$$

(ii) $AR(1)$ -Modell mit Drift, $\rho \neq 0$:

$$Z_{\hat{\rho}_N} = N(\hat{\rho}_N - 1) - \frac{1}{2} \frac{(s^2 - s_\varepsilon^2)}{N^{-2} \sum_1^N (X_{t-1} - \bar{X}_{-1})^2}$$

$$Z_t = \frac{s_\varepsilon}{s} t \hat{\rho}_N - \frac{1}{2} \frac{(s^2 - s_\varepsilon^2)}{s[N^{-2} \sum_1^N (X_{t-1} - \bar{X}_{-1})^2]^{1/2}}$$

mit $\bar{X}_{-1} = \sum_1^{N-1} X_t / (N-1)$

(iii) $AR(1)$ -Modell mit Drift $\rho \neq 0$ und linearem Trend $\delta \neq 0$

$$Z_{\hat{\rho}_N} = N(\hat{\rho}_N - 1) - \frac{N^6}{24D_Y} (s^2 - s_\varepsilon^2)$$

$$Z_t = \frac{s_\varepsilon}{s} t \hat{\rho}_N - \frac{N^3 (s^2 - s_\varepsilon^2)}{4\sqrt{3} D_Y^{1/2} s}$$

mit $D_Y = \det(YY')$ und die Regressoren sind $Y = (1, t, X_{t-1})$.

Die Grenzverteilungen der Statistiken $Z_{\hat{\rho}_N}$ und Z_t sind identisch zu den Grenzverteilungen der Statistik $F = N(\hat{\rho}_N - 1)$ und die t-Statistik gilt entsprechend, wenn $\sigma^2 = \sigma_\varepsilon^2$. Die asymptotischen kritischen Werte für die Tests sind die gleichen kritischen Werte wie bei Fuller (1976).

Auswahl der optimalen Anzahl von Lag in AR(p)-Modell

Schwert (1989) stellte fest, dass die Größe der Teststatistik im ADF-Test von der Anzahl der Lag-Termen p sehr stark abhängig ist. Mehrere Methoden zur Auswahl von p wurden seither vorgeschlagen. Ng and Perron (1995) überprüften diese Methoden im Detail¹⁸. Drei davon sind

1. Darstellung von p als eine Funktion der Anzahl der Beobachtungen N .

Als ein Verfahren zur Auswahl der Laganzahl im AR[p]-Modell für erweiterten Dickey-Fuller-Test schlägt Schwert (1989) folgendes vor:

$$p = \text{Int}\{c(N/100)^{1/d}\},$$

mit $c=12$ und $d=4$. Das Problem dieses Kriteriums liegt in seiner Nicht-Anwendbarkeit für einen ARMA[p,q]-Prozess.

2. Informationsbasierende Methode.

Ein informationsbasierendes Kriterium empfiehlt ein p zu wählen, das die Funktion

$$I_p = \ln \hat{\sigma}_p^2 + p \frac{C_N}{N}$$

minimiert. Das Akaike Informationskriterium (AIC) setzt $C_N = 2$; das Schwarz Bayesian Informationskriterium (BIC) empfiehlt $C_N = \ln(N)$ auszuwählen. Das sind die am meisten verwendeten Kriterien in Ökonometrie. Die weiteren Kriterien sind die Bias-korrigierte Version der AIC, AICC von Hurvich und Tsay:

$$I_p = \ln \hat{\sigma}_p^2 + \frac{2p}{N - p - 2},$$

sowie das Kriterium von Hannan-Quinn:

$$HQ_p = \ln \hat{\sigma}_p^2 + \frac{2p \cdot c \cdot \ln(\ln N)}{N},$$

bei dem $c > 1$ eine beliebig gewählte Konstante ist. Wie Ng und Perron (1995) zeigten, liefern AIC und BIC für ARMA[p,q]-Prozess asymptotisch gleiche Ergebnisse.

¹⁸Eine ausführliche Darstellung der Methoden findet sich bei Maddala (1998), Kapitel 3.

3. Sequentielle Methode von Hall

Hall in [54] diskutiert zwei mögliche sequentielle Regeln im Zusammenhang mit Autoregression. Die erste Regel ist die sogenannte *general to specific rule*. Man beginnt mit den grossen Werten von p (p_{max}), testet den letzten Koeffizient $\hat{\varphi}_{p_{max}}$ auf Signifikanz und reduziert p entsprechend iterativ, bis die signifikante Statistik auftritt. Die andere Regel ist, die sogenannte *specific to general rule*. Das Verfahren läuft in umgekehrter Folge. Man beginnt mit einem kleinen Wert p und erhöht ihn dann sukzessiv. Hall zeigte, dass die *specific to general rule* im Allgemeinen asymptotisch nicht aussagekräftig ist.

Ng und Perron (1995) haben die AIC, BIC und Halls *general to specific* Approximation mit Hilfe der Monte-Carlo-Simulation und ab $N=100$ für AR- und MA-Prozesse verglichen. Sie kamen zum Ergebnis, dass

- (i) AIC und BIC Kriterien beide sehr kleine Werte von p wählen, am häufigsten $p = 3$. Dies führt zur Verzerrung der Schätzung, insbesondere bei den MA-modellierten Fehlern.
- (ii) beide Kriterien nach Hall tendieren dazu, grosse Werte von p zu wählen. Je höher p_{max} ist, desto höher wird der Wert p vorgeschlagen, allerdings ist damit ein Verlust der Mächtigkeit des Tests verbunden.

Mit Bezug auf weitere Studien kommt Maddala (1998) zum Ergebnis, dass Halls Kriterien vorteilhaft sind.

2.2.3 Probleme

Die Einheitswurzel-Tests, im einzelnen die DF-, ADF- und PP-Tests, sind durch mehrere Probleme gekennzeichnet. Die Schwachstellen der Tests sind ihre niedrige Mächtigkeit und die Verzerrung der Testgrößen. Die zwei zentralen Fragen bezüglich Einheitswurzeltests sind in diesem Zusammenhang: Woran liegt dies und wie könnten die Probleme gelöst oder umgangen werden? Eine große Anzahl der zuletzt entstandenen Modifikationen der Standard-Einheitswurzeltests kann als Versuch einer Antwort auf diese Fragen interpretiert werden¹⁹.

Schwert (1989) hat als erster auf die Existenz des Größenverzerrung-Problems mit Hilfe der Monte-Carlo-Simulationen hingewiesen. Er argumentiert, dass sich die Verteilung der Teststatistik beim DF-Test weit von der von Dickey und Fuller eingeführten Verteilung unterscheidet, wenn eine MA-Komponente im Prozess eine starke Präsenz hat. Er behauptet ebenso, dass die PP-Tests an dieser Verzerrung leiden, wenn die

¹⁹Für eine ausführliche Diskussion des Problems stehe Stock (1994)

MA-Parameter sehr groß sind. Dies ist der Fall in Systemen mit vielen ökonomischen Zeitreihen. Als einen Ausweg kann der von Said und Dickey (1984) entwickelte t-Test angewendet werden, der für die autoregressiven Prozesse der höheren Ordnung entwickelt wurde. Schwert argumentiert, dass eine Entscheidung über die korrekte Struktur des ARIMA-Modells wichtig ist, bevor die Einheitswurzel-Tests durchgeführt werden können.

Im Gegensatz zu Schwert beschäftigen sich DeJong und andere (1992a, 1992b) mit der niedrigen Mächtigkeit der Einheitswurzel-Tests. Die Wissenschaftler argumentieren in ihrem ersten Artikel (1992a), dass die DF-, ADF- und PP-Tests eine niedrige Mächtigkeit besitzen, wenn ein trend-stationärer Prozess als eine Alternativhypothese zum differenz-stationären Prozess aufgestellt wird. Sie vergleichen in (1992b) mit Hilfe der Monte-Carlo-Simulationen unter anderem die Mächtigkeit der ADF- und PP-Tests mit dem Ergebnis, dass der PP-Test generell eine sehr niedrige Mächtigkeit (kleiner als 0,10) gegenüber der trend-stationären Alternative hat. Der ADF-Test ist sinnvoller in der Praxis zu benutzen, weil er eine höhere Mächtigkeit (approximativ ein Drittel) hat.

Beide Probleme haben einen weiteren Aspekt, nämlich dass die Fehlertermvarianz einen Effekt auf die Verteilung der Teststatistik hat. Die von Beveridge und Nelson (1981) vorgeschlagene Zerlegung zeigt, dass eine Zeitreihe mit einer Einheitswurzel genau in einen Random-Walk-Prozess mit Drift und eine stationäre Komponente zerlegt werden kann. Die trend-stationäre Zeitreihe hat eine triviale Beveridge-Nelson-Zerlegung, dabei ist die Varianz der Random-Walk-Komponente identisch Null. Liu und Praschnik (1993) untersuchten den Einfluss der Varianz der Störterme sowohl in der Random-Walk als auch in der stationären Komponente auf die Verteilungen der ADF- und PP-Statistik. Sie zeigten, dass die Verzerrung der Teststatistikgrößen sehr empfindlich auf das Verhältnis der beiden Varianzen reagiert, allerdings ist der ADF-Test weniger sensibel.

Ein weiteres Problem der Einheitswurzel-Tests wurde von Cochrane (1991) untersucht. Cochrane weist auf die Existenz der Einheitswurzel-Prozesse hin, dessen Likelihood-Funktion und Autokorrelationsfunktion den Funktionen der stationären Prozesse sehr ähnlich sind. Insbesondere gilt das für kleine Stichproben.

Als eine mögliche Lösung dieser Probleme sind vielfältige Varianten der Einheitswurzel-Tests entstanden²⁰, darunter sind der Durbin-Hausman-Test (oder Hausmans Spezifikation der Fehlerterme) von Choi (1992); die Modifikationen der PP-Tests von Perron und Ng (1996); die Vorwärts- und Rückwärts-DF-Regression von Leybourne (1995); der DF-GLS-Test von Elliot, Rothenberg und Stock (1996).

Zwei von vielen entstandenen Modifikationen der Einheitswurzel-Tests werden im Folgenden dargestellt²¹.

²⁰Eine ausführliche Darstellung und Diskussion der Tests findet sich bei Maddala (1998), Kapitel 4.

²¹Beide Tests wurden in dieser Arbeit angewendet.

Modifizierter PP-Test

Der original PP-Test berücksichtigt die MA-Fehler in seiner Schätzung. Obwohl der PP-Test eine höhere Mächtigkeit im Vergleich zu dem DF-Tests aufweist, liefert er stark verzerrte Statistiken bei den negativen MA-Fehlertermen. Perron und Ng (1996) schlagen folgende Änderungen der Z_ρ - und Z_t -Statistik vor:

$$MZ_{\hat{\rho}_N} = Z_\rho + \frac{N}{2}(\hat{\rho} - 1)^2 \quad (2.33)$$

und

$$MZ_t = (N^{-2} \sum X_{t-1}^2 / s^2)^{1/2} MZ_{\hat{\rho}_N}. \quad (2.34)$$

Die Varianz der Störterme ε_t wird wie bei den Standard PP-Test $s_\varepsilon^2 = \frac{1}{N} \sum_{t=1}^N \varepsilon_t^2$ geschätzt. Für die Schätzung der Varianz der Summe der Fehlerterme schlagen Perron und Ng einen autoregressiven Schätzer

$$s^2 = \frac{s_{e_k}^2}{(1 - \hat{b}(1))^2},$$

mit

$$s_{e_k}^2 = \frac{1}{T} \sum_{t=k+1}^T \hat{e}_{tk}$$

und

$$\hat{b}(1) = \sum_{j=1}^k \hat{b}_j$$

vor. \hat{e}_{tk} und \hat{b}_j sind aus der Regression

$$\Delta X_t = b_0 X_{t-1} + \sum_{j=1}^k b_j \Delta X_{t-j} + e_{tk}.$$

DF-GLS-t-Test

Der DF-GLS-t-Test von Elliot, Rothenberg und Stock (1996) wird bei der Überprüfung der Nullhypothese $H_0 : a_0 = 0$ für folgende Regression

$$\Delta X_t^d = a_0 X_{t-1}^d + \sum_{i=1}^p \Delta X_{t-i}^d + \varepsilon \quad (2.35)$$

ausgeführt. Dabei ist X_t^d ein lokal trendbereinigter Prozess. Die lokale Trendbereinigung ist von der Annahme des Modells (nur mit Drift oder mit eine linearen Trend) abhängig.

Wird von einem linearen Trend ausgegangen, sieht der trendbereinigte Prozess wie folgt aus:

$$X_t^d = X_t - \hat{\beta}_0 - \hat{\beta}_1 t.$$

Die Koeffizienten $(\hat{\beta}_0, \hat{\beta}_1)$ sind vor der Regression \bar{X} an \bar{Z} abgeleitet, mit

$$\bar{X} = [X_1 (1 - \bar{\alpha}L) X_2 \dots (1 - \bar{\alpha}L) X_T]$$

$$\bar{Z} = [Z_1 (1 - \bar{\alpha}L) Z_2 \dots (1 - \bar{\alpha}L) Z_T]$$

und

$$Z_t = (1 t)', \quad \bar{\alpha} = 1 + \frac{\bar{c}}{T}.$$

Dabei wird $\bar{c} = -7$ für ein Modell mit Drift und $\bar{c} = -13,5$ für ein Modell mit einem linearen Trend gesetzt.

ADF-Test mit GARCH(1,1)

Die Annahme der Normalverteilung der Daten ist eine wichtige Voraussetzung für die Anwendung nicht nur der Einheitswurzel-Tests. Weil in dieser Arbeit eine Untersuchung der Aktienmärkte durchgeführt wird, ist eine Überprüfung der Normalverteilung der Renditen notwendig. Außer, dass die Hypothese der Normalverteilung für viele Aktienmarktrenditen abgelehnt wird, stellen viele Untersuchungen eine linksschiefe Verteilungsdichte der Renditen fest²². Desweiteren sind die Renditen durch Leptokurtosis gekennzeichnet²³, d.h. dass die Verteilungsdichte der Renditen an den Rändern über der Normalverteilung liegt. Die 'dicken' Ränder der Verteilungsdichte implizieren, dass extreme Kursbewegungen wahrscheinlicher sind, als man unter der Normalverteilung erwarten würde. Ein weiteres Phänomen der Renditen ist die zeitvariierende Volatilität der Renditen. Es gibt eine enge Verbindung zwischen Leptokurtosis und Heteroskedastizität, weil genau die Zeitreihen mit einer zeitabhängigen Varianz eine Verteilung mit „dicken“ Rändern generieren²⁴. Eine Möglichkeit, um diese Art der Nicht-Normalität in den Daten zu berücksichtigen, ist eine Modellierung der Heteroskedastizität mit Hilfe von ARCH- (Autoregressive Conditional Heteroskedastic) und GARCH- (Generalized Autoregressive Conditional Heteroskedastic) Prozessen.

Eine weitere äquivalente Darstellung eines AR(p)-Prozesses lautet:

$$(1 - L)X_t = \beta X_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta X_{t-i} + \varepsilon_t. \quad (2.36)$$

Im Fall der Heteroskedastizität der Störterme ε_t kann die zeitabhängige Varianz $\sigma^2 = \sigma_t^2$ zusätzlich mit Hilfe eines GARCH(1,1) Prozesses modelliert werden:

$$\sigma_t^2 = \omega + \zeta \varepsilon_{t-1}^2 + \phi \sigma_{t-1}^2, \quad (2.37)$$

²²Vgl. Fraser und Power (1997), Shawsky und andere (1997), Scheicher (2001), Ahlgren und Antell (2002), Syriopoulos (2004)

²³Vgl. Seo (1999), Scheicher (2001))

²⁴Vgl. Scheicher (2001)

mit $\varepsilon_t = \sigma_t u_t$ und $u_t \sim i.i.d.(0, 1)$.

Die Nullhypothese des Einheitswurzel-Tests, die hier $H_0 : \beta = 0$ lautet, wird gegen die Alternativhypothese $H_1 : |\beta| < 1$ getestet. Im Gegensatz zum Standard ADF-Test wird hier die Maximum-Likelihood-Schätzer (MLE) angewendet:

$$\hat{\theta}_n = \max_{\theta \in \Theta} \mathfrak{L}_n(\theta) \quad (2.38)$$

mit dem Parametervektor $\theta = (\beta \gamma' \omega \phi \zeta)'$. Der Vektor $\gamma' = (\gamma_1 \dots, \gamma_{p-1})$ repräsentiert die Koeffizienten aus dem autoregressiven Teil des Modells. Die Likelihood-Funktion

$$\mathfrak{L}_n(\theta) = 1/n \sum_{t=1}^n \ell_t(\theta)$$

wird über die Parametermenge $\Theta \subset \mathbb{R}^{p+3}$ maximiert. Dabei wird $\ell_t(\theta)$ wie folgt dargestellt:

$$\ell_t(\theta) = -0,5 \log \sigma_t^2(\theta) - 0,5 u_t^2(\theta) / \sigma_t^2(\theta).$$

Die Grenzverteilung der Teststatistik unter der Nullhypothese ist zusätzlich von der deterministischen Komponente abhängig. Die asymptotische Verteilung der Teststatistik

$$\hat{\tau} = \frac{n \hat{\beta}_n}{\sqrt{\hat{v}ar(n \hat{\beta}_n)}} \quad (2.39)$$

stellt eine Kombination aus DF t-Verteilung und Standardnormalverteilung dar. Die Kombination hängt von dem relativen Gewichtungsfaktor ρ ab, der die Wichtigkeit des GARCH-Prozesses darstellt. Die Schätzung des Parameters ergibt sich wie folgt.

$$\hat{\rho} = \frac{1}{(\hat{\kappa} - 1) \hat{A}(\alpha, \zeta) + \hat{\sigma}^2 / \hat{\rho}^2} \quad (2.40)$$

dabei ist

$$\begin{aligned} \hat{\sigma}^2 &= 1/n \sum_{t=1}^n \hat{\sigma}_t^2, \\ \hat{\rho}^2 &= \left(1/n \sum_{t=1}^n 1/\hat{\sigma}_t^2 \right)^{-1}, \\ \hat{\kappa}^2 &= 1/n \sum_{t=1}^n (\hat{\varepsilon}_t / \hat{\sigma}_t)^4, \end{aligned}$$

und $\hat{A}(\alpha, \zeta)$ stellt einen Grenzwert dar, der von der geschätzten Varianz $\hat{\sigma}_t^2$ und deren Moving-Average-Zerlegung abhängig ist. Nach Seo (1999) gilt: der Koeffizient $\sigma^2 / \rho^2 > 1$, wenn der GARCH-Parameter nicht Null ist. Wenn dieser Koeffizient genau gleich Eins ist ($\sigma^2 / \rho^2 = 1$), dann ist der Einfluss des GARCH Prozesses statistisch nicht

relevant und die Verteilung der Teststatistik $\hat{\tau}$ ist mit der DF-t-Verteilung identisch. Die Verteilung und entsprechende Tabellen für die kritischen Werte sind aus Seo (1999) entnommen.

Bei der Berücksichtigung der zeitabhängigen Varianz der Störterme durch ein GARCH-Modell werden zusätzliche Annahmen im Modell getroffen. Das kann wieder zur einer Verschlechterung der Testeigenschaften führen. Allerdings zeigt Seo (1999), dass die Einführung des GARCH-Terms bei dem ADF-Test die Mächtigkeit des Tests erheblich erhöht.

2.3 Kointegrationskonzept

2.3.1 Kointegration. Granger-Repräsentationstheorem

Die Idee der Kointegration lässt sich sehr einfach im zweidimensionalen Fall präsentieren. Falls zwischen den trendstationären Variablen X_t und Y_t eine funktionale Gleichgewichtsbeziehung $Y_t = \alpha X_t$ besteht, so kann als eine Minimalbedingung erwartet werden, dass die Abweichung vom Gleichgewicht $Z_t = Y_t - \alpha X_t$ keinen Trend aufweist. Anders ausgedrückt, besitzt die Linearkombination $Y_t - \alpha X_t$ dann einen niedrigeren Integrationsgrad als die Ausgangsvariablen. Für den mehrdimensionalen Fall liefern Engle und Granger (1987) folgende Definition der Kointegration:

Definition 2.11 Sei (X_t) ein n -dimensionaler Vektor, dessen Komponenten alle vom Integrationsgrad d sind ($I(d)$). Existiert ein Vektor $\beta \neq 0$, $\beta \in \mathbb{R}^n$, so dass

$$Z_t = \beta' X_t \sim I(d - b), \text{ mit } b > 0,$$

so heißen die Komponenten von X_t kointegriert vom Grad (d, b) . β bezeichnet den Kointegrationsvektor.

Im Rahmen ökonomischer Fragestellungen ist der Fall $(X_t) \sim I(1)$ und $(Z_t) \sim I(0)$ von besonderem Interesse. Besteht (X_t) aus zwei Variablen, so existiert höchstens eine Gleichgewichtsbeziehung, die durch einen eindeutigen Kointegrationsvektor (bis auf einen multiplikativen Faktor) dargestellt werden kann. Zwischen n Variablen können bis zu $(n-1)$ unabhängige Gleichgewichtsbeziehungen bestehen und daher bis zu $(n-1)$ kointegrierte Vektoren. Das Eindeutigkeitsproblem wird an dieser Stelle deutlich. Werden die Kointegrationsvektoren in einer Matrix β zusammen gefasst, so gilt die Bedingung $Z_t = \beta' X_t \sim I(0)$ (jedes Element des Vektors (Z_t) ist integriert vom Grad 0) als die Bedingung für die Kointegration. Die Anzahl r der Kointegrationsvektoren stellt der Kointegrationsrang des Systems X_t dar.

Verschiedene äquivalente Darstellungsformen der kointegrierten Systeme werden in der Praxis benutzt:

(i) die Vektor-Moving-Average-Darstellung

Die Beschreibung der kointegrierten Zeitreihen mit Hilfe von Moving-Average-Prozessen basiert auf der sog. Wold (1938)²⁵ Darstellung und Beveridge-Nelson (1981) Zerlegung. Angenommen, (X_t) ist vom Integrationsgrad $I(1)$, damit ist nach der Definition des integrierten Prozesses (ΔX_t) ein stationärer Prozess. Definiere u_t durch

$$u_t = \Delta X_t - \mu.$$

Ist die Matrixfolge $(\Psi_t)_{t=0}^\infty$ absolut summierbar, so lässt sich u_t als $VMA[\infty]$ -Prozess darstellen

$$u_t = \Psi(L)\varepsilon_t, \tag{2.41}$$

$$\Psi(L) = I_n + \Psi_1 L^1 + \Psi_2 L^2 + \dots$$

(ε_t) ist ein Vektor-White-Noise-Prozess mit einer nicht singularen Kovarianzmatrix Λ .

Lemma 2.12 *Sei (X_t) ein n -dimensionaler stochastischer Prozess, mit $(X_t) \sim I(1)$ und $u_t = \Delta X_t - \mu$ hat eine Wold Darstellung (2.41) mit $\Psi_0 = I_n$ und $(t\Psi_t)_{t=0}^\infty$ absolut summierbarer Matrixfolge. (X_t) ist genau dann vom Rang r ($0 < r < n$) kointegriert, wenn eine Matrix $\beta \in \mathbb{R}^{n \times r}$ existiert, so dass $\beta'\Psi(1) = 0$ und $\beta'\mu = 0$ gilt. $\Psi(1) = I_n + \Psi_1 + \Psi_2 + \dots$*

(ii) die Phillips-Triangular-Darstellung (Philips (1991a))

Angenommen, dass die Spalten der Matrix $\beta \in \mathbb{R}^{n \times r}$ die Basis für den Raum der kointegrierten Vektoren präsentieren. Dann kann diese Matrix zur Trapezform transformiert werden:

$$\beta' = \begin{pmatrix} \beta_{1,1} & \cdots & \beta_{1,n} \\ \beta_{2,1} & \cdots & \beta_{2,n} \\ \vdots & & \\ \beta_{r,1} & \cdots & \beta_{r,n} \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & \cdots & 0 & -\gamma_{1,r+1} \cdots & -\gamma_{1,n} \\ 0 & 1 & \cdots & 0 & -\gamma_{2,r+1} \cdots & -\gamma_{2,n} \\ \vdots & & & & & \\ 0 & 0 & \cdots & 1 & -\gamma_{r,r+1} \cdots & -\gamma_{r,n} \end{pmatrix} = (I_r, -\Gamma'),$$

mit $\Gamma' \in \mathbb{R}^{r \times (n-r)}$.

Angenommen, dass der Prozess $Z_t = \beta'X_t$ stationär ist, dann wird zuerst $\nu^* = E(Z_t)$ und dadurch $Z_t^* = Z_t - \nu^*$ definiert. Zerlege

$$X_t = \begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix},$$

²⁵Vgl. Hamilton (1994), Kapitel 4.

mit $X_{1,t} \in \mathbb{R}^r$ und $X_{2,t} \in \mathbb{R}^{n-r}$. Mit den oben eingeführten Bezeichnungen gilt:

$$Z_t^* + \nu^* = (I_r - \Gamma') \begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix},$$

oder

$$X_{1,t} = \Gamma' X_{2,t} + Z_t^* + \nu^*. \quad (2.42)$$

Genau so wie in dem ersten Fall wird für $\Delta X_{2,t}$ definiert:

$$u_{2,t} = \Delta X_{2,t} - \mu_2 \quad (2.43)$$

wobei $u_{2,t}$ und μ_2 die letzten $(n-r)$ Elemente von μ und u_t aus \mathbb{R}^n repräsentieren. Die Gleichungen (2.42) und (2.43) werden als die Phillips-Triangular-Darstellung für ein kointegriertes System bezeichnet. Wenn der Vektor X_t genau r Kointegrationsbeziehungen hat und die beiden Gleichungen erfüllt sind, dann ist $X_{2,t} \sim I(1)$ ohne Kointegrationsbeziehungen.

(iii) die Vektor-Autoregressive Darstellung

Die Beschreibung der kointegrierten Zeitreihen mit Hilfe eines Vektor-Autoregressive-Prozesses basiert wieder auf der Wold Darstellung. Angenommen, dass X_t ein $VAR[p]$ -Prozess aus (2.5) ist:

$$X_t - \nu = \Phi_1 X_{t-1} + \dots + \Phi_p X_{t-p} + \varepsilon_t$$

oder

$$\Phi(L)X_t = \nu + \varepsilon_t,$$

mit

$$\Phi(L) = I_n - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_p L^p.$$

Nach der Definition der kointegrierten Zeitreihen muss ΔX_t ein stationärer Prozess sein. Angenommen, dass ΔX_t eine Wold Darstellung hat, also

$$\Delta X_t = (1 - L)X_t = \mu + \Psi(L)\varepsilon_t \quad (2.44)$$

mit

$$\Psi(L) = I_n + \Psi_1 L^1 + \Psi_2 L^2 + \dots$$

gilt. Wenn die Gleichung (2.44) mit $\Phi(L)$ multipliziert wird, so erhält man:

$$\begin{aligned} \Phi(L)(1 - L)X_t &= \Phi(L)\mu + \Phi(L)\Psi(L)\varepsilon_t \\ &\stackrel{L\mu=\mu}{=} (I_n - \Phi_1 - \dots - \Phi_p)\mu + \Phi(L)\Psi(L)\varepsilon_t \\ &= \Phi(1)\mu + \Phi(L)\Psi(L)\varepsilon_t. \end{aligned} \quad (2.45)$$

Setzt man $\Phi(L)X_t = \nu + \varepsilon_t$ in der Gleichung (2.45), so ist

$$(1 - L)(\nu + \varepsilon_t) = \Phi(1)\mu + \Phi(L)\Psi(L)\varepsilon_t.$$

Aus $(1 - L)\nu = 0$ folgt

$$(1 - L)\varepsilon_t = \Phi(1)\mu + \Phi(L)\Psi(L)\varepsilon_t.$$

Es ist offensichtlich, dass $\Phi(1)\mu = 0$, und die Polynome $(1 - L)I_n$ und $\Phi(L)\Psi(L)$ identische Polynome in L sein sollen, d.h. $(1 - z)I_n = \Phi(z)\Psi(z)$ für jedes $z \in \mathbb{C}$. Praktisch impliziert diese Bedingung, dass für $z = 1$ $\Phi(1)\Psi(1) = 0$, daher ist die Matrix $\Phi(1)$ singulär ($\det(\Phi(1)\Psi(1)) = \det(\Phi(1))\det(\Psi(1)) = 0$) und somit hat $|\Phi(z)| = 0$ eine Einheitswurzel.

Sei $\varphi' \in \mathbb{R}^n$ eine Zeile von $\Phi(1)$. Dann gilt $\varphi'\Psi(1) = 0'$ und $\varphi'\mu = 0$. Nach Lemma (2.12) ist φ ein Kointegrationsvektor. Wenn $\beta_1, \beta_2, \dots, \beta_r$ die Basis des Kointegrationsraumes ist, dann muss φ eine lineare Kombination von $\beta_1, \beta_2, \dots, \beta_r$ sein, d.h. es existiert ein Vektor $a \in \mathbb{R}^r$, so dass $\varphi = (\beta_1, \beta_2, \dots, \beta_r)a$ oder $\varphi' = a'\beta'$ ist. Wird das Verfahren auf jede Zeile der Matrix $\Phi(1)$ angewendet, so erhält man die Gleichung:

$$\Phi(1) = \alpha\beta'. \quad (2.46)$$

Als Ergebnis dieser Überlegungen lässt sich das folgende Lemma formulieren:

Lemma 2.13 *Sei (X_t) ein n -dimensionaler $VAR[p]$ -Prozess, mit $(X_t) \sim I(1)$ und $u_t = \Delta X_t - \mu$ hat eine Wold Darstellung (2.41) mit $\Psi_0 = I_n$ und $(t\Psi_t)_{t=0}^{\infty}$ absolut summierbarer Matrixfolge. (X_t) ist genau dann vom Rang r ($0 < r < n$) kointegriert, wenn eine Matrix $\beta \in \mathbb{R}^{n \times r}$ existiert, so dass $\Phi(1)$ die Darstellung $\alpha\beta'$ mit einer Matrix $\alpha \in \mathbb{R}^{n \times r}$ hat, und $\Phi(1)\Psi(1) = 0$ und $\Phi(1)\mu = 0$ gilt.*

(iv) das Fehlerkorrekturmodell

Es basiert auf der VAR-Darstellung. Der $VAR[p]$ -Prozess aus (2.5) wird in eine andere Form umgeschrieben, in der (X_t) durch verzögerte Differenzen und verzögerte Niveaugrößen erklärt wird:

$$\begin{aligned} X_t &= \mu + \Phi_1 X_{t-1} + \dots + \Phi_p X_{t-p} + \varepsilon_t \\ &= \xi_1 \Delta X_{t-1} + \dots + \xi_{p-1} \Delta X_{t-(p-1)} + \mu + \rho X_{t-1} + \varepsilon_t \end{aligned}$$

mit

$$\begin{aligned} \rho &= \Phi_1 + \Phi_2 + \dots + \Phi_p \\ \xi_i &= -(\Phi_{i+1} + \Phi_{i+2} + \dots + \Phi_p), \forall i = 1, 2, \dots, p-1. \end{aligned}$$

Man subtrahiert von beiden Seiten X_{t-1} :

$$\begin{aligned} \Delta X_t &\stackrel{\xi_0 = \rho - I_n}{=} \xi_1 \Delta X_{t-1} + \dots + \xi_{p-1} \Delta X_{t-(p-1)} + \mu + \xi_0 X_{t-1} + \varepsilon_t \\ &\stackrel{\rho - I_n = -\Phi(1)}{=} \xi_1 \Delta X_{t-1} + \dots + \xi_{p-1} \Delta X_{t-(p-1)} + \mu - \Phi(1)X_{t-1} + \varepsilon_t \\ &\stackrel{\Phi(1) = \alpha\beta'}{=} \xi_1 \Delta X_{t-1} + \dots + \xi_{p-1} \Delta X_{t-(p-1)} + \mu - \alpha\beta' X_{t-1} + \varepsilon_t \\ &\stackrel{Z_t = \beta' X_t}{=} \xi_1 \Delta X_{t-1} + \dots + \xi_{p-1} \Delta X_{t-(p-1)} + \mu - \alpha Z_{t-1} + \varepsilon_t \quad (2.47) \end{aligned}$$

Die Darstellung (2.47) wird die Fehlerkorrektur-Darstellung des kointegrierten Systems genannt. Die Fehlerkorrektur-Modelle kombinieren langfristige Gleichgewichtsbeziehungen, die auf der ökonomischen Theorie basieren, mit der kurzfristigen Dynamik in den Modellen.

Ein wichtiges Ergebnis im Bezug auf die vorgestellten Darstellungen eines kointegrierten Systems ist das Granger-Repräsentationstheorem²⁶.

Satz 2.14 (Granger-Repräsentationstheorem) *Sei X_t ein reellwertiger n -dimensionaler Prozess, so dass ΔX_t eine Wold Darstellung $\Delta X_t = \mu + \Psi(L)\varepsilon_t$ hat, und ε_t ist ein n -dimensionaler White-Noise-Prozess mit Λ eine positiv definite Varianz-Kovarianz Matrix, und sei die Folge $(t\Psi_t)_{t=0}^{\infty}$ absolut summierbar. Angenommen, dass X_t genau r ($0 < r < n$) Kointegrationsbeziehungen hat. Dann existiert eine Matrix $\beta \in \mathbb{R}^{n \times r}$, deren Spalten linear unabhängig sind, so dass der Vektor $Z_t \in \mathbb{R}^r$ definiert durch*

$$Z_t = \beta' X_t$$

stationär ist.

Die Matrix β hat die Eigenschaft:

$$\beta' \Psi(1) = 0.$$

Wenn der Prozess (X_t) als ein VAR[p]-Prozess dargestellt werden kann ($X_t - \mu = \Phi_1 X_{t-1} + \dots + \Phi_p X_{t-p} + \varepsilon_t$), dann existiert eine Matrix $\alpha \in \mathbb{R}^{n \times r}$, so dass

$$\Phi(1) = \alpha \beta',$$

und es existieren die Matrizen $\xi_1, \xi_2, \dots, \xi_{p-1}$ aus $\mathbb{R}^{n \times n}$, so daß ΔX_t die Fehlerkorrektur-Darstellung

$$\Delta X_t = \xi_1 \Delta X_{t-1} + \dots + \xi_{p-1} \Delta X_{t-(p-1)} + \mu - \alpha Z_{t-1} + \varepsilon_t$$

hat.

2.3.2 Schätzverfahren. Kointegrationstests

Kleinste-Quadrate-Methode

Der unbekannte Kointegrationsvektor eines Variablen-systems kann mit Hilfe der Kleinste-Quadrate-Methode geschätzt werden, wenn Kointegration zwischen den Prozessen besteht. Da die Linearkombinationen integrierter Variablen mit zunehmendem

²⁶Vgl. Granger (1987).

Stichprobenumfang eine stetig wachsende Varianz aufweisen, (eine Ausnahme stellen diejenigen, die durch Kointegrationsvektoren gebildet werden) sucht diese Methode gezielt den Kointegrationsparameter. Angenommen, dass $(X_t) \sim I(1)$ ein $I(1)$ -Prozess darstellt, und β ist ein Kointegrationsvektor. Definitionsgemäß und unter der Bedingung, dass $Z_t = \beta'X_t$ ein stationärer Prozess, ergodisch²⁷ für den zweiten Moment Prozess ist, gilt

$$N^{-1} \sum_{t=1}^N Z_t^2 = N^{-1} \sum_{t=1}^N (\beta'X_t)^2 \xrightarrow{p} E(Z_t^2). \quad (2.48)$$

Im Gegensatz hierzu, wenn β keinen Kointegrationsvektor darstellt (d.h. $Z_t = \beta'X_t \sim I(1)$) und (Z_t) nicht ergodisch ist, divergiert der Term aus (2.48):

$$N^{-1} \sum_{t=1}^N (\beta'X_t)^2 \xrightarrow{p} \infty.$$

Um einen konsistenten Schätzer für β zu erhalten, muss der Vektor β so ausgewählt werden, dass dieser Term minimiert wird. Dafür wird erst der Vektor β normiert. Angenommen, dass das erste Element des Kointegrationsvektors $\beta_1 \neq 0$, dann kann er gleich Eins gesetzt werden, $\beta_1 = 1$. Die restlichen Elemente werden als ein System unbekannter Parameter $(-\gamma_2 - \gamma_3 \dots - \gamma_n)$ dargestellt:

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} 1 \\ -\gamma_2 \\ \vdots \\ -\gamma_n \end{pmatrix}$$

Im zweiten Schritt werden die Koeffizienten $(-\gamma_2 - \gamma_3 \dots - \gamma_n)$ so gewählt, dass der Term

$$N^{-1} \sum_{t=1}^N (\beta'X_t)^2 = N^{-1} \sum_{t=1}^N (X_{1,t} - \gamma_2 X_{2,t} - \gamma_3 X_{3,t} - \dots - \gamma_n X_{n,t})^2$$

minimiert wird. Die Konsistenz der Kleinst-Quadrate-Schätzer des Kointegrationsvektors wurde von Stock (1994) in einem Satz hergeleitet, der hier in einer leicht abgewandelten Form dargestellt ist.

Satz 2.15 (Satz von Stock) Sei $X_t \in \mathbb{R}^n$ ein stochastischer Prozess der Ordnung Eins, $I(1)$. Der Vektor $X_t = (X_{1,t}, X_{2,t})'$, mit einem Skalar $X_{1,t}$ und einem Vektor $X_{2,t} \in \mathbb{R}^{n-1}$ sei durch genau eine Integrationsbeziehung charakterisiert. Der Kointegrationsvektor sei normiert, d.h. $\beta = (1, -\Gamma)$, mit $\Gamma = (\gamma_2, \gamma_3, \dots, \gamma_n)$. Nach der Definition der kointegrierten Zeitreihen gilt $Z_t = \beta'X_t = X_{1,t} - \Gamma'X_{2,t} \sim I(0)$. Seien zusätzlich die

²⁷Definition und Beweis der Ergodizität für den zweiten Moment der Varianz-Kovarianz stationären Prozessen findet sich in Hamilton (1994), Kapitel 7.

Komponenten des Vektors $X_{2,t}$ nicht kointegriert. Dann gilt für den Kleinste-Quadrate-Schätzer $\hat{\Gamma}$ für Γ

$$N^{1-\delta}(\hat{\Gamma} - \Gamma) \xrightarrow{p} 0$$

für alle $\delta > 0$.

Das Wesentliche an dem Satz von Stock besteht darin, dass die Residuen $(\hat{\Gamma} - \Gamma)$ erheblich schneller gegen Null als die Residuen bei stationären Variablen konvergieren. Aufgrund dieser schnelleren Konvergenz ist der Kleinste-Quadrate-Schätzer des Kointegrationsvektors von Stock (1994) superkonsistent genannt worden. Der Kleinste-Quadrate-Schätzer für den Kointegrationsvektor bleibt auch konsistent, falls die Fehlerterme autokorreliert sind, oder korreliert mit $\Delta X_{2,t}, \Delta X_{3,t}, \dots, \Delta X_{n,t}$ sind.

Bei der Existenz von mehr als einem Kointegrationsvektor stellt sich die Frage, welcher Vektor mittels Kleinste-Quadrate-Methode geschätzt wird, bzw. ob sich überhaupt eine eindeutige Schätzung ergibt. Die Antwort auf diese Frage wird durch den Satz von Wooldridge (1991) gegeben.

Satz 2.16 (Satz von Wooldrige) Sei (X_t) ein stochastischer Prozess mit Phillips-Triangular-Darstellung (2.42), der genau $r > 1$ Kointegrationsbeziehungen hat. Die $\beta_2, \beta_3, \dots, \beta_r$ stellen die Koeffizienten der linearen Projektionen von $Z_{1,t}^*$ auf $Z_{2,t}^*, Z_{3,t}^*, \dots, Z_{r,t}^*$ dar.

$$Z_{1,t}^* = \sum_{i=2}^r \beta_i Z_{i,t}^* + u_t$$

Angenommen, dass

$$\begin{pmatrix} Z_t^* \\ u_{2,t} \end{pmatrix} = \sum_{s=0}^{\infty} \Psi_s \varepsilon_{t-s},$$

mit (ε_t) einem n -dimensionalen White-Noise-Prozess mit endlichem vierten Moment und einer positiv definiten Varianz-Kovarianz Matrix $E(\varepsilon_t \varepsilon_t') = PP'$. Die Folge von $(n \times n)$ Matrizen $(s\Psi_s)_{s=0}^{\infty}$ sei absolut summierbar. Angenommen, dass die Zeilen von $\Psi(1)P$ linear unabhängig sind. Dann konvergieren die Kleinste-Quadrate-Schätzer aus der Regression

$$X_{1,t} = \eta + \gamma_2 X_{2,t} + \gamma_3 X_{3,t} + \dots + \gamma_n X_{n,t} + u_t$$

in Wahrscheinlichkeit gegen:

$$\hat{\eta}_N \xrightarrow{p} (1, -\beta') \nu_1^*$$

mit $\beta = (\beta_2, \beta_3, \dots, \beta_r)' \in \mathbb{R}^{r-1}$ und

$$\begin{pmatrix} \hat{\gamma}_{2,N} \\ \hat{\gamma}_{3,N} \\ \vdots \\ \hat{\gamma}_{n,N} \end{pmatrix} \xrightarrow{p} \begin{pmatrix} \beta \\ \gamma \end{pmatrix}$$

mit

$$\gamma = \Gamma \begin{pmatrix} 1 \\ -\beta \end{pmatrix}.$$

Wie man aus diesem Satz erkennen kann, konvergieren die Residuen in Wahrscheinlichkeit gegen

$$\widehat{u}_t \xrightarrow{p} (1, -\beta') Z_t^*.$$

Die Kleinste-Quadrate-Methode hat folgende Eigenschaften, die sie für die Praxis attraktiv erscheinen lassen. Der erste Vorteil besteht darin, dass dieses Verfahren es ermöglicht, die Kointegrationsparameter separat zu schätzen, ohne dass der gesamte stochastische Prozess für die untersuchten Variablen spezifiziert werden müsste. Der zweite Vorteil besteht in den günstigen asymptotischen Eigenschaften des Verfahrens. Andererseits entsteht bei der kleinen Stichprobe eine erhebliche Verzerrung des Schätzers²⁸. Zusätzlich war die Verteilung des Schätzers in dem von Stock untersuchten Modell asymmetrisch. Das zweite Problem des Verfahrens liegt in der Abhängigkeit des Schätzers von der gewählten Normierung der Koeffizienten. Im multivariaten Fall entsteht ein weiteres Problem der Instabilität des Schätzers. Damit ist gemeint, dass die relativ kleinen Änderungen des Schätzzeitraumes zu relativ großen Änderungen der Parameterschätzungen führen. Die Schätzung springt gewissermaßen von einem Kointegrationsvektor zum anderen. Zusammengefasst liegt die Stärke des Kleinste-Quadrate-Schätzverfahrens in der Behandlung des bivariaten Falls oder derjenigen multivariaten Modelle, in denen davon ausgegangen werden kann, dass es nur eine Kointegrationsbeziehung gibt, und aus theoretischen Gründen eine klare Aufteilung in abhängige und unabhängige Variablen vorliegt. In anderen Situationen ist mit Schwierigkeiten bei der Interpretation der Ergebnisse zu rechnen.

Maximum-Likelihood-Schätzverfahren

Das Maximum-Likelihood-Schätzverfahren wurde von Johansen (1988) entwickelt. Zusammen mit Juselius (1990) hat Johansen dieses Verfahren erweitert und auf verschiedene Datensätze angewendet.

Ausgangspunkt der Schätzung ist das n-dimensionale VAR[p]-Modell in Fehlerkorrekturform wie in (2.47).

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{p-1} \Delta X_{t-(p-1)} + \Pi X_{t-1} + \mu + \varepsilon_t,$$

mit $t = 1, 2, \dots, N$ und (ε_t) einem n-dimensionalen Gauss-White-Noise-Prozess, $E(\varepsilon_t) = 0$ und

$$E(\varepsilon_t \varepsilon_s') = \begin{cases} \mathbf{\Lambda} & \text{für } t = s \\ \mathbf{0} & \text{sonst.} \end{cases}$$

²⁸Vgl. Stock (1994).

Das Ziel des Verfahrens ist, festzustellen, ob die Koeffizientenmatrix Π Informationen über die langfristige Beziehungen zwischen den Variablen im Datenvektor enthält. Drei mögliche Fälle kommen dabei in Betracht. Gilt für Matrix Π , $rg(\xi_0) = r = n$, dann hat sie den vollen Rang, die Variablen des Vektorprozesses (X_t) sind im Niveau stationär und die Annahme $X_t \sim I(1)$ ist falsch. Im zweiten Fall: $r = 0$, ist die Matrix Π eine Nullmatrix, alle Variablen sind integriert, und es existieren keine stationären Linearkombinationen. Falls jedoch für den Rang von Π gilt, dass $0 < r < n$, dann besitzt das System r Kointegrationsbeziehungen und das System kann als Fehler-Korrektur-System interpretiert werden²⁹.

Interessant ist der dritte Fall. Nach dem Granger-Repräsentationstheorem (2.14) gilt, dass die Matrix Π in zwei $(n \times r)$ Matrizen $\xi_0 = -\alpha\beta'$ zerlegbar ist, wobei α die Matrix der Anpassungskoeffiziente und β die Kointegrationsmatrix ist, beide den Rang r haben und die Zeile von β' die r verschiedenen Kointegrationsvektoren angeben. β spannt damit den Kointegrationsraum auf, so dass für X_t als ein nicht stationärer Prozess gilt: $\beta'X_t$ ist stationär.

Im Folgenden soll wieder von der Annahme einer Ordnung $p = 1$ des $VAR[p]$ -Prozesses ausgegangen werden. Die Ergebnisse lassen sich leicht auf den Fall einer beliebigen Ordnung p verallgemeinern. Damit hat die Gleichung aus (2.47) folgende Gestalt:

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \Pi X_{t-1} + \mu + \varepsilon_t.$$

Für eine Stichprobe von Beobachtungswerten (X_0, \dots, X_t) ist die Likelihood-Funktion durch

$$\begin{aligned} \mathcal{L}(\Lambda, \xi_1, \xi_0, \mu) &= |\Lambda|^{-\frac{N}{2}} \exp\left\{-\frac{1}{2} \sum_{t=1}^N [(\Delta X_t - \xi_1 \Delta X_{t-1} - \xi_0 X_{t-1} - \mu)'\right. \\ &\quad \left. \times \Lambda^{-1} (\Delta X_t - \xi_1 \Delta X_{t-1} - \xi_0 X_{t-1} - \mu)]\right\} \end{aligned} \quad (2.49)$$

gegeben. Das Ziel ist es, die Likelihood-Funktion aus (2.49) zu maximieren, unter der Hypothese eines reduzierten Ranges von Π .

'Konzentrieren' der Likelihood-Funktion bedeutet das Maximum der Funktion für (μ, Γ_1) unter der Voraussetzung einer gegebenen Λ und Π zu finden. Die Maximumstellen der konzentrierten Likelihood-Funktion sind identisch mit den Maximumstellen einer Likelihood-Funktion aus (2.49). Für den geschätzten Vektor der Parameter $(\hat{\mu}, \hat{\Gamma}_1)$ kann die Likelihood-Funktion unter der Bedingung eines festen Π und einer festen Matrix Λ maximiert werden. Die Schätzwerte $\hat{\mu}$ und $\hat{\Gamma}_1$ werden durch die Regression erhalten:

$$\Delta X_t - \Pi X_{t-1} = \mu + \Gamma_1 \Delta X_{t-1} + \varepsilon_t. \quad (2.50)$$

Die Residuen dieser Schätzung

$$R = \Delta X_t - \Pi X_{t-1} - \hat{\mu} - \hat{\Gamma}_1 \Delta X_{t-1} \quad (2.51)$$

²⁹Vgl. Engle und Granger (1987).

sind definitionsgemäß orthogonal zu ΔX_{t-1} und haben einen Mittelwert von Null. Die Regression wird in zwei Schritten durchgeführt. Zuerst wird ΔX_t auf ΔX_{t-1} regressiert und in einem zweiten Schritt X_t auf ΔX_{t-1} . Dabei entstehen zwei Residuenvektoren, die orthogonal zu ΔX_{t-1} sind und Erwartungswerte von Null haben.

$$\begin{aligned} R_{0t} &= \Delta X_t - \tilde{\mu} - \psi \Delta X_{t-1} \\ R_{1t} &= X_{t-1} - \check{\mu} - \phi \Delta X_{t-1} \end{aligned}$$

Dann ist auch $R_{0t} - \Pi R_{1t}$ orthogonal zu ΔX_{t-1} und durch das Einsetzen einer Gleichung in die andere wird ersichtlich, dass $R_{0t} - \Pi R_{1t}$ die gleiche Form wie in (2.51) hat:

$$\begin{aligned} R_{0t} - \xi_0 R_{1t} &= \Delta X_t - \tilde{\mu} - \psi \Delta X_{t-1} - \xi_0 (X_{t-1} - \check{\mu} - \phi \Delta X_{t-1}) \\ &= \Delta X_t - \xi_0 X_{t-1} - \hat{\mu} - \hat{\xi}_1 \Delta X_{t-1}. \end{aligned}$$

Deshalb ist der Residuenvektor R gleich $R_{0t} - \Pi R_{1t}$ bei gegebenen Matrizen Π und Λ . Die konzentrierte Likelihood-Funktion wird wie folgt dargestellt:

$$\begin{aligned} \mathcal{L}(\Lambda, \Pi) &= \mathcal{L}(\Lambda, \hat{\Gamma}_1, \Pi, \hat{\mu}) \\ &= |\Lambda|^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2} \sum_{t=1}^N [(R_{0t} - \Pi R_{1t})' \Lambda^{-1} (R_{0t} - \Pi R_{1t})] \right\}. \end{aligned} \quad (2.52)$$

Die Beziehung $R_{0t} - \Pi R_{1t}$ lässt sich als eine um die kurzfristigen Effekte bereinigte Ausgangsgleichung interpretieren.

Bezeichne mit S_{ij} , $i, j = 0, 1$ die Momentenmatrizen der Residuen R_{0t} und R_{1t} .

$$\begin{aligned} S_{00} &= N^{-1} \sum_{t=1}^N R_{0t} R_{0t}' \\ S_{01} &= N^{-1} \sum_{t=1}^N R_{0t} R_{1t}' \\ S_{10} &= N^{-1} \sum_{t=1}^N R_{1t} R_{0t}' \\ S_{11} &= N^{-1} \sum_{t=1}^N R_{1t} R_{1t}' . \end{aligned}$$

Unter der Bedingung $\Pi = -\alpha\beta'$ kann ein Schätzer für α und Λ mittels einer Regression von R_{0t} auf $\beta' R_{1t}$ für die festen Matrix β erhalten werden. Die beiden Schätzer können als eine Funktion von β ausgedrückt werden:

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

Die geschätzten Werte $\hat{\alpha}$ und $\hat{\Lambda}$ werden in die Likelihood-Funktion (2.52) eingesetzt, so erhält man eine dreifach konzentrierte Likelihood-Funktion:

$$\begin{aligned} \mathcal{L}(\beta) &= \mathcal{L}(\hat{\Lambda}, \hat{\Gamma}_1, \hat{\alpha}\beta', \hat{\mu}) \\ &= |\hat{\Lambda}|^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2} \sum_{t=1}^N \left[(R_{0t} - \hat{\alpha}\beta' R_{1t})' \hat{\Lambda}^{-1} (R_{0t} - \hat{\alpha}\beta' R_{1t}) \right] \right\} \\ &= |S_{00} - S_{01}\beta(\beta' S_{11}\beta)^{-1}\beta' S_{10}|^{-\frac{N}{2}} \\ &\times \exp \left\{ -\frac{1}{2} \sum_{t=1}^N [(R_{0t} - S_{01}\beta(\beta' S_{11}\beta)^{-1}\beta' R_{1t})' \right. \\ &\times (S_{00} - S_{01}\beta(\beta' S_{11}\beta)^{-1}\beta' S_{10})^{-1} (R_{0t} - S_{01}\beta(\beta' S_{11}\beta)^{-1}\beta' R_{1t})] \left. \right\}. \end{aligned}$$

Das Maximum der Likelihood-Funktion, die neben Konstanten nur noch die Determinante von $\hat{\Lambda}$ enthält, ist proportional zu

$$\mathcal{L}_{\max}^{-\frac{2}{N}}(\beta) = |S_{00} - S_{01}\beta(\beta' S_{11}\beta)^{-1}\beta' S_{10}|. \quad (2.53)$$

Das resultierende Problem besteht jetzt in der Minimierung von $\mathcal{L}_{\max}^{-\frac{2}{N}}(\beta)$ bezüglich β . Das ist gleichbedeutend mit der Minimierung von

$$\frac{|S_{00}| |\beta' (S_{11} - S_{10} S_{00}^{-1} S_{01}) \beta|}{|\beta' S_{11} \beta|}.$$

Eine Gleichung dieser Form lässt sich als Eigenwertproblem lösen:

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0.$$

Das Eigenwertproblem hat genau n Lösungen, und $\hat{\lambda}_1 > \hat{\lambda}_2 > \dots > \hat{\lambda}_n$ sind genau die Eigenwerte von $S_{10} S_{00}^{-1} S_{01}$ bezüglich S_{11} ³⁰.

Unter der Hypothese $\Pi = -\alpha\beta'$, mit bekanntem $r < n$, wird der Maximum-Likelihood-Schätzer von β ermittelt, indem

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0$$

gelöst wird, so dass sich die $\hat{\lambda}_1 > \hat{\lambda}_2 > \dots > \hat{\lambda}_n$ und die Matrix $\hat{V} = (\hat{v}_1, \hat{v}_2, \dots, \hat{v}_n)$ der zugehörigen Eigenvektoren ergeben. Die r ersten Zeilen von \hat{V} werden die kanonischen Varianten genannt, und die Eigenwerte sind die quadrierten kanonischen Korrelationen der Niveau-Residuen R_{1t} bezüglich der Differenzen-Residuen R_{0t} .

Die r ersten Eigenvektoren \hat{v}_i ($i = 1, 2, \dots, r$) bestimmen die Linearkombinationen $v_i' X_t$, die am stärksten mit dem bedingten Prozess $(\Delta X_t | \Delta X_{t-1})$ korreliert sind.

³⁰Die Herleitung der oben dargestellten Gleichungen findet sich bei Johansen und Juselius (1990).

Dementsprechend ist die Größe der λ_i ein Maß dafür, wie stark die Kointegrationsrelation $v_i'X_t$ mit einer Linearkombination des stationären Prozesses ΔX_t korreliert ist. Die letzten $(n-r)$ Kombinationen definieren die Linearkombinationen, in denen der Prozess $(\Delta X_t|\Delta X_{t-1})$ nicht stationär ist, und sind daher auch nicht mit dem stationären Teil des Prozesses korreliert. Folglich setzt man $\lambda_i = 0$ ($i = r + 1, \dots, n$). Als Schätzer von β werden die r ersten Eigenvektoren

$$\hat{\beta} = (\hat{v}_1, \hat{v}_2, \dots, \hat{v}_r) \tag{2.54}$$

gewählt, korrespondierend zu den r größten Eigenwerten.

Die Schätzer der anderen Parameter ergeben sich dann als

$$\hat{\alpha} = S_{01}\hat{\beta}, \tag{2.55}$$

$$\hat{\Pi} = S_{01}\hat{\beta}\hat{\beta}', \tag{2.56}$$

$$\hat{\Lambda} = S_{00} - \hat{\alpha}\hat{\alpha}'. \tag{2.57}$$

wobei $\hat{\alpha}, \hat{\Pi}, \hat{\Lambda}$ von der Wahl des optimalen Schätzers $\hat{\beta}$ abhängen, während die maximierte Likelihood-Funktion für ein bekanntes r

$$\mathcal{L}_{max}^{-\frac{2}{N}} = |S_{00}| \prod_{i=1}^r (1 - \hat{\lambda}_i)$$

nicht mehr von der Wahl von β abhängt.

In der empirischen Praxis ist der Rang von β im allgemeinen nicht bekannt, sondern muss aus den Daten ermittelt werden. Das statistische Problem liegt darin, zwischen den Eigenwerten, die von Null signifikant verschieden sind und solchen die mit Null identisch sind, zu unterscheiden. Mit Hilfe eines Likelihood-Ratio-Test wird der Rang der Matrix β geschätzt.

Der Maximum-Likelihood-Schätzer identifiziert keine Kointegrationsvektoren, sondern stellt einen Schätzer des Kointegrationsraumes dar. Johansen und Juselius (1990) und Johansen(1991) schlagen verschiedene strukturelle Hypothesen hinsichtlich der Form der Parameter α und β vor, die mit Verhältnistests überprüft werden können. Durch eine Annahme oder eine Ablehnung unterschiedlicher Hypothesen wird bei den empirischen Anwendungen eine parametrische Interpretation der Kointegrationsvektoren angestrebt.

Engle-Granger-Tests

Ein zweistufiges Regressions-Verfahren zur Untersuchung von Variablen auf Kointegration wurde von Engle und Granger (1987) vorgestellt. Die Überprüfung der Existenz einer Kointegrationsbeziehung zwischen den Variablen X_t und Y_t erfolgt bei

dem Verfahren in zwei Schritten. Zuerst werden beide Zeitreihen (X_t) und (Y_t) auf ihren Integrationsgrad untersucht. Dazu werden oben beschriebene Einheitswurzeltests, beispielsweise der erweiterte Dickey-Fuller und der Phillips-Perron-Test verwendet. Im zweiten Schritt wird überprüft, ob die Residuen ε_t der Regression $X_t = \alpha + \beta Y_t + \varepsilon_t$ stationär sind und somit zwischen X_t und Y_t eine Kointegrationsbeziehung vorliegt. Anwendung findet die Kleinste-Quadrate-Methode für die Schätzung der Parameter α und β und wiederum der Einheitswurzel-Test. Das Engle-Granger-Verfahren ist ein anschauliches und einfaches Verfahren, allerdings ist es nicht ohne weiteres anwendbar. Erstens hängen die Testergebnisse von der Variablenspezifikation ab. Zweitens kann im mehrdimensionalen Fall der Kointegrationsvektor nicht eindeutig geschätzt werden, weil beispielsweise zwischen drei Variablen bis zu zwei Kointegrationsbeziehungen bestehen können.

Wenn ein System aus mehreren Zeitreihen auf Kointegration überprüft werden soll, so ist es üblich, mit dem Johansens Verfahren zu arbeiten. Eine Begründung für die Anwendung des Verfahrens ist, dass die Einheitswurzel-Tests eine niedrige Mächtigkeit haben. Das kann zur Schein-Kointegration führen. Im mehrdimensionalen Fall bringt die Kleinsten-Quadrate-Methode Probleme in Bezug auf die Mehrdeutigkeit der Schätzung. Ein weiteres Problem liegt in der Schwierigkeit bei der Identifizierung der abhängigen und unabhängigen Variablen.

Johansens Tests

Bei dem Maximum-Likelihood-Verfahren von Johansen lassen sich zwei Teststatistiken mit Hilfe der konzentrierten Likelihood-Funktion $L_{\max}^{-2/N}$ aus (2.53) herleiten.

Die erste wird beim sogenannten Trace-Test benutzt. Der Trace-Test prüft, ob die $(n-r)$ kleinsten Eigenwerte die maximierte Likelihood-Funktion signifikant beeinflussen. Ist dies nicht der Fall, liegen höchstens r Kointegrationsvektoren vor. Die Nullhypothese $H_0 : r = r_0$ wird gegen die Alternativhypothese $H_1 : r > r_0$ für $r_0 = n - 1, \dots, 0$ getestet. Die Trace-Teststatistik ist

$$-2 \ln(Q; r|n) = -N \sum_{i=r+1}^n \ln(1 - \hat{\lambda}_i). \quad (2.58)$$

Zu beachten ist, dass sich für $\hat{\lambda}_i$ mit $i = r + 1, \dots, n$ die Teststatistik $-2 \ln(Q; r|n)$ dem Wert Null annähern wird. Andererseits wird sie dann groß, wenn ein oder mehrere der $\hat{\lambda}_i$ nahe bei 1 sind.

Der zweite Test wird λ_{\max} -Test genannt. Er überprüft, ob der $(r+1)$ -te Eigenwert noch die maximierte Likelihood-Funktion signifikant verändert. In diesem Fall wird $H_0 : r \leq r_0$ gegen die Alternativhypothese $H_1 : r = r_0 + 1, r_0 = n - 1, \dots, 0$ getestet. Die Teststatistik des λ_{\max} -Tests ist

$$-2 \ln(Q; r|r+1) = -N \ln(1 - \hat{\lambda}_{r+1}). \quad (2.59)$$

Die Verteilungen dieser Teststatistiken sind keine Standard t-Verteilung. Johansen (1991) hat die Grenzverteilungen als Funktionale auf Standard Brownsche Bewegung dargestellt. Die asymptotischen Verteilungen der Teststatistiken und Schätzer sind davon abhängig, welche Annahmen bezüglich der Modellbildung getroffen werden. Kritische Werte für die jeweiligen Fälle wurden durch Simulation von Johansen und Juselius (1990) für $n - r = 1, \dots, 5$ ermittelt.

2.3.3 Probleme

Kointegrationsbeziehungen zwischen internationalen Aktienmärkten wurden bereits in mehreren Untersuchungen festgestellt. So identifizierten Taylor and Tonks (1989) eine Kointegrationsbeziehung für die Aktienmärkte von Großbritannien, Deutschland, den Niederlanden und Japan. Für ein System aus fünf Aktienmarktindizes (USA, Kanada, Deutschland, Japan und Großbritannien) fand Kasa (1992) bis zu vier Kointegrationsvektoren. Arshanapalli and Doukas (1993) berichteten von Kointegration für die G5-Länder und Corhay und andere (1993) identifizierten Kointegration für die europäischen Aktienmärkte. Zu ähnlichen Ergebnissen kamen auch Leachman und Francis (1995) oder Dickinson (2000).

Mehrere empirische Untersuchungen liefern aber auch andere Ergebnisse. Richards (1995) hat zum Beispiel die Arbeit von Kasa (1992) mit Hilfe verschiedener Methoden kritisch beleuchtet und kommt zu dem Ergebnis, dass die nationalen Aktienindizes nicht um eine gemeinsame stochastische Komponente kointegriert sind. Auch Godbout und van Norden (1997) und Ahlgren und Antell (2002) befassen sich mit ähnlichen Widersprüchen in der Literatur.

Ein Grund für solche Ergebnisse ist in der Vielfalt der relevanten Parameter und Freiheitsgrade bei der Auswahl eines geeigneten ökonometrischen Modells zu sehen. So werden mehrere Parameter immer wieder alternativ berücksichtigt: Zeithorizont, Anzahl der Verzögerungen (Lags) in dem Fehlerkorrektur-Modell, Datenfrequenz, Trendeinbeziehung, Strukturbrüche, Volatilität des Marktes, Saisoneffekte (vgl. Kasa (1992), Cheung und Lai (1993), Seo (1999), Hubrich und andere (2001), Lütkepohl (2004), Lütkepohl und andere (2004)).

Die Johansen-Methode bietet im Vergleich zu dem Engle-Granger-Verfahren einige Vorteile. Sie kann für mehrdimensionale Systeme angewendet werden. Die Methode ermöglicht sowohl die Ermittlung der Anzahl der signifikanten Kointegrationsvektoren als auch eine Schätzung des Kointegrationsraumes mit Hilfe der Likelihood-Funktion. Ein weiterer Vorteil gegenüber dem Engle-Granger-Verfahren besteht darin, dass die Johansen-Methode die Einheitswurzel-Tests für die Feststellung der Kointegration nicht nutzt. Der Hauptkritikpunkt des Engle-Granger-Verfahrens ist die geringere Mächtigkeit der angewendeten Einheitswurzel-Tests.

Allerdings haben auch die Johansen-Tests mehrere Nachteile. Sie reagieren sehr sensibel auf die Größe der Stichprobe oder auf die Verletzung der Annahme der Normalverteilung der Störterme. Zu beachten ist die Sensibilität der Tests bei einer Fehlspezifikation der Anzahl der Verzögerungen p im VEC[p]-Modell. Einen weiteren Aspekt stellt die Empfindlichkeit der Tests gegenüber der Auswahl des Trends in dem Modell dar. Eine Reihe von Tests wurde in der letzten zehn Jahren entwickelt, um die Schwäche der Johansen-Methode zu umgehen³¹.

Maddala (1998) diskutiert den Beitrag von Podivinsky (1990), der die Mächtigkeit des Johansen-Tests, des erweiterten DF-Tests und der Kointegrationsregression nach Durbin-Watson für die Stichproben mit einem kleinen Umfang untersucht hat. Dabei fand Podivinsky heraus, dass die von Johansen und Juselius tabellierten kritischen Werte nicht unbedingt für die Tests in den Kleinstichproben geeignet sind. Bei seinen Monte-Carlo-Simulationen kam Eitrheim (1992) zum ähnlichen Ergebnis. Er hat festgestellt, dass die Anzahl der Kointegrationsvektoren in Kleinstichproben nahe des Kointegrationsfalls zu niedrig geschätzt werden kann. Eine weitere Anmerkung von Podivinsky ist, dass alle Tests zur falschen Bestätigung der Nullhypothese führen können, wenn die Anzahl der in der Analyse verwendeten Variablen niedrig ist. Andererseits führt eine Überspezifikation zu weniger gravierenden Konsequenzen.

Cheung und Lai (1993) stellten fest, dass in den endlichen Stichproben die Verzerrung der Testgrößen bei dem Johansen-Test eine positive Funktion von $T/(T - kp)$ darstellt. Dabei ist T der Stichprobenumfang, k ist die Anzahl der Variablen und p die Anzahl der Lags in dem VAR[p]-Modell. Wenn $T/(T - kp) > 1$ für jeden endlichen Zeithorizont T ist, dann sind die Teststatistiken so verzerrt, dass die Kointegration 'zu oft' gefunden wird.

Eine richtige Spezifikation der Anzahl der Verzögerungen p ist bei den Kointegrationstests genauso wichtig wie bei den Einheitswurzel-Tests. Cheung und Lai (1993) stellten fest, dass das Akaike-Informationskriterium (AIC) und das Bayesian-Informationskriterium (BIC) für eine Auswahl von p nicht besonders geeignet sind. Dies überrascht nicht, wenn man sich an die dies bezügliche Diskussion in dem Kapitel über die Einheitswurzel-Tests erinnert. Für die Modelle mit der kleinen Ordnung p sind der trace- und der λ_{max} -Test nicht geeignet, weil die Teststatistiken stark verzerrt sind. Die Tests neigen eine scheinbare Kointegration zu identifizieren, das heißt sie identifizieren Kointegrationsbeziehungen, die gar nicht existieren. Boswijk und Franses (1992) kamen zum identischen Ergebnis. Sie stellten zusätzlich fest, dass eine Überspezifikation von p zu einer niedrigen Mächtigkeit der Tests führt.

Die späteren Untersuchungen von Ng und Perron (1995), Lütkepohl und Saikonen (1999) oder von Hübrich und anderen (2001) bestätigen diese Ergebnisse. Lütkepohl und Saikonen (1999) vergleichen unter anderem AIC und das Schwarz- Informations-

³¹Einen ausführlichen Vergleich der Tests bezüglich ihrer Mächtigkeit befindet sich bei Maddala (1998) und bei Hübrich, Lütkepohl, Saikonen (2001)

kriterium (SIC) für eine Auswahl von p in einem kointegrierten VAR[p]-System ohne eine deterministische Komponente. Einerseits wählt SIC die Anzahl der Verzögerungen p konsistent für einen endlichen datengenerierenden Prozess, dagegen überspezifiziert der AIC diese Anzahl mit positiver Wahrscheinlichkeit. Die asymptotische Theorie der Kointegrationstests hilft nicht bei einer Diskriminierung der Kriterien. In kleinen Stichproben sind die Kointegrationstests insbesondere von p abhängig. Wenn die Anzahl der Verzögerungen p zu klein ausgewählt wird, dann führt es zu einer starken Verzerrung der Statistiken und zu einer Reduzierung der Mächtigkeit der Tests. Das Gleiche gilt allerdings für große p , vgl. Hubrich und andere (2001).

Im allgemeinen ist eine Auswahl der Länge der Lags mit einem Selektionskriterium eine gute Strategie für die Anwender der Methodik, wenn dieses Kriterium eine Balance zwischen guter Approximation des datengenerierenden Prozesses und dem effizienten Nutzen der Information aus der Stichprobe herzustellen versucht, vgl. Lütkepohl (1999). Allerdings scheint es, dass kein optimales Kriterium dafür existiert, so Hubrich und andere (2001).

Ein weiterer Aspekt muss bei der Anwendung der Kointegrationstests berücksichtigt werden, nämlich der Einfluss der saisonalen Dummy-Variablen und Strukturverschiebungen. Johansen (1995) stellte fest, dass die Einführung der zentrierten saisonalen Dummy-Variablen in der Regression keine Auswirkung auf die asymptotischen Eigenschaften der Tests haben. Allerdings gilt das nicht für die Shift-Dummies³².

Die auf der Likelihood-Funktion basierten Tests wurden unter der Annahme der Gauss-Normalverteilung abgeleitet. Mehrere ökonomische Zeitreihen, insbesondere Finanzdaten, sind durch 'dicke Ränder' der Verteilung und sogenannte 'Varianz-Cluster' gekennzeichnet. Dies ist inkonsistent mit der Normalverteilungsannahme. Laut Cheung und Lai (1993) führt das zu signifikanten Effekten auf beide λ_{max} - und λ_{trace} -Tests, allerdings auf den Trace-Test in geringerem Maße. Lukas (1997, 1998) schlägt vor, eine Pseudo-Likelihood-Funktion für die entsprechenden Tests zu benutzen, diese soll zur Verbesserung der Mächtigkeit der Tests führen.

Hubrich und andere (2001) verglichen weitere Kointegrationstests miteinander, wie beispielsweise die auf der χ^2 -Verteilung basierten Tests, die auf der kanonischen Korrelation basierten Tests, die Stock-Watson- sowie nicht parametrische Tests von Bierens. Die genannten Tests, mit Ausnahme des auf der χ^2 -Verteilung basierten Tests, liefern eine schlechte Performance, sogar unter idealen Voraussetzungen. Diese Tests wurden von Hubrich und anderen für eine Anwendung in der Praxis nicht empfohlen.

³²Eine ausführliche Analyse der beiden Probleme befindet sich bei Hubrich und anderen (2001).

Kapitel 3

Empirische Analyse und Folgerungen

Für die Untersuchung der Gleichgewichtsbeziehungen zwischen internationalen Aktienmärkten wurde vorerst eine empirische Datenanalyse ausgewählt. Im vorherigen Kapitel 2 wurde ein ökonometrisches Konzept zur Beschreibung der Aktienmarktbeziehung und die Schätzmethode für die langfristigen Zusammenhänge der Aktienmärkte dargestellt. Die Kointegrationsmethode ist eine besonders geeignete Methode, um langfristige Gleichgewichtsbeziehungen festzustellen. Die Voraussetzung für die Kointegrationsuntersuchung ist ein identischer Integrationsgrad aller Aktienindizes. Das kann mittels Einheitswurzel-Tests festgestellt werden. Sowohl bivariate als auch mehrdimensionale Kointegrationsanalysen liefern interessante Erkenntnisse zum Verstehen der gemeinsamen Aktienbewegung auf internationalen Märkten.

Die Anzahl der Kointegrationsbeziehungen kann als der Abhängigkeitsgrad der Märkte in dem Gesamtsystem interpretiert werden. Allerdings, wenn nur eine einzige Kointegrationsbeziehung gefunden wird, erleichtert es die ökonomische Interpretation der Ergebnisse. Die geschätzte Kointegrationsgleichung stellt das langfristige Gleichgewicht zwischen den internationalen Aktienmärkten dar. Die langfristige Kointegrationsanalyse wird durch das Schätzen der kurzfristigen Anpassungen (bei VECM), die Varianz-Dekomposition und die Impuls-Antwort-Analyse ergänzt.

3.1 Datenbeschreibung

Die empirischen Daten der Aktienmärkte der G5-Länder dienen zur Überprüfung der vorgestellten ökonomischen Theorien. Fast 80% des gesamten Aktienhandels wird auf den Börsen dieser Länder abgewickelt (vgl. Grafik (3.1)). Die Bedeutung der größten europäischen, der amerikanischen und der japanischen Börsen kann durch die neu entstandenen Handelsplätze schnell wachsender Wirtschaften wie China nachlassen. Die Daten von 2003 zeigen allerdings, dass die Änderungen bisher nicht so gravierend sind,

wie zunächst gedacht wurde. Aus der Sicht der internationalen Diversifikation stellen die Aktienindizes der G5-Länder ein sehr breites Portfolio dar. Zusätzlich ist jeder nationale Aktienindex gut über die Wirtschaftsbranchen jedes Landes diversifiziert.

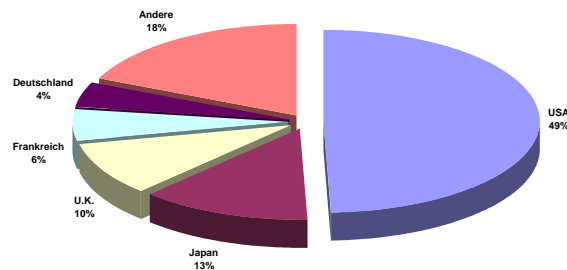


Abbildung 3.1: Vergleich des Handelsvolumens der internationalen Aktienmärkte der G5-Länder im Jahr 2000

Quelle: Morgan Stanley Capital International Perspectives, Juni 2000.

Für die Analyse wurden die Daten aus der OECD Datenbank verwendet. Die Zeitreihen wurden vom Juli 1964 bis April 2004 erfasst; unter Berücksichtigung der monatlichen Frequenz stehen 485 Beobachtungen zur Verfügung. Für die Untersuchung wurden die folgenden internationalen Aktienindizes verwendet:

- Amerikanischer Aktienindex, *NYSE Common Stocks*, ist ein zusammengesetzter Aktienindex aus allen bei der New York Stock Exchange registrierten Unternehmen. Unter anderem umfasst er Investmentfonds und ausländische Unternehmen. Der Index ist auf der Basis der täglichen Preise als ein gewichtetes arithmetisches Mittel berechnet. Die Gewichtung entspricht der Marktkapitalisierung der Unternehmen. Die Monatsdaten sind aus den täglichen Durchschnittswerten gewonnen. Quelle ist das US Bureau of Census.
- *CDAX*, der deutsche Aktienindex, wird von der Deutschen Börse AG in Frankfurt am Main gerechnet. Der Index zeigt Preisbewegungen aller bei der Deutschen Börse AG zugelassenen Unternehmen. Dies entspricht bis zu 727 Unternehmen¹ aus allen Branchen. Die Dividenden wurden von den Aktienpreisen abgezogen; die Preise sind vor Steuern ausgewiesen. Die monatlichen Werte sind als Durchschnitte aus täglichen Daten gewonnen, die die Aktienpreise am Börsenschluss wiedergeben. Der Index ist nach der *Laspeyres Formel* berechnet; dabei wird der

¹Seit 1993 beinhaltet der Index alle Unternehmen der BRD, inklusive aller Unternehmen der ehemaligen DDR.

Preis jeder Aktie mit der Anzahl zum Handel zugelassener Aktien gewichtet. Quelle ist die deutsche Zentralbank.

- Financial Times Stock Exchange Actuaries non financial Index, *FT-SE-A*, wird auf der London Stock Exchange notiert. Der britische Index ist ein marktgewichteter Preisindex vor Steuern. Er repräsentiert die Wertentwicklung heimischer und in Großbritannien angesiedelter internationaler Unternehmen. Der Index umfasst ca. 660 produzierende Unternehmen mit einem Kapitalwert von über 60 Mio. Pfund. Die täglichen Werte stellen die mit der Marktkapitalisation des Unternehmens gewichteten Durchschnitte dar. Für eine Berechnung der monatlichen Werte wurden die täglichen Durchschnitte genommen. Quelle ist die Bank of England.
- *TOPIX* stellt die Aktienpreisbewegung ausgewählter Unternehmen aus Japan dar. Der Index repräsentiert die Preise der Unternehmensaktien, die in der ersten Sektion der Tokyo Stock Exchange gehandelt werden. Die erste Sektion beinhaltet 1239 Unternehmen (Stand: Januar 1997); Investmentfonds sind ausgeschlossen. Die monatlichen Werte wurden aus den täglichen Werten berechnet; dabei wurde der Marktwert des Unternehmens bei der Bildung der gewichteten Durchschnitte verwendet. Quelle ist die japanische Zentralbank.
- Index of the Société des Bourses Françaises, *SBF 250*, wird von der French Stock Exchange Society bestimmt. Der Index ist konzipiert, um Änderungen sowohl auf dem Gesamtmarkt als auch auf seinen Komponenten zu reflektieren. Der *SBF 250* stellt einen gemeinsamen Index für zwölf Sektoren dar, die in drei Gruppen (Industrie, Service, Finanzen) aggregiert sind. Die Investmentfonds sind aus der Berechnung ausgeschlossen, allerdings können an der Paris Stock Exchange zugelassene ausländische Unternehmen in der Berechnung beinhaltet sein. Der Wert des Indizes ist um die ausgezahlten Dividenden bereinigt. Die monatlichen Werte sind Durchschnitte der täglichen Endwerte. Quelle ist die französische Zentralbank.

Entsprechend den Angaben der OECD stellen sämtliche Aktienindizes Preisindizes dar. Ausnahmen hiervon sind der französische Index *SBF 250* und der deutsche Index *CDAX*, die als Performance Indizes² zur Verfügung stehen. Das Jahr 2000 wird als Basisjahr genommen. (2000=100%).

Für ökonomische Untersuchungen, sowohl aus theoretischer als auch aus praktischer Sicht, sind reale Daten interessant. Bei der Transformation der nominalen in reale Größen wird Neutralität des Geldes unterstellt. Für eine Analyse langfristiger Zusammenhänge der verschiedenen Aktienmärkten ist die Neutralitätsannahme annäherungsweise richtig³. Für die Deflationierung der Aktienindizes wurden die Konsumentenpreisindizes benutzt (CPI), weil sie im Vergleich zu den Produzentenpreisindi-

²Ein Performance Index berücksichtigt die Dividendenzahlungen für die Aktien des in dem Index enthaltenen Unternehmens.

³Vgl. Mankiw (1994).

zes unter anderem die Preise der Finanzdienstleistungen enthalten. Folgende CPI stellt OECD zur Verfügung:

- Der CPI der USA stellt die durchschnittliche Preisänderung eines festgesetzten Warenkorbs inklusive der Dienstleistungen dar. Alle Steuern sind in den Preisen enthalten. Die Stichprobe umfasst nur die städtische Bevölkerung und reflektiert die Bevölkerungsverteilung von 2000. Ausgeschlossen sind Leute, die in ländlichen Regionen und auf Farmen leben, genauso die Personen, die militärischen Dienst leisten. Die Preise wurden in 87 Städten von ca. 23 000 Einzelhandelsunternehmen zusammengefasst. Die Preise für Eigentumswohnungen und für Miete umfassen 20 000 Haushalte. Quelle: US Bureau of Labor Statistics.
- Der deutsche Preisindex erfasst die monatliche Schwankung der Lebenskosten durch den Vergleich der Preisvariation eines festgelegten Warenkorbs. Sämtliche Steuern sind in den Preisen enthalten. Der Warenkorb umfasst 750 repräsentative Güter und Dienstleistungen, die über 350 000 Preisnotierungen liefern. Die Anzahl an der Umfrage beteiligten Einzelhandelsunternehmen liegt bei ca. 20 000; sie sind über 190 Regionen Deutschlands verteilt. Die Daten bis 1991 reflektieren die Preisänderungen in Westdeutschland, ab 1991 stellt der Index die Preisänderungen für das gesamte Deutschland dar. Quelle: Statistisches Bundesamt Deutschland.
- Die monatlichen Preisänderungen in Japan sind durch japanischen CPI repräsentiert und stellen im Wesentlichen die Transaktionskosten dar. Ähnlich zum amerikanischen CPI enthält der japanische Index Haus- und Mietkosten. Ausgeschlossen aus dem Index sind temporäre Preisermäßigungen und spezielle Verkaufspreise für Sonderangebote. Die Stichprobe umfasst fast alle Schichten der Bevölkerung; ausgeschlossen sind Beschäftigte der Agrarindustrie, Forstwirtschaft, Fischerei und Ein-Personen-Haushalte. Die Preise sind durch die Zählungen am Mittwoch, Donnerstag und Freitag der Woche erfasst, die den zwölften Tag jedes Monats beinhaltet. Ca. 580 repräsentative Artikel, die in 34.000 Gütern und Dienstleistungen widergespiegelt werden, ergeben 233.000 monatliche Preisnotierungen. Quelle: Statistics Bureau of Japan
- Der französische CPI gibt die Preisänderungen eines Warenkorbs mit 303 Kategorien der Güter und Dienstleistungen wieder. Aktionspreise für Kleidung usw. sind im Index eingeschlossen, wenn sie allgemein innerhalb von 24 Stunden gültig waren. Die Mietpreise sind im Index dagegen nicht enthalten. Der Index umfasst alle Arten der Haushalte und deckt 94,1% der Konsumkategorien dieser ab. Die Preise für den öffentlichen Verkehr, Tabakwaren, Autos, Elektrizität, Gas, Medikamente und medizinische Leistungen, Post und Telekommunikation wurden zentral erhoben. Die restlichen Preisänderungen wurden durch Stichproben bei 30.000 Verkaufsstellen in über 100 Städten erhoben, die über 160.000 Preisnotierungen generiert haben. Quelle: Statistical Office of France.

- Der britische Preisindex ist als Retail Prices Index (RPI) bekannt; er gibt die Messung der monatlichen durchschnittlichen Preisänderungen von Konsumgütern und Dienstleistungen wieder. Die Preise beinhalten alle Steuern. Der RPI deckt das Vereinigte Königreich, und damit England, Schottland, Wales und Nordirland ab. Die Bevölkerungsgruppe der Stichprobe beinhaltet fast alle Arten der Haushalte; ausgeschlossen sind Rentner und Haushalte mit einem Einkommen, das zu den Top 4% des Gesamteinkommens gehört. Der RPI umfasst 14 Gruppen, 85 Sektoren und 630 Artikel, die jeden Januar neu zusammengestellt werden. Der Index beinhaltet über 600 separate Güter und Dienstleistungen, die in 146 geographischen Regionen ermittelt werden. Monatliche Preise ergeben über 120 000 Preisnotierungen von 20.000 Verkaufsstellen. Die Haus- und Mietpreise sind im Index enthalten. Quelle: Office for National Statistics, UK.

Für alle fünf CPI wurde das Jahr 2000 als Basisjahr genommen; für die Berechnung wurde die *Laspeyres Formel* verwendet; die Preise sind saisonbereinigt.

Die Untersuchung der Aktienindizes wurde in dieser Arbeit zum Zweck der Vergleichsanalyse in einer gemeinsamen Währung (DM/Euro)⁴ und in der nationalen Währung durchgeführt. Die Konvertierung in die gemeinsame Währung berücksichtigt die Perspektive der europäischen Anleger⁵, die nicht gegen das Wechselkursrisiko abgesichert sind. Die Umrechnung in (DM/Euro) ist insbesondere relevant für die europäischen Anleger bei ihren internationalen Investitionsentscheidungen⁶.

Ein Vergleich der Ergebnisse der empirischen Tests in einer gemeinsamen Währung und in der nationalen Währung ist aus theoretischer Sicht ebenfalls nachvollziehbar. Bei einer Untersuchung der internationalen Märkte in gemeinsamer Währung wird die Gültigkeit der Kaufkraftparität angenommen. Es wird darauf verzichtet, wenn die Untersuchung in der inländischen Währung durchgeführt wird. Dabei muss die Kaufkraftparität nicht zu jedem Zeitpunkt der beobachteten Periode gelten.

Ein nächster Aspekt der Datenkonvertierung ist mit den statistischen Möglichkeiten der Modellierung verbunden. Es ist von Interesse, ob das Wechselkursrisiko bei der Modellierung genauer erfasst wird, wenn der Wechselkurs als eine unabhängige exogene Variable in das Gesamtsystem eingeführt wird. Eine eindeutige Antwort auf

⁴Der Wechselkurs \$/DM wurde bis 1999 in \$/Euro von OECD angepasst.

⁵Unter 'europäischen' Investoren sind die deutschen und die französischen Anleger zu verstehen. Im Folgenden werden allerdings alle Tests aus der Sicht der deutschen Investoren präsentiert. Aufgrund der Transformation der Daten in Euro könnten die entsprechenden Tests auch für die französischen Investoren durchgeführt werden. Für die britischen Investoren ist eine Umrechnung in Pfund Sterling notwendig.

⁶In einer großen Anzahl der Untersuchungen wurde eine Konvertierung in US \$ gemacht, z.B. Kasa (1992), Dickinson (2000); einige Analysen wurden auch aus der Sicht der englischen Anleger durchgeführt, z.B. von Taylor und Tonks (1989) oder von Fifield und anderen (1999) und (2002). Allgemein existieren kaum wissenschaftlichen Veröffentlichungen aus der Perspektive der deutschen Anleger, vgl. Gerke, Mager und Röhrs (2005)

diese Frage konnte bis jetzt nicht gegeben werden. Sie eröffnet ein weiteres Feld der Aktienmarktanalyse.

Für die Untersuchung wurden die entsprechenden Wechselkurse aus dem Datenpool des Internationalen Monetary Fonds (IMF) genommen. Die Notierung ist in amerikanischen Cent pro eine (DM/Euro), ein Yen, ein Frank/Euro und pro ein Pfund Sterling gegeben. Die Endmonatsdaten sind nicht saisonbereinigt. Mittels der Dreisatz-Umrechnung wurden alle Aktienindizes nach (DM/Euro) konvertiert.

Für die Renditen der Aktienmärkte wird eine stetige Verzinsung angenommen. Die zusammengesetzte Rendite eines Aktienmarktes wurde danach als Differenz der logarithmierten Werte berechnet:

$$r_t = \ln P_t - \ln P_{t-1} \quad (3.1)$$

3.1.1 Explorative Statistiken

In den nachfolgenden Tabellen (3.1) und (3.2) sind explorative Statistiken zu den realen Aktienindizes in (DM/Euro) und in der eigenen Währung zusammengefasst.

Tabelle 3.1: Explorative Statistiken der Aktienindizes in (DM/Euro)

	USA	Japan	France	Germany	UK
Mittelwert	3,504	3,626	3,382	3,518	3,589
Std. Abweichung	0,527	0,787	0,540	0,449	0,514
Schiefe	0,284	-0,032	0,078	0,537	-0,331
Wölbung	2,516	1,691	2,615	2,566	2,831
Jarque-Bera	11,198**	34,574**	3,468	27,030**	9,376**

Notiz: JB ist der Jarque-Bera-Test auf Normalverteilung.

* statistisch signifikant zum Niveau 5%; ** statistisch signifikant zum Niveau 1%.

Der Mittelwert der logarithmierten Indizes in (DM/Euro) ist niedriger im Vergleich zum Mittelwert der Indizes in eigener Währung. Eine Ausnahme stellt der französische Markt dar, bei dem eine umgekehrte Situation beobachtet wird. Der höchste Wert der Indizes in eigener Währung liegt bei 4,107 Punkten und ist vom japanischen Markt. Sein Unterschied zum Wert in (DM/Euro) liegt im Bereich von 0,5 Punkten und stellt damit auch die höchste Differenz in der Gruppe dar. Der britische Index verliert mehr Punkte durch die Konvertierung in (DM/Euro) als der amerikanische Index. Die Wirkung der Konvertierung auf den Mittelwert lässt sich nicht eindeutig interpretieren. Sowohl eine Abwertung des amerikanischen Dollars als auch eine Aufwertung des japanischen Yen zur deutschen Mark führt zu einer gleichgerichteten Reaktion der Werten im Niveau.

Tabelle 3.2: Explorative Statistiken der Aktienindizes in der eigenen Wahrung

	USA	Japan	France	Germany	UK
Mittelwert	3,673	4,107	3,331	3,518	3,803
Std. Abweichung	0,475	0,557	0,549	0,449	0,447
Schiefe	0,480	0,171	0,260	0,537	-0,275
Wolbung	2,333	2,006	2,623	2,566	2,356
Jarque-Bera	27,476**	22,259**	8,285*	27,030**	14,465**

Notiz: JB ist der Jarque-Bera-Test auf Normalverteilung.

* statistisch signifikant zum Niveau 5%; ** statistisch signifikant zum Niveau 1%.

Eine umgekehrte Situation wird bei der Volatilitat der Indizes beobachtet. Die Standardabweichung der Preise in (DM/Eur) ist hoher als die Standardabweichung der Preise in eigener Wahrung mit Ausnahme des franzosischen Marktes. Dies lasst sich so interpretieren, dass eine Berucksichtigung des Wechselkursrisikos durch die Konvertierung sich in der erhohten Volatilitat der Preise widerspiegelt.

Die Wirkung der Umrechnung der Aktienindizes in eine gemeinsame Wahrung lasst sich wie folgt zusammenfassen: Keine eindeutige anderungsrichtung des Mittelwertes der Aktienindizes lasst sich erkennen. Allerdings weisen fast alle Aktienindizes eine hohere Volatilitat nach der Konvertierung auf.

In den folgenden Tabellen (3.3) und (3.4) sind explorative Statistiken zu den realen Renditen der Aktienindizes in (DM/Euro) und in der eigenen Wahrung zusammengefasst.

Tabelle 3.3: Explorative Statistiken der realen Renditen der Aktienindizes in (DM/Euro)

	USA	Japan	France	Germany	UK
Mittelwert	0,001	0,003	0,001	0,001	0,001
Std. Abweichung	0,044	0,051	0,058	0,043	0,053
Schiefe	-0,661	-0,227	-0,510	-0,882	0,022
Wolbung	4,567	3,473	5,761	6,439	8,240
Jarque-Bera	84,541**	8,640**	174,360**	375,582**	552,573**
$Q(12)r^2$	25,090**	37,776**	41,17**	29,167**	41,190**

Notiz: JB ist der Jarque-Bera-Test auf Normalverteilung.

$Q(12)r^2$ ist die Ljung-Box Q-Statistik fur Autokorrelation der quadrierten Renditen bis Lag 12.

* statistisch signifikant zum Niveau 5%; ** statistisch signifikant zum Niveau 1%.

Tabelle 3.4: Explorative Statistiken der realen Renditen der Aktienindizes in eigener Wahrung

	USA	Japan	France	Germany	UK
Mittelwert	0,002	0,002	0,001	0,001	0,001
Std. Abweichung	0,036	0,043	0,057	0,043	0,047
Schiefe	-0,707	-0,275	-0,543	-0,882	0,184
Wolbung	4,774	3,681	6,017	6,944	11,154
Jarque-Bera	103,526**	15,413**	206,843**	375,582**	1340,855**
$Q(12)r^2$	28,034**	51,219**	38,93**	29,167**	48,896**

Notiz: JB ist der Jarque-Bera-Test auf Normalverteilung.

$Q(12)r^2$ ist die Ljung-Box Q-Statistik fur Autokorrelation der quadrierten Renditen bis Lag 12.

* statistisch signifikant zum Niveau 5%; ** statistisch signifikant zum Niveau 1%.

Der japanische Markt bringt die hochste durchschnittliche monatliche Rendite in (DM/Euro) von 0,3% pro Monat uber die Gesamtperiode von 1964 bis 2004. Das Wachstum der japanischen Wirtschaft, oder der sogenannte Aufholprozess, schlagt sich in den Aktienmarktdaten nieder. Der Einfluss dieses Prozesses ist sogar wesentlich starker als die Rezession der Wirtschaft nach der Asienkrise in den neunziger Jahren. Die anderen Aktienmarkte liefern identische Renditen von 0,1% pro Monat.

Der amerikanische Markt verliert durch die Konvertierung in (DM/Euro) 0,1% Punkte in der durchschnittlichen realen Rendite. Der Wert von 0,2% sinkt auf 0,1% pro Monat. Dies kann durch die tendenzielle Aufwertung der deutschen Mark in den letzten 35 Jahren⁷ und sich durch die Schwankung des Euro erklart werden. Die tendenzielle Abwertung der deutschen Mark zum japanischen Yen kann den Zuwachs der realen Rendite von 0,2% auf 0,3% pro Monat erklaren. Die Bewegung der Wechselkurse Frank/DM, Pfund Sterling/DM und Pfund Sterling/Euro bringt keine wesentliche Veranderung der realen Renditen der europaischen Markte. Die Renditen liegen bei 0,1% pro Monat im Durchschnitt wahrend der letzten 40 Jahre. Die Renditen der europaischen Aktienmarkte sind von der Konvertierung nicht beeinflusst. Analog zu den Ergebnisse fur die Niveaudaten ist die Standardabweichung der realen Renditen in der europaischen Wahrung hoher als die Standardabweichung der Renditen in der eigenen Wahrung. Die hochste Variation der Renditen bringt der franzosische Markt in Hohe von 0,058 pro Monat, gefolgt von Grobritannien mit 0,053 und Japan mit 0,051 pro Monat. Die hohe Volatilitat des Aktienmarktes spiegelt nicht zuletzt die entsprechenden Schwankungen in der Wechselkursbewegung wider. Die niedrigste Standardabweichung der realen Rendite aus der deutschen Sicht bringt der deutschen Aktienindex, weil hier keine direkte Wirkung⁸ des Wechselkurses vorhanden ist. Die Annahme der Normal-

⁷bis 1999.

⁸Da der CDAX auslandische Unternehmen beinhaltet, die bei der Deutschen Borse AG zugelassen

verteilung der Renditen ist eine notwendige Voraussetzung für die Anwendung vieler statistischer Verfahren bei der Untersuchung der Aktienmärkte. Der Jarque-Bera Test lehnt die Hypothese der Normalverteilung für alle Aktienmarktrenditen zum Signifikanzniveau 1% ab⁹. Gründe für dieses Ergebnis sind: Zum Ersten sind die Renditen durch eine linksschiefe Verteilungsdichte in allen Zeitreihen gekennzeichnet. Eine Ausnahme stellt der britische Index dar. Dies impliziert, dass die Verteilung der Zeitreihen um den Mittelwert einen langen linken Teil hat. Die negative Schiefe der Renditeverteilungsdichte von G5-Ländern wurde schon von Fraser und Power (1997), Shawsy und andere (1997), Scheicher (2001), Ahlgren und Antell (2002), Syriopoulos (2004) festgestellt. Die Ergebnisse sind für Tages-, Monats- und Quartalsdaten sowie für unterschiedliche Zeiträume robust.

Zum Zweiten sind die Renditen durch den sogenannten Leptokurtosis gekennzeichnet, vgl. Seo (1999), Scheicher (2001). Das bedeutet, dass die Verteilungsdichte der Renditen an den Rändern über der einer Normalverteilung liegt. Die Wölbung der Renditeverteilung liegt zwischen 3,473 beim japanischen und 8,240 beim britischen Markt, und ist wesentlich höher als die Wölbung der Normalverteilung. Die 'dicken' Ränder der Renditeverteilung implizieren, dass extreme Kursbewegungen wahrscheinlicher sind, als man unter der Annahme einer Normalverteilung erwarten würde.

Der dritte stilisierte Fakt ist die zeitvariierende Volatilität der Renditen. Die Heteroskedastizität der Renditen wird mit Hilfe der Ljung-Box Q(12)-Teststatistik für Autokorrelation der quadrierten Renditewerte nachgewiesen¹⁰. Es gibt eine enge Verbindung zwischen Leptokurtosis und Heteroskedastizität. Die Zeitreihe mit einer zeitabhängigen Varianz generiert eine Verteilung mit 'dicken' Rändern, vgl. Scheicher (2001).

In einigen Fällen kann eine Normalverteilung der Daten durch eine Box-Cox-Transformation der Datenreihen $X^{(\lambda)} = (X^\lambda - 1)/\lambda$ erreicht werden. Allerdings bringt die Transformation in diesem Fall keine wünschenswerten Ergebnisse, weil nicht alle Werte der Renditen r_t positiv sind¹¹. Eine weitere Möglichkeit, um diese Art der Nicht-Normalität in den Daten zu berücksichtigen, ist eine Modellierung der Heteroskedastizität mit Hilfe von ARCH- (Autoregressive Conditional Heteroskedastic) und GARCH- (Generalized Autoregressive Conditional Heteroskedastic) Prozessen.

sind, entsteht eine indirekte Wirkung des Devisenmarktes auf den Index.

⁹Die Bezeichnung JB steht in den Tabellen für den Jarque-Bera-Test auf Normalverteilung. H_0 : Daten sind mit Normalverteilung vereinbar; H_1 : Daten sind mit Normalverteilung nicht vereinbar.

¹⁰Die Bezeichnung $Q(12)r^2$ steht für die Ljung-Box Q-Statistik für Autokorrelation der quadrierten Renditen bis Lag 12. H_0 : 'keine Autokorrelation zwischen den quadrierten Renditen vorhanden' wird gegen H_1 : 'es besteht Autokorrelation mindestens zu einem quadrierten Renditewert mit dem Lag zwischen 1 und 12' getestet.

¹¹Um die Box-Cox-Transformation für alle λ definieren zu können, müssen alle Werte der Zeitreihe X_t positiv sein. Die Box-Cox-Transformation kann zu normalverteilten Daten im Fall $0 < \lambda < 1$ führen, vgl. W. Green (2003).

3.1.2 Korrelationsanalyse

Die Korrelationsanalyse der Aktienmärkte wird oft bei Untersuchungen zur Diversifikation der Märkte oder bei einer Überprüfung der MEH durchgeführt. Die Korrelationsanalyse stellt eine relativ einfache Methode für beide Fragestellungen dar. Die wichtigsten Arbeiten auf dem Bereich wurden unter anderen von Markovitz (1959), Solnik (1974), Solnik (1988), Fama und French (1993), King und andere (1994), Longin und Solnik (1995), Chou und Ng (1995), Fama und French (2001), Pan, Liu und Roth (2001), Elton und Gruber (2003), Reilly und Brown (2003) durchgeführt.

Als erstes wurde festgestellt, dass die Korrelation zweier internationaler Aktienmärkte von der Länge der Beobachtungsperiode abhängt. Die *'kurzfristige'* ist meist höher als die *'langfristige'* Korrelation¹². Longin und Solnik (1995) stellen zusätzlich fest, dass die Kovarianz der *'langfristigen'* internationalen Aktienmarktrenditen über die Zeit instabil ist. Solnik und andere (1996) zeigen unter anderem, dass auf den verschiedenen Finanzmärkten die Korrelation in den Perioden mit hoher Volatilität steigt¹³. Seit der Arbeit von Solnik wurde nach Ursachen gesucht, die die Instabilität der Korrelationsmatrizen verursachen könnten. Unter anderem wurden in den letzten Jahren Faktoren wie Zeithorizont (vgl. Pan und anderen (2001)), ex-post Daten (vgl. Shawsky und andere (1997), Fifield und anderen (1999) und (2002)), Volatilität des Marktes (vgl. Engle und Ng (1993), King und anderen (1994), Fraser und Power (1997), Longin und Solnik (2001)) untersucht.

In den Tabellen (3.5) und (3.6) sind die Korrelationsmatrizen der Aktienindizes über die Gesamtperiode 1964-2004 in eigener und gemeinsamer Währung dargestellt. Alle Werte der Matrizen sind positiv.

Tabelle 3.5: Korrelationsmatrix der Aktienindizes in (DM/Euro)

	USA	Japan	Frankreich	Deutschland	UK
USA	1,000				
Japan	0,293	1,000			
France	0,907	0,465	1,000		
Germany	0,837	0,660	0,909	1,000	
UK	0,907	0,558	0,909	0,898	1,000

Für die Aktienindizes in (DM/Euro) besteht eine sehr hohe Korrelation in der Gruppe der europäischen Länder. Ebenfalls sind sehr hohe Werte der Korrelationskoeffizienten zwischen dem US-amerikanischen und den europäischen Aktienmärkten beob-

¹²Bei der Berechnung der *'kurzfristigen'* Korrelation werden die Zeiträume von einem bis ein Paar Tagen in Betracht gezogen.

¹³"...correlation increases in periods of high market volatility".

Tabelle 3.6: Korrelationsmatrix der Aktienindizes in eigener Wahrung

	USA	Japan	Frankreich	Deutschland	UK
USA	1,000				
Japan	0,384	1,000			
France	0,948	0,494	1,000		
Germany	0,902	0,609	0,930	1,000	
UK	0,906	0,581	0,906	0,915	1,000

achtbar, von 0,907 fur USA/UK und USA/Frankreich bis 0,837 bei USA/Deutschland. Eine Ausnahme stellt der japanische Markt dar, dessen Korrelationskoeffizient mit USA wesentlich niedriger als bei den anderen Markten der Gruppe ist.

Die paarweise Korrelation der realen Aktienindizes in eigener Wahrung ist in den meisten Fallen um ca. 0,05 Punkte hoher als die vorherigen Werte. Die eventuell entgegengesetzte Bewegung der Wechselkurse zueinander verringert die entsprechenden Koeffizienten. Allgemein scheinen die Korrelationskoeffizienten eher zu hoch als erwartet zu sein. Einerseits deutet die hohe Korrelation der Aktienindizes auf eine signifikante lineare Verbindung der Aktienmarkte hin. Andererseits kann sie jedoch auch eine Scheinkorrelation aufgrund der Instationaritat der Zeitreihen sein.

Die Korrelationsmatrizen der Renditen in (DM/Euro) und in eigener Wahrung sind durch die Tabellen (3.7) und (3.8) dargestellt.

Tabelle 3.7: Korrelationsmatrix der Aktienrenditen in (DM/Euro)

	USA	Japan	Frankreich	Deutschland	UK
USA	1,000				
Japan	0,409	1,000			
Frankreich	0,438	0,318	1,000		
Deutschland	0,505	0,329	0,484	1,000	
UK	0,602	0,385	0,378	0,417	1,000

Die Korrelation der Aktienrenditen ist ebenfalls immer positiv, aber wesentlich niedriger als die Korrelation der Aktienindizes. Der hochste Korrelationswert liegt bei 0,602 und reprasentiert lineare Abhangigkeit des US-amerikanischen und britischen Marktes dar. Die Differenz der Korrelation der japanischen Rendite mit der Aktienrenditen den anderen Lander ist nicht so deutlich wie im Fall der Niveaugroen. Das Paar Japan/Frankreich liefert allerdings den kleinsten Korrelationswert von 0,318.

Die Korrelationskoeffizienten der Renditen in eigener Wahrung sind hoher als die

Tabelle 3.8: Korrelationsmatrix der Aktienrenditen in eigener Wahrung

	USA	Japan	Frankreich	Deutschland	UK
USA	1,000				
Japan	0,376	1,000			
Frankreich	0,452	0,314	1,000		
Deutschland	0,491	0,346	0,495	1,000	
UK	0,577	0,360	0,370	0,433	1,000

entsprechende Koeffizienten in gemeinsamer Wahrung nur bei europaischen Markten. Eine entgegengesetzte Wirkung hat der Wechselkurse auf die Koeffizienten europaischer mit dem amerikanischen Markt. Die Koeffizienten sind kleiner, wenn die Daten nicht konvertiert sind. Eine langfristige Aufwertung sowohl der deutschen Mark als auch des franzosischen Franc und des britischen Pfund gegenuber dem amerikanischen Dollar schlagt sich in der kleineren Korrelation dieser Markte zu USA nieder.

Die Wirkung der Konvertierung der Daten lasst sich somit wie folgt zusammenfassen. Die Reaktion der Mittelwerte der Indizes auf eine Konvertierung in die gemeinsame Wahrung lasst sich nicht eindeutig interpretieren. Fur die Renditen sinkt wie erwartet der Gewinn aus der auslandischen Investition aus Sicht der inlandischen Anleger, wenn die auslandische Wahrung aufgewertet wird. Bei der Abwertung der auslandischen Wahrung beobachtet man eine entgegengesetzte Reaktion. Die Volatilitat der Preise und der Renditen steigt bei der Berucksichtigung des Wechselkursrisikos. Die Korrelationskoeffizienten reagieren auch nicht eindeutig auf die Konvertierung.

Eine Korrelationsanalyse der Renditen ist das zentrale Element bei Standard Risiko/Rendite-Portfolio-Modellen. Aufgrund der kurzfristigen Ausrichtung der Korrelationsmethode ist sie jedoch weniger geeignet, um langfristige Beziehungen festzustellen. Die Methode ist genauso wenig fur instationare Zeitreihen geeignet. Von Granger und Engle (1987) wurde die Kointegrationsmethode entwickelt, die fur die Untersuchung der langfristigen Verbindungen und der gemeinsamen Trends der internationalen Aktienmarkte besser geeignet ist. Die Methode liefert eine Moglichkeit zur Modellierung sowohl kurzfristiger als auch langfristiger Beziehungen zwischen den Aktienpreisen in mehrdimensionalen Systemen.

Die Korrelationsanalyse und die Kointegrationsmethode sind von verschiedenen Konzepten abgeleitet. Eine hohe Korrelation der Renditen impliziert nicht unbedingt eine hohe Kointegration in Preisen. Die Korrelation reflektiert gemeinsame Bewegungen in den Renditen, welche uber die Zeit zur Instabilitat neigen. An sich ist die Korrelation eine kurzfristige Messung und die korrelationsbasierenden Portfolio-Strategien schlagen eine kurzfristige Rebalancierung von Investitionstiteln vor. Dagegen bestimmt die Kointegration die langfristigen Beziehungen, welche sich uber Perioden mit einer niedrigeren

Korrelation erstrecken. Ein auf kointegrierten Finanztiteln basierendes Portfoliomanagement kann für langfristige Investitionen effektiver sein als das korrelationsbasierende Management, aufgrund seiner schlechten langfristigen Performanz. Das Kointegrationskonzept bereichert die traditionellen Portfoliomodelle durch eine Einführung der mehrdimensionalen Analyse der Preisbewegungen.

3.2 Kointegrationsanalyse der Aktienmärkte

Eine der Voraussetzungen der Kointegrationsanalyse ist der identische Integrationsgrad der Zeitreihen. Mit einer Feststellung des Integrationsgrades wird bei der empirischen Analyse begonnen. Dann folgt die Analyse der paarweisen Abhängigkeit in der Gruppe der G5-Länder. Dies beinhaltet die paarweisen Kointegrationstests und Granger-Kausalitätstests der Renditen. Im weiteren wird eine mehrdimensionale Kointegrationsanalyse durchgeführt, mit einer anschließenden Schätzung des Vektorkorrekturmodells, einer Impuls-Antwort-Analyse und einer Varianz-Dekomposition für den deutschen Aktienmarkt. Alle Tests wurden für die realen Daten sowohl in DM/Euro als auch in eigener Währung durchgeführt. Die Ergebnisse unterscheiden sich im Wesentlichen nicht, deswegen sind hier nur die Tabellen für die Werte in gemeinsamer Währung dargestellt.

3.2.1 Einheitswurzel-Tests

Um die sogenannte Scheinkorrelation zwischen den Märkten auszuschließen, werden die Zeitreihen auf Stationarität überprüft. Die Feststellung der Integrationsordnung einer Zeitreihe ist der erste Schritt für die weiteren Untersuchungen der gemeinsamen Bewegung der Aktienmärkte mit Hilfe der Kointegrationsmethode. Zur Überprüfung der Stationaritätseigenschaft der Zeitreihen wurden hier die im vorherigen Kapitel beschriebenen Einheitswurzel-Tests angewendet. Darunter sind der modifizierte PP-Test von Perron und Ng (1996), der DF-GLS-t-Test von Elliot, Rothenberg und Stock (1996) und der modifizierte DF-Test mit einer GARCH(1,1)-Schätzung der Varianz der Störterme.

Der datengenerierende Prozess ist im allgemeinen unbekannt, deswegen muss desweiteren entschieden werden, welche deterministischen Terme in die Schätzgleichung aufgenommen werden. Es wurden zwei Modelle des datengenerierenden Prozesses untersucht, sowohl mit einem Achsenabschnitt als auch mit einem deterministischen Trend in den Datenreihen.

Die Entscheidung über die Laglänge in dem Modell kann das Testergebnis maßgeblich beeinflussen. Deshalb wurde in dieser Arbeit jeweils Modelle mit einer hoher und einer niedriger Lag-Ordnung getestet. Für die Tests wurden die Lags alternativ mit

Hilfe des Akaike-Informations-Kriteriums (AIC) und des 'general-to-specific'-Ansatzes von Hall bestimmt. Die Nullhypothese in jedem Test: 'Die Zeitreihe beinhaltet eine Einheitswurzel' wird gegen die Alternativhypothese: 'Die Zeitreihe ist stationär' getestet (d.h. $I(1)$ wird gegen $I(0)$ Prozess getestet). Die Hypothese wird abgelehnt, wenn die Teststatistik kleiner als der kritische Wert ist.

Eine niedrige Mächtigkeit der Einheitswurzel-Tests begründet eine Anwendung modifizierten Verfahren. Aufgrund der Vielfältigkeit dieser Modifizierungen wurde eine große Anzahl der Tests durchgeführt. Die Tests sind so ausgewählt, dass sie die Besonderheiten der Daten berücksichtigen. Der modifizierte DF- und PP-Tests wurden aufgrund der festgestellten Heteroskedastizität in den Zeitreihen durchgeführt. Das Ergebnis lässt sich wie folgt zusammenfassen: Alle Aktienindizes haben die Integrationsordnung Eins. Eine detaillierte Beschreibung der Testergebnisse ist in den folgenden Tabellen dargestellt.

Modifizierter Ng-Perron-Test

Tabelle 3.9: Modifizierter Ng-Perron-Test für Aktienindizes und Renditen. Modell mit einem Achsenabschnitt.

	Aktienindex		Rendite	
	MZ_ρ (Hall)	MZ_ρ (AIC)	MZ_ρ (Hall)	MZ_ρ (AIC)
USA	-2,440 (19)	-2,428 (1)	-145,666 (6)	-224,578 (0)
Japan	-0,215 (1)	-0,215 (1)	-29,815 (8)	-29,815 (8)
Frankreich	-2,149 (20)	-2,785 (3)	-216,333 (1)	-297,598 (0)
Deutschland	-4,139 (19)	-3,867 (1)	-45,953 (9)	-45,953 (9)
UK	-3,414 (19)	-3,692 (2)	-159,160 (11)	-269,935 (1)

Notiz: Der kritische Wert für die MZ_ρ -Statistik ist: -13,8 für das Signifikanzniveau 1%; -8,1 für das Signifikanzniveau 5%; -5,7 für das Signifikanzniveau 10%.

Die Ergebnisse des modifizierten Ng-Perron-Tests sind in den beiden Tabellen (3.9) und (3.10) zusammengefasst. In den Spalten stehen die entsprechenden Z_ρ -Statistiken für die Aktienindizes und ihre Renditen. Die asymptotischen kritischen Werte sind aus Ng-Perron (1996) genommen. Die Anzahl der Lags (k) wurde nach der 'general-to-specific'-Regel von Hall und nach dem modifizierten AIC selektiert. Die Anzahl der Verzögerungen liegt zwischen 20 und 1 bei den Aktienindizes und zwischen 11 und 0 bei deren Differenzen. Weder die Anzahl der Lags noch der unterstellte datengenerierende Prozess hat einen Einfluss auf das Ergebnis. Bei keiner Zeitreihe konnte die Nullhypothese einer Einheitswurzel zum Signifikanzniveau 1% abgelehnt werden. Die Nullhypothese konnte für die differenzierten Zeitreihen zum Signifikanzniveau 1% nicht beibehalten werden. Alle fünf Aktienindizes stellen instationäre Datenreihen der Integrationsordnung Eins dar.

Tabelle 3.10: Modifizierter Ng-Perron-Test für Aktienindizes und Renditen. Modell mit einem deterministischen Trend.

	Aktienindex		Rendite	
	MZ_ρ (Hall)	MZ_ρ (AIC)	MZ_ρ (Hall)	MZ_ρ (AIC)
USA	-2,668 (19)	-2,650 (1)	-220,630 (6)	-220,296 (0)
Japan	-5,100 (1)	-5,100 (1)	-108,459 (12)	-111,040 (3)
Frankreich	-2,312 (20)	-2,985 (3)	-80,563 (5)	-252,749 (0)
Deutschland	-5,708 (19)	-5,171 (1)	-299,904 (18)	-206,975 (0)
UK	-3,936 (19)	-4,201 (2)	-294,041 (19)	-270,495 (1)

Notiz: Der kritische Wert für die MZ_ρ -Statistik ist: -23,8 für das Signifikanzniveau 1%; -17,3 für das Signifikanzniveau 5%; -14,2 für das Signifikanzniveau 10%.

Elliott-Rothenberg-Stock DF-GLS Test

Tabelle 3.11: Elliott-Rothenberg-Stock DF-GLS-Test für Aktienindizes und Renditen. Modell mit einem Achsenabschnitt.

	Aktienindex		Rendite	
	t -Statistik	(k)	t -Statistik	(k)
USA	-0,996	(1)	-3,761	(24)
Japan	-0,788	(1)	-2,121	(22)
Frankreich	-1,082	(1)	-5,670	(10)
Deutschland	-1,310	(1)	-2,799	(23)
UK	-1,297	(2)	-15,522	(0)

Notiz: Der kritische Wert für t -Statistik ist: -2,57 für das Signifikanzniveau 1%; -1,94 für das Signifikanzniveau 5%; -1,62 für das Signifikanzniveau 10%.

Die Ergebnisse der Elliott-Rothenberg-Stock-Variante des DF-GLS-Tests sind in den beiden Tabellen (3.11) und (3.12) aufgeführt. In den Spalten stehen die entsprechenden t -Statistiken für die Aktienindizes und ihre Renditen. Es wurden beide datengenerierende Prozesse, mit einem deterministischen Trend und ohne Trend, in die Schätzgleichung einbezogen. Die kritischen Werte wurden aus der Arbeit von MacKinnon (1996) genommen. Die Darstellung der Tests ist in diesem Fall nur auf Selektion nach dem AIC beschränkt, weil die Ergebnisse von der Auswahl der Lags nicht beeinflusst sind. Die Nullhypothese einer Einheitswurzel konnte bei keinem Aktienindex zum Signifikanzniveau 1% abgelehnt werden. Bei den differenzierten Zeitreihen stellen die Tests die Stationarität der Renditen fest. Die Nullhypothese kann zum Niveau 1% nahezu bei keiner Rendite bestätigt. Dies gilt für beide Modelle des datengenerieren-

Tabelle 3.12: Elliott-Rothenberg-Stock DF-GLS-Test für Aktienindizes und Renditen. Modell mit einem Trend.

	Aktienindex		Rendite	
	t -Statistik	(k)	t -Statistik	(k)
USA	-1,070	(1)	-3,716	(24)
Japan	-1,435	(1)	-3,061	(24)
Frankreich	-1,121	(1)	-5,650	(10)
Deutschland	-1,602	(1)	-3,638	(23)
UK	-1,426	(2)	-15,551	(0)

Notiz: Der kritische Wert für t -Statistik ist: -3,480 für das Signifikanzniveau 1%; -2,890 für das Signifikanzniveau 5%; -2,570 für das Signifikanzniveau 10%.

den Prozesses.

ADF-Test mit GARCH(1,1)

Tabelle 3.13: Modifizierter Einheitswurzel-Test (ADF-Test) mit einem GARCH(1,1)-Prozess für Aktienindizes und Renditen

	Aktienindex		Rendite	
	$\hat{\sigma}^2/\hat{\rho}^2$	τ -Statistik	$\hat{\sigma}^2/\hat{\rho}^2$	τ -Statistik
USA	1,297	2,023	1,302	-13,933
Japan	1,152	1,918	1,141	-13,259
Frankreich	1,272	-0,246	1,401	-17,421
Deutschland	1,313	1,211	1,546	-13,354
UK	1,408	0,000	1,438	-12,919

In dieser Arbeit wurden die Ergebnisse auf den modifizierten ADF-Test mit einem GARCH(1,1)-Term und auf die Schätzung des Quotienten $\hat{\sigma}^2/\hat{\rho}^2$ beschränkt. Nach der Schätzung des Gewichtungsparmeters ρ konnten die genauen kritischen Werte ausgewählt werden. Allerdings zeigen die errechneten Werte bereits jetzt, dass auch dann keine qualitativ andere Ergebnisse zu erwarten sind. In der Tabelle (3.13) sind entsprechend die geschätzten Werte des Quotienten und die Teststatistik τ für Aktienindizes und deren Renditen zusammengefasst. Aus der Schätzung des Quotienten $\hat{\sigma}^2/\hat{\rho}^2$ folgt, dass bei allen Zeitreihen sowohl der Aktienindizes als auch bei deren Renditen die GARCH-Parameter sich von Null unterscheiden. Der Schätzer liegt zwischen 1,152 für den japanischen Aktienindex und 1,408 für den Aktienindex in Großbritannien. Der Wert des geschätzten Koeffizienten für die Renditen liegt etwas höher, zwischen 1,141 für Japan und 1,546 für Deutschland. Die kritischen Werte für ein Modell ohne einen

Achsenabschnitt und ohne einen linearen Trend für jedes ρ liegen zwischen $-2,33$ und $-2,55$ für das Signifikanzniveau 1%. Dadurch bestätigt der modifizierte ADF-Test in allen Fällen eine Präsenz des Einheitswurzel-Prozesses in den Niveaus der Daten und kann die Stationaritätshypothese für die ersten Differenzen nicht ablehnen. Damit liefert die verbesserte Methode die gleichen Ergebnisse wie der Standard ADF-Test.

Perron (1989) hat gezeigt, dass die Mächtigkeit eines Einheitswurzel-Tests bei der Ablehnung der Nullhypothese sinkt, wenn die Alternativhypothese (stationäre Zeitreihe mit einem Strukturbruch) richtig ist. Während der letzten vierzig Jahren könnten mehrere Ereignisse auf der nationalen und internationalen Ebene die strukturellen Änderungen eines Aktienmarktes verursachen. Dazu gehören zum Beispiel der erste und der zweite Ölpreis-Shocks, der Aktienmarktcrash im Oktober 1987, die asiatische und russische Finanzkrise in den späteren neunziger Jahren oder die Internetblase in 2000. Das falsche Beibehalten der Einheitswurzel-Hypothese in den Aktienpreisen bei den Standard Einheitswurzel-Tests kann die falsche Spezifikation des deterministischen Trends reflektieren. Beispielsweise hat Perron die Nullhypothese: *'Ein instationärer Prozess mit einem Drift und einem exogenen Strukturbruch'* gegen die Alternativhypothese: *'Ein stationärer Prozess um einen deterministischen Trend mit einem exogenen Strukturbruch'* mit dem Ergebnis getestet, dass die Tests zwischen den beiden Alternativen nicht unterscheiden können. Zusätzlich beinhaltet das Einschließen eines Strukturbruches in das Modell das genaue Wissen des Datums, wann der Strukturbruch stattgefunden hat. Zivot und Andrews (1992) endogenisierten den Prozess der Bestimmung eines Strukturbruches für die Modelle von Perron. Narayan und Smyth (2004) haben zwei Modelle von Zivot und Andrews für die G7-Länder und Australien mit beiden Arten des Einheitswurzel-Tests untersucht. Die Nullhypothese eines instationären Prozesses konnte bei keinem der acht Aktienmärkte abgelehnt werden. Die Untersuchung wurde mit dem Datensatz der OECD für die Periode von 1960 bis 2003 durchgeführt. Dementsprechend wird in dieser Arbeit auf eine Wiederholung der Tests dieser Art verzichtet, weil sie identische Ergebnisse wie die oben beschriebenen Einheitswurzel-Tests bringen würden.

Die Ergebnisse des modifizierten Ng-Perron-Tests, der Elliott-Rothenberg-Stock-Variante des DF-GLS-Tests und der modifizierten ADF-Tests sind in Bezug auf die Datenfrequenz (Quartals- oder Monatswerte), die Anzahl der Verzögerungen und auf den datengenerierenden Prozess robust. Bei dem Signifikanzniveau 1 % erkennen diese Tests in allen Fällen eine Präsenz des Einheitswurzel-Prozesses in den Niveaus der Daten. Die Nullhypothese kann nicht abgelehnt werden. Die Stationaritätshypothese für die ersten Differenzen wird fast bei allen Tests zum Signifikanzniveau von 1% angenommen.

3.2.2 Paarweise Kointegration versus Granger-Kausalität

Für die weitere Untersuchung der Aktienmärkte bietet sich zunächst die bivariate Analyse des Systems aus fünf Aktienindizes an. Sie wird zum Zweck einer Vergleichsanalyse sowohl mit den existierenden Forschungsarbeiten als auch mit der multivariaten Analyse durchgeführt. Die paarweise Untersuchung beinhaltet die zweidimensionale Kointegrationsanalyse und den Granger-Kausalität-Test. Beide Aspekte, das langfristige Gleichgewicht und die kurzfristige Dynamik, werden dabei berücksichtigt.

Die Voraussetzungen für eine Untersuchung der Aktienmärkte mit Hilfe der Kointegrationsmethode sind gegeben. Alle Aktienindizes haben den Integrationsgrad Eins. Die Existenz langfristiger Gleichgewichte in den Paaren der Aktienmärkte wird mittels des Johansen-Kointegrationstests überprüft. Bei der Schätzung der Kointegrationsbeziehung wurde ein Modell des datengenerierenden Prozesses mit einem deterministischen Trend in den Zeitreihen der Indizes unterstellt. Es werden beim Test vier Lags im autoregressiven Term angenommen. Die Ergebnisse der Tests sind in der Tabelle (3.14) zusammengefasst.

Tabelle 3.14: Paarweise Kointegrationstest für Aktienindizes. Modell mit einem deterministischen Trend in den Daten

	USA	Japan	Frankreich	Deutschland	UK
(A) λ_{trace} -Statistik					
USA	-				
Japan	6,92	-			
Frankreich	9,79	11,07	-		
Deutschland	7,25	8,30	15,69	-	
UK	9,46	6,64	21,02	12,05	-
(B) λ_{max} -Statistik					
USA	-				
Japan	5,55	-			
Frankreich	9,04	7,56	-		
Deutschland	6,32	5,32	12,69	-	
UK	7,99	4,54	18,64	10,82	-

Notiz: Der kritische Wert für λ_{trace} -Statistik ist: 20,04 für das Signifikanzniveau 1%; 15,41 für das Signifikanzniveau 5%.

Der kritische Wert für λ_{max} -Statistik ist: 18,63 für das Signifikanzniveau 1%; 14,07 für das Signifikanzniveau 5%.

Die kritischen Werte sind von Osterwald-Lenum (1992). Anzahl der Lags ist 4.

Eine einzige gleichgewichtige Verbindung zwischen dem französischen und dem britischen Aktienindex wird mittels der beiden Johansen-Tests festgestellt. Die Null-

hypothese: *'keine Kointegration'* konnte bei den beiden λ_{max} - und λ_{trace} -Tests zum Signifikanzniveau 1% nicht beibehalten werden. Die Nullhypothese: *'Existenz mindestens einer Kointegrationsbeziehung'* kann zu dem Signifikanzniveau nicht abgelehnt werden. Eine zweite Kointegrationsbeziehung wird zwischen dem französischen und dem deutschen Aktienindex mittels des λ_{trace} -Tests festgestellt. Die Nullhypothese: *'keine Kointegration'* wird zum Signifikanzniveau 5% abgelehnt.

Die oft gefundene Kointegration zwischen dem amerikanischen Index und den europäischen Aktienindizes wird in diesem Fall nicht bestätigt¹⁴. Die MEH kann eine Erklärung für die schwache Rolle des amerikanischen Marktes für die europäischen Aktienmärkte liefern. Es könnte ein Einfluss der Änderungen und Shocks auf dem US-Aktienmarkt auf Europa existieren, allerdings ist dieser Effekt relativ kurzfristig. Dementsprechend spiegelt er sich nicht in einer langfristige Beziehung über vierzig Jahre wider. Die früheren Untersuchungen konzentrierten sich im wesentlichen auf diesen Einfluss in einer *Nach-Crash-Periode*. Beispielsweise konnte der Zusammenhang zwischen USA und Europa in der Aufschwungphase nach dem Crash von 1987 festgestellt werden. Wenn man allerdings die Periode mit zwei *'bubbles'* betrachtet, den Aktien-crash von 1987 und die *'Internetblase'* von 2000, ist die Verbindung nicht mehr so deutlich.

Überraschend ist eine schwache Präsenz der Kointegration im europäischen Raum. Es wurden zwar zwei Verbindungen zwischen Frankreich und UK sowie Deutschland und Frankreich festgestellt, allerdings konnte keinem der Märkte eine führende Rolle zugewiesen werden. Eine Erklärung der schwachen Abhängigkeit könnte die Konvertierung der Indizes in die gemeinsame Währung sein. Dadurch wurde der Abhängigkeitseffekt möglicherweise schon berücksichtigt. Die europäischen Aktienmärkte könnten während kurzer Perioden voneinander abhängig sein. Allerdings sind die Investoren eventuell im wesentlichen auf das Geschehen im eigenen Land konzentriert. Sie informieren sich ausschließlich über einheimische Unternehmen. Der Informationsverarbeitungsprozess im eigenen Land findet statt; die Information über ausländische Unternehmen wird nicht berücksichtigt. Dementsprechend könnte keine langfristigen Gleichgewichtsbeziehungen in der europäischen Region festgestellt werden.

Zusätzlich konnte aus den Ergebnissen hergeleitet werden, dass der französische Aktienmarkt von den anderen europäischen Märkten eventuell stärker beeinflusst wird. Die Größe des Marktes, gemessen an dem Handelsvolumen, kann keine Erklärung dafür liefern, weil diesbezüglich der deutsche Markt den kleinsten Aktienmarkt in der Gruppe darstellt. Eine Erklärung könnte in den real wirtschaftlichen Verflechtungen der französischen mit den europäischen Firmen liegen. Für die Bestätigung der Hypothese ist eine ausführlichere Untersuchung der internationalen Verbindungen notwendig.

Der Granger-Kausalitätstest der Aktienmarktrenditen erweitert die paarweise Ana-

¹⁴Vgl. zum Beispiel Kasa (1992), Malliaris und Urrutia (1992), Kwan und anderen (1995), Gerrits und Yüce (1999), Scheicher (2001).

Tabelle 3.15: Granger-Kausalität-Test für Aktienrenditen. F-Statistik

Nullhypothese	Lag Spezifikation	
	p=1	p=4
(a) Japan nicht kausal für USA	0,11	1,29
USA nicht kausal für Japan	3,04	1,42
(b) Frankreich nicht kausal für USA	1,80	1,03
USA nicht kausal für Frankreich	9,29**	2,90*
(c) Deutschland nicht kausal für USA	1,33	1,59
USA nicht kausal für Deutschland	11,28**	3,14**
(d) UK nicht kausal für USA	0,51	0,51
USA nicht kausal für UK	1,02	1,02
(e) Frankreich nicht kausal für Japan	2,91	2,03
Japan nicht kausal für Frankreich	4,10*	1,71
(f) Deutschland nicht kausal für Japan	0,68	0,91
Japan nicht kausal für Deutschland	1,19	1,14
(g) UK nicht kausal für Japan	2,69	1,47
Japan nicht kausal für UK	1,13	0,53
(h) Deutschland nicht kausal für Frankreich	3,60	2,57*
Frankreich nicht kausal für Deutschland	15,67**	4,27**
(i) UK nicht kausal für Frankreich	4,12*	3,27**
Frankreich nicht kausal für UK	16,48**	4,27**
(j) Deutschland nicht kausal für UK	2,35	0,99
UK nicht kausal für Deutschland	1,72	1,11

Notiz: F-Statistik für den Test der Nullhypothese: *Die Rendite des Landes A ist nicht kausal für die Rendite des Landes B.*

* statistisch signifikant zum Niveau 5%; ** statistisch signifikant zum Niveau 1%.

lyse der Aktienmärkte. Bei diesem Test wird allerdings vorerst von einer zeitlichen und keiner kausalen Abhängigkeit ausgegangen. Die zeitliche Kausalität ist zwar eine notwendige aber nicht hinreichende Bedingung für eine funktionale Kausalität. Der Test wurde unter Berücksichtigung von einem und vier Lags durchgeführt, weil die Testergebnisse stark von der Lag-Spezifikation abhängen. Die Ergebnisse sind in der Tabelle (3.15) dargestellt.

Drei wichtige Feststellungen müssen hier erwähnt werden. Erstens konnte ein wesentlicher Unterschied zum Ergebnis der Kointegrationstests festgestellt werden. Eine führende Rolle des amerikanischen Marktes wird mit dem Granger-Kausalitätstest bestätigt. Zwar werden nicht alle europäischen Aktienmärkte laut Test vom amerikanischen Markt beeinflusst, dennoch haben Änderungen auf dem US-Aktienmarkt einen signifikanten Einfluss auf die deutschen und französischen Märkte. Beide Nullhypothesen: *'USA ist nicht kausal für Frankreich'* und *'USA ist nicht kausal für Deutschland'*

werden zum Signifikanzniveau von 1% (einmal zum Niveau 5%) für beide Lagspezifikationen abgelehnt. Eine Ausnahme stellt der britische Aktienmarkt dar. Die Nullhypothese: *'USA ist nicht kausal für UK'* kann dagegen beibehalten werden. Dies ist mit den Ergebnissen von Taylor und Tonks (1989) vereinbar, widerspricht aber dem Resultat von Kwan und anderen (1995) oder Gerrits und Yüce(1999).

Zweitens bestätigt der Granger-Kausalitätstest die Verbindung zwischen Frankreich und UK, sie wird sogar als zweiseitig identifiziert. In dem Modell mit vier Monats-Lags wird die Kausalität zwischen Frankreich und Deutschland ebenfalls als zweiseitig zum Signifikanzniveau 5% festgestellt. Die Nullhypothese, dass Frankreich Deutschland zeitlich nicht beeinflusst, wird zum Signifikanzniveau von 1% abgelehnt. Das Ergebnis ist bezüglich der Lag-Spezifikation robust.

Drittens besteht analog zum Kointegrationstest für den japanischen Aktienmarkt keine kausale Beziehung weder mit dem amerikanischen noch mit den europäischen Märkten. Allerdings kann die Nullhypothese: *'Japan nicht kausal für Frankreich'* in dem Modell mit einem Lag nicht beibehalten werden. Somit beeinflusst zeitlich der japanische den französischen Aktienmarkt.

Allgemein wurde mittels der Granger-Kausalitätstests deutlich mehr Abhängigkeit der Aktienmärkte voneinander festgestellt. Dies widerspricht nicht den Ergebnissen der Kointegrationsanalyse, sondern bestätigt die Unterschiede, die hinter den beiden Konzepten stehen. Es existieren kurzfristige zeitlichen Verbindungen zwischen verschiedenen Aktienmärkten, die mittels Granger-Kausalitätstests festgestellt werden konnten. Sie verschwinden allerdings aus der langfristigen Sicht der bivariaten Kointegrationsanalyse.

3.2.3 Multivariate Kointegration

Bivariate Analysen zur Aufdeckung multivariater Zusammenhänge sind problematisch und im besten Fall ein erster Schritt der empirischen Analyse. Speziell kann die Analyse scheinbare Beziehungen feststellen, weil bestimmte Einflussfaktoren nicht berücksichtigt werden sind, vgl. Lütkepohl (1982). Deswegen wird im weiteren eine umfassende Untersuchung des Systems bestehend aus fünf Aktienmärkten durchgeführt. Die mehrdimensionale Kointegrationsanalyse wird ähnlich zum zweidimensionalen Fall gestaltet. Als erstes wird der Kointegrationsrang des Gesamtsystems mit Hilfe der Johansen-Tests geschätzt. Der zweite Schritt beinhaltet die Schätzung der Kointegrationsbeziehung oder des Kointegrationsraumes nach Johansen-Verfahren. Eine Fehler-Korrektur-Darstellung des geschätzten Systems ist hilfreich für die Analyse der kurzfristigen Anpassungen, die im dritten Schritt durchgeführt wird. Zusätzliche Erkenntnisse zu kurzfristigen Reaktionen der Märkte liefern die Impuls-Antwort-Funktionen und die Varianz-Dekomposition-Analyse. Die Gesamtuntersuchung wird aus der Sicht der deutschen Investoren durchgeführt.

Die Rangschätzung des Gesamtsystems

Um den Einfluss eines möglichen deterministischen Trends in den Daten zu berücksichtigen, wurden drei unterschiedliche VECM für die Schätzung der Anzahl der Kointegrationsbeziehungen verwendet. Das erste Modell hat keinen deterministischen Trend und keinen Achsenabschnitt weder in den Datenreihen selbst noch in dem zu schätzenden Kointegrationsraum:

$$\Delta X_t = \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + \alpha \beta' X_{t-1} + \theta D_t + \varepsilon_t \quad (3.2)$$

Das zweite Modell beinhaltet einen linearen deterministischen Trend in den Daten und einen Achsenabschnitt im Kointegrationsraum:

$$\Delta X_t = \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + \alpha(\beta' X_{t-1} + \mu_0) + \alpha_{\perp} \mu_1 + \theta D_t + \varepsilon_t \quad (3.3)$$

Das dritte Modell hat einen linearen deterministischen Trend in den beiden Komponenten, sowohl in den Datenreihen als auch im Kointegrationsraum:

$$\Delta X_t = \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + \alpha(\beta' X_{t-1} + \mu_0 + \gamma_0 t) + \alpha_{\perp} \mu_1 + \theta D_t + \varepsilon_t \quad (3.4)$$

Die Kointegrationsanalyse wurde sowohl für die Gesamtperiode von Juli 1964 bis Juli 2004, als auch für Subperioden mit unterschiedlicher Anzahl von Lags (von 1 bis 24), mit unterschiedlicher Anzahl von zentrierten Dummi-Variablen (von 1 bis 12) und für Monats- und Quartalsdaten durchgeführt. Die Monats- oder Quartalsdaten liefern trotz unterschiedlicher Frequenz fast identische Ergebnisse. Deswegen wird die Darstellung der Ergebnisse auf Monatsdaten beschränkt.

Eine ausführliche Untersuchung des Einflusses der Zeitperiodenlänge liefert folgende Resultate. Bei keinem der drei Modelle konnte in den Subperioden¹⁵ Kointegration festgestellt werden. Die gefundene Kointegrationsbeziehung aus der vierzigjährigen Perspektive wird anscheinend durch die verschiedenen kurzfristigen oder mittelfristigen Ereignisse, wie der 'Börsencrash' in 1987, oder die 'Asienkrise' und das 'Internetbubble' überlagert. Die Existenz einer Kointegrationsbeziehung aus vierzigjähriger Sicht entspricht der Idee des langfristigen ökonomischen Gleichgewichts, auf die die empirische Kointegrationsmethode basiert. Deswegen wurde für die weitere Untersuchung die gesamte Periode von vierzig Jahren gewählt.

¹⁵Die Gesamtperiode von Juni 1964 bis Oktober 2004 wurde zuerst bis Oktober 1987, Aktiencrash, verkürzt. Für die weitere Tests wurde diese Periode immer um 2 Jahre verlängert.

Die saisonalen Schwankungen wurden in der Modellierung durch die Einführung der zwölf saisonalen zentrierten Dummi-Variablen berücksichtigt. Die Ergebnisse zeigen, dass nur die Dummi-Variable für Januar für die Modellierung statistisch signifikant ist. Das bestätigt die Existenz des sogenannten 'Januar'-Effekts der Weltbörse. Die zentrierten Dummi-Variablen beeinflussen die Verteilung der Teststatistik nicht, vgl. Madalla (1998), deswegen dürften im weiteren die kritischen Werte aus Osterwald-Lenum (1992) verwendet werden.

Die Ergebnisse für Modelle mit vielen Lags können nicht deutlich erklärt werden und sind eher auf die Schwächen der Johansen-Methode zurückzuführen. Es wurde im theoretischen Teil schon diskutiert (Kapitel 2.2.3), dass es kein der beiden Informationskriterien (AIC und BIC) allein für die Auswahl der Laganzahl geeignet ist. Alle drei Modelle wurden benutzt, um mit unterschiedlicher Laganzahl die Existenz von Kointegration zu testen. Es wurde in keinem der drei Modelle Kointegration gefunden, wenn die Anzahl der Lags kleiner als zwei war. Bei der Laganzahl drei und vier wurde mit dem zweiten Modell eine Kointegrationsbeziehung festgestellt. Für das erste Modell existiert eine langfristige stabile Beziehung nur bei der Laganzahl von vier. Wenn die Anzahl der Lags auf 23 erhöht wird¹⁶, wird keine Kointegration in dem System gefunden. Allerdings wird mit einer Anzahl der Verzögerungen von 24 (genau 2 Jahre) wieder eine einzige Kointegrationsbeziehung mit dem zweiten Modell festgestellt. Es stellt sich die Frage, ob die Kointegrationsbeziehung nur aufgrund der niedrigen Mächtigkeit der Johansen-Tests gefunden wird, oder hier ein ökonomisches Phänomen auf den Aktienmärkten, wie Konjunktur oder Lagerhaltungszyklen, vorliegt. Wie es schon im theoretischen Teil angesprochen wurde, ist die Mächtigkeit des Johansen-Tests bei der hohen Anzahl von Lags niedrig. Allerdings wird keine Kointegration mit einer Anzahl der Verzögerungen von höher als 24 gefunden. Von der ökonomischen Seite kann eine Argumentation, dass die Börsen Informationsshocks bei einer hohen Kapitalmobilität innerhalb von zwei Jahren verarbeiten, nicht vertretbar sein. Deswegen wurden für die weiteren Untersuchungen die Modelle mit vier Lags verwendet.

Tabelle 3.16: Johansens Kointegrationstest für die Aktienindizes (Modell 1)

Eigene Werte	H_0	λ_{max}	Kritische Werte	H_0	λ_{trace}	Kritische Werte
0,045	r=0	22,20	30,04	$r = 0$	59,53	59,46
0,041	r=1	20,10	23,8	$r \leq 1$	37,33	39,89
0,025	r=2	12,25	17,89	$r \leq 2$	17,23	24,31
0,010	r=3	4,85	11,44	$r \leq 3$	4,98	12,53
0,000	r=4	0,13	3,84	$r \leq 4$	0,13	3,84

Notiz: Die kritischen Werte zum Niveau 5% sind von Osterwald-Lenum (1992). Das Modell beinhaltet eine zentrierte saisonale Dummi-Variable. Die Anzahl der Lags ist 4.

¹⁶Das entspricht einer Vergrößerung der Periode von einem Quartal bis zu zwei Jahren

Tabelle 3.17: Johansens Kointegrationstest für die Aktienindizes (Modell 2)

Eigene Werte	H_0	λ_{max}	Kritische Werte	H_0	λ_{trace}	Kritische Werte
0,053	r=0	25,91	33,46	$r = 0$	69,31	68,52
0,042	r=1	20,40	27,07	$r \leq 1$	43,40	47,21
0,026	r=2	12,69	20,97	$r \leq 2$	23,00	29,68
0,018	r=3	8,65	14,07	$r \leq 3$	10,31	15,41
0,003	r=4	1,66	3,76	$r \leq 4$	1,66	3,76

Notiz: Die kritischen Werte zum Niveau 5% sind von Osterwald-Lenum (1992). Das Modell beinhaltet eine zentrierte saisonale Dummi-Variable. Die Anzahl von Lags ist 4.

Tabelle 3.18: Johansens Kointegrationstest für die Aktienindizes (Modell 3)

Eigene Werte	H_0	λ_{max}	Kritische Werte	H_0	λ_{trace}	Kritische Werte
0,063	r=0	31,18	37,52	$r = 0$	81,57	87,31
0,042	r=1	20,48	31,46	$r \leq 1$	50,39	62,99
0,030	r=2	14,51	25,54	$r \leq 2$	29,91	42,44
0,018	r=3	8,78	18,96	$r \leq 3$	15,40	25,32
0,014	r=4	6,62	12,25	$r \leq 4$	6,62	12,25

Notiz: Die kritischen Werte zum Niveau 5% sind von Osterwald-Lenum (1992). Das Modell beinhaltet eine zentrierte saisonale Dummi-Variable. Die Anzahl von Lags ist 4.

In den Tabellen (3.16), (3.17) und (3.18) finden sich die Testergebnisse für alle drei Modelle während der Gesamtperiode (1964-2004), in der monatlichen Datenfrequenz, mit einer saisonalen zentrierten Januar-Dummi-Variable D_t und mit vier Lags.

Der λ_{trace} -Test stellt die Existenz einer Kointegrationsbeziehung im ersten und im zweiten Modell zum Signifikanzniveau von 5% fest. Die Nullhypothese, dass die fünf Aktienindizes von USA, Japan, Deutschland, Frankreich und UK nicht kointegriert sind ($r = 0$), wird gegen die Alternativhypothese einer oder mehrerer Kointegrationsbeziehungen ($r \geq 1$) getestet. Die Nullhypothese kann nicht beibehalten werden. Der λ_{trace} -Test kann allerdings nicht mehr als einen Kointegrationsvektor feststellen, die Nullhypothese $r \leq 1$ kann zum Signifikanzniveau von 5% nicht abgelehnt werden. Mit dem λ_{max} -Test wurde keine Kointegrationsbeziehung identifiziert¹⁷. Für das dritte Modell wurde weder mit λ_{trace} -Test noch mit λ_{max} -Test die Existenz der Kointegration festgestellt.

¹⁷Wie schon im theoretischen Teil besprochen wurde, ist der λ_{trace} -Test weniger durch 'Nichtnormalität' der Störtermen beeinflusst.

Die Präsenz eines einzigen Kointegrationsvektors impliziert, dass die Aktienindizes der G5-Länder sich in einem langfristigen Gleichgewicht befinden, das eindeutig spezifiziert werden kann. Ein wichtiger Aspekt der internationalen Globalisierung der Finanzmärkte ist ein Zusammenhang zwischen den Preisen. Erwartungsgemäß können die zukünftigen Fluktuationen in den Preisen eines Marktes von den Änderungen auf einem anderen Markt durch den perfekten Informationsaustausch, den einfachen Zugang zu den Märkten und das zeitliche Portfoliomanagement ausgelöst werden. Der Kointegrationsvektor beschreibt genau diese Verbindung für die G5-Länder. Die Präsenz einer Gleichgewichtsbeziehung zwischen den Preisen führt zu einer hohen Korrelation der langfristigen Renditen, vgl. Taylor und Tonks (1989), Kasa (1992), Gerrits und Yüce (1999), Gilmore and McManus (2002), Syriopoulos (2004). Dies liefert eine mögliche Erklärung zu den bescheidenen Gewinnen aus der internationalen Diversifikation während der letzten 30 Jahre, vgl. Malkiel (2004).

Die letzte Entscheidung für die Auswahl eines geeigneten Modells für weitere Untersuchungen des Systems der Aktienmärkte der G5-Länder liegt bei der Auswahl des deterministischen Trends. Mit dem dritten Modell, das die Trendkomponente in den Datenreihen und in dem Kointegrationsraum beinhaltet, wurde bei keiner Konstellation der Einflussfaktoren Kointegrationsbeziehungen festgestellt. Möglicherweise kann diese Resultat auf eine Überparametrisierung des Modells zurückgeführt werden. Das zweite Modell hat sich als ein robustes Modell erwiesen, deswegen wird für die Schätzung des Kointegrationsvektors und der kurzfristigen Anpassungen dieses Modell verwendet.

Schätzen des VECM

Der normalisierte Kointegrationsvektor β und die entsprechenden Anpassungskoeffizienten zum Gleichgewicht α_i sind in der Tabelle (3.19) für das Modell 1 zusammengefasst.

Tabelle 3.19: Der normalisierte Kointegrationsvektor und die Anpassungskoeffizienten (Modell 1)

	USA	Japan	Frankreich	Deutschland	UK
β	-2,109	-0,651	0,410	1,000	1,356
α	0,007	0,006	0,001	-0,019	-0,007

Die relative Größe¹⁸ der Koeffizienten des Kointegrationsvektors β liefert Informationen über die individuelle Rolle der einzelnen Märkte in der Gruppe. Die gleichgewichtige Beziehung ist durch die Gleichung gegeben:

$$SP_D = 2,109SP_{USA} + 0,651SP_{Jap} - 0,410SP_{Fr} - 1,356SP_{UK}$$

¹⁸Die Koeffizienten sind normalisiert, so dass bei Deutschland Eins steht.

Das Ergebnis weist auf eine positive langfristige Verbindung zwischen Deutschland, USA, Japan und eine negative Verbindung für Deutschland zu Frankreich und zu UK. Der Preis des deutschen Aktienindizes wird am stärksten von den amerikanischen und britischen Märkten beeinflusst. Eine unterproportionale Reaktion auf die Shocks auf dem japanischen und französischen Märkten wird ebenfalls festgestellt. Die Reaktion des Aktienpreises ist durch die Fluktuation des Wechselkurses verstärkt. Dadurch könnten die hohen Koeffizienten bei den amerikanischen und britischen Aktienpreisen erklärt werden, weil die Wechselkurse DM/\$ und DM/Pfund eine höhere Fluktuation im Vergleich zu DM/Yen und DM/Franc während der Gesamtperiode aufweisen.

Für das Modell 2 ist der normalisierte Kointegrationsvektor β und die Anpassungskoeffiziente α_i in der Tabelle (3.20) aufgeführt.

Tabelle 3.20: Der normalisierte Kointegrationsvektor und die Anpassungskoeffizienten (Modell 2)

	USA	Japan	Frankreich	Deutschland	UK
β	-1,068	-0,387	-0,272	1,000	0,904
α	0,012	0,013	0,032	-0,026	-0,007

Der Unterschied zum ersten Modell liegt hier in den niedrigeren relativen Koeffizienten in der Gleichgewichtsbeziehung, beispielsweise ist der Koeffizient vor SP_{USA} gleich 1,068 im Vergleich zum Koeffizient aus dem ersten Modell von 2,109. Allerdings entspricht die Gesamtreaktion der Zusammensetzung des Kointegrationsvektors aus dem ersten Modell. Die höchste Wirkung auf den deutschen Index haben der amerikanische und britische Indizes. Frankreich und Japan weisen unterproportionale Koeffizienten auf in Höhe von 0,272 und 0,387.

Im Vergleich zur zweidimensionalen Analyse wird die führende Rolle des amerikanischen Aktienmarktes aus der langfristigen Sicht bei der Gesamtbetrachtung bestätigt. Als eine mögliche Erklärung dafür könnte die Berücksichtigung eines Einflusses der weiteren Aktienmärkte liefern.

Die kurzfristigen Anpassungen

Die beide Tabellen (3.21) und (3.22) ergeben ein einheitliches Bild bezüglich kurzfristiger Anpassungen nach einem Monat. Die Änderungen auf dem eigenen Aktienmarkt sind statistisch relevant für die Änderungen der Renditen im folgenden Monat. Eine Ausnahme stellt der französische Markt dar. Die starke Rolle des amerikanischen Marktes spiegelt sich in den bei deutschen und französischen Märkten stehenden Koeffizienten wider. Es wird kein Einfluss von der amerikanischen Seite auf den britischen Aktienindex beobachtet. Der deutsche Aktienindex wird zusätzlich von den Änderun-

Tabelle 3.21: Die Matrix der kurzfristigen Anpassungen nach einem Monat. (Modell 1)

	$\Delta SP_{Ger,t}$	$\Delta SP_{USA,t}$	$\Delta SP_{Jap,t}$	$\Delta SP_{Fr,t}$	$\Delta SP_{UK,t}$
$\Delta SP_{Ger,t-1}$	0,18 (3,26)	-0,09 (-1,48)	-0,17 (-2,50)	0,05 (0,61)	-0,05 (-0,75)
$\Delta SP_{USA,t-1}$	0,14 (2,32)	0,32 (5,18)	0,12 (1,60)	0,19 (2,27)	0,06 (0,80)
$\Delta SP_{Jap,t-1}$	-0,04 (-0,87)	-0,03 (-0,60)	0,30 (5,88)	0,07 (1,21)	0,03 (0,51)
$\Delta SP_{Fr,t-1}$	0,13 (3,47)	0,07 (1,81)	0,07 (1,40)	0,03 (0,54)	0,18 (3,64)
$\Delta SP_{UK,t-1}$	-0,02 (0,33)	0,06 (1,24)	0,08 (1,42)	0,00 (0,07)	0,29 (5,00)

Notiz: In Klammern steht die F-Statistik für den geschätzten Koeffizient.

gen der französischen Renditen beeinflusst. Entsprechend sind der amerikanische und der japanische Aktienmärkte von den Änderungen auf den anderen Märkten der G5-Länder-Gruppe weitgehend unabhängig.

3.2.4 Dynamische Analyse

Ein Teil der Strukturanalyse besteht aus dynamischen Simulationen für das System der zu untersuchenden Variablen. Die dynamischen Simulationen geben einen Einblick in die ökonomische Relevanz der Variablen im Gesamtsystem. Die dynamische Analyse verwendet die Berechnung der Impuls-Antwort-Funktionen und der Varianz-Dekomposition der Vorhersagefehler aller Variablen. Die Impuls-Antwort-Analyse stellt die zukünftige Reaktion des Gesamtsystems mittels des aufgestellten Modells bei einer Änderung der unabhängigen Variablen dar. Die Varianz-Dekomposition der Vorhersagefehler liefert zuerst eine prozentuale Aufteilung der Varianz der Störteme nach Einflussfaktoren. Im Folgenden wird die Änderung dieser Zerlegung innerhalb der bestimmten Periode simuliert.

Impuls-Antwort-Analyse

Die Grafik (3.2) stellt die Reaktion des deutschen Aktienindizes auf die Einheits-shocks in den einzelnen nationalen Märkten dar. Der Zeithorizont umfasst 10 Monate. Die Shocks auf den internationalen Märkten haben einen permanenten Einfluss auf den deutschen Index, der auch noch nach dem zehnten Monat beobachtbar ist. Am stärksten reagiert der deutsche Aktienindex auf Änderungen im eigenen Land. Die Reaktion auf die Änderung bleibt nach 10 Monaten erhalten und führt zum neuen Niveau des

Tabelle 3.22: Die Matrix der kurzfristigen Anpassungen nach einem Monat (Modell 2).

	$\Delta SP_{Ger,t}$	$\Delta SP_{USA,t}$	$\Delta SP_{Jap,t}$	$\Delta SP_{Fr,t}$	$\Delta SP_{UK,t}$
$\Delta SP_{Ger,t-1}$	0,20 (3,53)	-0,09 (-1,56)	-0,17 (-2,55)	0,05 (0,66)	-0,05 (-0,67)
$\Delta SP_{USA,t-1}$	0,14 (2,40)	0,32 (5,18)	0,12 (1,62)	0,21 (2,49)	0,06 (0,84)
$\Delta SP_{Jap,t-1}$	-0,04 (-0,82)	-0,03 (-0,61)	0,30 (5,86)	0,08 (1,31)	0,03 (0,52)
$\Delta SP_{Fr,t-1}$	0,12 (3,24)	0,08 (1,87)	0,07 (1,42)	0,02 (0,38)	0,17 (3,57)
$\Delta SP_{UK,t-1}$	-0,02 (-0,36)	0,06 (1,19)	0,08 (1,34)	-0,02 (-0,25)	0,29 (4,93)

Notiz: In Klammern steht die F-Statistik für den geschätzten Koeffizient.

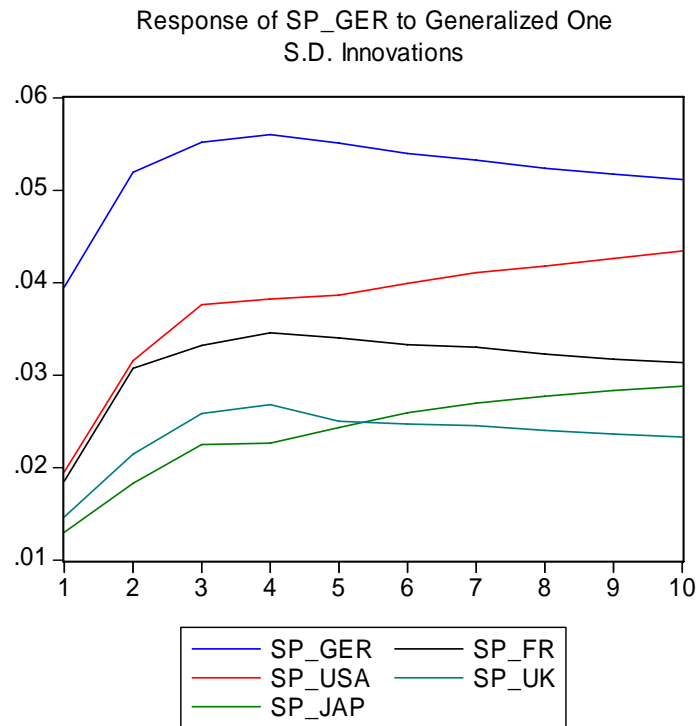
Aktienindizes. Ähnlich reagiert der deutsche Index auf Schocks in anderen Märkten, allerdings ist die Reaktion wesentlich schwächer. Eine Ausnahme stellt hier wieder der amerikanische Markt dar. Die Reaktion des deutschen Indizes auf die Änderung auf dem amerikanischen Markt ist im ersten Monat stärker im Vergleich zu den anderen ausländischen Märkten, allerdings wird sie in den weiteren Monaten nicht wesentlich schwächer sein, als es bei dem französischen oder dem britischen Aktienmarkt der Fall ist. Der deutsche Markt reagiert am schwächsten auf Schocks auf dem japanischen Aktienmarkt. Die Reaktion ist ähnlich zu der Reaktion auf Änderungen auf dem amerikanischen Aktienindex.

Varianz-Dekomposition

Die Ergebnisse der Varianz-Dekomposition der Störterme bestätigen die Impuls-Antwort-Analyse im Hinblick auf den grössten Einfluß des eigenen Landes auf prognostizierte Werte. Sie sind in den Tabellen (3.23) und (3.24) zusammengefasst. Die Ergebnisse lassen sich als eine Prognose für zwölf kommende Monate darstellen.

Sowohl im ersten als auch im zweiten Modell wächst die Varianz des Prognosefehlers von 4% auf 20% pro Monat innerhalb eines Jahres. Die Bewegung des deutschen Marktes spielt die größte Rolle für eine Vorhersage dieser Varianz in beiden Modellen. Der Einfluss sinkt allerdings von 100% im erstem Monat auf 89,01% nach einem Jahr im ersten Modell und auf 91,98% nach einem Jahr im zweiten Modell. Dagegen wächst die Wirkung des amerikanischen Marktes für die Varianz der Prognosefehler bis auf 9,49% nach einem Jahr im ersten Modell und bis auf 5,38% nach zwölf Monaten im zweiten Modell. Im ersten Modell sinkt die Wichtigkeit des französischen Marktes von 0,74% bis auf 0,35%. Im zweiten Modell verändert sich der Einfluss in eine andere Richtung,

Abbildung 3.2: Impuls-Antwort-Funktionen für den deutschen Aktienindex



er wächst von 0,86% nach einem Monat bis auf 1,4% nach einem Jahr. Allerdings ist der Einfluss der europäischen und des japanischen Marktes auf die Vorhersage des Prognosefehlers für die Entwicklung des deutschen Aktienmarktes vernachlässigbar. Der Gesamtwert liegt unter 2% im ersten Modell und unter 3% im zweiten Modell.

Tabelle 3.23: Varianz-Dekomposition für Deutschland (Modell 1)

Periode	Varianz	Deutschland	USA	Japan	Frankreich	UK
1	0,04	100,00	0,00	0,00	0,00	0,00
2	0,07	98,16	1,48	0,00	0,74	0,05
3	0,09	96,87	2,44	0,03	0,73	0,03
4	0,10	95,95	3,73	0,07	0,59	0,02
5	0,12	95,15	4,23	0,11	0,48	0,02
6	0,13	94,37	5,03	0,17	0,39	0,05
7	0,14	93,56	5,79	0,22	0,33	0,10
8	0,15	92,71	6,55	0,27	0,29	0,17
9	0,16	91,82	7,30	0,33	0,28	0,27
10	0,17	90,90	8,04	0,39	0,29	0,38
11	0,18	89,96	8,77	0,45	0,31	0,50
12	0,19	89,01	9,49	0,51	0,35	0,64

Tabelle 3.24: Varianz-Dekomposition für Deutschland (Modell 2)

Periode	Varianz	Deutschland	USA	Japan	Frankreich	UK
1	0,04	100,00	0,00	0,00	0,00	0,00
2	0,07	98,15	0,93	0,01	0,86	0,05
3	0,09	96,86	2,11	0,03	0,97	0,03
4	0,11	96,29	2,50	0,03	1,17	0,02
5	0,12	95,84	2,79	0,08	1,24	0,05
6	0,14	95,27	3,19	0,17	1,28	0,09
7	0,15	94,67	3,60	0,28	1,32	0,14
8	0,16	94,09	3,99	0,38	1,34	0,20
9	0,17	93,54	4,36	0,48	1,35	0,27
10	0,18	92,99	4,71	0,58	1,37	0,35
11	0,19	92,47	5,06	0,66	1,38	0,43
12	0,20	91,98	5,38	0,74	1,40	0,51

Kapitel 4

Ökonomische Konzepte

Eine philosophische Feststellung von Immanuel Kant ist:

„Theorie ohne Empirie ist leer. Empirie ohne Theorie ist blind.“

Die wissenschaftlichen Untersuchungen aus mehreren Jahren der Aktienmarktgeschichte bestätigen diese Aussage und unterstreichen die Notwendigkeit einer Auseinandersetzung mit den ökonomischen Theorien. Im zweiten Kapitel der Arbeit wurden empirische Konzepte für die Schätzung der Gleichgewichtsbeziehung der Aktienmärkte beschrieben. Bei der empirischen Datenanalyse der Aktienmärkte der G5-Länder wurde mittels Kointegrationsanalyse eine einzige langfristige Gleichgewichtsbeziehung festgestellt. Für eine Interpretation dieser Gleichgewichtsbeziehung könnten zwei Ansätze ökonomischer Theorie in Betracht genommen werden. Die Markteffizienz-Hypothese und die Portfolio-Theorie stellen die wichtigsten Ideen der Aktienmarktforschung des zwanzigsten Jahrhunderts dar. Im vorliegenden Kapitel 4 werden vorerst die beiden Theorien kurz dargestellt. Ein großer Teil des Kapitels widmet sich einer ausführlichen Diskussion bezüglich schwacher empirischer Hinweise für die Gültigkeit dieser Theorien. Ein Versuch, das empirische Kointegrationskonzept mit den beiden Theorien zu verbinden, schließt das Kapitel ab.

4.1 Die Markteffizienz-Hypothese

4.1.1 Die MEH und der Aktienmarkt

Die Entstehung der Markteffizienz-Hypothese (MEH) kann auf den ersten theoretischen Beitrag von Bachelier (1900) und auf die empirische Untersuchung von Cowles (1933)

zurückgeführt werden¹. Als ein erster Einsatz der neueren Literatur zum Thema kann die Arbeit von Samuelson (1965) bezeichnet werden, dessen Beitrag sehr geschickt durch den Titel seines Artikels formuliert ist: *„Proof that Properly Anticipated Prices Fluctuate Randomly“*. Die Idee der effizienten Märkte formuliert Samuelson (1965) im Bezug zu vorhandenen Informationen auf dem Markt, so dass im Weiteren an erster Stelle über Informationseffizienz gesprochen wird. Dabei ist der informationseffizienten Markt nicht mit einem Pareto-effizienten Markt zu verwechseln². Sie wurde von Fama (1970), (1991), (1998), Fama und French (1988), Campbell und Shiller (1988), Malkiel (1992), (2003), (2005), Engel (1996), Campbell und anderen (1997), Shiller (2003) und vielen anderen erweitert und empirisch überprüft.

Eine Formulierung der MEH kann wie folgt zusammengefasst werden: Die Preisänderungen auf einem informationseffizienten Markt können ausschließlich zufällig sein, wenn die Preise richtig antizipiert wurden. Dies impliziert, dass die Aktienpreise die Preiserwartungen und die möglichen Informationen aller Marktteilnehmer schon beinhalten. Fama (1970) fasste die Idee der Informationseffizienz des Marktes in seiner klassischen Studie zusammen:

„A market in which prices always ‘fully reflect’ available information is called ‘efficient’.“

Da der Ausdruck *‘fully reflect’* von Fama nicht weiter verdeutlicht wird, entsteht den Bedarf einer detaillierten Erklärung dieses Begriffs. Später bietet Malkiel (1992) eine mehr explizite Definition eines effizienten Marktes an:

„A capital market is said to be efficient if it fully and correctly reflects all relevant information in determining security prices. Formally, the market is said to be efficient with respect to some information set ... if security prices would be unaffected by revealing that information to all participants. Moreover, efficiency with respect to an information set ... implies that it is impossible to make economic profits by trading on the basis of [that information set].“

Der erste Satz von Malkiel wiederholt die Definition von Fama. Sein zweiter und dritter Satz erweitern die Definition mit zwei Alternativen. Der zweite Satz schlägt vor, dass die Markteffizienz beim Aufdecken einer zusätzlichen Information für Marktteilnehmer durch eine Messung der Preisreaktion getestet werden kann. Wenn sich die Preise beim Zugang einer neuen Information nicht ändern, dann ist der Markt effizient bezüglich dieser Information. Obwohl diese Interpretation leicht nachvollziehbar ist,

¹Vgl. Campbell und andere (1997)

²Bei einer pareto-effizienten Allokation ist es nicht möglich, irgendein Individuum besser zu stellen, ohne dass mindestens ein anderes schlechter gestellt wird.

stellt sie jedoch eine Herausforderung für die praktischen Tests dar. Der dritte Satz von Malkiel interpretiert die Markteffizienz auf andere Weise. Seine Idee liefert die Grundlage für die meisten empirischen Untersuchungen der Markteffizienz, wobei zwei unterschiedlichen Wege der Verwirklichung gegangen werden. Zunächst wird versucht, Gewinne zu messen, die die professionellen Manager auf den Aktienmärkten erzielen konnten. Wenn die Manager tatsächlich die größeren Renditen erwirtschaften, so ist der Markt ineffizient bezüglich der Information, die diese Manager besitzen. Der Vorteil dieser Approximation besteht in der Realitätsnähe, d.h. sie basiert auf dem realen Handel der realen Marktteilnehmer. Allerdings kann in den meisten Fällen keine direkte Information und die Handelsstrategien der Manager erhalten werden³. Die zweite alternative Approximation, die eher auf einem hypothetischen Handel basiert, spezifiziert eine exakte Informationsmenge. Die darauf basierten Prognosen können zu keinen Überschussgewinnen auf einem effizienten Markt führen. Die Implementation dieser Approximation benötigt eine Definition der entsprechenden Informationsmenge.

Eine Klassifizierung der Markteffizienzform bezüglich der Informationsmenge liefert Roberts (1967):

- Bei der schwachen Form der Markteffizienz enthält die Informationsmenge nur die historischen Preisdaten oder Renditen.
- Bei der halbstarken Form der Markteffizienz enthält die Informationsmenge alle für die Marktteilnehmer bekannten (öffentlich zugänglichen) Informationen.
- Bei der starken Form der Markteffizienz enthält die Informationsmenge alle bekannten Informationen (einschließlich der privaten Informationen).

Die MEH wird den verschiedenen Märkten unterstellt, wie zum Beispiel dem Bondmarkt oder dem Devisenmarkt. Für die Aktienmärkte ist es üblich⁴, die Hypothese mit Hilfe der Überschussrendite eines Wertpapiers zu formulieren. Die Überschussrendite stellt eine Differenz zwischen der tatsächlichen Rendite des Wertpapiers und ihrer normalen Rendite dar. An dieser Stelle ist eine Spezifikation der *'normalen'* Rendite einer Aktie von entscheidender Bedeutung. Die klassische Annahme lautet, dass die *'normale'* Rendite eines Wertpapiers über die Zeit konstant bleibt. In den letzten Jahren wächst das Interesse der Forscher an den Gleichgewichtsmodellen, die die zeitvariierende *'normale'* Rendite beschreiben (vgl. Campbell und andere (1997)). Eine Vorhersage der Überschussrendite kann auf verschiedenen Informationsmengen basieren. Beispielsweise darf die Informationsmenge entsprechend der Definition der schwachen Effizienz nur aus den historischen Preisdaten bestehen. Wenn keine Prognose für die Überschussrendite gemacht werden kann, dann ist die Rendite in diesem Sinne zufällig. Die MEH kann in diesem Fall nicht abgelehnt werden.

³vgl. Fama (1991)

⁴Vgl. Campbell und andere (1997).

Malkiel (2003) benutzt eine weitere Definition der Markteffizienz in seinem neuesten Artikel. Er interpretiert die MEH, so dass kein Investor eine überdurchschnittliche Rendite ohne Akzeptanz eines überdurchschnittlichen Risikos auf den informationseffizienten Märkten erwirtschaften kann. Ein bekanntes Beispiel des Gesprächs zwischen einem Professor und einem Studenten illustriert die Meinung, die einige Finanzökonomien über die Markteffizienz haben. Als der Student anhält, um einen 100\$ Schein von der Strasse aufzuheben, sagt der Professor zu ihm, er solle dies nicht tun. Die Begründung ist eben, dass es sich um einen falschen Geldschein handelt. Wenn es ein richtiger Geldschein wäre, wäre er nicht hier, meint der Professor. Ähnlich ist die Interpretation der MEH, die mit einem angemessenen Rendite/Risiko-Verhältnis eines Wertpapiers verbunden ist. Wenn auf einem Aktienmarkt eine Arbitrage-Möglichkeit existiert hätte, dann würde sie sofort von den Marktteilnehmern antizipiert. Dementsprechend kann keine Möglichkeit existieren, 'kostenlose' Gewinne auf den informationseffizienten Märkten zu erwirtschaften.

Entsprechend der letzten Interpretation der MEH können die Märkte effizient sein, obwohl sie einige falsche Bewertungen der Information aufweisen, wie zum Beispiel der 'Aktiencrash' von 1987 und die 'Internet Blase' von 2000. Die Märkte können auch effizient sein, obwohl manche Teilnehmer ein irrationales Verhalten aufweisen. Ebenfalls kann die Effizienz der Märkte bestehen, wenn die Aktien eine größere Volatilität aufweisen, als durch eine Fundamentalebewertung erklärt werden kann. Viele Ökonomen glauben an Effizienz der Finanzmärkte, weil sie keine überdurchschnittlichen Renditen ohne zusätzliches Risiko erlauben, mit anderen Worten, weil sie an die Echtheit des 100\$ Geldscheins auf der Strasse nicht glauben.

4.1.2 Die Random-Walk-Hypothese

Eine Formalisierung der MEH bringt einige Vereinfachungen und Einschnitte ihrer ökonomischen Idee mit sich. Bei der Modellierung wird versucht, die zukünftigen Änderungen der Aktienpreisen allein mit Hilfe vergangener Preisänderungen zu bestimmen. Die auf diesem Wege konstruierte Prognose ist sehr beschränkt. Allerdings kann diese Strategie einer Vereinfachung des Modells einige Erkenntnisse zu dem Verhalten der Aktienpreise bringen. Die Martingal- und Random-Walk-Theorie, die zwei wichtigsten Theorien in der Wahrscheinlichkeitsrechnung, liegen dieser relativ einfachen Idee zu Grunde.

Die notwendigen Bezeichnungen und Definitionen werden im Weiteren für die Modelle eingeführt.

Definition 4.1 Sei P_t der Preis einer Aktie zu dem Zeitpunkt t . Mit R_t wird die Nettorendite einer Aktien in einer Zeitperiode $[t - 1, t]$ bezeichnet.

$$R_t = \frac{\Delta P_t}{P_{t-1}} = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1. \quad (4.1)$$

Die Rendite einer Aktie wird dann durch $(1 + R_t)$ definiert.

Aus theoretischen und praktischen Gründen ist es insbesondere sinnvoll, für die Aktienmärkte den Begriff der stetigen Rendite einzuführen.

Definition 4.2 Die stetige Rendite oder logarithmierte Rendite einer Aktie wird als \log von der Rendite $(1 + R_t)$ definiert:

$$r_t \equiv \log(1 + R_t) = \log\left(\frac{P_t}{P_{t-1}}\right) = p_t - p_{t-1}, \quad (4.2)$$

mit $p_t \equiv \log P_t$.

Mit Ω_t wird im Weiteren eine Informationsmenge zu dem Zeitpunkt t bezeichnet, die für die Vorhersage der Aktienrenditen relevant ist. $E(X_{t+1}|\Omega_t) = E_t(X_{t+1})$ ist die mathematische bedingte Erwartung einer Zufallsvariable X zu dem Zeitpunkt $t+1$ mit Einbeziehung der Informationen aus der Vergangenheit, also aller Informationen bis zu dem Zeitpunkt t .

Ein rationales Verhalten aller Marktteilnehmer stellt eine starke Annahme bei der Formalisierung der MEH dar. Die Bildung der mathematischen Erwartung spiegelt die Annahme des rationalen Verhaltens der Teilnehmer wieder. Mit Hilfe der mathematischen Erwartung lässt sich die MEH wie folgt formulieren. Auf einem informationseffizienten Markt hängt der zukünftige Preis einer Aktie nur von dem rational erwarteten Aktienpreis und von einer unerwarteten Störung ab, die einen unerwarteten Informationsschock repräsentiert:

$$P_{t+1} = E(P_{t+1}|\Omega_t) + \varepsilon_{t+1}. \quad (4.3)$$

Dabei wird für den Vorhersagefehler ε_{t+1} ein White-Noise-Prozess unterstellt. Die ökonomische Interpretation dieser Annahme ist, dass der durchschnittliche Fehler bei einer rationalen Prognose über die Zeit zu Null tendiert:

$$E(\varepsilon_{t+1}|\Omega_t) = E(P_{t+1} - E(P_{t+1}|\Omega_t)|\Omega_t) = 0.$$

Bei der Formalisierung der MEH ist die Annahme, dass der Erwartungsfehler mit allen vergangenen Informationen bis zu dem Zeitpunkt t unkorreliert ist, besonders wichtig:

$$E(\varepsilon'_{t+1}\Omega_t|\Omega_t) = 0. \quad (4.4)$$

Die Annahme ist unter der Orthogonalitätseigenschaft der bedingten Erwartungen bekannt. Eine intuitive Vorstellung für die Gültigkeit der Gleichung (4.4) liefert folgende Überlegung. Wenn die Informationsmenge Ω_t für eine Vorhersage der Fehler ε_{t+1} benutzt werden könnte, dann könnte diese Informationsmenge zu einer Verbesserung der Vorhersage von P_{t+1} führen. Das bedeutet aber, dass nicht alle relevanten Informationen

bei der Vorhersage von P_{t+1} verwendet wurden. Daraus folgt, dass der Vorhersagefehler bei einer optimalen Preisbildung unabhängig von den vergangenen Information sein muss.

Die MEH kann mit Hilfe des sogenannten 'Martingals' formuliert werden. Ein stochastischer Prozess (P_t) , für den gilt:

$$E(P_{t+1}|\Omega_t) = P_t, \quad (4.5)$$

wird als ein Martingal bezeichnet. Dementsprechend bleiben die erwarteten Änderungen des Martingals im Durchschnitt Null über die Zeit, $E(P_{t+1} - P_t|\Omega_t) = 0$. Wenn der Preis einer Aktie einem stochastischen Prozess mit der Martingal-Eigenschaft folgt, dann gilt für die Preisänderung, also für die Aktienrendite, dass sie über die Zeit konstant bleibt, $E(R_{t+1}|\Omega_t) = R_0$. Unter R_0 wird eine 'normale' Rendite verstanden. Die 'normale', oder die gleichgewichtige Rendite, repräsentiert die Opportunitätskosten der Aktienhändler. Als eine statistische Approximation für die 'normale' Rendite wird in der Regel die Rendite der festverzinslichen Wertpapiere benutzt. Entsprechend gilt die Martingal-Eigenschaft für den logarithmierten Prozess $p_t = \log P_t$.

Für den Störterm des logarithmierten Prozesses mit der Martingal-Eigenschaft $\varepsilon_{t+1} = p_{t+1} - E(p_{t+1}|\Omega_t)$ gilt entsprechend, dass der erwartete Fehler im Durchschnitt über die Zeit verschwindet $E_t(\varepsilon_{t+1}) = E_t(p_{t+1}) - E_t(p_t) = p_t - p_t = 0$.

Aus der ökonomischen Sicht lässt die Martingal-Eigenschaft eines stochastischen Prozesses folgende Interpretation zu. Die beste Vorhersage für einen Aktienpreis $E_t(p_{t+1})$ ist der aktuelle Wert der Aktie p_t . Äquivalent für die Aktienrendite bleiben die erwarteten Änderungen des Aktienpreises über die Zeit unverändert, wenn die Erwartungen nur durch das Einbeziehen der Information bezüglich vergangener Preise gebildet wurden. Dementsprechend ist der Markt bezüglich der Informationsmenge Ω_t effizient, wenn die Aktienpreise dem stochastischen Prozess mit der Martingal-Eigenschaft folgen. Alle Informationen bis zum Zeitpunkt t wurden von dem Preisbildungsprozess (p_t) berücksichtigt. Im Durchschnitt werden keine Fehler bei den Vorhersagen gemacht.

Die Martingal-Hypothese für den Preisbildungsprozess hat zusätzlich folgendes Charakteristikum. Die erwarteten Preisänderungen $(\varepsilon_t)_{t \in T}$ sind unkorrelierte Variablen, damit gilt für die Kovarianz: $Cov(\varepsilon_t, \varepsilon_{t+k}) = E_t(\varepsilon_t \varepsilon_{t+k}) = 0$ für jedes $k \neq 0$. Dies impliziert die Wirkungslosigkeit der linearen Vorhersageregeln für die zukünftigen Preisänderungen.

Ein Random-Walk-Prozess ist ein stochastischer Prozess mit der Martingal-Eigenschaft. Die einfachste Version der Random-Walk-Hypothese für den Preisbildungsprozess kann wie folgt formuliert werden: Der logarithmierte Aktienpreis $p_t = \log P_t$ folgt einem Random-Walk-Prozess mit dem Driftparameter μ , dabei hat der Prozess die

unabhängigen und identisch verteilten Störterme⁵:

$$p_{t+1} = \mu + p_t + \varepsilon_{t+1}, \varepsilon_t \sim iid(0, \sigma^2). \quad (4.6)$$

Die Unabhängigkeit der Störterme $(\varepsilon_t)_{t \in T}$ impliziert ihre Unkorreliertheit. Für den Driftparameter $\mu = 0$ gilt, dass der logarithmierte Prozess (p_t) einen Martingal darstellt. Ist $\mu \neq 0$, wird der Prozess (p_t) um den Trend μt bereinigt. Der Prozess $q_{t+1} = p_{t+1} - \mu t$ ist dann ein Martingal.

Aufgrund der Tatsache, dass ein Random-Walk-Prozess die Martingal-Eigenschaft hat, liefert die Annahme des Random-Walk-Prozesses für die Preisbildung einer Aktie eine Darstellung der MEH für den Aktienmarkt.

Die mathematische Interpretation der ökonomischen Idee der MEH ist mit vielen Annahmen verbunden, die an erster Stelle die Beschreibung der Ergebnisse stark erschweren. Die häufigsten Annahmen bezüglich der Verteilung der Störterme (interpretiert als die unvorhersehbaren Nachrichten) sind die Annahmen der unabhängigen und normalverteilten Zuwächse. Beide Annahmen vereinfachen zwar die statistischen Tests, allerdings entsprechen sie kaum der Realität. Unter der Annahme der Normalverteilung für die Störterme besteht immer eine positive Wahrscheinlichkeit, dass der Preis der Aktie negativ werden kann. Dies entspricht allerdings nicht dem realen Handel.

Eine weitere Annahme bezüglich der Störterme, nämlich dass sie unabhängig und identisch verteilt sind (*iid*), ist ebenfalls für die langfristigen Finanztitel nicht plausibel. Ein Beispiel liefert die zweihundertjährige Geschichte der New Yorker Börse. Die unzähligen Änderungen kann man in den ökonomischen, sozialen, technologischen und institutionellen Entwicklungen des Landes während der letzten zwei Jahrhunderte beobachten. Die Faktoren, die die Aktienpreise eventuell beeinflussen, wurden mehrere Male geändert. Die *iid*-Annahme kann deswegen aus ökonomischen Gründen vereinfacht werden. Die Zuwächse können unabhängig über die Zeit bleiben, brauchen aber beispielsweise nicht mehr identisch verteilt sein. Allerdings bringt diese Abschwächung alle Probleme für die statistischen Tests, die mit der '*Heteroskedastizität*' der Störterme verbunden sind.

4.1.3 Die MEH und empirische Nachweise

Eine verbleibende Frage ist, ob die Markteffizienz-Hypothese testbar ist. Obwohl die empirische Methodologie sich gut bewährt hat, existieren immer einige Schwierigkeiten bei der Interpretation der Ergebnisse. Erstens, muss bei jedem statistischen Test der Markteffizienz eine Annahme bezüglich eines Gleichgewichtsmodell, das die normale Rendite eines Wertpapiertitels beschreibt, getroffen werden. Wenn die MEH abgelehnt wird, kann es wegen tatsächlicher Ineffizienz des Marktes passieren oder weil

⁵Im Weiteren notiert *iid* die *identically and independently distributed* Zufallsshocks mit dem Mittelwert Null und der konstanten Varianz σ^2 .

das Gleichgewichtsmodell falsch spezifiziert wurde. Dieses 'gemeinsame Hypothesen'-Problem zeigt, dass die MEH letztendlich nie abgelehnt werden kann. Zweitens, eine perfekte Effizienz der Märkte ist eine unrealistische Bezugsgröße, die beim realen Handel nicht vorkommen kann. Grossman und Stiglitz (1980) zeigen, dass die Überschussgewinne dann existieren können, wenn es Kosten für das Aufbereiten der Informationen geben würden. Diese Überschussgewinne sind eben notwendig, um den Investoren die Kosten des Suchens und der Verarbeitung der Informationen zu kompensieren. Die Gewinne sind nicht länger überproportional, wenn die Informationskosten richtig ausgewiesen werden. In den großen und liquiden Märkten können diese Informationskosten nur sehr kleine Überschussgewinne rechtfertigen. Es ist auch sehr schwierig zu sagen, wie klein die Gewinne sind, gerade weil die Kosten nicht präzise gemessen werden können.

Die Notation der relativen Markteffizienz, d.h. die Effizienz eines Marktes gemessen an einem anderen Markt, beispielsweise New Yorker Stock Exchange versus Paris Bourse, eine Terminbörse im Vergleich zu einem Kassamarkt, oder eine Auktion im Vergleich mit einem Handelsmarkt, kann ein eher verwendbares Konzept sein. Die Vorteile der relativen gegenüber der absoluten Markteffizienz können durch folgende Analogie gesehen werden. Die Bewertung der Effizienz eines physischen Systemes basiert auf der Relation des Inputs zum Output des Systems. Das Input/Output-Verhältnis oder der sogenannte Wirkungsgrad eines Motors gibt die Auskunft in Prozent darüber, wie viel nutzbare Energie durch das Verbrennen des Kraftstoffs freigesetzt wird. Der Rest geht auf die Wärmeverluste zurück. Ein System, das zum Beispiel ohne Wärmeverluste funktioniert, wird als 100%ig effizient betrachtet.

Die Ansicht der Ingenieure würde eher zu dem Konzept der statistischen Tests passen, um zu entscheiden, ob ein Markt im Vergleich zu einem perfekten Markt effizient ist. Das Messen der relativen Effizienz relativ zum reibungslosen Ideal ist bei vielen statistischen Untersuchungen angewendet worden. Analog zum technischen Fall ist die perfekte Markteffizienz eine Idealisierung, die in der Ökonomie nicht vorkommt, aber dennoch eine sinnvolle Bezugsgröße für die Messung der relativen Markteffizienz darstellt.

Seit dreißig Jahren spielt die MEH eine große Rolle, sowohl für das Verständnis der Aktienmarktbeziehung als auch für die Vorhersagen der Renditen⁶. Malkiel (2003) stellt viele empirische Untersuchungen der MEH aus den letzten Jahren vor und diskutiert ihre Ergebnisse ausführlich. Eine Verbindung der Informationseffizienz zur Renditenprognosen steht im Mittelpunkt seiner Kritik.

Eine der Interpretationen der MEH ist, dass keine überdurchschnittlichen Gewinne ohne zusätzliches Risiko möglich sind, d.h. ein effizienter Markt keine Arbitrage erlaubt. Die gefundenen linearen Abhängigkeiten der Renditen oder der Aktienpreise können

⁶Vgl. die Arbeiten von Campbell (1987), Campbell und Shiller (1988) und (1998), Fama (1998), Fama und French (1988, 1993, 1997, 2001, 2002), Lo und MacKinlay (1999), Shiller (2000, 2003).

nach der Auffassung der Autoren eine Arbitrage-Möglichkeit für die Marktteilnehmer bieten. Bestimmte saisonale oder Tageseffekte, wie der sogenannte *'Januar Effekt'*⁷ oder der *'Montagseffekt'*⁸ sollen ebenso *'kostenlose'* Gewinne versprechen. Die Prognose der Renditen, die beispielsweise die ausgezahlten Dividenden oder das Preis-Gewinn-Verhältnis benutzen, liefern ebenfalls den Hinweis für Ineffizienz der Märkte. Einige Untersuchungen zeigen, dass die sogenannten *'value'*-Aktien höhere Renditen als die *'growth'*-Aktien bringen⁹. Für einen Investor ist diese Regel ein Signal für eine mögliche Arbitrage und gegen die Effizienz des Marktes.

Die Argumentation der Kritiker der MEH kommt aus der Sicht der Marktgeschichte der letzten Jahre, die keine plausible Erklärung der Bewegung der Aktienmärkte anbietet. Eher eine emotionale als die rationale Überlegung vieler Investoren soll eine dominante Rolle bei der Aktienbewertung spielen. Als ein Beispiel einer falschen oder irrationalen Marktbewertung wird der *'Aktiencrash'* von Oktober '87 genannt, als die internationalen Aktienmärkte über 30% des eigenen Wertes im Vergleich zu wichtigen ökonomischen Indikatoren verloren haben. Die *'Internetblase'* in den späteren 90er Jahren kann ebenfalls nur durch ein irrationales Verhalten der Investoren begründet werden.

Ein Vergleich der von professionellen Managern geführten Aktienfonds mit der Performanz des Marktes liefert entsprechend den direkten Test der MEH. Wenn die Manager höhere Gewinne als der Markt mit ihren Portfolios erwirtschaften können, bedeutet das, dass nicht alle relevanten Informationen des Preisbildungsprozesses auf dem Markt berücksichtigt wurden. Die Manager nutzen dieses zusätzliche Wissen, um entsprechend die überdurchschnittlichen Rendite zu erzielen.

Für mehrere gefundene und vermutete Phänomene auf dem Aktienmarkt fand Malkiel (2003, 2005) eine Erklärung, die der Grundidee der MEH nicht widerspricht.

Bei den kurzfristigen Renditen¹⁰ wird oft eine positive Autokorrelation festgestellt. Die nachgewiesene Existenz der Autokorrelation bringt den Hinweis für eine lineare Abhängigkeit zwischen den zukünftigen und dem vergangenen Preis. Lo und MacKinlay (1999) meinen, dass es *'zu viel'* Gemeinsamkeit in der Aktienpreisbewegung auf dem Markt bedeutet, die zu einer Ablehnung der MEH in der Form eines Random-Walk-Prozesses führt. Darüber hinaus stellen Lo und andere (2000) fest, dass sich bestimmte Bildmuster in dem Aktienverlauf mit Hilfe nicht parametrischer statistischer Methoden erkennen lassen. Beispielsweise liefern solche Formationen wie *'head-and-shoulders'*

⁷Unter dem Januar Effekt wird die überproportionale Rendite der Aktien in den ersten fünf Tagen jedes Januars im Vergleich zu Dezember verstanden. Sieh Haugen und Lakonishok (1988).

⁸Eine überdurchschnittliche Rendite der Aktien am Montag, vgl. French (1980).

⁹Die *'value'*-Aktie unterscheidet sich von der *'growth'*-Aktie durch den Wert des Preis-Gewinn-Multiplikators und des Preis-Buchwert-Multiplikators. Der Preis-Buchwert-Multiplikator ist als der Wert aller Aktien der Firma abzüglich der Verbindungen dividiert durch die Anzahl der ausgegebenen Aktien definiert. Die *'value'*-Aktien sind durch niedrigen sowohl den Preis-Gewinn- als auch den Preis-Buchwert-Multiplikator charakterisiert.

¹⁰In diesem Fall wird die Aktie innerhalb einiger Tage oder Wochen beobachtet.

oder 'double-bottoms' eine Möglichkeit zu Prognosen¹¹. 'Januar-' oder 'Montageffekt' wurden ebenso bei den kurzfristigen Renditen beobachtet.

Bei der Diskussion dieser Effekte (die Autokorrelation oder das Muster in den Preisen) werden solche Aspekte berücksichtigt wie die psychologischen Mechanismen, die statistische und ökonomische Signifikanz, die Dauer und die Konsistenz der Effekte und die Anwendbarkeit dieser Regel für den Handel.

Ökonomen und Psychologen aus dem Feld der Verhaltenstheorie der Finanzmärkte¹² bringen die kurzfristigen Effekte in Zusammenhang mit einem psychologischen 'feedback' Mechanismus. Individuen beobachten die zunehmenden Preise auf dem Aktienmarkt und steigen in den Markt ein, um sich der erfolgsversprechenden Sache anzuschließen. Dieses Verhalten produziert eine Art des 'bandwagon' Effekts. Shiller (2000) erklärt die steigenden Aktienkurse in den USA in den späteren Neunzigern als ein Resultat der 'psychologischen Ansteckung', die letztendlich zu einer irrationalen Bewertung der Preise führte.

Es ist wichtig, zwischen der statistischen und ökonomischen Signifikanz zu unterscheiden, weil ein Aktienmarkt keinen mathematisch perfekten Random-Walk-Prozess darstellen kann. Einerseits wurde die kurzfristige Autokorrelation in mehreren Untersuchungen statistisch signifikant nachgewiesen, andererseits ist sie relativ klein, um bei einer gewinnversprechenden Handlungsstrategie Anwendung zu finden, weil die Transaktionskosten die Profitabilität solcher Handelsstrategien zunichte machen¹³.

Die Beständigkeit der Effekte spielt eine entscheidende Rolle bei der Diskussion über die MEH. Beispielsweise begutachtet Fama (1998) sämtliche empirische Untersuchungen, die unter anderem eine ökonomische Erklärung der Reaktion der Aktienmärkte auf Informationsschocks zu finden versuchen. Zu den News gehören solche Bekanntgaben wie die Dividendenausschüttung oder die Aktiensplits. Fama (1998) zeigt, dass viele der Renditenanomalien ausschließlich in einem Kontext mit dem spezifischen Modell gefunden werden konnten. Die Anomalien tendieren zu verschwinden, wenn ein anderes Modell der normalen Rendite, die verschiedenen Methoden der Risikoanpassung oder die anderen statistischen Messmethoden bei den Untersuchungen angewendet werden. Die Anomalien sind auch nicht lange beobachtbar nach ihrer Entdeckung. Eine Erklärung dafür bringt Schwert (2001). Er meint, dass die Marktteilnehmer relativ

¹¹Die 'head-and-shoulders' Struktur (HS) des Verlaufs von einem Aktienpreis ist ein technisches Bildmuster, das sehr oft bei der technischen Analyse für die Prognosen verwendet wird. Das HS Muster besteht aus den fünf lokalen Extremwerten zu den Zeitpunkten t_1, \dots, t_5 .

$$HS \equiv \begin{cases} E_1 \text{ ist ein Maximum} \\ E_3 > E_1, E_3 > E_5 \\ E_1, E_5 \text{ liegen innerhalb des } 1,5\% \text{ Intervalls um ihren Durchschnitt} \\ E_2, E_4 \text{ liegen innerhalb des } 1,5\% \text{ Intervalls um ihren Durchschnitt} \end{cases}$$

¹²Behavioral Finance

¹³Vgl. Odean (1999).

schnell mögliche Bildmuster für die Prognosen erkennen und sie bei eigenen Strategien nutzen. Dies führt dazu, dass diese so aufgebauten Prognosen nicht mehr profitabel werden. Malkiel (2003) ist der Meinung, dass die Anomalien in kurzfristigen Renditen nie signifikant groß oder stabil waren, um einen überdurchschnittlichen Gewinn aus der Aktieninvestition zu erwirtschaften.

Im Gegensatz zu kurzfristigen Renditen wird eine negative Autokorrelation der langfristigen Renditen nachgewiesen (vgl. Fama und French (1988) oder Poterba und Summers (1988)). Dieses Phänomen kann mit Hilfe der Argumentation von DeBond und Thaler (1985) eventuell erklärt werden. Bei DeBond und Thaler sind Investoren den Wellen vom Optimismus und Pessimismus unterworfen. Dies führt zu permanenten Überreaktionen der Preise und entsprechend zu einer Abweichung von den Fundamentalwerten. Letztendlich ist eine Rückwärtsbewegung der Preise beobachtbar, die in der negativen Korrelation widerspiegelt wird. Eine andere Erklärung für die negative Korrelation der langfristigen Renditen wäre die Verbindung zum Zinssatz. Wenn der Zinssatz steigt, gehen die Aktien und Bondspreise nach unten, und bei sinkenden Zinsen steigen die Preise für die riskanten Anlagen. Wenn der Zinssatz seinerseits eine negative Autokorrelation aufweist, dann soll auch die Rendite des Aktienmarktes eine negative Autokorrelation haben. Deswegen widerspricht diese Feststellung nicht der MEH (vgl. Malkiel (2003)). Allerdings sind die Ergebnisse zur Autokorrelation der langfristigen Rendite längst nicht einheitlich. Der negative Effekt ist in einigen Perioden stärker und wiederum in anderen schwächer (vgl. Fluck, Malkiel und Quandt (1997)).

Die statistischen Untersuchungen von Fama und French (1988) und von Campbell und Shiller (1988) zeigen eine Möglichkeit, wie die Renditen mit Hilfe der Dividenden prognostiziert werden können. Sie behaupten, dass über 40% der Varianz der zukünftigen Aktienmarktrenditen auf Basis der Dividendenauszahlungen vorhergesagt werden kann. Dieses Ergebnis ist nicht notwendigerweise inkonsistent mit der MEH. Die Dividendenauszahlung für eine Aktie tendiert hoch zu sein, wenn der Marktzins hoch ist. Deswegen kann die gefundene Dividenden-Preis-Verbindung als eine Anpassung des Aktienmarktes an den allgemeinen ökonomischen Bedingungen interpretiert werden. Allerdings kann die gefundene Dividenden/Preis-Verbindung, wie das folgende Beispiel zeigt, nicht bestätigt werden. Ein stark wachsender Aktienmarkt (bis zu 15% pro Jahr) in den Neunzigern folgt dem Markt mit niedrigen Dividendenzahlungen in den achtzigern (ca 3% pro Jahr). Die niedrigen Dividenden sind aber ein Indikator für die niedrigen zukünftigen Preise. Fama und French (2001) erklären das damit, dass die modernen Firmen die erwirtschafteten Gewinne eher in die Aktien reinvestieren, anstatt sie in Form von Dividenden auszuschütten.

Eine andere Variante der Prognose, die auf dem Preis-Gewinn-Multiplikator basiert, bringt ebenfalls weder wesentlich bessere noch eindeutige Ergebnisse. Die weiteren statistischen Regeln wie die Relation des kurzfristigen Zinses zur zukünftigen Rendite, die Termstruktur des Zinses, die Spreads zwischen den hochverzinslichen Bonds und dem kurzfristigen Zins, besitzen eine gewisse Prognosekraft. Allerdings erwiesen sich

sehr viele Handelsstrategien, die auf solchen Erkenntnissen basieren, als nicht profitabel aufgrund der hohen Transaktionskosten. Ebenso bringen die Starategien wenig Erfolg, die eine technische Analyse für die Prognosen benutzen. Das Capital-Asset-Pricing-Model (CAPM), das auch eine Verbindung zwischen der normalen Rendite und der Überschussrendite darstellt, kann für die Prognosen angewendet werden. Lokonishok, Shleifer und Vishny (1994) heben das Modell als ein Gegenargument zur MEH hervor. Malkiel (2003) meint allerdings, dass das CAPM nur ein Teil des Risikos abfangen kann und nicht das Gesamtrisiko des Marktes. Dies widerspricht der Idee der Informations-effizienz wiederum nicht.

Eine kurze Zusammenfassung der Ergebnisse kann wie folgt formuliert werden: Bestimmte quantitative Effekte in dem Verhalten der Aktienrenditen wurden in den letzten Jahren gefunden. Diese Effekte können der Idee der MEH widersprechen. Allerdings lassen sie sich relativ einfach entweder durch die Psychologie der Marktteilnehmer oder durch die ökonomischen Konzepte erklären. Die Vorhersagen der Aktienrenditen durch Gewinne oder Dividenden können beispielsweise die Anpassung der Aktienmärkte zu ökonomischen Bedingungen reflektieren. Diese Effekte sind nicht von langer Dauer, und sie verschwinden sehr oft nach ihrer Bekanntmachung. Mehrere von den Effekten sind ein Ergebnis des Data-Mining. Die Prognosen, die auf solchen Effekten basieren, sind weniger robust und nicht profitabel. Ein Zitat von Roll (1992) illustriert das noch deutlicher:

„I have personally tried to invest money, my client’s money and my own, in every single anomaly and predictive device that academics have dreamed up ... I have attempted to exploit the so-called year-end anomalies and a whole variety of strategies supposedly documented by academic research. And I have yet to make a nickel on any of these supposed market inefficiencies ... a true marketinefficiency ought to be an exploitable opportunity. If there’s nothing investors can exploit in a systematic way, time in and time out, then it’s very hard to say that information is not being properly incorporated into stock prices.“

Die zwei Aspekte in dem Zusammenhang mit der MEH sollen noch angesprochen werden, nämlich die überproportionalen Crashes der Aktienmärkte und die Performanz der professionellen Manager. Die Aktiencrashes von 1987 und 2000 könnten als ein Widerspruch zur MEH angesehen werden. Es stellt sich die Frage, wie die Märkte effizient sein konnten, wenn die Aktienpreise kurzfristig ein Drittel ihres Wertes verloren haben (z.B. vom Anfang bis Mitte Oktober 1987). Der Crash im Oktober ’87 und die *’Internetblase’* im Jahr 2000 sind eher eine Ausnahme aus der Regel und sie stellen den Preis für die Flexibilität der Märkte dar¹⁴.

¹⁴Vgl. Malkiel (2003):

...’bubble’ periods are the exception rather than the rule, and acceptance of such occasional mistakes is the necessary price of a flexible market system that usually does a

In seinem letzten Artikel (2005) stellt Malkiel den amerikanischen Aktienindex S&P 500 und eine große Anzahl der aktiv gemanagted Aktienfonds gegenüber. In den letzten 20 Jahren (bis 31.12.2003) überbot der Aktienindex fast 90% aller Aktienfonds in den USA. Seine durchschnittliche Rendite war 12,78% pro Jahr im Vergleich zu der durchschnittlichen Rendite der Aktienfonds von 10,54% pro Jahr. Insbesondere in der 'bubble'-Zeit verlor der S&P 500 durchschnittlich nur 5,34% pro Jahr (Periode von 31.12.99 bis 31.12.2003). Nur zwei aktive Fonds konnten kleinere Verluste von 0,5% pro Jahr von 'Legg Mason value Tr' und 4,09% pro Jahr von 'Fidelity new millenium' ausweisen. Die hohen Transaktionskosten und ebenso die unangemessenen Gehälter der Manager vernichten die versprochenen hohen Gewinne aus den aktiven Fonds. Dementsprechend überbieten die Manger nicht die Renditen der 'benchmarks' und liefern dadurch einen Hinweis, dass der Markt eben alle notwendige Information verarbeitet und wiedergibt.

Das Unterstellen des rationalen Verhaltens der Investoren prägt die MEH als die wichtigste Annahme. 'Behavioral Finance' ist die Theorie der Finanzmärkte, die das psychologische Verhalten der Investoren und den soziologischen Aspekt bei der Preisbildung auf Aktienmärkten berücksichtigt. 'Behavioral Finance' hat eine widersprüchliche Sicht zu MEH. Shiller (2003) kritisiert den Literaturüberblick zu dieser Theorie von Fama (1998). Fama weist auf zwei Schwachpunkte der 'Behavioral Finance'-Theorie hin, nämlich, dass die entdeckten Anomalien auf dem Aktienmarkt entweder durch eine Überreaktion oder eine Unterreaktion der Investoren argumentiert werden, oder sie verschwinden nach ihrer Entdeckung. Nach Ansicht von Shiller existiert kein fundamentales psychologisches Prinzip, das genau beschreibt, wann die Leute zu einer Überreaktion oder einer Unterreaktion neigen. Deswegen ist es nicht überraschend, dass die Untersuchungen der Anomalien auf den Finanzmärkten keine eindeutigen Reaktionen feststellen konnten. Der zweite Kritikpunkt von Fama ist ebenfalls nach Meinung von Shiller schwach, weil bestimmte Effekte, beispielsweise die ansteigende Volatilität der Märkte vor den 'Blasen', durch das Vorwissen doch nicht eliminiert werden konnten. Darüber hinaus, der Fakt, dass einige Anomalien manchmal verschwinden oder ihre Richtung mit der Zeit ändern, bringt keinen Nachweis, dass die Märkte voll rational sind. Genau das zeichnet eher einen irrationalen Markt aus. Die grundlegende Relation nämlich, die bei Ineffizienz der Märkte vorgeschlagen wird, ist keine Verbindung, die einfach zu testen ist und die über alle Zeitperioden halten soll. Der Fundamentalwert einer Aktie ist ebenfalls mühevoll zu messen, darüber hinaus, wenn eine 'spekulative Blase' für eine lange Zeit hält, dann kann die Fundamentalbewertung nicht während der langen Stichprobenperioden beobachtet werden.

very effective job of allocating capital to its most productive uses.

4.2 Portfolio-Theorie

Die Informationseffizienz der Aktienmärkte kann die eventuelle Unabhängigkeit der Preisbewegung der Finanztitel begründen. Die Portfolio-Theorie spiegelt die einfache Idee der Risikostreuung durch eine Investition in die verschiedenen Wertpapiere wieder. Beispielsweise stellt der Branchen-Aktienindex wie 'Automobile' oder der allgemeine Aktienindex wie 'DAX' ein breit gestreutes Portfolio dar. Eine Investitionsentscheidung kann von der Bewertung des Ausgleichs zwischen dem Risiko und der erwarteten Rendite des Portfolios abhängig sein. Die theoretischen Ansätze dieser Bewertung stellen wichtige Aufgabenfelder der modernen Portfolio-Theorie dar.

Die Idee, dass ein hoch eingegangenes Risiko auch eine höhere Rendite verspricht, ist ein fester Bestandteil der Portfolio-Theorie¹⁵. Die risikoscheuen Investoren konzipieren ihr Portfolio nach dem Minimum-Risiko-Prinzip. Die risikofreudigen Investoren werden eher versuchen, maximale Gewinne mit ihren Portfolios zu erwirtschaften, wenn die Gewinne dem Nutzen der Investoren entsprechen. Ein risikoreiches oder risikoarmes Portfolio kann aus den inländischen Wertpapieren zusammengesetzt werden. Dementsprechend kann das Portfoliorisiko beispielsweise durch eine Branchen-Diversifikation reduziert werden. In dieser Arbeit wird nach den Erfolgsaussichten der internationalen Diversifikation in den Industrieländern nachgefragt. Entsprechend beschäftigt sich der nachfolgende Teil mit den theoretischen Ansätzen und empirischen Untersuchungen auf diesem Teilgebiet der Portfolio-Theorie.

Die Rendite eines Portfolios, das ausschließlich aus Aktien besteht, ist ein gewichteter Durchschnitt der Renditen von einzelnen Aktien $R_P = \sum_{i=1}^N (g_i R_i)$. Mit R_i wird die Rendite einer einzelnen Aktie bezeichnet. Das Gewicht der Aktie i wird entsprechend mit g_i bezeichnet. Die erwartete Rendite eines Portfolios besteht dann aus den erwarteten Renditen der einzelnen Aktien:

$$E_{t-1}(R_{P,t}) = \sum_{i=1}^N E_{t-1}(g_i R_{i,t}) = \sum_{i=1}^N g_i E_{t-1}(R_{i,t}). \quad (4.7)$$

Unter dem Risiko eines Portfolios versteht man die Standard-Abweichung der Portfoliorendite vom erwarteten Wert, $\sigma_{P,t}$:

$$\sigma_P = [E(R_P - E(R_P))^2]^{1/2} = \left[\sum_{i=1}^N (g_i^2 \sigma_i^2) + \sum_{i=1}^N \sum_{\substack{k=1 \\ k \neq i}}^N (g_i g_k \sigma_{ik}) \right]^{1/2}. \quad (4.8)$$

Die σ_i^2 bezeichnet die Varianz der Renditen der einzelnen Aktien und σ_{ik} stellt daher die Kovarianz zwischen diesen Renditen dar. Wenn die Aktienrenditen nicht korreliert¹⁶

¹⁵Vgl. Elton und Gruber (2003), Kapitel 1.

¹⁶Es besteht keine lineare Abhängigkeit der Renditen, damit ist der Korrelationskoeffizient $\rho = \frac{\sigma_{ik}}{\sigma_i \sigma_k}$ identisch Null.

sind, verkleinert sich das Risiko des Portfolios um den Kovarianzterm:

$$\sigma_P = \left[\sum_{i=1}^N (g_i^2 \sigma_i^2) \right]^{1/2}.$$

Im Fall einer perfekten positiven (oder negativen) Korrelation¹⁷ zwischen den Renditen der einzelnen Aktien steht vor dem Summand der Kovarianzen ein Plus- (oder Minus-) Zeichen:

$$\sigma_P = \left[\sum_{i=1}^N (g_i^2 \sigma_i^2) + (-) \sum_{i=1}^N \sum_{\substack{k=1 \\ k \neq i}}^N (g_i g_k \sigma_i \sigma_k) \right]^{1/2}. \quad (4.9)$$

Folglich verkleinert sich das Risiko eines Aktienportfolios, wenn die Kovarianzterme überwiegend ein negatives Vorzeichen haben.

Die nächste wichtige Entscheidung, die ein Investor treffen muss, ist die optimale Gewichtung des Portfolios. Es entsteht dadurch ein Optimierungsproblem, das sich wie folgt formulieren lässt:

es wird das Minimum der Risikofunktion

$$\sum_{i=1}^N (g_i^2 \sigma_i^2) + \sum_{i=1}^N \sum_{\substack{k=1 \\ k \neq i}}^N (g_i g_k \sigma_i \sigma_k) \quad (4.10)$$

unter den folgenden Nebenbedingungen gesucht:

1. $\sum_{i=1}^N g_i = 1$
2. $\sum_{i=1}^N g_i E_{t-1}(R_{i,t}) = E_{t-1}(R_{P,t})$
3. $g_i \geq 0, i = 1, \dots, N.$

Die Variation der erwarteten Rendite $E(R_P)$ zwischen der Rendite des Minimum-Varianz-Portfolios und der Rendite des Maximum-Rendite-Portfolios liefert einen optimal ausgewählten Satz der Gewichte g_i . Das aufgezeichnete Problem ist aus der Klasse quadratischer Optimierung, weil in den Gleichungen die quadratischen Terme g_i^2 und $g_i g_k$ vorkommen.

¹⁷Die Formel für das Risiko eines Portfolios kann wie folgt umgeschrieben werden:

$$\sigma_P = \left[\sum_{i=1}^N (g_i^2 \sigma_i^2) + \sum_{i=1}^N \sum_{\substack{k=1 \\ k \neq i}}^N (g_i g_k \rho_{ik} \sigma_i \sigma_k) \right]^{1/2}.$$

Im Fall einer perfekten positiven Korrelation ist jeder Korrelationskoeffizient $\rho_{ik} = 1$, damit hat man die Gleichung (4.9) mit einem positiven Vorzeichen beim Kovarianzterm.

Aufgrund der Definition des effizienten Portfolios verbindet man die Portfolioanalyse eher mit der Korrelationsanalyse. Die Nachteile dieser Analyse liegen allerdings in der rein statistischen Ausrichtung, die ebenfalls auf rein rationalen mathematischen Erwartungen basiert. Weder eine ökonomische Theorie noch ein psychologischer Aspekt der Prognosebildung werden in die Analyse einbezogen. Die effizienten Portfolios werden ausschließlich auf der Basis der ex-post Daten gebildet. Zahlreiche Beispiele mit eher bescheidenen Gewinnen aus der internationalen Diversifikation kann man in Elton und Gruber (2003), bei Reilly und Brown (2003) oder bei Malkiel (2005) finden.

4.2.1 Die internationale Diversifikation

Seit Jahrzehnten empfehlen Investmentberater den Investoren, ihre Portfolios durch internationale Titel zu ergänzen. Die Portfolio-Manager aus Frankreich, Deutschland und USA investieren immer mehr in ausländische Märkte¹⁸. Nach wie vor bleibt für Investoren und Manager die Frage interessant, weshalb über die inländischen Grenzen hinaus investiert wurde. Mögliche Gründe können die Minderung der Verlusten sein, die aus inländischen Investitionen resultieren, oder aber die Absicht, höheren Renditen zu erzielen.

Die Möglichkeit, das Risiko zu reduzieren, ohne auf die Rendite zu verzichten, ist zum Hauptziel des modernen Portfolio-Managements geworden. Nach der Portfolio-Theorie soll ein Investor in sein Portfolio eine Klasse internationaler Aktien aufnehmen, die eine niedrigere Korrelation zu den einheimischen Aktien aufweist. Die Risikoreduktion beinhaltet hier eine Kompensation der inländischen Verluste mit den ausländischen Gewinnen und umgekehrt. Trotz vieler Argumente haben die Investoren kein Bedürfnis, beispielsweise aus einem *'bullischen'*¹⁹ Markt auszusteigen, um sich durch die Diversifikation jeder Art gegen einen eventuellen zukünftigen Absturz des Marktes abzusichern. Es scheint, die Investoren würden die steigenden Gewinne dem sinkenden Risiko vorziehen. Unglücklicherweise ist eine genaue Vorhersage der unterschiedlichen Phasen des Aktienmarktes unmöglich. Daher bedeutet ein spekulatives Spiel ein hohes Verlustrisiko. Insgesamt vermeidet die internationale Diversifikation große Verluste, wenn der inländische Markt fällt. Allerdings führt sie auch zu kleineren Gewinnen bei steigenden einheimischen Aktien.

Die traditionelle Meinung über die internationale Diversifikation ist identisch mit dem Gedanken der Investoren, die die festverzinslichten Wertpapiere mit einem niedrigen Zinsniveau in ihren Portfolios halten. Es handelt sich also um eine Risikoreduktion oder eine Verlustreduzierung. Für die Investoren, die nach großen Gewinnen aus einem Aktiengeschäft streben, ist ein Portfolio-Management völlig ungeeignet. In seinem Artikel aus siebziger Jahren schreibt Solnik zu diesem Thema:

¹⁸Vgl. Elton und Gruber (2003), Kapitel 12.

¹⁹Über einen *'bullischen'* Aktienmarkt spricht man, wenn die Aktien während einer bestimmten Periode steigen.

“...primary motivation in holding a portfolio of stocks is to reduce risk“.

Er demonstrierte schon vor 30 Jahren, wie das Niveau des systematischen Portfoliorisikos sinkt, wenn die internationalen Aktien in dieses Portfolio aufgenommen werden. Solnik schätzt die Volatilität der Renditen der international diversifizierten Portfolios und kommt zu dem Ergebnis:

„an internationally well-diversified portfolio would be half as risky as a portfolio of US stocks...“²⁰.

Michaud, Bergstrom, Frashure und Wolahan (1996) bewerteten die Daten der letzten zwanzig Jahren und kamen zum Ergebnis:

„...international diversification increases return per unit of risk relative to a comparable US-only portfolio.“

Die letzten zehn Jahre der Aktienmarktgeschichte passen den Forscher und Investoren allerdings nicht in das Konzept. Hanna, McCormack und Perdue (1999) zeigen, dass jedes aus dem „S&P 500“-Index und einem internationalen Index der G7-Länder konstruierte Portfolio allein durch „S&P 500“ dominiert wird. Für die neu entstandenen Aktienmärkte, die laut Portfolio-Theorie ein hohes Potenzial für Gewinne aus internationaler Diversifikation versprechen, haben Fifield, Power und Singlar (2002) festgestellt:

„... a strategy based on forecasted means, standard deviations and correlations, achieved very few of the gains attained in ex-post analyses of emerging market diversification.“

Die widersprüchlichen Ergebnisse bezüglich der internationalen Diversifikation unterstreichen die Intention dieser Arbeit.

Aus Sicht eines inländischen Investors berechnet sich die Rendite einer ausländischen Anlage aus der Rendite, die diesen Investitionstitel auf dem ausländischen Markt erwirtschaftet, und aus den Änderungen des entsprechenden Wechselkurses der ausländischen Währung:

$$1 + R_d = (1 + R_f)(1 + R_e) \quad (4.11)$$

Mit R_d wird die inländische Rendite bezeichnet. R_f ist die Rendite der Anlage im Ausland und R_e ist eine Änderung des Wechselkurses. Eine Vereinfachung dieser Gleichung bringt das folgende Verhältnis zwischen den Renditen

$$R_d = R_f + R_e + R_f R_e.$$

²⁰Die Untersuchung wurde aus der Sicht eines amerikanischen Anlegers gemacht.

Der letzte Term in der Gleichung ist wesentlich kleiner als die anderen beiden Terme, so dass die Rendite für den inländischen Investor approximativ aus der Rendite des Titels auf dem ausländischen Markt und dem Wechselkursgewinn oder dem Verlust bestehen kann. Für die erwartete Rendite wird die folgende Gleichung aufgestellt:

$$E_t(R_d) = E_t(R_f) + E_t(R_e). \quad (4.12)$$

Entsprechend ist das Risiko der ausländischen Anlage durch die Standard-Abweichung der Rendite von der erwarteten Rendite gegeben:

$$\sigma_d = [\sigma_f^2 + \sigma_e^2 + 2\sigma_{f,e}]^{1/2} \quad (4.13)$$

Die Rendite eines internationalen Portfolios lässt sich wie folgt berechnen:

$$E_t(R_{P,d}) = E_t(R_{P,f}) + E_t(R_e) = \sum_{i=1}^N g_i E(R_i) + \sum_{j=1}^k E(R_{e_j}). \quad (4.14)$$

Die Anzahl der verschiedenen Länder, in denen die Diversifikation durchgeführt wurde, wird mit k bezeichnet. Um eine Portfolioentscheidung durchführen zu können, sind geschätzte Werte der erwarteten Rendite, der Standard-Abweichung der Rendite und des Korrelationskoeffizienten zwischen der Rendite und dem Wechselkurs notwendig.

Der Effekt des Wechselkursrisikos auf die Volatilität der Rendite einer ausländischen Anlage kann wie folgt beschrieben werden²¹: Die Volatilität der Aktien in der Fremdwährung steigt üblicherweise durch das Wechselkursrisiko. Das Risiko der internationalen Aktienportfolios lässt sich allerdings durch die Absicherung der Währung grundsätzlich minimieren. Beispielsweise kann parallel zum Kauf einer Auslandsaktie ein Kontrakt für den Verkauf der entsprechenden ausländischen Währung abgeschlossen werden. Die einzige Schwierigkeit dabei ist, den genauen Verkaufszeitpunkt der Aktie festzulegen, weil der Optionskontrakt immer zeitlich determiniert ist. Eine Auswirkung hat die Absicherung des Wechselkursrisikos auf die entsprechenden Korrelationskoeffizienten, nämlich dass sie im Durchschnitt bis zu 4%, wie Elton und Gruber (2003) zeigen, und sogar bis zu 13% bei Kaplans und Schaefer (2001) sinken. Der Effekt der Minderung des Wechselkursrisikos bei der erwarteten Rendite ist weniger eindeutig. Er variiert von Land zu Land und von Periode zu Periode. Zum Vergleich steht das Ergebnis der inländischen ex-post-Untersuchung im vierten Kapitel der Arbeit.

Der Korrelationskoeffizient zwischen den internationalen Aktien ist relativ gesehen zur Korrelation der inländischen Aktien niedrig. Die Integration der europäischen Märkte und die wachsende Abhängigkeit der Weltmärkte spiegelt sich in den steigenden Korrelationskoeffizienten wieder²². Allerdings haben die Ländergruppen mit relativ

²¹Die entsprechenden Tabellen sind für die Renditen und die Varianz der internationalen Anlagen, die in amerikanischer Währung und in französischen Franc konvertiert sind, in Elton und Gruber (2003), Kapitel 12 dargestellt.

²²Vgl. Elton und Gruber (2003), Kapitel 12.

hochintegrierten Ökonomien wie USA und Kanada, Benelux oder die skandinavischen Länder nach wie vor eine relativ niedrige Korrelation zwischen den Gruppen im Vergleich zu den inländischen Koeffizienten. Aus theoretischer Sicht stellt dies eine Möglichkeit zur Risikoreduktion in den internationalen Portfolios dar, die Aktien aus diesen Ländern beinhalten. Bezüglich der Erträge gibt es bis jetzt immer noch keine ökonomische Argumentation, weshalb die Renditen der ausländischen Märkte höher als der inländischen Märkte sein sollten.

Einen weiteren Hinweis für die Güter der international diversifizierten Portfolios liefern die Untersuchungen zur Performanz der internationalen Aktienfonds²³. Als ein Bezugswert werden die inländischen Indizes wie 'S&P 500' für den amerikanischen Markt oder 'MSCI Europa index' für Europa beispielsweise bei Malkiel(2005) genommen. Malkiel (2005) kommt zum Ergebnis:

„... past performance does not predict future returns. ... Index funds tend to outperform actively managed funds in international as well as in domestic markets.“

Als einen der Gründe der schlechten Performanz der Fonds nennt Carhard (1997):

„...the investment cost of expense ratios, transaction costs, and load fees all have a direct, negative impact on performance“.

Für Malkiel (2005) ist die Erklärung die Gültigkeit der MEH.

4.2.2 Das Capital-Asset-Pricing-Model

Die grundlegende Annahme der Portfolio-Theorie basiert auf der Risikoaversion der Anleger. Eine geeignete Handelsstrategie für diese Anleger ist der Erwerb von Aktien mit unterschiedlichen Risiken. Eine Diversifikationsstrategie ist fast immer vorteilhaft für risikoscheue Anleger, weil sie das Gesamtrisiko der Anlage reduziert. Dennoch bringt dies im Extremfall, wenn die Erträge der Aktien perfekt korrelieren, keinen Vorteil für den Anleger. Allerdings kann das systematische Risiko²⁴ durch eine Diversifikationsstrategie nicht beseitigt werden. Das Capital-Asset-Pricing-Model (CAPM) gibt Aufschluss über die Bewertung des systematischen Risikos. Das Modell erklärt

²³Vgl. Malkiel (1995), Gruber (1996), Elton, Gruber und Blake (1996), Carhart (1997), Carhart, Carpenter, Lynch, und Musto (2002), Malkiel (2003), Daeves (2004), Malkiel (2005).

²⁴Das Gesamtrisiko einer Anlage besteht aus zwei Komponenten, dem systematischen und dem nicht systematischen Risiko. Die erste Komponente kann durch fundamentale Faktoren bestimmt werden. Die zufälligen Schocks auf den Aktienmärkten sowie Nachrichten über Bilanzen der Unternehmen können zu dem nicht systematischen Risiko zugeordnet werden.

die Größe der Risikoprämie, die sich durch die Differenz zwischen einem erwarteten Ertrag einer Aktie und einem Ertrag einer risikolosen Anlage definiert. Das CAPM gehört zu einer Klasse der allgemeinen Gleichgewichtsmodelle. Diese ermöglichen eine Verbindung zwischen der erwarteten Rendite und dem Risiko, wenn die Märkte in einem Gleichgewicht sind. Darüber hinaus haben Gleichgewichtsmodelle Auswirkungen auf die Charakteristika eines optimalen Portfolios.

Das Standard-CAPM²⁵

Die Grundlage für das CAPM wurde von Markowitz (1959) entwickelt. Sharpe (1964) und Lintner (1965) erweiterten die Arbeit von Markowitz mit dem Ergebnis, dass die Investoren mit homogene Erwartungen und den Mittelwert-Varianz-effizienten Portfolios führen dazu, dass das Marktportfolio selbst zu einem effizienten Portfolio wird. So kann keine Arbitragemöglichkeit durch Diversifikation jeder Art entstehen. Damit lässt sich die CAPM-Gleichung als eine direkte Folge der Mittelwert-Varianz-Effizienz des Marktportfolios interpretieren.

Zu den wichtigsten Faktoren, die die Nachfrage der Investoren nach Aktien beeinflussen, gehören das Einkommen oder Vermögen der Anleger, die erwartete Rendite der Aktie, das mit der Aktie verbundene Risiko und die Liquidität der Aktie. Sie werden am meisten mit den Eigenschaften einer alternativen Anlage verglichen. Die Aktiennachfrage kann positiv oder negativ auf die Änderungen dieser Größen reagieren. Die Nachfrage reagiert beispielsweise negativ auf das mit der Aktie verbundene Risiko. Steigt dieses Risiko im Vergleich zu dem Risiko einer alternativen Anlage an, fällt die Nachfrage nach dieser Aktie.

Die folgenden Annahmen bilden die Voraussetzungen für das Berechnen der Risikoprämie. Die erste Annahme basiert auf der Existenz eines vollkommenen Kapitalmarktes im Gleichgewicht und der Tobin-Separation²⁶. Die zweite Annahme betrifft die Anleger, deren Verhalten als risikoscheu angenommen wird, die eine Vermögensmaximierung erzielen wollen, und homogene Erwartungen, risikolose sowie unbeschränkte Anlage- und Verschuldungsmöglichkeiten haben. Es wird unterstellt, dass die Wertpapiere beliebig teilbar sind. Zusätzlich werden keine Transaktionskosten und keine Steuern erhoben.

Die Risikoprämie wird folgendermaßen berechnet:

$$E_t(R_i) - R_f = \beta_{i,m}(E_t(R_m) - R_f). \quad (4.15)$$

Wobei $E_t(R_i) = E(R_{i,t+1}|\Omega_t)$ eine erwartete Rendite der Aktie i zum Zeitpunkt $(t+1)$ unter Berücksichtigung aller Informationen über die Preisentwicklung bis zum Zeit-

²⁵Vergleiche Standard-CAPM und Zero-Beta-CAPM aus Campbell (1997)

²⁶Die Tobin-Separation unterstellt, dass die Struktur des Portfolios unabhängig von den Präferenzen der Anleger ist. Es gibt aber eine Unterscheidung zwischen einem Portfolio und einer alternativen risikolosen Anlage.

punkt t ist. $E_t(R_m)$ beschreibt eine erwartete Rendite des Marktportfolios, die durch zum Beispiel einen zusammengesetzten Aktienindex dargestellt werden kann. R_f ist die Rendite einer risikolosen langfristigen Anleihe.

Der Koeffizient

$$\beta_{i,m} = \frac{Cov(R_i, R_m)}{Var(R_m)}$$

beschreibt das systematische Risiko, das durch Diversifikation nicht eliminiert werden kann. Der Koeffizient zeigt, wie unabhängig die Aktienrendite von der Bewegung des Marktes ist. Die Bedingung $\beta_{im} > 1$ impliziert eine risikobehaftete Aktie (aggressive) und $\beta_{im} < 1$ eine defensive Aktie.

Statt für eine einzelne Aktie gilt das CAPM für ein Portfolio aus mehreren Aktien. Dadurch wird die Risikoprämie eines Portfolios $E_t(R_P) - R_f$ durch $\beta_{P,m}(E_t(R_m) - R_f)$ erklärt. Das Optimierungsproblem des Portfolios bleibt, wie in die Gleichung (4.10) eingeführt ist, erhalten.

Das CAPM begründet, dass die Kalkulation der erwarteten Rendite einer risikobehafteten Anlage von zwei Faktoren abhängig ist. Erstens, spielt bei ihrer Bewertung die Risikoprämie des Marktes $E_t(R_m) - R_f$ eine entscheidende Rolle. Zweitens ist sie durch ihre Komponenten determiniert, die mit dem Markt perfekt korreliert sind. Die Komponenten der Risikoprämie der Aktienanlage, die nicht mit dem Markt korrelieren, können diversifiziert werden, so dass sie diese Prämie nicht mehr beeinflussen.

Das CAPM führt zu einer Anzahl der wichtigen Schlussfolgerungen. Ein erster Aspekt ist, dass die erwartete Rendite von dem Risiko der Aktienanlage unabhängig ist. Es ist richtig, dass die Anlagen mit einem hohen β -Koeffizient auch ein hohes eigenes Risiko haben, weil ein Teil ihrer Volatilität durch den β -Koeffizient determiniert wird. Allerdings, um eine hohe Volatilität zu haben, braucht die Aktie keinen hohen β -Koeffizient zu besitzen. Die Aktie oder das Aktienportfolio mit einem hohen Risiko wird dann eine hohe erwartete Rendite haben, wenn dieses Risiko auf das gesamte Aktienmarktrisiko sehr empfindlich reagiert. Zweitens ermöglicht der β -Koeffizient in der CAPM eine Messung des Aktienrisikos, das nicht diversifiziert werden kann. Beispielsweise, wenn zwei Aktien die entsprechenden Koeffizienten von 0,8 und 1,4 haben, dann hat das 50/50 Portfolio den β -Koeffizient von 1,1. Allerdings, wenn das Verfahren für alle Aktien ausgeweitet wird und noch die Kapitalisierung des Unternehmens mit in die Gewichtung genommen wird, dann ist der Marktkoeffizient $\beta = 1$. Wenn der Markt sich in einem Gleichgewicht befindet, liegen alle einzelnen Aktien auf der 'Securities Market Line'. Wenn das nicht der Fall ist, werden die Investoren versuchen, den Markt zu überbieten, indem sie einen höheren Sharpe-Quotient²⁷ erzielen. Der Sharpe-Quotient der Anlage ist allerdings von dem Sharpe-Quotienten des Marktes abhängig²⁸.

²⁷Der Sharpe-Quotient ist durch $(E_t(r_i) - R_f)/\sigma_i$ definiert.

²⁸Aus der CAPM-Gleichung (4.15) folgt

$$E_t(R_i - R_f)/\sigma_i = Cov(E_t(R_i), E_t(R_m))/(\sigma_m^2 \sigma_i) E_t(R_m - R_f) = \rho E_t(R_m - R_f)/\sigma_m$$

Dadurch geht in der Bewertung der erwarteten Rendite nur ein Teil des eigenen Risikos ein, nämlich $\rho\sigma_i$, das mit dem Marktportfolio korreliert.

Ein weiteren Aspekt beim CAPM ist, dass die erwartete Rendite der Aktie nicht von der Wachstumsrate des erwarteten zukünftigen Zahlungsstroms abhängig ist. Wenn das CAPM für die Analyse der Aktienbewegung einer Firma angewendet wird, wird ausschließlich nach dem spezifischen β für die Firma gesucht. Dieser Parameter ist wesentlich leichter zu schätzen als der zukünftige Zahlungsstrom eines Unternehmens (vgl. Perold (2004)).

Die Nicht-Standard-Form des CAPMs

Beim Standard-CAPM wurden tiefgreifende Annahmen getroffen. Das Abschwächen dieser Annahmen führt entsprechend zu verschiedenen Varianten des CAPMs. Dabei werden beispielsweise der Einfluss von Steuern oder heterogenen Erwartungen der Investoren berücksichtigt. Eine einfache Erweiterung des Standard-CAPMs liefert die Black-Variante, die nicht die risikolosen unbeschränkten Anlage- und Verschuldungsmöglichkeiten im Modell unterstellt.

Im Gegensatz zu der Sharpe-Lintner-Variante des CAPMs vergleicht die Black-Version die erwartete Rendite der Aktie i mit der erwarteten Rendite des sogenannten Zero-Beta-Portfolios $E_t(R_z)$.

$$E_t(R_i) - E_t(R_z) = \beta_{im}(E_t(R_m) - E_t(R_z)). \quad (4.16)$$

Das Zero-Beta-Portfolio ist definiert als ein Portfolio mit der kleinsten Varianz von allen Portfolios, die nicht mit dem Marktportfolio korrelieren. Der Koeffizient β ist in diesem Modell identisch mit dem Koeffizienten aus der Sharpe-Lintner-Variante des CAPMs.

Die wichtigsten Charakteristika des Zero-Beta-Modells können wie folgt zusammengefasst werden. Wenn Kredite nicht zu einem risikofreien Zinssatz R_f aufgenommen oder ausgeliehen werden können, kann jeder Investor sein eigenes Aktienportfolio entsprechend seiner Präferenzen bilden. Das Portfolio setzt sich aus der Kombination eines beliebigen effizienten Portfolios (M) und des Zero-beta-Portfolios (Z) zusammen. In dem Zero-Beta-Portfolio werden die risikobehafteten Aktien in einem bestimmten Verhältnis kombiniert, so dass die erwartete Rendite dieses Portfolios $E_t(R_z) = \sum x_i E_t(R_i)$ mit der erwarteten Rendite des Portfolios M nicht korreliert ($Cov(R_z, R_m) = 0$), und eine minimale Varianz besitzt. Die Kombination aus den Portfolios (M) und (Z) ist nicht eindeutig. Trotzdem kann der gleichgewichtige Ertrag für jede Aktie als eine lineare Funktion von $E_t(R_z)$ und $E_t(R_m)$ beschrieben werden.

Die Berücksichtigung der Steuer in dem Modell kann wie folgt aufgefasst werden:

$$E_t(R_i) - R_f - \tau(\delta_i - R_f) = \beta_{i,m}(E_t(R_m) - R_f - \tau(\delta_M - R_f)). \quad (4.17)$$

Dabei sind δ_i und δ_M die Dividendenzahlungen für die Aktie i und für das Marktportfolio. Die Variable τ bezeichnet den Steuersatz, gemessen an dem relevanten Satz für Kapitalgewinne und der Einkommensteuer²⁹.

Beim konsumorientierten CAPM wird in das Optimierungskalkül die Nutzenfunktion der Investoren einbezogen. In einem Modell mit zwei Perioden wird der Nutzen aus dem Konsum mit Hilfe folgender Nutzenfunktion dargestellt:

$$u(c_t, c_{t+1}) = u(c_t) + \beta E_t(u(c_{t+1})).$$

Dabei ist β ein subjektiver Diskontierungsfaktor. Man bezeichnet mit x die Menge der zu kaufenden Aktien und mit a_t das Geld, das dem Haushalt in der Periode t zur Verfügung steht. Dabei setzt sich der Konsum zum Zeitpunkt t aus der Differenz von der Anfangsausstattung und dem Geldbetrag zusammen, der für den Kauf von Aktien ausgegeben wird.

$$c_t = a_t - p_t x$$

Die Variable p_t ist der (logarithmierte) Preis der Aktie. In der zweiten Periode werden die Aktien verkauft, dadurch ist der Konsum zum Zeitpunkt $t+1$:

$$c_{t+1} = a_{t+1} + (p_{t+1} + d_{t+1})x$$

d_t bezeichnet die Dividenden für die Aktien zum Zeitpunkt $t+1$. Der Haushalt optimiert seinen Konsum über zwei Perioden. Das Optimierungsproblem lässt sich wie folgt formulieren:

$$\max_x \{u(c_t) + \beta E_t(u(c_{t+1}))\}$$

Aus der Bedingung erster Ordnung erhält man die Aktienpreisgleichung:

$$p_t = E_t \left(\beta \frac{u'(c_{t+1})}{u'(c_t)} (p_{t+1} + d_{t+1}) \right) = E_t(m_{t+1}(p_{t+1} + d_{t+1})) \quad (4.18)$$

Dabei ist die Variable $m_{t+1} \equiv \beta \frac{u'(c_{t+1})}{u'(c_t)}$ auch als 'stochastischer Diskontierungsfaktor' oder 'Pricing-Kernel' bekannt. In diesem Modell ist der Pricing-Kernel äquivalent zur abdiskontierten intertemporalen Substitutionsrate des Konsums.

Die Bewertung der Risikoprämie lässt sich wie folgt zusammenfassen:

$$E_t(R_i) - R_f = - \frac{\text{cov}(u'(c_{t+1}), R_{i,t+1})}{E_t(u'(c_{t+1}))} \quad (4.19)$$

Das konsumbasierte CAPM sagt aus, dass die Überschussrendite einer riskanten Anlage proportional zur Kovarianz von der zukünftigen Aktienrendite und dem zukünftigen Konsum ist³⁰.

²⁹Eine genaue Herleitung befindet sich bei Elton und Gruber (2003), Kapitel 14.

³⁰Die ausführliche Herleitung der Formel befindet sich bei Semmler (2003).

Voerst wurde bei mehreren Untersuchungen eher auf heterogene anstelle von homogenen Erwartungen bei CAPM gestützt³¹. All diese Modelle führen zu einer Form des gleichgewichtigen Preises, die den schon besprochenen Modellen ähneln. Der gleichgewichtige Preis der Aktie kann mit den Termen der erwarteten Rendite, der Kovarianz und der Varianz ausgedrückt werden, allerdings liegt der Unterschied in einem komplexeren Gewichtungssystem der Komponenten in dieser Gleichung. Die Gewichtung ist komplizierter, weil sie eine Information über die Nutzenfunktion der Investoren enthält. Ein Ausgleich³² zwischen dem Preis und dem Wohlstand der Investoren soll dabei ermittelt werden.

Ungedeckte Zinsparität, Inflation und CAPM

In dem praktischen Teil dieser Arbeit werden zwei Komponenten von Aktienpreisbewegungen berücksichtigt, nämlich das Wechselkursrisiko und die Inflation. Eine Einbindung der Komponente in das Standard-CAPM kann auf folgende Weise erreicht werden.

In einer Welt mit hoher Kapitalmobilität können Inländer und Ausländer genauso leicht mit den heimischen und ausländischen Aktien in eigener und ausländischer Währung handeln. Haben inländische und ausländische Aktien das ähnliche Risiko und die Liquidität, weil es nur wenige Hindernisse für internationale Kapitalbewegung gibt, ist es nachvollziehbar, dass diese Aktien perfekte Substitute sind. Unter der Annahme, dass die erwartete Rendite der inländischen Anleihen größer als die der ausländischen ist, werden inländische und ausländische Anleger in inländische Aktien investieren. Analog, wenn die erwartete Rendite der ausländischen Anleihen größer ist, dann werden beide Gruppen in ausländische Aktien investieren. Um die existierenden Angebote beider Arten von Anleihen festzuhalten, müssen die erwarteten Renditen der beiden Anleihen gleich sein. Diese Voraussetzung kann folgendermaßen ausgedrückt werden:

$$\frac{1 + R_t^{inl}}{1 + R_t^{ausl}} = \frac{E_t(S_{t+1})}{S_t}. \quad (4.20)$$

oder approximativ:

$$R_t^{inl} - R_t^{ausl} = E_t(s_{t+1}) - s_t. \quad (4.21)$$

In der Gleichung bezeichnet $s_t = \log S_t$ den logarithmierten Wechselkurs S_t . $\log(1 + R) \approx R$, wegen $0 < R < 1$. Diese Gleichung ist die Ungedeckte-Zinsparität-Bedingung, die besagt, dass die Differenz zwischen der Rendite der inländischen Anlagen und der Rendite der ausländischen Anlagen durch eine erwartete Wertsteigerung der Fremdwährung erklärbar ist.

Unter der Annahme der Ungedeckten-Zinsparität beträgt der erwartete Gewinn

³¹Vgl. Lintner (1969), Sharpe(1970), Fama(1976).

³²Damit ist *marginal rate of substitution* gemeint.

aus einer Inlandsanlage für die Anleger:

$$1 + E_t(R_i) = \frac{E_t(S_{t+1})}{S_t}(1 + R^{ausl}) \quad (4.22)$$

Die Sharpe-Lintner-Version des CAPM liefert dann eine Bewertung der Risikoprämie bezüglich des risikofreien Zinssatzes im Inland R_f^{inl}

$$E_t(R_i) - E_t(R_f^{inl}) = \beta_{im}(E_t(R_m) - E_t(R_f^{inl})), \quad (4.23)$$

wobei β_{im} abhängig von der Kovarianz zwischen dem Marktportfolio und dem ausländischen Portfolio ist, weil die erwartete Rendite der Aktie durch die Ungedeckten-Zinsparität-Gleichung (4.22), also durch eine Änderung des Wechselkurses und durch die ausländische risikolose Markttrendite, festgelegt wird. Die Bezeichnung $E_t(R_m)$ steht für die erwartete Rendite des Marktportfolios von Aktien, die in unterschiedlichen Währungen gehalten werden.

Eine weitere Variante für die Einbindung des Wechselkurses in CAPM ist die sogenannte Mittelwert-Varianz-Regel ($\mu - \sigma$ -Regel). Wenn R_i, R_f^{inl}, R_m die nominalen Renditen bezeichnen, dann sieht die CAPM-Gleichung mit Berücksichtigung des Wechselkurses folgendermaßen aus:

$$E_t(R_i) - E_t(R_f^{inl}) - \sigma_{i\hat{s}} = \frac{[\sigma_{im} - \frac{\sigma_{i\hat{s}}}{\alpha}]}{[\sigma_m^2 - \frac{\sigma_{m\hat{s}}}{\alpha}]}(E_t(R_m) - E_t(R_f^{inl}) - \sigma_{m\hat{s}}) \quad (4.24)$$

Dabei bezeichnet $\sigma_{i\hat{s}}$ die Kovarianz zwischen der Aktienrendite R_i und der geschätzten Wechselkursänderung \hat{s} , $\sigma_{m\hat{s}}$ die Kovarianz zwischen der Markttrendite R_m und Wechselkursänderung \hat{s} , und die Variable α bezeichnet den Anteil der risikobehafteten Aktien zu der Gesamtanzahl von Aktien.

Wenn Wechselkursänderungen \hat{s} mit der Markttrendite oder mit der Rendite der Aktien R_i nicht korrelieren, dann gilt $\sigma_{i\hat{s}} = \sigma_{m\hat{s}} = 0$, so dass sich die Gleichung (4.24) auf die Sharpe-Lintner-Version des CAPMs (4.15) reduziert.

Eine einfache Erweiterung des CAPMs mit Berücksichtigung der Inflation wäre ein Übergang von nominalen zu realen Renditen mit Hilfe der Fischer-Gleichung:

$$R^{real} = R^{nominal} - E_t(\pi), \quad (4.25)$$

mit $E_t(\pi)$ der erwarteten Inflationsrate. Allerdings bringt es weitere Schwierigkeiten, weil eine Hypothese bezogen auf die Bildung der Erwartungen bei Inflation aufgestellt werden soll. Beispielsweise ist die Annahme der Martingal-Eigenschaft für Inflation ökonomisch nicht vertretbar.

Eine weitere Möglichkeit, Inflation bei CAPM zu berücksichtigen, basiert wieder auf der $\mu - \sigma$ -Regel. Die Aktienkurse sollen sich proportional zur Inflation des

entsprechenden Landes bewegen und nominale Dividenden wachsen mit der Inflationsrate. Weil nominale Renditen proportional zu nominalen Dividenden steigen, sollen die nominalen Aktienpreise sich mit dem Preisniveau in die gleiche Richtung bewegen. Das bedeutet, dass der reale Aktienindex mit der Inflation nicht korrelieren kann. Allerdings wurde bei mehreren Untersuchungen eine negative Korrelation zwischen der Inflation und dem realen Aktienindex nachgewiesen³³. Eine ausführliche Arbeit über Aktienpreis-Inflation-Puzzle ist von Kim (2003) verfasst worden.

In ihrer Arbeit von (1976) erweiterten Friend, Landskroner und Losq das CAPM durch das Einbeziehen der Inflation mit Hilfe der $\mu - \sigma$ -Regel, welches einen spezifischen Fall eines mehrperioden Gleichgewichtsmodells darstellt. Bei der Nutzenfunktion in dem Modell wurde eine konstante Risikoaversion der Investoren unterstellt. Die Definition des Aktienrisikos und des Markttrisikos wurde ebenfalls modifiziert. Wenn R_i , R^f und R_m die nominalen Renditen bezeichnen, dann sieht die CAPM-Gleichung mit Berücksichtigung der Inflation folgendermaßen aus:

$$E_t(R_i) - E_t(R_f) - \sigma_{i\pi} = \frac{[\sigma_{im} - \frac{\sigma_{i\pi}}{\alpha}]}{[\sigma_m^2 - \frac{\sigma_{m\pi}}{\alpha}]} (E_t(R_m) - E_t(R_f) - \sigma_{m\pi}) \quad (4.26)$$

Hierbei bezeichnet $\sigma_{i\pi}$ die Kovarianz zwischen der Aktienrendite R_i und der Inflation π . Die Variable $\sigma_{m\pi}$ beschreibt die Kovarianz zwischen der Marktrendite R_m und der Inflation π und α den Anteil der risikobehafteten Aktien im Verhältnis zur Gesamtanzahl von Aktien. Friendl und andere (1976) zeigen, dass das Markttrisiko höher sein wird, als das Standard-CAPM vorschlägt, solange die Korrelation zwischen der Marktrendite und die Inflationsrate positiv ist. Sie zeigten ebenfalls, dass das Risiko der Aktien mit der Kovarianz zwischen der Aktienrendite und die Inflationsrate stark zusammenhängt. Im Falle einer positiven Korrelation der beiden Größen kommt es zu einer Überbewertung des Aktienrisikos beim Standard-CAPM. Für die Aktienrendite zeigen sie, dass das Standard-CAPM die erwartete Aktienrendite überschätzt, wenn die Korrelation zwischen der Aktienrendite und der Inflation größer als das Produkt der Korrelation zwischen der Aktienrendite und der Marktrendite und die Korrelation zwischen der Marktrendite und der Inflation ist.

Wenn die Inflation π nicht mit der Marktrendite oder mit der Rendite der Aktie i korreliert, dann gilt $\sigma_{i\pi} = \sigma_{m\pi} = 0$, so dass sich die Gleichung (4.26) auf das Standard-CAPM (4.15) reduziert.

4.2.3 Das CAPM und Empirische Nachweise

Die Tests des CAPMs basieren auf drei Folgerungen aus der Relation zwischen der erwarteten Rendite der Aktie und der β -Prämie des Marktes. Erstens hat die erwartete Rendite eine einfach lineare Abhängigkeit zu dem Koeffizienten β , keine andere Variable

³³Vgl. Gauthier (2001).

hat einen erklärenden Einfluss auf die Rendite. Zweitens ist die β -Prämie positiv, d.h. die erwartete Rendite des Marktportfolios $E_t(R_m)$ übersteigt die erwartete Rendite des Aktienportfolios $E_t(R_P)$, wenn die Aktien aus diesem Portfolio mit dem Marktportfolio nicht korrelieren. Drittens, in der Sharpe-Lintner-Variante des CAPMs haben die mit dem Markt nicht korrelierenden Aktien eine erwartete Rendite, die identisch mit der risikofreien Zinsrate ist. Die meisten Tests zu diesen Vorhersagen benutzen entweder die Querschnitts- oder die Zeitreihenregressionen.

Einige empirische Untersuchungen des CAPMs wendeten die Querschnittsregressionen an, um den Achsenabstand und die Neigung in der Sharpe-Lintner-Modellvariante zu schätzen. Die Methode beinhaltet die Regression der durchschnittlichen Aktienrenditen gemessen an den geschätzten β -Koeffizienten der Aktien. Nach diesem Modell soll der Abstand von der x-Achse zu dem risikofreien Zinssatz R_f gleich sein und die Neigung der Gerade identisch mit der Risikoprämie des Marktes $E_t(R_m) - R_f$. Zwei offensichtliche Probleme sind dabei entstanden. Erstens, die Schätzung der einzelnen β -s war unpräzise. Zweitens, die Regressionsresiduen weisen eine Quelle der gemeinsamen Variation auf, beispielsweise die industriellen Effekte. Die Verbesserung des Verfahrens kam von Blume (1970), Friend und Blume (1970) und Black, Jensen und Scholes (1972), die bei der Schätzung des β -s mit gesamten Portfolios statt einzelnen Aktien gearbeitet haben.

Als erster erfasste Jensens (1968) das Sharpe-Lintner-CAPM in Form einer Zeitreihenregression:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{i,m}(R_{m,t} - R_{f,t}) + \varepsilon_{i,t} \quad (4.27)$$

Nach der Regressionsschätzung soll ein 'Jensens α ' signifikant identisch Null sein.

Alle Tests der beiden Regressionsarten lehnten die Nullhypothese ab, d.h. dass die überdurchschnittliche Rendite mit Hilfe der Sharpe-Lintner-Version des CAPMs nicht erklärt werden kann. Ein positiver Zusammenhang bestand zwar zwischen den Variablen, allerdings war er zu 'pauschal'. Bei den Querschnittsregressionen war der Achsenabstand größer als der durchschnittliche risikofreie Zinssatz³⁴. Der Koeffizient vor β wurde kleiner geschätzt als die durchschnittliche Risikoprämie des Marktes³⁵. Die Ergebnisse stammen von Douglas (1968), Black, Jensens und Scholes (1972), Miller und Scholes (1972), Blume und Freind(1973), Fama und MacBeth (1973) und Fama und French (1992)³⁶. Der Achsenabstand bei den Zeitreihenregressionen war nach der Schätzung signifikant von Null unterschiedlich, das wurde bei der Sharpe-Lintner-Version des CAPMs nicht erwartet. Der geschätzte Achsenabstand war signifikant positiv bei den Aktien mit kleinen β -Koeffizienten und negativ bei den Aktien mit großen β -s.

Seit den späten 70er Jahren wurde als Referenz bei den CAPM-Tests die Version

³⁴Normalerweise wurde er durch die Rendite des einmonatigen Schatzbriefs approximiert.

³⁵Sie wurde durch die Differenz zwischen der Rendite des Aktienportfolios der USA und der Rendite des einmonatigen Schatzbriefs approximiert.

³⁶Vgl. Fama und French (2004).

von Black benutzt. Zuerst stellte Basu (1977) fest, dass die Aktien mit einem hohen Gewinn-Preis-Verhältnis (G/P) höhere zukünftige Renditen aufweisen, als das CAPM vorschlägt. Banz (1981) dokumentiert in seiner Arbeit den sogenannten '*Größe-Effekt*', der besagt, dass bei den nach der Marktkapitalisierung der Unternehmen sortierten Aktien die durchschnittlichen Renditen der kleinen Firmen höher sind, als das CAPM behauptet. Bhandari (1988) fand heraus, dass der hohe Schulden-Eigenkapital-Quotient der Firmen mit höheren Renditen relativ zur Höhe ihrer β -Koeffizienten assoziiert wird. Letztendlich stellten Statmann (1980) und Rosenberg, Reid und Lanstein (1985) fest, dass Aktien mit einem hohen Buchwert-Marktwert-Quotient (B/M) des Unternehmens die durchschnittlich höhere Renditen haben, die nicht mit ihren β -s aus dem CAPM kompatibel sind.

Fama und French (1992) wiederholten und synthetisierten die vorherigen empirischen Untersuchungen zum CAPM. Unter Anwendung der Querschnittsdaten stellten sie fest, dass die Größe des Unternehmens, das G/P-Verhältnis, der Schulden/Eigenkapital- und Buchwert/Marktwert-Quotient eine signifikante Verbesserung bei der Schätzung der Regressionen zwischen den erwarteten Renditen und den β -Koeffizienten liefern. In einer weiteren Arbeit kamen Fama und French (1996) zu einem ähnlichen Ergebnis bei der Anwendung der Zeitreihenregressionen. Sie fanden heraus, dass verschiedene Preisquotienten viele Informationen über die erwarteten Renditen haben.

Aus Untersuchungen der letzten Jahre weiß man, dass das CAPM weder in seiner klassischen Form noch in der Black-Variante ein schwerwiegendes Problem hat. Es gab keinen empirischen Hinweis für die Gültigkeit des Modells. Die erwarteten Renditen der Aktien konnten nicht nur mit dem Markt- β und der Risikoprämie des Marktes erklärt werden. Weitere Faktoren sind für eine Beschreibung der Renditen notwendig.

Zwei Erklärungen der irrationalen Bewertung der Renditen haben sich während der letzten Jahre herauskristallisiert. Einerseits steht die sogenannte '*behavioral*'-Sicht zur Debatte. Nach Meinung der Behavioristen besitzen diejenigen Firmen Aktien mit einem hohen B/M-Verhältnis, deren Aktien in schlechten Zeiten schon gefallen sind, obwohl ein niedriger B/M-Quotient eher mit Wachstumsunternehmen assoziiert wird. Dementsprechend argumentieren die Behavioristen, dass eine Aufteilung der Unternehmen nach ihrem B/M-Verhältnis eine übermäßige Investorenreaktion in '*schlechte*' und '*gute*' Zeiten an der Börse widerspiegelt. Die Investoren überschätzen die historische Performance des Marktes. Als Folge sind die Preise für die '*growth*'-Aktien (kleines B/M) zu hoch und die Preise für die '*value*'-Aktien (hohes B/M) zu niedrig³⁷.

Eine weitere Begründung für die Diskrepanz zwischen der Theorie und den empirischen Untersuchungen zur CAPM liegt in der Notwendigkeit einer Erweiterung des CAPMs. Das CAPM basiert auf einer Reihe sehr realitätsfremder Annahmen. Zum Beispiel ist die Beschränkung, dass die Investoren nur mit Hilfe des Mittelwertes und der Varianz des einperioden Portfolios über zukünftige Renditen entscheiden können,

³⁷Vgl. Fama und French (1996), (2004)

sehr extrem. Es wäre sinnvoller, zu untersuchen, wie die Renditen mit dem Arbeitseinkommen oder mit den zukünftigen Investitionsmöglichkeiten korrelieren. Die Varianz der Portfoliorendite misst einen wichtigen Teil des Risikos, allerdings beschreibt sie es nicht vollständig. Eine Lösung bietet Merton (1973) in Form des intertemporalen CAPMs (ICAPM). Hier können die Investoren eine Entscheidung über die Zeit hinweg optimieren und berücksichtigen dabei mindestens drei Faktoren, wie das Arbeitseinkommen, die Konsumpreise und die Investitionsalternativen. Als ein Ergebnis kann formuliert werden, dass ein optimales Portfolio nun 'multifaktor'-effizient sein soll.

Eine Variante des ICAPMs wurde von Fama und French (1993)[34], (1996) in Form eines Drei-Faktor-Modells vorgestellt:

$$E(R_{i,t}) - R_{f,t} = \beta_{i,m}(R_{m,t} - R_{f,t}) + \beta_{i,s}E(SMB_t) + \beta_{i,h}E(HML_t) \quad (4.28)$$

Die Variable SMB_t (small minus big) stellt eine Differenz zwischen den Portfoliorenditen der großen und der kleinen Firmen dar. Hingegen beschreibt die Variable HML_t (hight minus low) die Differenz zwischen den Portfoliorenditen der Firmen mit hohem und niedrigem B/M-Verhältnis. Bei der Querschnittsschätzung kamen während der Periode von 1927 bis 2003 für den amerikanischen Markt signifikant von Null unterschiedliche Werte für die Marktprämie $R_{m,t} - R_{f,t}$ von 8,3% pro Jahr, für die Faktoren SMB_t und HML_t 3,6% und 5,0% pro Jahr. Alle drei Prämien zeigen hohe Volatilität von 21% für $(R_{m,t} - R_{f,t})$, 14,6% (SMB_t) und 14,2% (HML_t) pro Jahr. Auch wenn die durchschnittlichen Prämien relativ groß sind, impliziert eine hohe Volatilität eine Unsicherheit über die richtig erwarteten Prämien. Für die Zeitregression

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{i,m}(R_{m,t} - R_{f,t}) + \beta_{i,s}SMB_t + \beta_{i,h}HML_t + \varepsilon_{i,t} \quad (4.29)$$

wurde der Koeffizient α_i signifikant als Null geschätzt. Fama und French (1998) zeigten zusätzlich, dass das internationale ICAPM besser als das internationale CAPM die durchschnittlichen Renditen des Aktienportfolios beschreibt, die aus den 13 wichtigsten Aktienmärkten zusammengestellt worden sind.

Eine Entscheidung über die Gültigkeit des CAPMs kann durch empirische Tests nicht eindeutig getroffen werden. Einerseits behaupten die Behavioristen, dass die Ablehnung des CAPMs durch ein irrationales Verhalten begründbar ist, andererseits zeigen beispielsweise Fama und French, dass eine Erweiterung des CAPMs um weitere Faktoren dennoch eine rationale Bewertung der Überschussrendite ermöglicht.

Ein weiteres Problem der Schätzung des CAPMs ist die Approximation des Marktes. Theoretisch ist es nicht eindeutig, welche Aktien das Marktportfolio enthält. Beispielsweise kann das Humankapital auch als ein Asset aufgefasst werden. Daher basieren die Tests des CAPMs auf einer Proxi für das Marktportfolio. Roll (1977) behauptet:

„...because the tests use proxies, not true market portfolio, we learn nothing about the CAPM“.

Für die CAPM-Version von Sharpe (1964) und Lintner (1965) konnte bis jetzt kein empirischer Beweis festgestellt werden. Eine Erklärung dafür kann das irrationale Verhalten der Investoren sein. Eine weitere Erklärung könnte sein, dass das CAPM in dieser Form nur einen Teil des Gesamtrisikos beschreibt. Deswegen ist es notwendig, weitere Faktoren wie das G/P-Verhältnis oder den B/M-Quotient in das Modell aufzunehmen.

Das CAPM wird oft zum Vergleich der offenen und aktiv gemanagten Fonds angewendet werden. Weil das Standard-CAPM das G/P- oder B/M-Verhältnis nicht berücksichtigt, lässt sich erklären, wieso die auf den kleinen β -s oder 'value'-Aktien basierten Portfolios im Vergleich zur Sharpe-Lintner-Version des CAPMs überdurchschnittliche Renditen produzieren. Fama und French (2004) meinten:

„... the CAPM's empirical problems probably invalidate its use in applications.“

4.3 Relevanz der Kointegration für Aktienpreise

Eine Reihe der Publikationen der letzten Jahren³⁸ konnte in den Aktienmärkten führender Industrieländerin mit Hilfe der Kointegrationsmethode einen einzigen gemeinsamen stochastischen Trend identifizieren. Dieses Ergebnis impliziert, dass die Renditen der internationalen Aktienmärkte zwar kurzfristig unterschiedlichen Pfaden folgen können, langfristig aber sehr dicht beieinander liegen. Die Nähe wird in diesem Fall als durchschnittliche Distanz zwischen der Zeitreihen und nicht als ähnlicher Muster der Reihen verstanden. Die Existenz einer solcher Kointegrationsbeziehung kann als ein langfristiger Gleichgewicht des Systems interpretiert werden.

Mit Hilfe der Gleichung (4.6), wobei $\mu = E_t(R_{t+1})$ ist, kann der logarithmierte Preis eines Aktienindex wie folgt ausgedrückt werden:

$$p_{t+1} = p_0 + \sum_{j=1}^t E_j(R_{j+1}) + \sum_{j=1}^t \varepsilon_{j+1}. \quad (4.30)$$

Der Aktienindex ist äquivalent zu seinem Startwert, plus die akkumulierten erwarteten Änderungen des Aktienpreises, plus ein zufälliger Term (Random-Walk Komponente), der unerwartete stochastische Störungen präsentiert.

Unter der Annahme, dass die Überschussrendite (die Differenz zwischen erwarteter Aktienrendite und dem risikofreien Zinssatz) mit Hilfe des CAPM beschrieben wird, folgt aus der Gleichung (4.15):

$$E_t(R_{t+1}) - R_{f,t+1} = \beta_t E_t(R_{m,t+1} - R_{f,t+1}). \quad (4.31)$$

³⁸Vgl. Kasa (1992), Leachman und Francis (1995) oder Dickenson (2000).

Mit $R_{m,t+1}$ ist die Weltmarktrendite gemeint; der Risikofaktor β muss nicht unbedingt konstant sein. Die Kointegration der zwei Aktienindizes ($i=a,b$), ausgedrückt durch den Kointegrationsvektor $(1, -\alpha)^{39}$, verlangt:

$$(\beta_{a,t} - \alpha\beta_{b,t}) \sum_{j=1}^t E_{j-1}(R_{m,t+1} - R_{f,t+1}) + \sum_{j=1}^t (\varepsilon_{a,j} - \alpha\varepsilon_{b,j}) \sim I(0). \quad (4.32)$$

Unter der Annahme, dass die Preiszeitreihen instationär und die kumulierte Risikoprämie sowie unerwartete Störungen $\varepsilon_{i,t}$ unabhängig voneinander sind, hat die Kointegration zwei Implikationen. Wenn die Risikofaktoren β_a und β_b von der Zeit unabhängig sind, dann impliziert Kointegration eine Beziehung zwischen den langfristig erwarteten Überschussrenditen zweier kointegrierter Aktienindizes:

$$E_t(R_{a,t+1} - R_{f,t+1}) - \alpha E_t(R_{b,t+1} - R_{f,t+1}) \sim I(0),$$

und eine Beziehung zwischen langfristig erwarteten Renditen der Aktienmärkte:

$$E_t R_{a,t+1} - \alpha E_t R_{b,t+1} \sim I(0).$$

Die Anwendung des Standard CAPMs und die Annahme der Kointegration zwischen der Aktienindizes schließt die Existenz einer langfristiger Beziehung zwischen erwarteten Renditen der zweier Aktienindizes nicht aus.

Wenn aber die Risikofaktoren nicht konstant sind, soll der Term $(\beta_{a,t} - \alpha\beta_{b,t})^{40}$ stationär und sein Mittelwert gleich Null sein. Damit stehen die Risiken der entsprechenden Aktienmärkte in einer gleichgewichtigen Beziehung. Sie brauchen zwar nicht identisch in jede Zeitperiode sein, aber müssen sich ausgleichen.

Desweiteren sollen die akkumulierten Fehlerterme $\sum_{j=1}^t \varepsilon_{a,j}$ und $\sum_{j=1}^t \varepsilon_{b,j}$ auch kointegriert mit einem Vektor $(1, -\alpha)$ sein. Dies ließe auf die Existenz einer explizit beschriebenen Beziehung zwischen den unerwarteten Störungen der beiden Aktienindizes schließen. Es ist nachvollziehbar, dass $\varepsilon_{a,t}$ und $\varepsilon_{b,t}$ in Zeiten steigender Kurse eventuell korrelieren. Die Kointegration der kumulierten Fehlertermen würde aber implizieren, dass alle firmen- oder länderspezifischen Schocks zwar unerwartet sind, sich aber exakt ausgleichen. Alternativ können beide Fehlerterme $\varepsilon_{a,t}$ und $\varepsilon_{b,t}$ voneinander unabhängige Prozesse sein. Diese Interpretation ist mehr mit der ökonomischen Idee der Unabhängigkeit der Informationsschocks auf den Aktienmärkte vereinbar.

Es wird häufig die These aufgestellt, dass die Kointegration zwischen nationalen Märkten ein Zeichen vom hohen Grad der Kapitalmarktintegration ist⁴¹. Die Globalisierung und Liberalisierung der Finanzmärkte kann für die Kointegration der internationalen Aktienindizes verantwortlich sein. Die Integration der Märkte impliziert die

³⁹Die Preiszeitreihen $p_{a,t}$ und $p_{b,t}$ sind instationär, aber es existiert ein Vektor $(1, -\alpha)$, so dass die Linearkombination $p_{a,t} - \alpha p_{b,t} \sim I(0)$ stationär wird.

⁴⁰Der Koeffizient β ist ein stochastischer Term, weil er von der Kovarianz der Weltmarktrendite mit der Rendite des entsprechenden Aktienindizes abhängig ist.

⁴¹Vgl. Richards (1995), Leachman und Francis (1995), Dickenson (2000).

Verwendung des gleichen CAPMs für alle Aktien, unabhängig davon, wo sie gehandelt werden. Wenn das gleiche CAPM für alle Aktien benutzt wird, kann die mögliche Kointegrationsbeziehung zwischen Aktienindizes für jedes Aktienpaar mit konstanten β durch den Kointegrationsvektor (1,-1) definiert sein.

Zusätzlich könnte die Idee der Kointegration als ein Widerspruch der schwachen Form der MEH interpretiert werden⁴². Zwei Preissereien, die von effizienten Märkten stammen, können nicht kointegriert sein, weil sie keine Verbindung irgenwelcher Art haben dürfen. Ein anderen Argument ist, dass die Preise in einem effizienten Markt nicht vorhersagbar sind. Allerdings könnte die Kointegrationsbeziehung kaum als eine Vorhersageregeln für überproportionale Gewinne auf dem Markt verwendet werden. Solnik (1991) meint dazu:

„Some investigators have attempted to find leads or lags between markets. However, no evidence of a systematic delayed reaction of the national market to another has ever been found. The existence of such simple market inefficiencies is, indeed, unlikely since it would be so easy to exploit them to make an abnormal profit“.

Demensprechend kann die Existenz der Kointegration kaum als ein Hinweis auf Marktineffizienz interpretiert werden.

⁴²Vgl. Taylor und Tonks (1989), Richards (1995).

Kapitel 5

Zusammenfassung der Arbeit und Ausblick

Das Ziel dieser Arbeit war eine empirische Untersuchung langfristiger struktureller Gleichgewichtsbeziehungen zwischen internationalen Aktienmärkten durchzuführen. Bei den Fragen, sowohl nach langfristigen Beziehungen als auch nach kurzfristigen Anpassungen, wurde nachgegangen. Dabei wurden empirische Methoden angewendet, die teilweise ausschließlich in neueren Aufsätzen der wissenschaftlichen Journals zu finden sind.

Die Erklärung der empirischen Ergebnisse benötigt dennoch die ökonomischen Vorstellungen und Ideen. Deswegen wurde die Frage, ob sich ein gemeinsamer empirisch festgestellter Fundamentalprozess auf den internationalen Aktienmärkten als ein ökonomisches Gleichgewicht interpretieren lässt, ebenfalls verfolgt. Ein weiteres Ziel der Arbeit war, aus theoretischer Sicht nach ökonomischen Theorien zu suchen, die mit Fundamentalgleichgewichten kompatibel sind und als Erklärungshilfe der empirischen Ergebnisse herangezogen werden können. Zwei Theorien wurden in diesem Zusammenhang ausführlich diskutiert, Markteffizienz-Hypothese und Portfolio-Theorie mit dem zugehörigen Gleichgewichtsmodell, CAPM.

Die anschließende Frage bezüglich praktischer Implikation für die international agierenden Anleger, die aus den empirischen Ergebnissen und deren ökonomischen Interpretationsmöglichkeiten hervor kommt, wurde versucht zu beantworten.

Die Aktienindizes der G5-Länder wurden dabei anhand realer Daten für den Zeitraum Juni 1964 bis Oktober 2004 betrachtet. Für die Deflationierung der Aktienindizes wurden die Konsumentenpreisindizes der G5-Länder benutzt. Sowohl die Aktienindizes in der nationalen Währung als auch in einer gemeinsamen Währung (DM/Euro) wurden zum Zweck einer Vergleichsanalyse verwendet. Die Konvertierung in die gemeinsame Währung berücksichtigt die Perspektive deutscher Anleger, die nicht gegen

das Wechselkursrisiko abgesichert sind. Aus theoretischer Sicht wird die Gültigkeit der Kaufkraftparität bei einer Untersuchung der internationalen Märkte in gemeinsamer Währung angenommen.

Die Differenzen der logarithmierten Aktienpreise approximieren die Aktienmarktrenditen. Nach einer Überprüfung der klassischen Annahmen der empirischen Tests folgte die Korrelationsanalyse für die Aktienindizes und Renditen der G5-Länder. Im weiteren Verlauf wurden die Aktienindizes auf Stationarität, Integrationsordnung und Kointegration geprüft. Die Fehler-Korrektur-Form für verschiedene Modellvarianten wurde geschätzt und eine Strukturanalyse mit Hilfe des Granger-Kausalität-Tests, der Impuls-Antwort-Funktionen und der Varianz-Dekomposition durchgeführt.

Die Konvertierung der Aktienindizes in eine einheitliche Währung wirkt auf die Mittelwerte und die Varianz der Aktienportfolios wie folgt: Der Ertrag aus ausländischen Wertpapieren sinkt aus der Sicht der inländischen Anleger bei einer Aufwertung der inländischen Währung. Die Volatilität der Preise und der Renditen steigt durch das Wechselkursrisiko. Die Korrelationskoeffizienten reagieren nicht eindeutig auf die Konvertierung der Daten in einheitlicher Währung.

Der Jarque-Bera Test lehnt die Hypothese der Normalverteilung für alle Aktienmarktrenditen zum Signifikanzniveau 1% ab. Gründe für dieses Ergebnis sind: Zum Ersten sind die Renditen durch eine linksschiefe Verteilungsdichte in allen Zeitreihen gekennzeichnet. Eine Ausnahme stellt der britische Index dar. Zum Zweiten wurde das Phänomen des Leptokurtosis ebenfalls bei den Renditen aller Zeitreihen festgestellt. Der dritte stilisierte Fakt ist die zeitvariierende Volatilität der Renditen. Dem Rechnung tragend wurde bei den Einheitswurzel-Tests die zeitabhängige Varianz durch Modellierung eines GARCH(1,1)-Prozesses berücksichtigt. Zum Vergleich wurden der modifizierte Ng-Perron-Test und der Elliott-Rothenberg-Stock DF-GLS-Test durchgeführt. Die Ergebnisse der Einheitswurzel-Tests bestätigen die Integrationsordnung von Eins in allen Aktienindizes. Dies führt zur Bestätigung der Idee, dass ein Trend in den Datenreihen der Aktienindizes eher eine stochastische Komponente hat.

Die Ergebnisse der Einheitswurzel-Tests ermöglichen den zweiten Schritt der gesamten Untersuchung: die zweidimensionale Analyse, die eine Kointegrationsanalyse und den Granger-Kausalität-Test beinhaltet. Dabei wurden zwei Aspekte, sowohl die langfristigen Beziehungen als auch die kurzfristige Dynamik, berücksichtigt. Die oft mit unterschiedlichen Methoden festgestellte Verbindung zwischen dem amerikanischen und den europäischen Aktienindizes wurde im Fall der bivariaten Kointegration nicht bestätigt. Überraschend ist auch die schwache Kointegration im europäischen Raum. Es wurden zwei Verbindungen zwischen Frankreich und UK sowie Deutschland und Frankreich festgestellt, allerdings konnte keinem der Märkte eine führende Rolle zugewiesen werden.

Die Granger-Kausalität-Tests lieferten auf den ersten Blick andere Ergebnisse im Vergleich zur Kointegrationsanalyse. Zum einen wurde eine führende Rolle des ame-

rikanischen Marktes mit dem Granger-Kausalität-Test bestätigt. Zum anderen, sind die Verbindungen der europäischen Märkten hier wesentlich stärker. Die Abhängigkeit zwischen Frankreich und UK wurde sogar als gegenseitig identifiziert. In dem Modell mit vier Monats-Lags wurde die Kausalität zwischen Frankreich und Deutschland ebenfalls als gegenseitig festgestellt. Eine Analogie zu Ergebnissen der Kointegrationsanalyse bleibt allerdings in Bezug auf den japanischen Aktienmarkt. Es wurden keine kausalen Beziehungen weder mit dem amerikanischen noch mit den europäischen Märkten entdeckt. Allgemein wurde mittels des Granger-Kausalität-Tests weit mehr Abhängigkeit der Aktienmärkte festgestellt. Allerdings widerspricht dies nicht den Ergebnissen der Kointegrationsanalyse, sondern bestätigt die Unterschiede, die in beiden Konzepten angelegt sind. Es existierten kurzfristige zeitliche Verbindungen zwischen den Aktienmärkten der G5-Länder, die mittels des Granger-Kausalität-Tests festgestellt werden konnten. Mit der Zeit verschwinden diese Abhängigkeiten, wie mit der bivariaten Kointegrationsanalyse bestätigt wurde.

Die weitere Kointegrationsanalyse wurde für das komplette System aus fünf Aktienmärkten durchgeführt, um einen Vergleich für die zweidimensionale Analyse zu ermöglichen. Mehrere Aspekte wurden bei der Auswahl des zu schätzenden VECM berücksichtigt, zum Beispiel die Länge des Zeithorizonts, die Anzahl der Verzögerungen, die saisonalen Einflüsse, das Einbeziehen eines deterministischen Trends. Zwei Modelle wurden aufgrund der Untersuchungen gewählt, ein Modell ohne deterministischen Trend weder in den Daten noch im Kointegrationsraum, und ein weiteres mit einem deterministischen Trend in den Daten und einem Achsenabstand im Kointegrationsraum. Die entsprechenden VECMs, die die langfristige Gleichgewichtsbeziehung und kurzfristigen Anpassungen wiedergeben, wurden geschätzt.

Das Ergebnis der Schätzung ist die Existenz einer Kointegrationsbeziehung zwischen den Aktienmärkten der G5-Länder in beiden Modellen. Im Gegensatz zur zweidimensionalen Kointegrationsanalyse wurde im mehrdimensionalen Fall die führende Rolle des amerikanischen Aktienmarktes bestätigt. Die Erklärung dafür könnte in der Berücksichtigung mehrerer Variablen liegen, die einen Einfluss auf die Schätzung haben. Die Schätzung der Gleichgewichtsbeziehung in den beiden Modellen zeigt, dass der amerikanische Aktienmarkt den deutschen Markt am stärksten beeinflusst.

Die Präsenz eines einzigen Kointegrationsvektors impliziert, dass die Aktienindizes der G5-Länder sich in einem langfristigen Gleichgewicht befinden, das eindeutig spezifiziert werden kann. Ein wichtiger Aspekt der internationalen Globalisierung der Finanzmärkte ist ein Zusammenhang zwischen den Preisen. Erwartungsgemäß können die Fluktuationen in den Preisen eines Marktes von den Änderungen auf einem anderen Markt durch den perfekten Informationsaustausch, den einfachen Zugang zu den Märkten und das zeitliche Portfoliomanagement ausgelöst werden. Der Kointegrationsvektor beschreibt genau diese Verbindung für die G5-Länder. Diese Verbindung impliziert, dass die Renditen der internationalen Aktienmärkte zwar kurzfristig unterschiedlichen Pfaden folgen können, langfristig aber sehr dicht beieinander liegen. Die Nähe wird in

diesem Fall als durchschnittliche Distanz zwischen den Zeitreihen und nicht als ähnliches Muster der Reihen verstanden. Die Existenz einer solcher Kointegrationsbeziehung kann als ein langfristiges Gleichgewicht des Systems interpretiert werden.

Unter der Annahme, dass die Preiszeitreihen instationär und die kumulierte Risiko­prämie sowie unerwartete Störungen unabhängig voneinander sind, hat die Kointegration zwei Implikationen. Wenn die Risikofaktoren der internationalen Aktienmärkte, die bei CAPM durch verschiedene β -s dargestellt werden, von der Zeit unabhängig sind, dann schließt Kointegration eine langfristige Gleichgewichtsbeziehung zwischen den erwarteten Überschussrenditen kointegrierter Aktienindizes nicht aus.

Wenn aber die Risikofaktoren nicht konstant sind, soll der Term, der die Kointegrationbeziehung zwischen den Risikofaktoren darstellt, stationär mit einem Mittelwert identisch Null sein. Damit sollen die Risiken der entsprechenden Aktienmärkte in einer gleichgewichtigen Beziehung stehen. Sie brauchen zwar nicht identisch in jeder Zeitperiode sein, müssen sich jedoch langfristig ausgleichen.

Desweiteren sollen die akkumulierten Fehlerterme auch kointegriert sein. Dies ließe auf die Existenz einer explizit beschriebenen Beziehung zwischen den unerwarteten Störungen der beiden Aktienindizes schließen. Es ist nachvollziehbar, dass die Fehlerterme in Zeiten steigender Kurse eventuell korrelieren. Die Kointegration der kumulierten Fehlerterme würde aber implizieren, dass alle unternehmens- oder länderspezifischen Schocks zwar unerwartet sind, sich aber exakt ausgleichen. Alternativ können Fehlerterme voneinander unabhängige Prozesse sein. Diese Interpretation ist mehr mit der ökonomischen Idee der Unabhängigkeit der Informationsschocks auf den Aktienmärkten vereinbar.

Aus der Schätzung des VECMs folgt für kurzfristige Anpassungen, dass Änderungen auf dem eigenen Aktienmarkt generell für die Änderungen der Renditen im nachfolgenden Monat statistisch relevant sind. Die starke Rolle des amerikanischen Marktes spiegelt sich in den Koeffizienten wider, die in dem geschätzten VECM vor den deutschen und französischen Märkten stehen. Es wird von amerikanischer Seite kein Einfluss auf den britischen Aktienindex beobachtet. Der deutsche Aktienindex ist zusätzlich von den Änderungen der französischen Renditen beeinflusst. Der amerikanische und der japanische Aktienmarkt sind unabhängig von den Änderungen auf den anderen Märkten der G5-Länder-Gruppe.

Ein Teil der Strukturanalyse besteht aus dynamischen Simulationen für das System der zu untersuchenden Variablen. Die dynamischen Simulationen geben einen Einblick in die ökonomische Relevanz der Variablen im Gesamtsystem. Die dynamische Analyse verwendet die Berechnung der Impuls-Antwort-Funktionen und der Varianz-Dekomposition der Vorhersagefehler aller Variablen. Die Impuls-Antwort-Analyse stellt die zukünftige Reaktion des Gesamtsystems mittels des aufgestellten Modells bei einer Änderung der unabhängigen Variablen dar. Die Varianz-Dekomposition der Vorhersagefehler liefert zuerst eine prozentuale Aufteilung der Varianz der Störterme nach

Einflussfaktoren. Im Folgenden wird die Änderung dieser Zerlegung innerhalb der bestimmten Periode simuliert.

Die Impuls-Antwort-Analyse stellt die Reaktion des deutschen Aktienindex auf die Einheitsschocks in den einzelnen nationalen Märkten dar. Der Zeithorizont umfasst 10 Monate. Die Schocks auf den internationalen Märkten haben einen permanenten Einfluss auf den deutschen Index, der auch noch nach dem zehnten Monat beobachtbar ist. Am stärksten reagiert der deutsche Aktienindex auf Änderungen im eigenen Land. Die Reaktion bleibt nach zehn Monaten erhalten und führt zum neuen Niveau des Aktienindex. Ähnlich reagiert der deutsche Index auf Schocks in anderen Märkten, allerdings ist die Reaktion wesentlich schwächer. Eine Ausnahme stellt hier wieder der amerikanische Markt dar. Die Reaktion des deutschen Index auf die Änderung auf dem amerikanischen Markt ist im ersten Monat stärker im Vergleich zu den anderen ausländischen Märkten, allerdings wird sie in den weiteren Monaten nicht wesentlich schwächer sein, als es bei dem französischen oder dem britischen Aktienmarkt der Fall ist. Der deutsche Markt reagiert am schwächsten auf Schocks auf dem japanischen Aktienmarkt. Die Reaktion ist ähnlich zu der Reaktion auf Änderungen auf dem amerikanischen Aktienindex.

In beiden Modellen bei der Varianz-Dekomposition wächst die Varianz des Prognosefehlers von 4% auf 20% pro Monat innerhalb eines Jahres. Die Bewegung des deutschen Marktes spielt die größte Rolle für eine Vorhersage dieser Varianz. Der Einfluss sinkt allerdings von 100% im ersten Monat bis auf 89,01% nach einem Jahr. Dagegen wächst die Wirkung des amerikanischen Marktes für die Varianz der Prognosefehler bis auf 9,49% nach einem Jahr. Der Einfluss der europäischen Märkte und des japanischen Marktes auf die Vorhersage des Prognosefehlers ist für die Entwicklung des deutschen Aktienmarktes vernachlässigbar. Der Gesamtwert liegt unter 3%.

Die Existenz des ökonomisch begründeten globalen Gleichgewichts auf den internationalen Aktienmärkten kann als einen Hinweis auf eine relativ sichere langfristige Anlagestrategie gesehen werden. Zum einen sind die Erträge aus dem Aktiengeschäft höher als im Fall der privaten Rentenversicherung¹. Sie wird im weiteren unattraktiv, wenn die Steuervorteile für das Versicherungsgeschäft abgeschafft werden². Zum anderen trägt die Stagnation des Immobilienmarktes seit zehn Jahren zur Attraktivität des Aktiengeschäftes bei. Dementsprechend kann eine langfristige international diversifizierte Anlagestrategie als eine mögliche zusätzliche Altersvorsorge gesehen werden.

Eine Anwendung der Kointegrationsmethode für die Untersuchung ökonomischer gleichgewichtiger Beziehungen ist nicht unproblematisch. Kointegrationstheoretische Ansätze basieren auf dem funktionalen Gleichgewichtskonzept der ökonomischen Theorie. Es lässt sich ein Unterschied in den Konzepten der Gleichgewichte feststellen. Ein

¹Die Rendite der privaten Rentenversicherung lag im Durchschnitt bei 3% in den letzten 30 Jahren.

²Die immer neu entstehende Diskussion bezüglich der Steuervorteile der Versicherungsverträge führte zu einer Anzahl der neuen Produkte, wie Rister-Rente oder Fond-gebundene Policen.

Gleichgewicht ist als ein Zustand eines Systems definiert, bei dem keine Tendenz zu einer Änderung vorliegt. Ein Ungleichgewicht ist durch eigene Destruktion charakterisiert. Ein gleichgewichtiger Zustand des Systems kann (muss aber nicht) die Eigenschaft einer lokalen oder globalen Stabilität haben. Damit ist gemeint, dass das System nach einem exogenen Schock zu seinem stabilen Zustand zurückkehrt. Generell sind in der Ökonomie die stabilen Gleichgewichte von Interesse. Es ist ebenfalls vorstellbar, dass bestimmte politische oder ökonomische Entscheidungen das ökonomische System zu dem gewünschten Gleichgewicht führen können.

Ein allgemeines oder ein partielles Gleichgewicht kann mit Hilfe einer Funktion $f(x_1, x_2, \dots, x_n) = 0$ dargestellt werden, die eine Verbindung zwischen n ökonomischen Variablen beschreibt. Der Begriff eines *'langfristigen Gleichgewichts'* wird zur Notation einer gleichgewichtigen Verbindung verwendet, zu der das System über die Zeit konvergiert. Eine langfristige gleichgewichtige Beziehung zieht eine systematische gemeinsame Bewegung der ökonomischen Variablen nach sich, die beispielsweise durch eine lineare Gleichung $x_1 = \beta_2 x_2 + \dots + \beta_n x_n$ präzisiert werden kann.

Die genannte Definition eines Gleichgewichts ist keine Definition im Sinne der Markträumung. Selbstverständlich existiert eine Verbindung zwischen den beiden Definitionen. Wenn die Märkte geräumt sind, dann kann ein Gleichgewicht in einer funktionalen Form auftreten. Die Markträumung kann ein ökonomisches System zu einem Zustand führen, bei dem eine funktionale Verbindung zwischen den Variablen eingehalten wird. Die eingeführte funktionale Form eines Gleichgewichts ist allgemeiner als die Markträumungsbedingung.

Ein statistisches Konzept für ein funktionales Gleichgewicht basiert auf der Stationaritätseigenschaft der stochastischen Prozesse. Das Gleichgewicht, in einer funktionalen Form $f(x_1, x_2, \dots, x_n) = 0$, wird zwischen n Variablen eingehalten, wenn die Abweichung von dem Gleichgewicht $\varepsilon = (f(x_1, x_2, \dots, x_n) - 0)$ einen stationären stochastischen Prozess mit einem Mittelwert von Null entspricht. Der Fehler oder die Diskrepanz zwischen der Beobachtung und dem festgelegten Gleichgewicht hat eine festgesetzte Verteilung, die um Null zentriert ist und sich mit der Zeit nicht ändert.

Ein weiterer Faktor, der bei der Kointegrationsanalyse berücksichtigt werden muss, ist die zeitabhängige Varianz der Renditen. Dies kann durch eine Modellierung des ARCH / GARCH-Prozesses bei Johansen's Kointegrationstests und einer entsprechenden Schätzung des VECM geschehen. Die Existenz mehrerer Aktiencrashes während der letzten 40 Jahre liefert Hinweise, dass Strukturbrüche in der langfristigen Beziehung zwischen den Aktienmärkten der G5-Länder existieren könnten. Daher muss ein weiterer Schritt der Untersuchung sein, die Stabilität des Systems zu testen. Es bestehen zwei Möglichkeiten, die entsprechenden Tests unter Berücksichtigung der Strukturbrüche durchzuführen. Entweder weist man zunächst die Zeitpunkte der möglichen Strukturbrüche aus und führt die Kointegrationsanalyse unter Berücksichtigung dieser Zeitpunkte durch, oder man führt die Kointegrationsuntersuchung simultan unter unbekanntem Bruchpunkten durch.

Für eine Interpretation der in dieser Arbeit gefundenen Kointegrationsbeziehungen wäre es sinnvoll nachzuprüfen, ob die Beziehungen durch Fundamentalfaktoren erklärbar sind. Dafür sind weitere Untersuchungen für die wichtigsten gesamtwirtschaftlichen Indikatoren wie BIP, Zinssatz oder Wechselkurs nötig. Signifikante langfristige Beziehungen zwischen unterschiedlichen Aktienmärkten können mehrere Ursachen haben, inklusive strenger ökonomischer Ursachen und politischer Koordination, Reaktion auf internationale Handelsbedingungen, Finanzinnovationen und technologischer Fortschritt, Markt deregulation und Liberalisierung, internationale Aktivität der Unternehmen, Zinssatzbewegung und internationale Kapitalbewegung, Finanzkrisen und Ansteckungseffekte. Langfristige Verbindungen zwischen den Aktienmärkten haben wichtige regionale und internationale Implikationen, weil eine Isolation des inländischen Kapitalmarktes von externen Schocks unmöglich ist.

Literaturverzeichnis

- [1] N. Ahlgren and J. Antell. Testing for cointegration between international stock prices. *Applied Financial Economics*, 12:851–861, 2002.
- [2] B. Arshanapalli and J. Doukas. International stock market linkages: Evidence from the pre- and post-october 1987 period. *Journal of Banking and Finance*, 17:193–208, 1993.
- [3] A. Banerjee, J. Dolado, J.W. Galbraith, and D.F. Hendry. *Co-Integration, Error-Correction, and the Econometric Analysis of Non-Stationary Data*. Oxford University Press, 1997.
- [4] Rolf W. Banz. The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9:1:3–18, 1981.
- [5] Sanjay Basu. Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *Journal of Finance*, 12:3:129–56, 1977.
- [6] S. Beveridge and C.R. Nelson. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics*, 7:151–174, 1981.
- [7] Fischer Black, Michael C. Jensen, and Myron Scholes. The capital asset pricing model: Some empirical tests. In *Studies in the Theory of Capital Markets*, pages 79–121. Michael C. Jensen ed. New York: Praeger, 1972.
- [8] H.P. Boswijk and P.H. Franses. Dynamic specification and cointegration. *Oxford Bulletin of Economics and Statistics*, 54:369–381, 1992.
- [9] J. Y. Campbell. Stock returns and the term structure. *Journal of Financial Economics*, 18:373–399, 1987.
- [10] John Y. Campbell, Andrew W. Lo, and A. Craig MacKinlay. *The Econometrics of Financial Markets*. Princeton University Press, Princeton, New Jersey, 1997.

- [11] John Y. Campbell and Robert J. Shiller. Stock prices, earnings, and expected dividends. *The Journal of Finance*, XLIII(NO 3):661–676, 1988.
- [12] John Y. Campbell and Robert J. Shiller. Valuation ratios and the long-run stock market outlook. *Journal of Portfolio Management*, Winter 24:11–26, 1998.
- [13] Mark M. Carhart, Jenifer N. Carpenter, Anthony W. Lynch, and David K. Musto. Mutual fund survivorship. *Review of Financial Studies*, Vol. 15 Issue 5:1439–1464, 2002.
- [14] Y. Cheung and K. Lai. Finite sample sizes of johansen’s likelihood ratio tests for cointegration. *Oxford Bulletin of Economics and Statistics*, 55:313–328, 1993.
- [15] I. Choi. Durbin-hausman tests for unit root. *Oxford Bulletin of Economics and Statistics*, 54:289–304, 1992.
- [16] R.Y. Chou and W.K. Ng. Correlations structure of the permanent and temporary components of international stock market proces. Working Paper, georgia Institute of Technology, 1995.
- [17] J.H. Cochrane. A critique of tne application of unit root tests. *Journal of Economic Dynamics and Control*, 15:275–284, 1991.
- [18] A. Corhay, Rad A. Tourani, and J.-P. Urbain. Common stochastic trends in european stock markets. *Economics Letters*, 42:385–390, 1993.
- [19] R. Deaves. The comparative performance of load and no-load mutual funds in canada. *Canadian Journal of Administrative Sciences*, 21(4):326–333, 2004.
- [20] D.N. DeJong, J.C. Nankervis, N.E. Svin, and C.H. Whiteman. Integration versus trend stationarity in time series. *Econometrica*, 60:423–433, 1992.
- [21] D.N. DeJong, J.C. Nankervis, N.E. Svin, and C.H. Whiteman. The power problems of unit root tests in time series with autoregressive errors. *Journal of Econometrics*, 53:323–343, 1992.
- [22] David A. Dickey and Wayne A. Fuller. Distribution of the estimators for autoregressive time series with unit root. *Journal of the American Statistical Association*, 74:427–431, 1979.
- [23] David G. Dickinson. Stock market integration and macroeconomic fundamentals: An empirical analysis 1980-1995. *Applied Financial Economics*, 10:261–276, 2000.
- [24] G. Elliot, T.J. Rothenberg, and J.H. Stock. Efficient tests for an autoregressive unit root. *Econometrica*, 64:813–836, 1996.
- [25] Edwin J. Elton, Martin J. Gruber, Stephen J. Brown, and William N. Goetzmann. *Modern Portfolio Theory and Investment Analysis*. John Wiley & Sons, inc., sixth edition, 2003.

- [26] C. Engel. A note on cointegration and international market efficiency. *Journal of International Money and Finance*, 15:657–60, 1996.
- [27] C. Engle and V. K. Ng. Measuring and testing the impact of news on volatility. *Journal of Finance*, 5:1749–78, 1993.
- [28] R.F. Engle and C.W.J. Granger. Co-integration and error correction: Representation, estimation and testing. *Econometrica*, 55:251–276, 1987.
- [29] E. Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25:383–417, 1970.
- [30] E. Fama. Efficient capital markets:II. *Journal of Finance*, 46:1575–1618, 1991.
- [31] E. F. Fama. Market efficiency, long-term returns and behavioral finance. *Journal of Financial Economics*, 49:283–306, 1998.
- [32] Eugen F. Fama and Kenneth R. French. The capital asset pricing model: Theory and evidence. *Journal of Economic Perspectives*, 18(3):25–46, 2004.
- [33] Eugene Fama and Kenneth French. Business conditions and expected return on stocks and bonds. *Journal of Financial Economics*, 25:23–50, 1993.
- [34] Eugene Fama and Kenneth French. Common risk factors in the returns on Stocks and bonds. *Journal of Financial Economics*, 33:1:3–56, 1993.
- [35] Eugene Fama and Kenneth French. Disappearing dividends: Changign firm characteristics or lower propensity to pay. *Journal of Financial Economcs*, 60:1:3–43, 2001.
- [36] Eugene F. Fama and Keneth R. French. The cross-section of expected stock returns. *Journal of Finance*, 47:2:427–65, 1992.
- [37] Eugene F. Fama and Kenneth R. French. Permanent and temporary components of stock prices. *Journal of Political Economy*, 96:246–272, 1988.
- [38] Eugene F. Fama and Kenneth R. French. Multifactor explanations of asset pricing annomalies. *Journal of Finance*, 51:1:55–84, 1996.
- [39] Eugene F. Fama and Kenneth R. French. Forecasting profitability and earnings. *Journal of Business*, 73(2):161– 176, 2000.
- [40] S. G. M. Fifield, A. A. Lonie, D. M. Power, and C. D. Singlair. Emerging markets: A diasaggregated perspective on the gains from investing internationally. *Review of Pacific Basin Financial Markets and Policies*, 2(1):99–124, 1999.
- [41] S. G. M. Fifield, D. M. Power, and C. D. Sinclair. Emerging stock markets: A more realistic assessment of the gains from diversifcation. *Applied Financial Economics*, 12:213–229, 2002.

- [42] Zsuzsanna Fluck, Burton Malkiel, and Richard Quandt. The predictability of stock returns: A cross-sectional simulation. *Review of Economics and Statistics*, 79:2:176–83, 1997.
- [43] Patricia Fraser and David Power. Stock return volatility and information: An empirical analysis of pacific rim, UK and US equity markets. *Applied Financial Economics*, 7:241–254, 1997.
- [44] I. Friend, Y. Landskroner, and E. Losq. The demand for risky assets under uncertain inflation. *Journal of Finance*, 11, No.31:1287–1297, 1976.
- [45] Wayne A. Fuller. *Introduction to Statistical Time Series*. John Wiley, New York, 1976.
- [46] C. Gauthier. The fair value of the u.s. stock market: A structural VECM approach. Department of Finance Working Paper, 2001.
- [47] W. Gerke, F. Mager, and A. Röhrs. Twenty years of international diversification from a german perspective. *Schmalenbach Business Review*, 57:86–102, 2005.
- [48] Robert-Jan Gerrits and Ayse Yüce. Short- and long-term links among european and US stock markets. *Applied Financial Economics*, 9:1–9, 1999.
- [49] Claire G. Gilmore and Ginette M. McManus. International portfolio diversification: US and central european equity markets. *Emerging Markets Review*, 3:69–83, 2002.
- [50] M.-J. Godbout and S. Van Norden. Reconsidering cointegration in international finance: Three case studies of size distortion in finite samples. Working Paper, 1997.
- [51] William H. Green. *Econometric Analysis*. Pearson Education International, fifth edition, 2003.
- [52] S. Grossman and J. Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70:393–408, 1980.
- [53] M. Gruber. Another puzzle: The growth in actively managed mutual funds. *Journal of Finance*, 51:783–810, 1996.
- [54] A. Hall. Testing for a unit root in time series with pretest data based model selection. *Journal of Business and Economic Statistics*, 12:461–470, 1994.
- [55] James D. Hamilton. *Time Series Analysis*. Princeton University Press, New Jersey, 1994.
- [56] Michael E. Hanna, Joseph P. McCormack, and Grady Perdue. A nineties perspective on international diversification. *Financial Services Review*, 8:37–45, 1999.

- [57] Kirstin Hubrich, Helmut Lütkepohl, and Pentti Saikkonen. A review of systems cointegration tests. *Econometrics Reviews*, 20(3):247–318, 2001.
- [58] S. Johansen. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press: Oxford, 1995.
- [59] Soren Johansen. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12:231–254, 1988.
- [60] Soren Johansen. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59:1551–1580, 1991.
- [61] Soren Johansen and Katarina Juselius. Maximum likelihood estimation and inference on cointegration - with application to the demand for money. *Oxford Bulletin of Economics and Statistics*, 52:169–210, 1990.
- [62] K. Kasa. Common stochastic trends in international stock markets. *Journal of Monetary Economics*, 29:95–124, 1992.
- [63] Jeong-Ryeol Kim. The stock return-inflation puzzle and the asymmetric causality in stock returns, inflation and real activity. *Economics Letters*, 80:155–160, 2003.
- [64] Mervyn King, Enrique Sentana, and Sushil Wadhvani. Volatility and links between national stock markets. *Econometrica*, 62(4):901–933, 1994.
- [65] A.C.C Kwan, A.-B. Sim, and J.A. Cotsomitis. The causal relationship between equity indices on world exchanges. *Applied Economics*, 27:33–37, 1995.
- [66] L.L. Leachman and B. Francis. Long-run relations among the g-5 and g-7 equity markets: Evidence on the plaza and louvre accords. *Journal of Macroeconomics*, 17:551–77, 1995.
- [67] S.J. Leybourne. Testin for unit roots using forward and reverse dickey-fuller regressions. *Oxford Bulletin of Economics and Statistics*, 57:559–571, 1995.
- [68] J. Lintner. Security prices, risk and maximal gains from diversification. *Journal of Finance*, 20:587–615, 1965a.
- [69] P.C. Liu and J.Praschnik. The size of the nonstationary component and its effect on tests for unit roots. *Statistical Papers*, 34:83–88, 1993.
- [70] Andrew W. Lo and A.Craig MacKinlay. *A Non-Random Walk Down Wall Street*. Princenton: Princenton University Press, 1999.
- [71] Andrew W. Lo, Harry Mamaysky, and Jiang Wang. Foundation of technical analysis: Computational algorithmus,statistical inference, and empirical implementation. *Journal of Finance*, 55:4:17051756, 2000.

- [72] F. Longin and B. Solnik. Is the correlation in international equity returns constant: 1970-1990. *Journal of International Money and Finance*, 13:3–26, 1995.
- [73] François Longin and Bruno Solnik. Extreme correlation of international equity markets. *Journal of Finance*, 56, 2001.
- [74] H. Lütkepohl. *Forecasting Aggregated Vector ARMA Processes*. Springer-Verlag, Berlin, 1987.
- [75] H. Lütkepohl and P. Saikkonen. Order selection in testing for the cointegrating rank of a VAR process. In R.F. Engle and H. White, editors, *Cointegration, Causality, and Forecasting. A Festschrift in Honour of Clive W.J. Granger*, pages 168–199. Oxford University Press: Oxford, 1999.
- [76] Helmut Lütkepohl. *Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin, 1991.
- [77] Helmut Lütkepohl. Recent advances in cointegrations analysis. In *EUI Working Paper ECO No. 2004/12*, pages 1–40. European University Institute Department of Economics, Badia Fiesolana, I-50016 San Domenico Italy, 2004.
- [78] Helmut Lütkepohl, Pentti Saikkonen, and Carsten Trenkler. Testing for the cointegrating rank of a VAR process with level shift at unknown time. *Econometrica*, 72(2):647–663, 2004.
- [79] James G. Mackinnon. Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics*, 11(6):601–619, 1996.
- [80] G. S. Maddala and In-Moo Kim. *Unit Roots, Cointegration, and Structural Change*. Cambridge University Press, Cambridge, 1998.
- [81] B. Malkiel. Efficient market hypothesis. In P. Newman, M. Milgate, and J. Eatwell, editors, *New Palgrave Dictionary of Money and Finance*. Macmillan, London, 1992.
- [82] B. Malkiel. Returns from investing in equity mutual funds 1971 to 1991. *Journal of Finance*, 50:549–572, 1995.
- [83] Burton G. Malkiel. The efficient market hypothesis and its critics. *The Journal of Economic Perspectives*, 17 Number 1:59–82, 2003.
- [84] Burton G. Malkiel. Reflections on the efficient market hypothesis: 30 years later. *The Financial Review*, 40:1–9, 2005.
- [85] H. Markowitz. *Portfolio Selection: Efficient Diversification of Investments*. John Wiley, New York, 1959.

- [86] R. Meese and J. Geweke. A comparison of autoregressive univariate forecasting procedures for macroeconomic time series. *Journal for Business and Economic Statistics*, 2:191–200, 1984.
- [87] R.O. Michaud, G.L. Bergstrom, R.D. Frashure, and B.K. Wolahan. Twenty years of international equity investing: Still a route to higher returns and lower risks? *Journal Portfolio Management*, 23(1):9–22, 1996.
- [88] P. K. Narayan and R. Smyth. Modelling the linkages between the Australian and G7 stock markets: Common stochastic trends and regime shifts. *Applied Financial Economics*, 14:991–1004, 2004.
- [89] C.R. Nelson and C.I. Plosser. Trends and random walks in macroeconomic time series. *Journal of Monetary Economics*, 10:139–162, 1982.
- [90] S. Ng and P. Perron. Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of American Statistical Association*, 90:268–281, 1995.
- [91] Terrance Odean. Do investors trade too much? *American Economic Review*, 89:5:1279–298, 1999.
- [92] M. Osterwald-Lenum. A note with quartiles as the asymptotic distribution of the maximum likelihood cointegration rank test statistics. *Oxford Bulletin of Economics and Statistics*, 54(3):461–72, 1992.
- [93] Ming-Shiun Pan, Y. Angela Liu, and Herbert J. Roth. Term structure of return correlations and international diversification: Evidence from European stock markets. *The European Journal of Finance*, 7:144–164, 2001.
- [94] Anre F. Perold. The capital asset pricing model. *Journal of Economic Perspectives*, 18(3):3–24, 2004.
- [95] P. Perron. The great crash, the oil price shock and the unit root hypothesis. *Econometrica*, 57:1361–1401, 1989.
- [96] P. Perron and S. Ng. Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. *Review of Economic Studies*, 63:435–365, 1996.
- [97] P.C.B. Phillips. Time series regression with a unit root. *Econometrica*, 55:277–301, 1987.
- [98] P.C.B. Phillips. Optimal inference in cointegrated systems. *Econometrica*, 59:238–306, 1991.
- [99] P.C.B. Phillips. To criticize the critics: An objective Bayesian analysis of stochastic trends. *Journal of Applied Econometrics*, 6:333–364, 1991.

- [100] P.C.B. Phillips and P.Perron. Testing for a unit root in time series regression. *Biometrika*, 75:335–346, 1988.
- [101] Rank K. Reilly and Keith C. Brown. *Investment Analysis and Portfolio Management*. Thomson South-Western, 2003.
- [102] A.J. Richards. Comovements in national stock market returns: Evidence of predictability, but not cointegration. *Journal of monetary economics*, 36:631–654, 1995.
- [103] H Roberts. Statistical versus clinical prediction of the stock market. Center for Research in Security Prices, University of Chicago, 1967.
- [104] Richard Roll. A critique of the asset pricing theory's tests' part i: On past and potential testability of the theory. *Journal of Financial Economics*, 4:2:129–76, 1977.
- [105] Richard Roll and Robert J. Shiller. Comments: Symposium on volatility in u.s. and japanese stock markets. *Journal of Applied Corporate Finance*, 5:1:25–29, 1992.
- [106] P. Samuelson. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6:41–49, 1965.
- [107] M. Scheicher. The comovement of stock markets in hungary, poland and the czech republic. *International Journal of Finance and Economics*, 6:27–39, 2001.
- [108] Rainer Schlittgen and Bernd H. J. Streitberg. *Zeitreihenanalyse*. Oldenbourg, 1999.
- [109] G. William Schwert. Anomalies and market efficiency. In *Handbook of the Economics of Finance*, chapter 17. G. Constantinides et al., Amsterdam: North Holland, 2001.
- [110] G.W. Schwert. Tests for unit roots: A monte carlo investigation. *Journal for Business and Economic Statistics*, 7:147–159, 1989.
- [111] W. Semmler. *Asset Prices, Booms and Recessions*. Springer-Verlag, Berlin Heidelberg New York, 2003.
- [112] Byeongseon Seo. Distribution theory for unit root tests with conditional heteroskedasticity. *Journal of Econometrics*, 91:113–144, 1999.
- [113] W. Sharpe. Capital asset prices: A theory of market equilibrium under conditional of risk. *Journal of Finance*, 19:425–442, 1964.
- [114] Hany A. Shawky, Rolf Kuenzel, and Azmi D. Mikhail. International portfolio diversification: A synthesis and an update. *Journal of International Financial Markets, Institutions and Money*, 7:303–327, 1997.

- [115] Robert J. Shiller. *Irrational Exuberance*. Princeton: Princeton University Press, 2000.
- [116] Robert J. Shiller. From efficient market theory to behavioral finance. *Journal of Economic Perspectives*, 17, Number 1:83–104, 2003.
- [117] B. Solnik, C. Boucrelle, and Y. Le Fur. International market correlation and volatility. *The Journal Financial Analysts*, 52(5):57–67, 1996.
- [118] B. H. Solnik. Why not diversify internationally rather than domestically? *Financial Analysts Journal*, 20(4):48–54, 1974.
- [119] Bruno Solnik. *International Investment*. Addison-Wesley, 1988.
- [120] J.H. Stock. Unit roots, structural breaks, and trends. In R.F. Engle and D.L. McFadden, editors, *Handbook of Econometrics*, page vol. IV. Elsevier, Amsterdam, 1994.
- [121] T. Syriopoulos. International portfolio diversification to central european stock markets. *Applied Financial Economics*, 14:1253–1268, 2004.
- [122] M.P. Taylor and I. Tonks. The internationalisation of stock markets and the abolition of UK exchange controls. *Review of Economics and Statistics*, 71:332–336, 1989.
- [123] E. Elton und M. Gruber and C. Blake. The persistence of risk-adjusted mutual fund performance. *Journal of Business*, 69:133–157, 1996.
- [124] K.D. West. Asymptotic normality, when regressors have a unit root. *Econometrica*, 56:1397–1417, 1988.
- [125] Jeffrey Wooldridge. *Notes on Regression with Difference-Stationary Data*. Michigan State University, 1991.
- [126] E. Zivot and D. Andrews. Further evidence of great crash, the oil-price shock and uni-root hypothesis. *Journal of Business and Economic Statistics*, 10:251–70, 1992.