

Dissertation  
zur Erlangung des Grades  
eines Doktors der Naturwissenschaften  
der Technischen Universität Dortmund

---

**Sequenzielle Erkennung von  
Strukturbrüchen in der  
Varianzstruktur von multivariaten  
Zeitreihen**

---

Der Fakultät Statistik  
der Technischen Universität Dortmund  
vorgelegt von  
**Katharina Pape**

Eingereicht am: 8. Juni 2017

Mündliche Prüfung am: 2. August 2017

Erstgutachter: Prof. Dr. Dominik Wied

Zweitgutachter: Prof. Dr. Walter Krämer



# Zusammenfassung

In dieser Arbeit werden zwei sequenzielle Verfahren zur Detektion von Strukturbrüchen in den Varianzen von Zufallsvektoren vorgestellt. Um Referenzwerte zu erhalten, nutzen beide Verfahren die Information aus einer historischen Stichprobe, die als frei von Strukturbrüchen angenommen wird und die schon zu Beginn der eigentlichen Überwachungsphase vorliegt. Die Referenzwerte können mit Schätzwerten verglichen werden, die auf dem stückweise anwachsenden Datensatz des Überwachungszeitraums beruhen. Somit basieren die Detektorwerte auf der Information des historischen Zeitraums und werden für jeden Zeitpunkt des Überwachungszeitraums mit Hilfe der neu hinzugewonnenen Information aktualisiert. Anschließend können die Detektorwerte mit einer passend skalierten Grenzfunktion verglichen werden, wobei ein Überschreiten der Grenzfunktion durch die Detektorwerte das Vorliegen eines Strukturbruchs in der jeweils betrachteten Größe anzeigt.

Das erste Verfahren ist ein nichtparametrischer Strukturbruchtest, dessen Detektor auf den Differenzen von Varianzschätzern aus der historischen Stichprobe und aus den Daten des Überwachungszeitraums basiert. Unter der Nullhypothese, dass sich der Vektor der Varianzen während des gesamten Überwachungszeitraums nicht ändert, ist davon auszugehen, dass sich aus den Daten des Überwachungszeitraums bestimmte Varianzschätzer nur geringfügig von aus der historischen Stichprobe ermittelten Schätzwerten unterscheiden. Starke Fluktuationen hingegen, wie sie durch sehr unterschiedliche Schätzwerte hervorgerufen werden, sprechen für das Vorliegen eines Strukturbruchs in der betrachteten Zeitreihe.

Das zweite Verfahren ist ein parametrischer Test auf eine Änderung des Parametervektors im multivariaten GARCH-Modell mit dynamischer bedingter Korrelation (DCC-Modell) von Engle. Strukturbrüche in den Modellparametern können dabei nicht nur Änderungen in den bedingten und unbedingten Varianzen hervorrufen, sondern auch in der bedingten und unbedingten Kovarianz- und Korrelationsstruktur. Dies erlaubt eine weitreichende Detektion von strukturellen Änderungen in dieser Art von Modell, die beispielsweise für die Modellierung von Aktienrenditen gut geeignet ist. Das Verfahren basiert auf den Scores der Gaußschen Quasi-Loglikelihoodfunktion. Unter der Annahme eines konstanten Parametervektors sollten die Beiträge zum Gradienten der Quasi-Loglikelihoodfunktion für den aus der historischen Stichprobe ermittelten Parameterschätzer und für Beobachtungen aus dem Überwachungszeitraum nicht viel stärker schwanken als dies für die Werte aus dem historischen Zeitraum beobachtet wurde.

Da für beide betrachteten Methoden die Stoppzeit in der Regel nicht mit dem eigentlichen Bruchpunkt zusammenfällt, werden außerdem zwei Schätzer vorgestellt, mit denen nach dem Nachweis der bloßen Präsenz eines Strukturbruchs auch dessen Lage geschätzt werden kann. Für beide betrachteten Verfahren beziehungsweise die verwendeten Detektoren werden asymptotische Eigenschaften unter der Nullhypothese und der Alternative eines Strukturbruchs zu einem unbekanntem Zeitpunkt im Überwachungszeitraum hergeleitet. In beiden Fällen streben die Folgen der Detektorwerte unter der Nullhypothese gegen multivariate Gaußprozesse. Außerdem wird das Verhalten in endlichen Stichproben mit Hilfe von Simulationen und Anwendungen auf echte Daten, in diesem Fall logarithmierte Eintaesrenditen der Kurswerte von verschiedenen Aktien, untersucht.



# Inhaltsverzeichnis

<b>Kapitel 1 Einleitung und Motivation</b>	<b>1</b>
<b>Kapitel 2 Grundlagen und Notation</b>	<b>5</b>
2.1 Sequenzielle Überwachung von Zeitreihen . . . . .	5
2.2 Univariate und multivariate GARCH-Modelle . . . . .	7
2.3 Operatoren und Transformationsmatrizen . . . . .	9
<b>Kapitel 3 Nichtparametrische Detektion von Varianzänderungen</b>	<b>11</b>
3.1 Das Überwachungsszenario . . . . .	11
3.2 Das Überwachungsverfahren . . . . .	12
3.3 Simulationen . . . . .	22
3.4 Anwendung auf Aktienrenditen . . . . .	35
<b>Kapitel 4 Das DCC-Modell nach Engle</b>	<b>41</b>
4.1 Die Einführung des DCC-Modells . . . . .	41
4.2 Die Schätzung der Modellparameter . . . . .	44
4.3 Die Existenz einer stationären Lösung . . . . .	46
<b>Kapitel 5 Detektion von Parameteränderungen im DCC-Modell</b>	<b>49</b>
5.1 Das Überwachungsszenario . . . . .	49
5.2 Das Überwachungsverfahren . . . . .	52
5.3 Simulationen . . . . .	58
5.4 Anwendung auf Aktienrenditen . . . . .	64
<b>Kapitel 6 Fazit und Ausblick</b>	<b>69</b>
<b>Anhang A Beweise der Aussagen aus Kapitel 3</b>	<b>73</b>
A.1 Beweis von Theorem 3.1 . . . . .	73
A.2 Beweis von Theorem 3.2 . . . . .	74
A.3 Beweis von Theorem 3.3 . . . . .	75
A.4 Beweis von Lemma 3.1 . . . . .	76
<b>Anhang B Simulationsergebnisse aus Kapitel 3.3</b>	<b>77</b>
B.1 U.i.v. Zufallsvektoren unter der Nullhypothese . . . . .	77
B.2 U.i.v. Zufallsvektoren unter verschiedenen Alternativen . . . . .	78
B.3 GARCH-Prozesse unter der Nullhypothese . . . . .	80
B.4 GARCH-Prozesse unter verschiedenen Alternativen . . . . .	81
B.5 Vergleich von multivariatem und univariaten Verfahren . . . . .	83
B.6 Alternative, dass sich die Innovationsvarianzen ändern . . . . .	85
<b>Anhang C Darstellung der DCC-Modellgleichung als Markovkette</b>	<b>87</b>

<b>Anhang D Ableitungen der Beiträge zur Quasi-Loglikelihoodfunktion</b>	<b>91</b>
D.1 Rechenregeln für Matrizen . . . . .	91
D.2 Die ersten Ableitungen nach den Varianzparametern . . . . .	93
D.3 Die ersten Ableitungen nach den Korrelationsparametern . . . . .	94
D.4 Die zweiten Ableitungen nach den Modellparametern . . . . .	96
<b>Anhang E Beweise der Aussagen aus Kapitel 5</b>	<b>101</b>
E.1 Beweis von Proposition 5.1 . . . . .	101
E.2 Beweis von Proposition 5.2 . . . . .	103
E.3 Beweis von Theorem 5.1 . . . . .	131
E.4 Beweis von Theorem 5.2 . . . . .	133
<b>Anhang F Miscellanea</b>	<b>135</b>
<b>Tabellenverzeichnis</b>	<b>141</b>
<b>Abbildungsverzeichnis</b>	<b>143</b>
<b>Symbolverzeichnis</b>	<b>145</b>
<b>Abkürzungsverzeichnis</b>	<b>151</b>
<b>Literatur</b>	<b>153</b>
<b>Eidesstattliche Versicherung</b>	<b>157</b>





# Kapitel 1

## Einleitung und Motivation

Varianzen und Volatilitäten sind in ökonometrischen und finanzmathematischen Anwendungsbereichen von großem Interesse, da sie genutzt werden können, um das Risiko einer Vielzahl von Finanzinstrumenten zu messen. In Bezug auf das Streuungsverhalten herrscht im Bereich der Finanzökonometrie weitgehend Einigkeit darüber, dass bedingte Varianzen dynamisch modelliert werden sollten, siehe beispielsweise Engle (1982) oder Bollerslev (1986). Viele empirische Arbeiten weisen zudem ein stark variierendes Verhalten von Volatilitäten in Finanzzeitreihen nach, siehe zum Beispiel Wilson et al. (1996), Dolde und Tirtiroglu (2002) oder Kang et al. (2009), sodass die Möglichkeit einer zeitvariablen Modellierung auch für unbedingte Varianzen stets geprüft werden sollte. Zumindest aber muss berücksichtigt werden, dass das Streuungsverhalten vieler Finanzinstrumente nicht als über die Zeit konstant angenommen werden kann. Insbesondere gilt dies für die Varianzen von Aktienrenditen, die dazu neigen, in Krisenzeiten deutlich anzusteigen, wie zum Beispiel Schwert (1989 und 2011) oder Charles und Darné (2014) belegen.

Sind die betrachteten Finanzzeitreihen multivariat, so kann auch die gemeinsame Kovarianz- und Korrelationsstruktur betrachtet werden, um die in den Daten enthaltene Information optimal zu nutzen. Dass auch die Korrelationen nicht als konstant über die Zeit angenommen werden können, belegen unter anderem Longin und Solnik (1995), Karolyi und Stulz (1996) sowie De Santis und Gerard (1997). Außerdem zeigen Sandoval Jr. und De Paula Franca (2012), dass in Krisenzeiten sogar die Korrelation zwischen unterschiedlichen Finanzmärkten in Form der Renditen verschiedener Aktienindizes ansteigt.

In den letzten Jahren sind nicht nur viele Forschungsarbeiten auf den Gebieten der Modellierung von Volatilitäts- und Korrelationsstrukturen sowie der Strukturbrucherkenung entstanden, sondern auch im Spannungsfeld dieser beiden Disziplinen. Hinsichtlich der Tatsache, dass Änderungen der Volatilitäts- und Korrelationsstruktur gerade in finanziellen Anwendungen weitreichende Konsequenzen haben können, ist vor allem dieser Überschneidungsbereich von hohem Interesse. Beispielsweise benötigen Analysten Informationen über aufgetretene Strukturbrüche in den niedrigeren Momenten zur Portfoliooptimierung oder um auf Anzeichen aufkommender Krisen reagieren zu können. Wenn die Varianzen und Korrelationen beziehungsweise die Parameter, die diese Größen steuern, im betrachteten Zeitraum nicht konstant sind, kann dies zu verzerrten Prognosen oder falschen Risikoeinschätzungen führen. Deshalb ist es wünschenswert, entsprechende Änderungen so schnell wie möglich nach ihrem Auftreten zu erkennen, um gegebenenfalls auf die veränderte Situation reagieren zu können, wie zum Beispiel durch Portfolioumschichtungen oder durch die Anpassung von Prognosen.

Vor dem Hintergrund ständig wachsender Zeitreihen, die gerade in ökonomischen und finanzwirtschaftlichen Kontexten beobachtet werden können, scheinen sequenzielle Methoden in der Praxis deutlich besser geeignet zu sein als retrospektive Tests. Können die

klassischen retrospektiven Testverfahren die Präsenz potenzieller Bruchpunkte erst anzeigen, nachdem der gesamte Datensatz erhoben wurde, so bieten sequenzielle Verfahren die Möglichkeit, Strukturbrüche schon kurz nach deren Auftreten anzuzeigen, da sie neue Information in Form von „frischen“ Datenpunkten nutzen, um den verwendeten Detektor zu aktualisieren. Zur Überwachung der Volatilitäts- und Korrelationsstruktur wurden bisher einige univariate sequenzielle Verfahren vorgestellt, zum Beispiel durch Wied und Galeano (2013) oder Golosnoy (2007), wobei sich letztere Arbeit direkt auf die Überwachung optimaler Portfoliogewichte konzentriert, die wiederum von den Varianzen der betrachteten Finanzwerte abhängen. Im Gegensatz dazu beziehen sich multivariate Verfahren in der Regel auf retrospektive Szenarien, wie zum Beispiel in Aue et al. (2009b) oder Wied (2015).

Mit dem Ziel, Änderungen in den Varianzen und Korrelationen zu erkennen, werden in dieser Arbeit der sequenzielle und der multivariate Ansatz miteinander verknüpft und zwei verschiedene Überwachungsverfahren vorgestellt. Dabei handelt es sich um ein nichtparametrisches Verfahren zur Überwachung der Varianzen sowie ein parametrisches Verfahren zur Detektion von Strukturbrüchen im Parametervektor des von Engle (2002) vorgeschlagenen multivariaten GARCH-Modells mit dynamischer bedingter Korrelation (DCC-Modell). Beide Methoden basieren in ihrer Grundidee auf der CUSUM (*cumulative sum*)-Methode von Chu et al. (1996), die voraussetzen, dass vor dem Einsetzen des Überwachungszeitraums bereits eine gewisse historische Datengrundlage vorliegt. Diese erlaubt nicht nur eine stabilere Parameterschätzung, sondern auch eine zuverlässigere Überwachung. Dieser Ansatz wurde unter anderem von Horváth et al. (2004), Aue et al. (2006), Aue et al. (2009b) und Groen et al. (2013) weiterentwickelt. Während Berkes et al. (2004) und Aue et al. (2012) den Anwendungsbereich der von Chu et al. (1996) vorgeschlagenen Methode auf univariate GARCH(p,q)-Modelle beziehungsweise Hochfrequenz-Portfolio-Betas erweitern, stellen Wied und Galeano (2013) ein modellunabhängiges Überwachungsverfahren zur Erkennung von Korrelationsänderungen in bivariaten Zeitreihen vor.

Im ersten Teil dieser Arbeit wird ein nichtparametrischer Strukturbruchttest vorgestellt, der ausschließlich den  $p$ -dimensionalen Vektor der Varianzen einer Folge von Zufallsvektoren überwacht. Dabei stellt sich die Frage, warum ein multivariates Verfahren der Anwendung von  $p$  univariaten Tests auf die einzelnen Komponentenprozesse vorgezogen werden sollte. Hierbei ist zu erwarten, dass die Zufallsvektoren deutlich mehr Information enthalten als dies für eine Sammlung univariater Prozesse der Fall wäre. Deshalb sollte untersucht werden, ob die zusätzliche Information auch von diesem Verfahren ausgenutzt werden kann, das sich nur auf die Streuungs-, nicht aber auf die Korrelationsstruktur konzentriert. In diesem Fall sollten Strukturbrüche unter der Verwendung eines multivariaten Verfahrens deutlich zuverlässiger erkannt werden als durch eine Gruppe univariater Verfahren, deren Niveau beispielsweise mit Hilfe der Bonferronikorrektur adjustiert werden müsste.

Die Beschränkung der vorgestellten Methode auf die Überwachung des Vektors der Varianzen basiert auf mehreren Argumenten. Zunächst einmal enthalten die Varianzen in finanzwirtschaftlichen Kontexten grundlegende Information über das spezifische Risiko einzelner Finanzinstrumente. Außerdem zeigen Bissantz et al. (2011), dass der Einfluss von Fluktuationen der Volatilitäten auf Risikokennzahlen wie den *Value at Risk* deutlich höher ist als der Einfluss von Korrelationsschwankungen. Des Weiteren würde die Inklusion der Kovarianzen schon für relativ niedrige Dimensionen der betrachteten Zufallsvektoren bewirken, dass die Dimension des Vektors der zu überwachenden Momente unbrauchbar hoch wird, da mit dieser auch die Dimension einer Variationsmatrix steigt, die zur Skalierung des Detektors geschätzt werden muss. In weiterführenden Simulationen, die nicht Teil dieser Arbeit sind, konnte beobachtet werden, dass die daraus resultierende wachsende Schätzunsicherheit zu einer beachtlichen Schwächung des Verfahrens hinsichtlich der

Niveau- und Güteeigenschaften führt. Ein weiterer Punkt besteht darin, dass für die Überwachung der Korrelation die Konstanz der Varianzen häufig eine wichtige, wenn auch zweifelhafte Grundannahme bildet. So werden beispielsweise in Wied et al. (2012a) nur sehr eingeschränkte Variationen der Varianzen zugelassen, während ein Test im Sinne von Dette et al. (2015) hinsichtlich dieser Annahme deutlich flexibler ist und nur die lokale Stationarität des betrachteten Prozesses voraussetzt. Als Konsequenz der nichtparametrischen Schätzung von Mittelwert und Varianz lässt dies allerdings eine niedrigere Güte als für das Verfahren von Wied et al. (2012a) erwarten. Ausgehend von diesen Argumenten empfiehlt es sich somit in diesem Fall davon abzusehen, die gesamte Kovarianzmatrix zu überwachen, wie es von Aue et al. (2009a) im retrospektiven Fall vorgeschlagen wurde, und im Gegensatz dazu allein nach Varianzänderungen zu suchen.

Konkret wird der auf der Arbeit von Chu et al. (1996) basierende sequenzielle Ansatz von Wied und Galeano (2013) mit dem multivariaten Verfahren von Aue et al. (2009b) kombiniert. Die Basis des verwendeten Detektors bilden skalierte Fluktuationen von Vektoren einfacher Varianzschätzer, die aus verschiedenen Teilstichproben bestimmt werden. Genauer gesagt werden die Abweichungen der Vektoren einfacher Varianzschätzer, die sich aus dem wachsenden Datensatz des Überwachungszeitraums bestimmen lassen, von aus einer historischen Stichprobe bestimmten Referenzschätzwerten betrachtet.

Bei der Anwendung dieses Verfahrens auf echte Daten, deren Struktur auf das Zugrundeliegen eines GARCH-Modells hinweist, ergaben sich einige Probleme, deren Lösung in Abschnitt 3.4 im Detail vorgestellt wird. Aus diesem Grund konzentriert sich der zweite Teil der Arbeit auf ein parametrisches Verfahren, das gezielt zur Überwachung der Parameter in einem multivariaten GARCH-Modell genutzt werden kann.

Im Forschungsbereich der Strukturbruchererkennung in univariaten GARCH-Modellen wie sie von Bollerslev (1986) eingeführt wurden, haben sich in den letzten Jahren einige neue Entwicklungen ergeben. So schlagen unter anderem Amado und Teräsvirta (2014) sowie Silvennoinen und Teräsvirta (2016) *Lagrange Multiplier*-Ansätze zur retrospektiven Strukturbruchsuche vor, während Berkes et al. (2004) ein dem Vorgehen von Chu et al. (1996) ähnelndes sequenzielles Verfahren präsentieren, das auf Quasi-Loglikelihoodscores beruht. Als multivariate Erweiterung des Verfahrens von Berkes et al. (2004) präsentiert diese Arbeit eine Methode zur Überwachung der Parameter im DCC-Modell von Engle (2002). Während andere Verfahren sich häufig auf die Strukturbruchererkennung in Varianzen oder Korrelationen konzentrieren, erlaubt dieser Ansatz eine gemeinsame Überwachung beider Größen, da die betreffenden Parameter sowohl die bedingte als auch unbedingte Varianz- und Korrelationsstruktur des Prozesses steuern. Dabei werden analog zum Vorgehen von Berkes et al. (2004) der historische Datensatz zur Schätzung der Modellparameter genutzt und die Beiträge zum Gradienten der Gaußschen Quasi-Loglikelihoodfunktion an der Stelle dieses Schätzers ausgewertet. Falls sich der Vektor der Parameter im Laufe des Überwachungszeitraums ändert, so ist damit zu rechnen, dass die absolute Summe der individuellen Beiträge nach dem Bruchpunkt mit wachsendem Stichprobenumfang divergiert und somit auf das Vorliegen eines Strukturbruchs hinweist.

Das vorgestellte Vorgehen entspricht zwar einer direkten Erweiterung des Verfahrens von Berkes et al. (2004), allerdings ist eine Methodenübertragung von univariaten auf multivariate GARCH-Modelle häufig deutlich komplizierter als sie auf den ersten Blick wirken mag. Einerseits werden die multivariaten Modelle durch eine deutlich höhere Anzahl an Parametern charakterisiert, andererseits weisen Modelle, die sowohl eine dynamische Modellierung der Varianz- als auch der Korrelationsstruktur erlauben, in der Regel eine wesentlich komplexere Struktur auf als andere mögliche multivariate Erweiterungen. Der

Umgang mit den einzelnen Bestandteilen der Quasi-Loglikelihoodfunktion und den Beiträgen zum Gradienten dieser Funktion wird dabei durch die multiplikative Struktur der bedingten Kovarianzmatrix noch weiter erschwert wie die Ausführungen in den Kapiteln 4 und 5 belegen. Des Weiteren sind DCC-Modelle und ihre Eigenschaften weit weniger gut untersucht als univariate GARCH-Modelle und insbesondere das Grundmodell von Bollerslev (1986), das in Berkes et al. (2004) betrachtet wird. Für Modelle mit dynamischer bedingter Korrelation hingegen wurden grundlegende Eigenschaften erst kürzlich oder noch gar nicht nachgewiesen. So liefern Fermanian und Malongo (2017) erstmals Bedingungen für die Existenz und Eindeutigkeit einer stationären Lösung, während Bedingungen für die Existenz unbedingter Momente höherer Ordnung noch nicht allgemein nachgewiesen wurden. Diese Punkte gestalten den Umgang mit dieser Modellart in Anwendungen als besonders anspruchsvoll.

Unabhängig davon, welches der beiden Verfahren zur Strukturbruchererkennung angewendet wird, stellt sich nach der bloßen Entdeckung eines Strukturbruchs die Frage nach dessen zeitlicher Lage. Beide verwendeten Detektoren weisen eine additive Struktur auf, die dafür sorgt, dass nach dem Auftreten eines Strukturbruchs einige Zeit vergeht, bevor der Einfluss der Beobachtungen nach dem Bruchpunkt auf den Detektor so stark ist, dass dieser die verwendete Grenzfunktion überschreitet und dadurch eine Änderung der Varianzen oder Parameter anzeigt. Es muss somit davon ausgegangen werden, dass der Zeitpunkt der ersten Überschreitung der Grenzfunktion in der Regel nicht mit dem eigentlichen Zeitpunkt des Strukturbruchs übereinstimmt. Deshalb werden für beide Überwachungsmethoden Verfahren vorgestellt, mit denen die Lage des tatsächlichen Bruchpunkts geschätzt werden kann.

Die weitere Arbeit ist wie folgt strukturiert: Kapitel 2 bietet eine Einführung in die wesentlichen Grundlagen und die verwendete Notation, die die Basis für die gesamte Dissertation bilden. Kapitel 3 präsentiert das nichtparametrische Verfahren zur Überwachung des Vektors der Varianzen. Außerdem werden die asymptotischen Eigenschaften des dort verwendeten Detektors erläutert. Zur Beurteilung des Verhaltens der Methode in endlichen Stichproben runden eine ausführliche Simulationsstudie sowie eine Anwendung auf die log-Renditen der Aktien einiger DAX-Unternehmen dieses Kapitel ab. Anschließend wird in Kapitel 4 das von Engle (2002) vorgeschlagene DCC-Modell mit grundlegenden Annahmen an die Modellparameter vorgestellt. Aufbauend darauf präsentiert Kapitel 5 das parametrische Überwachungsverfahren zur Ermittlung von Strukturbrüchen im Parametervektor des Modells. Darüber hinaus wird eine parametrische Bootstrapmethode für die Anwendung des Überwachungsverfahrens in endlichen Stichproben vorgestellt und deren Wirksamkeit mit Hilfe einer Simulationsstudie und einer Anwendung auf die log-Renditen der Aktien verschiedener europäischer Versicherungsunternehmen untersucht. Den Abschluss des Kernteils dieser Arbeit bildet Kapitel 6 mit einem Fazit der vorgestellten Methoden sowie einem Ausblick auf weiterführende Forschungsansätze.

Weitere Details, einige Ergebnisse und die Beweise zu den Resultaten aus Kapitel 3 und 5 finden sich im Anhang, der wie folgt geordnet ist: Anhang A enthält die Beweise der Aussagen aus Abschnitt 3.2, während einige zusätzliche Simulationsergebnisse aus Abschnitt 3.3 im Sinne einer besseren Lesbarkeit und Übersichtlichkeit in Anhang B verschoben wurden. Anhang C bietet Details zur Darstellung der DCC-Modellgleichungen als Markovkette. Diese Schreibweise wird für den Nachweis einer eindeutigen stationären Lösung benötigt. Außerdem liefert Anhang D eine detaillierte Herleitung der individuellen Quasi-Loglikelihoodscores im DCC-Modell, die die Grundlage für den in Kapitel 5 entwickelten Detektor bilden. Die asymptotischen Eigenschaften dieses Detektors werden in Anhang E nachgewiesen. Abschließend bietet Anhang F eine Sammlung verschiedener Sätze und Aussagen, die in der Arbeit Anwendung finden.

## Kapitel 2

# Grundlagen und Notation

Dieses Kapitel dient als kurze Einführung in die beiden Kernbereiche der Arbeit und beschäftigt sich mit der Entwicklung von sequenziellen Testverfahren, die zur Überwachung von Zeitreihen genutzt werden können. Der Anwendungsbereich dieser Methoden wiederum liegt beispielsweise in der Analyse von Zeitreihen, die unter Berücksichtigung einer dynamischen Varianzstruktur modelliert werden können, wie etwa durch multivariate GARCH-Modelle. Dazu sei angemerkt, dass in der überwiegenden Mehrzahl der Literaturquellen, die sich mit GARCH-Modellen befassen, keine Notationsunterscheidung zwischen den betrachteten Zufallsvariablen oder -vektoren und ihren Realisationen unternommen wird. Beide werden in der Regel, und so auch in dieser Arbeit, mit Kleinbuchstaben gekennzeichnet. Im Gegensatz dazu werden für allgemeine Zeitreihen ohne unterstelltes Modell die zugrundeliegenden Zufallsvariablen und -vektoren mit großen und die realisierten Werte mit kleinen Buchstaben gekennzeichnet.

### 2.1 Sequenzielle Überwachung von Zeitreihen

Da viele Verfahren aus dem Bereich der Zeitreihenanalyse auf mehr oder weniger starken Stationaritätsannahmen beruhen, ist die Detektion von Strukturbrüchen ein Bereich, der schon lange die Aufmerksamkeit der Forschung auf sich gezogen hat. Beispielhaft sei hier die Arbeit von Page (1954) genannt, in der dieser das CUSUM-Verfahren zur Qualitätskontrolle einführt. Diese Methode bildet seitdem die Grundlage für die Entwicklung zahlreicher Überwachungsverfahren, die etwa in den Bereichen der Qualitätskontrolle, der Biometrie und der Ökonometrie Anwendung finden.

Über eine lange Zeit wurden hauptsächlich retrospektive Verfahren verwendet, in denen ein komplett erhobener Datensatz betrachtet und auf strukturelle Stabilität untersucht wurde. In der Praxis, beispielsweise in klinischen Studien oder bei der Qualitätskontrolle, sind mit der Durchführung jedes einzelnen Experiments häufig Kosten und ein hoher Aufwand verbunden. Deshalb ist es naheliegend, Verfahren zu entwickeln, die eine zweifelhafte Nullhypothese nicht erst dann ablehnen, wenn die gesamte Stichprobe untersucht worden ist, sondern bereits sobald aus den bisherigen Versuchen genug Information gesammelt wurde, die die Nullhypothese hinreichend zweifelhaft erscheinen lässt. Den Grundstein für diese sequenziellen Testverfahren liefert Wald (1948) mit dem *Sequential Probability Ratio*-Test. Obwohl sich Ökonomen und Ökonometriker häufig mit stetig wachsenden Zeitreihen wie den Bewertungen von Aktien und anderen Finanzinstrumenten konfrontiert sehen, wurden Verfahren, die Änderungen in der Struktur aufdecken sollen, wie beispielsweise die klassischen CUSUM-Tests, sehr lange nur im retrospektiven Kontext angewandt. Ein möglicher Ansatz, auf eine stetig anwachsende Datenbasis zu reagieren, geht auf Chu et al. (1996) zurück. Diese argumentieren aufbauend auf Robbins (1970), dass die wiederholte Anwendung der bekannten retrospektiven Methoden auf jeweils stückweise erweiterte und somit nur

marginal geänderte Datensätze nicht praktikabel ist, da dieses Vorgehen dazu führt, dass auch eine gültige Nullhypothese mit Wahrscheinlichkeit Eins abgelehnt wird, wenn ein Prozess nur lange genug überwacht wird. Als Alternative zu diesem Vorgehen schlagen Chu et al. (1996) ein sequenzielles CUSUM-Verfahren mit Niveauekontrolle zur Überwachung des Parametervektors im linearen Regressionsmodell vor. Konkret nehmen sie dabei an, dass sich der Datensatz in einen historischen Teil und einen Überwachungsbereich aufteilen lässt. Die historischen Datenpunkte liegen schon zum Start der Überwachung vor und werden als frei von Strukturbrüchen angenommen. Somit kann die in ihnen enthaltene zusätzliche Information in vielen Fällen zur Stabilisierung des Detektors beziehungsweise der verwendeten Schätz- oder Referenzwerte genutzt werden. In ihrer Arbeit verwenden Chu et al. (1996) Partialsummen von rekursiven Residuen, die sich bei der Anpassung eines linearen Modells ergeben. Dabei können die Daten des historischen Zeitraums zur Verbesserung der Parameterschätzung genutzt werden. Dieses Vorgehen verhindert ungewöhnlich starke Fluktuationen der rekursiven Residuen zu Beginn des Überwachungszeitraums, insofern diese auf eine mangelhafte Schätzung der Modellparameter aus zu wenigen Datenpunkten zurückzuführen ist. Dies gilt insbesondere, wenn der historische Zeitraum lang genug ist. Im Überwachungszeitraum wird der vorliegende historische Datensatz dann schrittweise um einzelne Datenpunkte erweitert. Diese neue Information kann genutzt werden, um einen anstatt einer klassischen Teststatistik verwendeten Detektor  $S_k$ ,  $k \in \mathbb{N}$ , oder im Fall von Chu et al. (1996) die Parameterschätzer zu aktualisieren. Die Detektorwerte werden mit den Werten einer Grenzfunktion  $g(\cdot)$  verglichen, die durch Multiplikation einer Konstanten so skaliert wird, dass die Wahrscheinlichkeit, einen Fehler erster Art zu begehen, gerade dem Signifikanzniveau  $\tilde{\alpha}$  entspricht. Dabei liegt ein Fehler erster Art vor, wenn der Detektor die Grenzfunktion überschreitet obwohl kein Strukturbruch vorliegt. Tritt nun im Überwachungszeitraum ein Strukturbruch auf, so ist damit zu rechnen, dass die Beobachtungen nach dem Bruchpunkt den Detektor mit der Zeit so stark beeinflussen, dass dieser die Grenzfunktion überschreitet. Aus diesem Abbruchkriterium ergibt sich die Stoppzeit für eine historische Stichprobe der Länge  $m$  und einen Überwachungszeitraum der Länge  $T$  als

$$\tau_m = \min \{k \in \{1, \dots, T\} : S_k > g(k)\}.$$

Entsprechend der Notation in Aue et al. (2012) bezeichne  $\tau_m < \infty$  den Fall, dass der Detektor die Grenzfunktion im Laufe eines Überwachungszeitraums der Länge  $T$  überschreitet. Dies impliziert eine Ablehnung der Nullhypothese. Im Fall von Chu et al. (1996) geschieht dies, sobald sich die aus den Daten des Überwachungszeitraums bestimmten Parameterschätzer zu sehr von den Referenzschätzern aus der historischen Stichprobe unterscheiden. Weiter bezeichne  $\tau_m = \infty$  den Fall, dass während der Überwachungsphase die Grenzfunktion kein einziges Mal durch den Detektor überschritten wird, siehe Aue et al. (2012). Der Ansatz von Chu et al. (1996) motivierte viele weitere Arbeiten wie beispielsweise Berkes et al. (2004), Horváth et al. (2004), Horváth et al. (2006) oder Wied und Galeano (2013), sowie die eigenen Arbeiten Pape et al. (2016) und Pape et al. (2017), die die Grundlage für diese Dissertation bilden.

## 2.2 Univariate und multivariate GARCH-Modelle

Modelle, die autoregressive Strukturen in der bedingten Varianz zulassen, sind inzwischen weitläufig bekannt und gut erforscht. Die Modelle mit (generalisierter) autoregressiver bedingter Heteroskedastizität (*generalized autoregressive conditional heteroscedasticity*) oder kurz ARCH- beziehungsweise GARCH-Modelle, siehe Engle (1982) beziehungsweise Bollerslev (1986), erlauben eine geeignete Modellierung der Volatilitätscluster wie sie typischerweise in Zeitreihen von Aktienrenditen und anderen Finanzwerten auftreten. Betrachte den Prozess  $\{y_t, t \in \mathbb{Z}\}$ . Im univariaten Fall wird zunächst angenommen, dass sich die Zufallsvariable  $y_t$ , deren Realisation von Interesse ist, als Produkt der zeitvariablen, auf die Prozessvergangenheit bedingten Standardabweichung und einer Innovation  $\epsilon_t$  darstellen lässt:

$$y_t = \sigma_t \epsilon_t \quad \text{mit} \quad \sigma_t^2 = \text{Var}(y_t | \mathcal{F}_{t-1}). \quad (2.1)$$

Der Innovationsprozess wird meist als starkes Standard-Weißes Rauschen angenommen, das heißt es gilt:

$$\mathbb{E}(\epsilon_t) = 0 \quad \forall t \in \mathbb{Z} \quad \text{und} \quad \text{Var}(\epsilon_t) = 1 \quad \forall t \in \mathbb{Z},$$

während die Innovationen untereinander paarweise unabhängig sind, siehe Douc et al. (2014) mit  $\sigma^2 = 1$ . Außerdem wird analog zu Klenke (2006) die an den Prozess  $\{y_t, t \in \mathbb{Z}\}$  adaptierte Filtration mit  $\mathbb{F}$  und die zugehörige Informationsmenge zum Zeitpunkt  $t$  mit  $\mathcal{F}_t$  bezeichnet, das heißt es gilt

$$\mathbb{F} := \{\mathcal{F}_t, t \in \mathbb{Z}\} \quad \text{mit} \quad \mathcal{F}_t = \sigma(y_t, y_{t-1}, y_{t-2}, \dots) \quad \text{und} \quad \mathcal{F}_t \subset \mathcal{F}_{t+1} \quad \forall t \in \mathbb{Z}.$$

Die bedingte Varianz  $\sigma_t^2$  kann nun auf vielfältige Art und Weise modelliert werden. Engle (1982) schlägt vor,  $\sigma_t^2$  in linearer Abhängigkeit von Quadraten vergangener Realisationen  $y_{t-1}, y_{t-2}, \dots$  zu definieren. Diese Struktur erlaubt eine praxisnahe Modellierung von Zeitreihen, die durch wechselnde Phasen von höherer und niedrigerer Volatilität geprägt sind. Ein solches Verhalten kann häufig bei Zeitreihen von Aktienrenditen beobachtet werden. Eine Umstellung der Terme in der Modellgleichung, wie sie sich beispielweise in (2.1) in Francq und Zakoïan (2010) findet, erlaubt dann eine Darstellung der quadrierten Variablen  $y_t$  als AR-Prozess, siehe (2.4) in Francq und Zakoïan (2010), falls für den ARCH-Fall  $\beta(z) = 0$  gesetzt wird.

Inspiriert durch dieses Vorgehen schlägt Bollerslev (1986) vor, bei der Modellierung der bedingten Varianz nicht nur vergangene quadrierte Beobachtungen, sondern auch vergangene Ausprägungen der bedingten Varianz zu berücksichtigen. Die Modellordnung  $(p, q)$  gibt dabei an, dass  $\sigma_t^2$  von  $\sigma_{t-1}^2, \dots, \sigma_{t-p}^2$  und  $y_{t-1}, \dots, y_{t-q}$  abhängt. Mit Hinblick auf (2.4) in Francq und Zakoïan (2010) erlaubt dies eine Darstellung der quadrierten Beobachtungswerte als ARMA-Prozess. Analog zum Übergang von AR- auf ARMA-Modelle, erlauben GARCH-Modelle einen sparsameren Umgang mit den Modellparametern. So zeigt Bollerslev (1986), dass bei gewöhnlichen GARCH-Modellen schon eine Modellordnung von  $(1,1)$  ausreicht, um einige der stilisierten Fakten von Aktienrenditen hinreichend gut abbilden zu können.

Weiterführend wurden zahlreiche Modellierungsarten für die bedingte Varianz vorgeschlagen. Die bekanntesten sind das integrierte GARCH (IGARCH)-Modell von Engle und Bollerslev (1986), das eine Einheitswurzel im autoregressiven Teil der ARMA-Darstellung (2.4) in Francq und Zakoïan (2010) zulässt, das exponentielle GARCH (EGARCH)-Modell von Nelson (1991), das die fast sichere Nichtnegativität der bedingten Varianz für alle  $t \in \mathbb{Z}$  im

Gegensatz zum gewöhnlichen GARCH-Modell auch für negative Modellparameter erlaubt, oder das *Threshold* GARCH (TGARCH)-Modell von Zakoïan (1994), das die bedingte Varianz nicht nur in Abhängigkeit vom Betrag, sondern auch vom Vorzeichen vergangener Beobachtungen modelliert, was eine Berücksichtigung verschiedener Asymmetrien, wie sie in der Praxis häufig zu beobachten sind, erlaubt. Eine Übersicht verschiedener ARCH- und GARCH-Modelle findet sich beispielsweise in Bollerslev (2009). Zudem liefern Francq und Zakoïan (2010) Details zu vielen der dort vorgestellten Modelle.

Die naheliegende Erweiterung der univariaten GARCH-Modelle auf den multivariaten Fall entspricht der Darstellung

$$y_t = H_t^{\frac{1}{2}} \epsilon_t \quad \text{mit} \quad H_t = \text{Cov}(y_t | \mathcal{F}_{t-1}), \quad (2.2)$$

wobei  $y_t$  im Gegensatz zu (2.1) ein  $p$ -dimensionaler Zufallsvektor ist und die Informationsmenge entsprechend  $\mathcal{F}_t = \sigma(y_t, y_{t-1}, y_{t-2}, \dots)$  mit zugehöriger Filtration  $\mathbb{F}$  ist. Der Prozess der  $p$ -dimensionalen Innovationsvektoren wird in der Regel als starkes multivariates Standard-Weißes Rauschen angenommen, d.h. es gilt

$$E(\epsilon_t) = \mathbf{0}_p \quad \text{und} \quad \text{Cov}(\epsilon_t) = \mathbb{I}_p,$$

während die Vektoren  $\epsilon_s$  und  $\epsilon_t$  für  $s \neq t$  jeweils paarweise unabhängig sind, siehe Florens et al. (2007), indem die dort verwendete Kovarianzmatrix der Innovationen hier als  $\Sigma = \mathbb{I}_p$  gewählt wird.

Wie im univariaten Fall wurden zahlreiche verschiedene Arten der Modellierung der bedingten Kovarianzmatrix vorgeschlagen. Die intuitivste Art der Spezifikation ist diejenige von Bollerslev et al. (1988) in Form von Vektor-GARCH-Modellen, bei denen sich die Parameter direkt interpretieren lassen. Allerdings beinhaltet sie eine hohe Anzahl an Parametern, weshalb in der Praxis meist stark vereinfachte Varianten dieses Modells betrachtet werden, bei denen beispielsweise ganze Parametergruppen auf Null oder gleich gesetzt werden. Einen Spezialfall des Vektor-GARCH-Modells stellt das von Engle und Kroner (1995) vorgeschlagene BEKK-Modell dar, das unter relativ milden Annahmen an die Modellparameter die nichtnegative Definitheit der bedingten Kovarianzmatrix für alle  $t \in \mathbb{Z}$  garantiert. Ein Sonderfall dieser BEKK-Modelle wiederum ist das skalare BEKK-Modell, das in Kapitel 3 zum Einsatz kommt und dort auch genauer beschrieben ist. Weitere beliebte Modellierungstypen sind das *Factor* GARCH (FGARCH)-Modell, das in hochdimensionalen Modellen eine Dimensionsreduktion ermöglicht, sowie Modelle, die die Varianzen der einzelnen Vektorkomponenten durch univariate GARCH-Prozesse modellieren und sie anschließend über eine zeitlich konstante (CCC (*constant conditional correlation*)-Modell) oder eine zeitvariable (DCC (*dynamic conditional correlation*)-Modell) Korrelationsmatrix verbinden. Die letztere Art von Modellen wird in den Kapiteln 4 und 5 eingehend behandelt. Ausführliche Übersichten zu multivariaten GARCH-Modellen finden sich beispielsweise in Bauwens et al. (2006) oder in Bollerslev (2009). Außerdem finden sich für viele multivariate GARCH-Modelle weitere Details und Eigenschaften in Francq und Zakoïan (2010).

## 2.3 Notation für die weiteren Kapitel

Die in den weiteren Kapiteln verwendeten Operatoren und Transformationsmatrizen und ihre Eigenschaften werden in Lütkepohl (1996) und Hafner und Herwatz (2008) präsentiert.

### 2.3.1 Operatoren

- $\text{vec}(\cdot)$ , **vec-Operator**: Fasst alle Einträge einer Matrix spaltenweise in einem Vektor zusammen.
- $\text{vech}(\cdot)$ , **vech-Operator**: Fasst alle Einträge auf und unter der Hauptdiagonalen einer quadratischen Matrix spaltenweise in einem Vektor zusammen.
- $\text{vecl}(\cdot)$ , **vecl-Operator**: Fasst alle Einträge unter der Hauptdiagonalen einer quadratischen Matrix spaltenweise in einem Vektor zusammen.
- Für eine beliebige symmetrische  $(p \times p)$ -Matrix  $A$  weisen die Vektoren  $\text{vec}(A)$ ,  $\text{vech}(A)$  beziehungsweise  $\text{vecl}(A)$  Dimensionen von  $p^2$ ,  $p^+ := \frac{1}{2}p(p+1)$  beziehungsweise  $p^- := \frac{1}{2}p(p-1)$  auf.

### 2.3.2 Transformationsmatrizen

- $K_{mn}$ , **Kommutationsmatrix**: Für eine beliebige  $(m \times n)$ -Matrix  $A$  gilt:  $\text{vec}(A') = K_{mn} \cdot \text{vec}(A)$ . Außerdem gilt:  $K_{mn} \sim (mn \times mn)$ .
- $L_p$ , **Eliminationsmatrix**: Für eine beliebige symmetrische  $(p \times p)$ -Matrix  $A$  gilt:  $L_p \cdot \text{vec}(A) = \text{vech}(A)$ . Außerdem gilt:  $L_p \sim (p^+ \times p^2)$ .
- $D_p$ , **Duplikationsmatrix**: Für eine beliebige symmetrische  $(p \times p)$ -Matrix  $A$  gilt:  $D_p \cdot \text{vech}(A) = \text{vec}(A)$ . Außerdem gilt:  $D_p \sim (p^2 \times p^+)$ .
- $D_p^+$ , **Moore-Penrose-Inverse von  $D_p$** : Im Allgemeinen entspricht diese Inverse nicht der Eliminationsmatrix, die eine generalisierte Inverse ist. Außerdem gilt:  $D_p^+ \sim (p^+ \times p^2)$ .
- $D_{p,-}$ : Die Matrix, die sich ergibt, wenn diejenigen Spalten aus  $D_p$  gelöscht werden, deren Indizes den Positionen der Hauptdiagonaleinträge von  $A$  in  $\text{vech}(A)$  entsprechen. Außerdem gilt:  $D_{p,-} \sim (p^2 \times p^-)$ .
- $D_{p,-}^+$ : Die Moore-Penrose-Inverse von  $D_{p,-}$ . Diese Matrix ergibt sich, wenn diejenigen Zeilen aus  $D_p^+$  entfernt werden, deren Indizes den Positionen der Hauptdiagonaleinträge von  $A$  in  $\text{vech}(A)$  entsprechen. Für eine symmetrische  $(p \times p)$ -Matrix  $A$  gilt außerdem:  $\text{vecl}(A) = D_{p,-}^+ \cdot \text{vec}(A)$ . Außerdem gilt:  $D_{p,-}^+ \sim (p^- \times p^2)$ .



## Kapitel 3

# Nichtparametrische Detektion von Varianzänderungen

### 3.1 Das Überwachungsszenario

Sei  $\{X_t, t \in \mathbb{Z}\}$  eine Folge von  $p$ -dimensionalen Zufallsvektoren, deren einzelne Einträge endliche vierte Momente und Kreuzmomente besitzen. O.B.d.A. kann angenommen werden, dass  $E(X_t) = 0$  für alle  $t \in \mathbb{Z}$  gilt. In finanzwirtschaftlichen Anwendungen, wo häufig die Renditen von Finanzinstrumenten oder ähnliche Zeitreihen untersucht werden, ist dies eine übliche Annahme. Sollte  $E(X_t) = \mu$  für alle  $t \in \mathbb{Z}$  und eine Konstante  $\mu \in \mathbb{R}$  gelten, so würde für die folgenden Ausführungen die Folge  $\{X_t - \mu, t \in \mathbb{Z}\}$  betrachtet werden. Konkret ist hier der potenziell zeitvariable Vektor der Varianzen der einzelnen Komponenten des Zufallsvektors  $X_t = (X_{1t}, \dots, X_{pt})'$  von Interesse. Dieser ist gegeben durch

$$\sigma_t^2 = (\sigma_{1t}^2, \dots, \sigma_{pt}^2)' \quad \text{mit} \quad \sigma_{jt}^2 = \text{Var}(X_{jt}) = E(X_{jt}^2), \quad \text{für } j = 1, \dots, p.$$

Da in der Praxis häufig nicht davon ausgegangen werden kann, dass  $\sigma_t^2$  über die Zeit konstant bleibt, wird ein Verfahren gesucht, das den Vektor der Varianzen überwacht und potenzielle Änderungen so schnell wie möglich nach ihrem Auftreten meldet. Dazu werden die Varianzen aus wachsenden Teilstichproben geschätzt und mit Schätzwerten verglichen, die zuvor aus einer Referenzstichprobe bestimmt wurden. Eine entscheidende Annahme ist dabei, dass im historischen Zeitraum der Länge  $m$  kein Strukturbruch in den Varianzen vorliegt:

**Annahme 3.1**  $\sigma_1^2 = \dots = \sigma_m^2$  mit  $m \in \mathbb{N}$ .

Die Gültigkeit dieser Annahme ist zentral für die Funktionsfähigkeit von Methoden wie sie in Abschnitt 2.1 vorgestellt wurden. Die Annahme kann überprüft werden, indem retrospektive Verfahren zum Aufspüren von Strukturbrüchen auf den historischen Datensatz angewandt werden. Ein Beispiel für eine solche Methode, die im Gegensatz zu sequenziellen Verfahren nicht davon ausgeht, dass der vorliegende Datensatz stückweise erweitert wird, sondern von fester Länge ist, wäre ein Pendant zum von Wied et al. (2012b) vorgeschlagenen Verfahren, das sich nur auf den Vektor der Varianzen konzentriert. In der Praxis ist es dank der hohen Frequenz, in der gerade im Finanzbereich Datenpunkte erhoben werden können, in der Regel nicht schwierig, eine ausreichend große historische Startstichprobe mit konstanter Varianzstruktur zu finden. Strukturbrüche können dann potenziell im Überwachungsbereich auftreten, der nach dem Ende des historischen Zeitraums einsetzt. Wie in Abschnitt 2.1 erläutert, wird der Datensatz Stück für Stück um einzelne  $p$ -dimensionale

Datenvektoren ergänzt. Die neugewonnene Information kann dann genutzt werden, um den verwendeten Detektor zu aktualisieren.

Die Nullhypothese der Konstanz der Varianzvektoren auch über den historischen Zeitraum hinaus lässt sich formulieren als

$$H_0 : \sigma_1^2 = \dots = \sigma_m^2 = \sigma_{m+1}^2 = \dots$$

Die Menge der Alternativen umfasst alle Szenarien, in denen sich  $\sigma_t^2$  zu mindestens einem Zeitpunkt im Überwachungszeitraum ändert. Um Alternativen konkret formulieren zu können, sei analog zum Vorgehen in Wied et al. (2012b)  $\mathcal{F}_B$  die Menge der Funktionen  $f : [0, \infty) \rightarrow \mathbb{R}$ , die beschränkt und auf dem Intervall  $[0, B + 1]$  integrierbar sind mit

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{t=1}^{\lfloor m(B+1) \rfloor} f\left(\frac{t}{m}\right) = \int_0^{B+1} f(z) dz.$$

Die Variable  $B$  wird dabei genutzt, um die Länge des Überwachungszeitraums als Vielfaches der Länge des historischen Zeitraums  $m$  anzugeben. Eine mögliche Alternative  $H_1$  ergibt sich dann, indem die Varianzen der einzelnen Vektorkomponenten dargestellt werden als

$$\text{Var}(X_{jt}) = \sigma_{jt}^2 = \bar{\sigma}_j^2 + g_j\left(\frac{t}{m}\right), \quad j = 1, \dots, p, \quad t \in \mathbb{Z}, \quad (3.1)$$

wobei  $\bar{\sigma}_j^2$ ,  $j = 1, \dots, p$ , zeitinvariante Konstanten und  $g_j(\cdot) \in \mathcal{F}_B$ ,  $j = 1, \dots, p$ , Funktionen sind, die die strukturelle Stabilität der einzelnen Varianzfunktionen bestimmen. Mit Annahme 3.1 gilt offensichtlich:

$$g_j(z) = 0 \quad \forall z \in [0, 1] \quad \text{und} \quad \forall j = 1, \dots, p.$$

Unter der Alternativen gilt allerdings auch für mindestens ein  $j \in \{1, \dots, p\}$ :

$$\int_1^{B+1} |g_j(z)| dz > 0,$$

was bedeutet, dass die Varianz der  $j$ -ten Vektorkomponente von einem Strukturbruch innerhalb des Überwachungsbereichs betroffen ist.

## 3.2 Das Überwachungsverfahren

Um asymptotische Aussagen bezüglich des Niveaus und der Güte der Prozedur herleiten zu können, müssen zunächst einige Annahmen aufgestellt werden. Diese sind die Gegenstücke zu den Annahmen (A1)-(A3) in Wied et al. (2012a) bzw. zu (A2)-(A4) in Wied und Galeano (2013). Wie in den genannten Arbeiten seien zunächst die folgenden Hilfsternme definiert:

### Definition 3.1

$$U_t := \begin{pmatrix} X_{1t}^2 - \mathbb{E}(X_{1t}^2) \\ \vdots \\ X_{pt}^2 - \mathbb{E}(X_{pt}^2) \end{pmatrix} \quad \text{und} \quad S_j := \sum_{t=1}^j U_t, \quad \text{für } j \in \mathbb{N} \quad \text{und } t \in \mathbb{Z}.$$

**Annahme 3.2**

$$\lim_{m \rightarrow \infty} \mathbb{E} \left( \frac{1}{m} S_m S_m' \right) =: \mathbf{D}_p,$$

wobei  $\mathbf{D}_p$  eine endliche und positiv definite Matrix ist.

**Annahme 3.3** Die  $r$ -ten absoluten Momente der einzelnen Komponenten von  $U_t$  sind gleichmäßig beschränkt für ein  $r > 2$ .

Im Folgenden werden die Detektorwerte aus den Realisationen von Zufallsvektoren bestimmt. In der Praxis können diese, gerade im Bereich von finanziellen Anwendungen, kaum als unabhängig angenommen werden. Deshalb wird ein Konzept benötigt, das ein moderates Maß an zeitlicher Abhängigkeit zwischen aufeinanderfolgenden Zufallsvariablen zulässt. Eine mögliche Lösung stellen Prozesse dar, die mischend oder *near-epoch dependent* sind. Beide Varianten lassen zeitliche Abhängigkeit in stochastischen Prozessen zu, insofern diese mit zunehmendem zeitlichen Abstand ausreichend schnell verschwindet.

**Definition 3.2 Stark mischender Prozess** (Davidson (1994), S.206ff)

Sei  $\{X_t, t \in \mathbb{Z}\}$  ein potenziell vektorwertiger Prozess auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P)$ . Sei der starke Mischungskoeffizient für zwei Unteralgebren  $\mathcal{G}$  und  $\mathcal{H}$  von  $\mathcal{F}$  gegeben durch:

$$\alpha(\mathcal{G}, \mathcal{H}) = \sup_{G \in \mathcal{G}, H \in \mathcal{H}} |\mathbb{P}(G \cap H) - \mathbb{P}(G)\mathbb{P}(H)|.$$

Seien außerdem die folgenden Unteralgebren von  $\mathcal{F}$  definiert:

$$\mathcal{F}_{-\infty}^t := \sigma(X_t, X_{t-1}, X_{t-2}, \dots) \quad \text{und} \quad \mathcal{F}_t^\infty := \sigma(X_t, X_{t+1}, X_{t+2}, \dots).$$

Betrachte nun die Folge  $\{\alpha_n, n \in \mathbb{N}\}$  mit

$$\alpha_n := \sup_{t \in \mathbb{Z}} \alpha(\mathcal{F}_{-\infty}^t, \mathcal{F}_{t+n}^\infty).$$

Der Prozess  $\{X_t, t \in \mathbb{Z}\}$  wird als stark mischend (oder  $\alpha$ -mischend) bezeichnet, wenn gilt:

$$\lim_{n \rightarrow \infty} \alpha_n = 0.$$

**Definition 3.3  $\mathcal{L}_p$ -near-epoch dependence** (Davidson (1994), S.261)

Ein potenziell vektorwertiger Prozess  $\{X_t, t \in \mathbb{Z}\}$  wird als in der  $\mathcal{L}_p$ -Norm near-epoch dependent oder kurz  $\mathcal{L}_p$ -near-epoch dependent bezeichnet, falls für  $p > 0$  ein stochastischer Prozess  $\{Y_t, t \in \mathbb{Z}\}$  existiert sowie eine Folge positiver Konstanten  $\{c_t, t \in \mathbb{Z}\}$  und eine nichtstochastische Nullfolge  $\{v_l, l \in \mathbb{N}_0\}$ , für die gilt:

$$\|X_t - \mathbb{E}(X_t | \sigma(Y_{t-l}, \dots, Y_{t+l}))\|_p \leq c_t v_l \quad \forall t \in \mathbb{Z}, l \in \mathbb{N}_0.$$

Dabei ist  $\|X\|_p = (\mathbb{E}|X|^p)^{1/p}$ , die  $\mathcal{L}_p$ -Norm für  $p > 0$ .

Während beim Konzept mischender Prozesse die Realisationswahrscheinlichkeit der Zufallsvariablen  $X_t$  auf die Informationsmenge zum Zeitpunkt  $t$  bedingt wird, wird im Fall der *near-epoch dependence* neben  $\{X_t, t \in \mathbb{Z}\}$ , dem eigentlichen Prozess von Interesse, ein zweiter stochastischer und potenziell multivariater Prozess  $\{Y_t, t \in \mathbb{Z}\}$  betrachtet. Die Erwartung von  $X_t$  wird dann auf diejenige  $\sigma$ -Algebra bedingt, die durch die Realisationen derjenigen Zufallsvariablen  $Y_t$  erzeugt wird, die einen gegebenen maximalen zeitlichen Abstand von  $X_t$  einhalten. Da der Prozess  $\{Y_t, t \in \mathbb{Z}\}$  mischend sein darf, es aber nicht sein muss und außerdem  $\{X_t, t \in \mathbb{Z}\}$  über ein Funktional von dem zugrundeliegenden mischenden Prozess abhängen kann, ist intuitiv klar, dass es sich bei der *near-epoch dependence* um ein deutlich flexibleres und allgemeineres Konzept handelt als bei einem einfachen mischenden Prozess.

Um quantifizieren zu können, wie schnell die Abhängigkeit mit wachsender zeitlicher Entfernung abklingt, wird der Prozess  $\{X_t, t \in \mathbb{Z}\}$  als  $\mathcal{L}_p$ -*near-epoch dependent* der Ordnung  $-\varphi_0$  bezeichnet, wenn für die Nullfolge  $\{v_l, l \in \mathbb{N}_0\}$  gilt:  $v_l = \mathcal{O}(l^{-\varphi})$  für alle  $\varphi > \varphi_0$ , siehe Davidson (1994), S.262. Analog ist ein Prozess stark mischend der Ordnung  $-\varphi^*$ , wenn für die Folge der Mischungskoeffizienten  $\{\alpha_l, l \in \mathbb{N}\}$  gilt:  $\alpha_l = \mathcal{O}(l^{-\varphi})$  für alle  $\varphi > \varphi^*$ , vergleiche Davidson (1994), S.210.

Um Aussagen über das asymptotische Verhalten des Detektors treffen zu können, wird im Folgenden angenommen, dass der zugrundeliegende Prozess  $\{Y_t, t \in \mathbb{Z}\}$  stark mischend ist, wobei die Ordnung der *near-epoch dependence* des Prozesses  $\{X_t, t \in \mathbb{Z}\}$  sowie die Ordnung des mischenden Prozesses von der Konstanten  $r$  aus Annahme 3.3 abhängen.

**Annahme 3.4** Der Prozess  $\{X_t, t \in \mathbb{Z}\}$  ist  $\mathcal{L}_2$ -*near-epoch dependent* der Ordnung  $-\varphi_0$  mit  $\varphi_0 = \frac{r-1}{r-2}$  auf einem stark mischenden Prozess  $\{Y_t, t \in \mathbb{Z}\}$  der Ordnung  $-\varphi^*$  mit  $\varphi^* = \frac{r}{r-2}$ .

Die Eigenschaft der  $\mathcal{L}_2$ -*near-epoch dependence* konnte bereits für verschiedene Typen von Prozessen nachgewiesen werden, so beispielsweise durch Hansen (1991) für GARCH(1,1)-Prozesse oder durch Gallant und White (1988) für ARMA(p,q)-Prozesse, bei denen die Nullstellen von  $f_\alpha(\cdot)$  außerhalb des Einheitskreises liegen. Dabei bezeichnet  $f_\alpha(\cdot)$  die  $z$ -Transformierte des Filters  $\alpha(L)$  im Lagoperator  $L$ , der für die Modellierung des autoregressiven Anteils des Prozesses verwendet wird. Dennoch ist es in der Praxis kaum möglich, zu überprüfen, ob ein Prozess *near-epoch dependent* ist, ohne ein Modell zu unterstellen. In vielen Fällen kann die Gültigkeit von Annahme 3.4 somit nur als erfüllt angenommen, nicht aber nachgewiesen werden.

Die Kernidee des Verfahrens basiert auf dem Fluktuationstest von Wied und Galeano (2013) zur Erkennung von Änderungen in der Korrelation zwischen zwei Zeitreihen. Eine der Stärken dieser Methode ist, dass es sich um ein nichtparametrisches und somit auch modellunabhängiges Verfahren handelt, was vielfältige Anwendungsmöglichkeiten eröffnet. Die in diesem Kapitel verwendeten Fluktuationen rühren von einem Vergleich von Varianzschätzungen aus verschiedenen Teilstichproben der verfügbaren Daten her. Wird die Schätzung für den Vektor der Varianzen, die sich aus den Realisationen der Zufallsvektoren  $X_k$  bis  $X_l$  für  $k < l$  ergibt, durch  $[\hat{\sigma}^2]_k^l$  bezeichnet, so gilt:

$$[\hat{\sigma}^2]_k^l = \left( [X_1^2]_k^l, \dots, [X_p^2]_k^l \right)' \quad \text{mit} \quad [X_j^2]_k^l = \frac{1}{l-k+1} \sum_{t=k}^l X_{tj}^2, \quad \text{für } j = 1, \dots, p.$$

Unter der Hypothese konstanter Varianzen sollten sich Varianzschätzungen aus wachsenden Teilstichproben des Überwachungszeitraums nicht zu stark von Schätzungen aus dem historischen Datensatz unterscheiden. Sei deshalb zunächst der Grundbaustein des Detektors gegeben als

$$V_k = \frac{k}{\sqrt{m}} \widehat{D}_{p,m}^{-1/2} \left( [\hat{\sigma}^2]_{m+1}^{m+k} - [\hat{\sigma}^2]_1^m \right) = \frac{k}{\sqrt{m}} \widehat{D}_{p,m}^{-1/2} \begin{pmatrix} [\overline{X}_1^2]_{m+1}^{m+k} - [\overline{X}_1^2]_1^m \\ \vdots \\ [\overline{X}_p^2]_{m+1}^{m+k} - [\overline{X}_p^2]_1^m \end{pmatrix}, \quad k \in \mathbb{N},$$

wobei  $\widehat{D}_{p,m}$  ein auf den ersten  $m$  Datenvektoren basierender Kernschätzer für die Matrix  $\mathbf{D}_p$  aus Annahme 3.2 ist. Zur Bestimmung der Schätzung werden analog zum Vorgehen in Wied et al. (2012b) und Wied und Galeano (2013) die Abweichungen der quadrierten Realisationen von den aus den ersten  $m$  Beobachtungen ermittelten Varianzschätzern betrachtet:

$$\tilde{V}_t = \frac{1}{\sqrt{m}} \tilde{U}_t \quad \text{mit} \quad \tilde{U}_t = \begin{pmatrix} X_{1t}^2 - [\overline{X}_1^2]_1^m \\ \vdots \\ X_{pt}^2 - [\overline{X}_p^2]_1^m \end{pmatrix}.$$

Somit ergibt sich ein Kernschätzer für die Matrix  $\mathbf{D}_p$  als

$$\widehat{D}_{p,m} = \sum_{t=1}^r \sum_{u=1}^r k\left(\frac{t-u}{\delta_r}\right) \tilde{V}_t \tilde{V}_u' \quad \text{mit} \quad k(x) = \begin{cases} 1 - |x|, & |x| \leq 1 \\ 0, & \text{sonst} \end{cases}.$$

Dabei ist  $k(\cdot)$  der Bartlettkern, siehe Hayashi (2000), und  $\delta_r$  die Bandbreite, die angibt, bis zu welchem maximalen Zeitabstand dyadische Produkte der Vektoren von mittelwertbereinigten quadrierten Realisationswerten in die Schätzung einbezogen werden. Die Wahl des Kerns wird durch die Argumentation von Wied et al. (2012a) motiviert. Allerdings wird mit  $\delta_a = \lceil a^{1/4} \rceil$ , wobei  $\lceil \cdot \rceil$  die untere Gaußklammer darstellt, eine andere Bandbreite gewählt. Diese Entscheidung beruht darauf, dass weiterführende Simulationen, die nicht Teil dieser Arbeit sind, gezeigt haben, dass diese Wahl aus einer Gruppe möglicher Bandbreiten die besten Ergebnisse bezüglich der Schätzung der Matrix  $\mathbf{D}_p$  liefert. Die Konsistenz des Schätzers für  $m \rightarrow \infty$ , die mit De Jong und Davidson (2000) gilt, ist ausschlaggebend für die asymptotische Verteilung des Detektors, die im Folgenden hergeleitet wird.

In Wied und Galeano (2013) können als Fluktuationen die absoluten Differenzen zweier Schätzer für die Korrelation verglichen werden. Im Gegensatz dazu werden in diesem Kapitel zwei Vektoren von Schätzwerten miteinander verglichen, was die Frage aufwirft, wann zwei Vektoren als nah beieinander betrachtet werden. Intuitive Lösungen für dieses Problem sind verschiedene Vektornormen oder eine quadratische Form im Vektor der gewichteten Schätzdifferenzen  $V_k$ . Letztere Variante wird beispielsweise von Aue et al. (2009b) im retrospektiven Zusammenhang verwendet. Im Gegensatz dazu wird in dieser Arbeit die euklidische Norm des Vektors  $V_k$  als Detektor genutzt. Konkret wird  $\|V_k\|_2$  mitlaufend für jeden Punkt  $k$  im Überwachungszeitraum bestimmt. Die Normierung des Vektors der Differenzen bietet dann neben der Quantifizierung des Abstands der Vektoren auch den Vorteil, dass eindimensionale Detektorwerte generiert werden, die mit den Werten einer univariaten Grenzfunktion verglichen werden können. Das Überwachungsverfahren bricht dann ab, wenn die Detektorwerte eine skalierte Grenzfunktion  $w(\cdot)$  überschreiten. Sobald

dies geschieht, kann die Nullhypothese konstanter Varianzen nicht länger beibehalten werden und muss abgelehnt werden. Dementsprechend ist das Abbruchkriterium definiert als

$$\tau_m = \min \left\{ k \leq [mB] : \|V_k\|_2 > c \cdot w \left( \frac{k}{m} \right) \right\}, \quad (3.2)$$

wobei  $w(\cdot)$  eine stetige positive Funktion ist. Außerdem ist  $c$  eine Konstante, die so gewählt wird, dass unter der Hypothese konstanter Varianzen

$$\lim_{m \rightarrow \infty} \mathbb{P}(\tau_m < \infty) = \tilde{\alpha} \in (0, 1)$$

das Signifikanzniveau des Tests ist. Um eine Verwechslung mit Modellparametern auszuschließen, wird das Signifikanzniveau hier und im Rest der Arbeit mit  $\tilde{\alpha}$  bezeichnet.

### Asymptotisches Verhalten des Detektors unter der Nullhypothese

Die bisher eingeführte Notation kann nun zur Beschreibung des asymptotischen Verhaltens des Detektors in Abhängigkeit von der Stoppzeit  $\tau_m$  in (3.2) genutzt werden.

**Definition 3.4** Sei  $\{\mathcal{G}_p(b) = (G_1(b), \dots, G_p(b))', b \in [0, B]\}$  ein  $p$ -dimensionaler stochastischer Prozess, dessen einzelne Komponentenprozesse  $p$  unabhängige Gaußprozesse  $\{G_j(b), b \in [0, B]\}$  sind mit Erwartungswertfunktion  $\mathbb{E}[G_j(k)] = 0$  und Kovarianzfunktion  $\mathbb{E}[G_j(k)G_j(l)] = \min(k, l) + kl$ , für  $j = 1, \dots, p$ .

**Theorem 3.1** Unter  $H_0$  und den Annahmen 3.1-3.4 gilt für alle  $B > 0$  und eine stetige positive Funktion  $w(\cdot)$ :

$$\lim_{m \rightarrow \infty} \mathbb{P}(\tau_m < \infty) = \lim_{m \rightarrow \infty} \mathbb{P} \left( \sup_{b \in [0, B]} \frac{\|V_{[m \cdot b] + 2}\|_2}{w(b)} > c \right) = \mathbb{P} \left( \sup_{b \in [0, B]} \frac{\|\mathcal{G}_p(b)\|_2}{w(b)} > c \right). \quad (3.3)$$

Wie detailliert in Aue et al. (2012) und Wied und Galeano (2013) dargelegt wird, kann die Grenzwahrscheinlichkeit in (3.3) anhand des Verhaltens von  $p$  unabhängigen Standard-Brownschen Bewegungen  $\{W_j(b) : b \in [0, 1]\}$ ,  $j = 1, \dots, p$ , angegeben werden. Diese Eigenschaft ist darauf zurückzuführen, dass  $\{G_j(b) : b \in [0, B]\}$  dieselbe Verteilung besitzt wie  $\{(1+b)W_j(b/(1+b)) : b \in [0, B]\}$ , für  $j = 1, \dots, p$ . Da beide Prozesse Gaußprozesse sind, kann dies durch einfaches Ausrechnen der Erwartungswert- und Kovarianzfunktion belegt werden. Somit gilt

$$\sup_{b \in [0, B]} \frac{\|\mathcal{G}_p(b)\|_2}{w(b)} = \sup_{b \in [0, B]} \frac{\sqrt{\sum_{j=1}^p [G_j(b)]^2}}{w(b)} \stackrel{d}{=} \sup_{b \in [0, B]} \frac{(1+b) \sqrt{\sum_{j=1}^p \left[ W_j \left( \frac{b}{1+b} \right) \right]^2}}{w(b)}, \quad (3.4)$$

wobei  $A_1 \stackrel{d}{=} A_2$  besagt, dass zwei Zufallsobjekte  $A_1$  und  $A_2$  die gleiche Verteilung besitzen.

### Wahl der Grenzfunktion

Die einzigen Regularitätsbedingungen für die Grenzfunktion  $w(\cdot)$  sind Stetigkeit und ein streng positiver Verlauf. Bei der Wahl einer passenden Funktion können verschiedene Kriterien herangezogen werden. Beispielsweise wählen Segen und Sanderson (1980) die Funktion  $g(t) = 2t \log \log(t)$ . Diese Funktion besitzt das langsamste Wachstum, das gerade noch die vorausgesetzten Annahmen erfüllt, und garantiert somit die geringste Verzögerung bei der Detektion eines Strukturbruchs. Chu et al. (1996) wiederum kritisieren, dass diese Wahl der Grenzfunktion in Simulationen schwankende Wahrscheinlichkeiten für einen Fehler erster Art liefert. Um dieses aus analytischer Sicht unbefriedigende Ergebnis zu verbessern, empfehlen sie die Verwendung der Funktion  $g(t) = \sqrt{t} [a^2 + \ln(t)]^{1/2}$  für eine passend gewählte Konstante  $a$ . Diese Funktion wiederum wird von Horváth et al. (2004) abgeändert in

$$g(t) = c\sqrt{m}(1+t) \left( \frac{t}{1+t} \right)^\gamma$$

für eine Konstante  $c$  und einen Justierungsparameter  $\gamma \in [0, 1/2)$ . Aufbauend darauf modifizieren Wied und Galeano (2013) die letztere Variante leicht, um für die Grenzfunktion ausschließlich streng positive Werte zu garantieren. Die Grenzfunktion  $g(\cdot)$  kann dann gewählt werden als  $g(t) = \sqrt{mc}w(t)$  mit

$$w(t) = (1+t) \cdot \max \left\{ \left( \frac{t}{1+t} \right)^\gamma, \varepsilon \right\}. \quad (3.5)$$

Dabei ist  $\gamma \in [0, 1/2)$  der Justierungsparameter und  $\varepsilon > 0$  eine Konstante, die allein dazu dient, die Dividierbarkeit durch Werte der Grenzfunktion zu garantieren. Dieser Wert kann in Anwendungen beliebig klein gewählt werden. Der Parameter  $\gamma$  wird genutzt, um das Verfahren zu adjustieren und ein optimales Verhalten in einer konkreten erwarteten Situation zu erreichen. Wie in Wied und Galeano (2013) eingehend diskutiert wurde, tritt ein Konflikt auf zwischen dem Ziel, Strukturbrüche so schnell wie möglich nach ihrem Auftreten nachzuweisen, und dem Bedürfnis, die Wahrscheinlichkeit für das fälschliche Ablehnen der Nullhypothese auf das theoretische Signifikanzniveau zu beschränken. Wird der Justierungsparameter  $\gamma$  nahe der Obergrenze  $\frac{1}{2}$  gewählt, neigt das Verfahren dazu, die Nullhypothese zügig abzulehnen. Dies ist wünschenswert, wenn ein möglicher Strukturbruch zu Beginn des Überwachungszeitraums erwartet werden kann. Leider fördert eine solche Wahl des Parameters in der Praxis die Neigung des Verfahrens zu Fehlern erster Art, also der fälschlichen Ablehnung einer gültigen Nullhypothese. Im Gegensatz dazu reduziert ein kleinerer Wert des Parameters  $\gamma$  eher die Auftrittshäufigkeit von Fehlern erster Art in Anwendungen, was allerdings gleichzeitig eine verminderte Fähigkeit zur zügigen Entdeckung von Varianzänderungen, die kurz nach Beginn des Überwachungszeitraums auftreten, mit sich bringt. Weiterführende Simulationen, die nicht Teil dieser Arbeit sind, haben gezeigt, dass größere Werte für  $\gamma$  schnell zu untragbar hohen Anteilen zu Unrecht abgelehnter Nullhypothesen führen, besonders wenn die betrachteten Zufallsvektoren von höherer Dimension sind. Als Konsequenz dieser Eigenschaft werden in diesem Kapitel keine  $\gamma$ -Werte größer als 0.25 berücksichtigt.

### Simulation von kritischen Werten

Wird nun die Grenzfunktion in (3.4) als  $w(\cdot)$  aus (3.5) festgelegt und  $u = \frac{b}{1+b}$  sowie  $s = \frac{u(1+B)}{B}$  gewählt, so ergibt sich analog zum Vorgehen in Wied und Galeano (2013):

$$\sup_{b \in [0, B]} \frac{\|\mathcal{G}_p(b)\|_2}{w(b)} \stackrel{d}{=} \sup_{u \in [0, \frac{B}{1+B}]} \frac{\sqrt{\sum_{i=1}^p [W_i(u)]^2}}{\max\{u^\gamma, \varepsilon\}} \stackrel{d}{=} \sup_{s \in [0, 1]} \left(\frac{B}{1+B}\right)^{\frac{1}{2}-\gamma} \frac{\sqrt{\sum_{i=1}^p [W_i(s)]^2}}{\max\{s^\gamma, \varepsilon \left(\frac{1+B}{B}\right)^\gamma\}}.$$

Da unter den Annahmen von Theorem 3.1

$$\lim_{m \rightarrow \infty} \mathbb{P}(\tau_m < \infty) = \mathbb{P}\left(\sup_{s \in [0, 1]} \left(\frac{B}{1+B}\right)^{\frac{1}{2}-\gamma} \frac{\sqrt{\sum_{i=1}^p [W_i(s)]^2}}{\max\{s^\gamma, \varepsilon \left(\frac{1+B}{B}\right)^\gamma\}} > c\right)$$

gilt, können Monte-Carlo-Simulationen genutzt werden, um Konstanten  $c = c(\tilde{\alpha})$  zu bestimmen, so dass für ein bestimmtes  $\tilde{\alpha} \in (0, 1)$  gilt:

$$\mathbb{P}\left(\sup_{s \in [0, 1]} \left(\frac{B}{1+B}\right)^{\frac{1}{2}-\gamma} \frac{\sqrt{\sum_{i=1}^p [W_i(s)]^2}}{\max\{s^\gamma, \varepsilon \left(\frac{1+B}{B}\right)^\gamma\}} > c(\tilde{\alpha})\right) = \tilde{\alpha}.$$

Unter der Verwendung der simulierten Skalierungskonstanten  $c(\tilde{\alpha})$ , die in diesem Fall als kritischer Wert fungiert, beträgt die Wahrscheinlichkeit für das Begehen eines Fehlers erster Art somit gerade  $\tilde{\alpha}$ , wenn  $m$  groß genug beziehungsweise der historische Zeitraum lang genug ist. Simulierte kritische Werte  $c(\tilde{\alpha})$  für alle Parameterkombinationen aus  $p \in \{2, 5, 10\}$ ,  $B \in \{0.5, 1, 2\}$ ,  $\gamma \in \{0, 0.25\}$  und Signifikanzniveaus von  $\tilde{\alpha} \in \{0.01, 0.05, 0.1\}$  können Tabelle 3.1 entnommen werden. Für jeden der betrachteten Fälle wurden jeweils 10.000 Brownsche Bewegungen auf Netzen von 10.000 äquidistanten Punkten simuliert.

$\gamma$	$B$	$\tilde{\alpha} = 0.01$			$\tilde{\alpha} = 0.05$			$\tilde{\alpha} = 0.1$		
		$p = 2$	$p = 5$	$p = 10$	$p = 2$	$p = 5$	$p = 10$	$p = 2$	$p = 5$	$p = 10$
0	0.5	1.906	2.327	2.846	1.551	2.027	2.580	1.399	1.882	2.415
	1	2.292	2.865	3.522	1.904	2.466	3.154	1.700	2.312	2.944
	2	2.625	3.337	4.021	2.192	2.870	3.638	1.974	2.645	3.401
0.25	0.5	2.523	3.158	3.890	2.144	2.776	3.439	1.943	2.587	3.260
	1	2.812	3.488	4.274	2.388	3.036	3.805	2.163	2.846	3.605
	2	2.985	3.746	4.582	2.535	3.293	4.132	2.300	3.052	3.872

TABELLE 3.1: Simulierte kritische Werte  $c(\tilde{\alpha})$ .

Außerdem sind die Grenzfunktionen  $c(\tilde{\alpha})w(\cdot)$ , die sich für ein Signifikanzniveau von  $\tilde{\alpha} = 0.05$  und mit Hilfe der simulierten kritischen Werte für  $B = 2$  in Tabelle 3.1 ergeben, in Abbildung 3.1 illustriert. Die Grafik verdeutlicht, dass die Grenzfunktion für höhere Werte des Justierungsparameters und für kleine  $t$  niedrigere Funktionswerte annimmt und somit früh auftretende Strukturbrüche zuverlässiger nachweisen kann, als unter Verwendung eines kleineren  $\gamma$ -Werts.

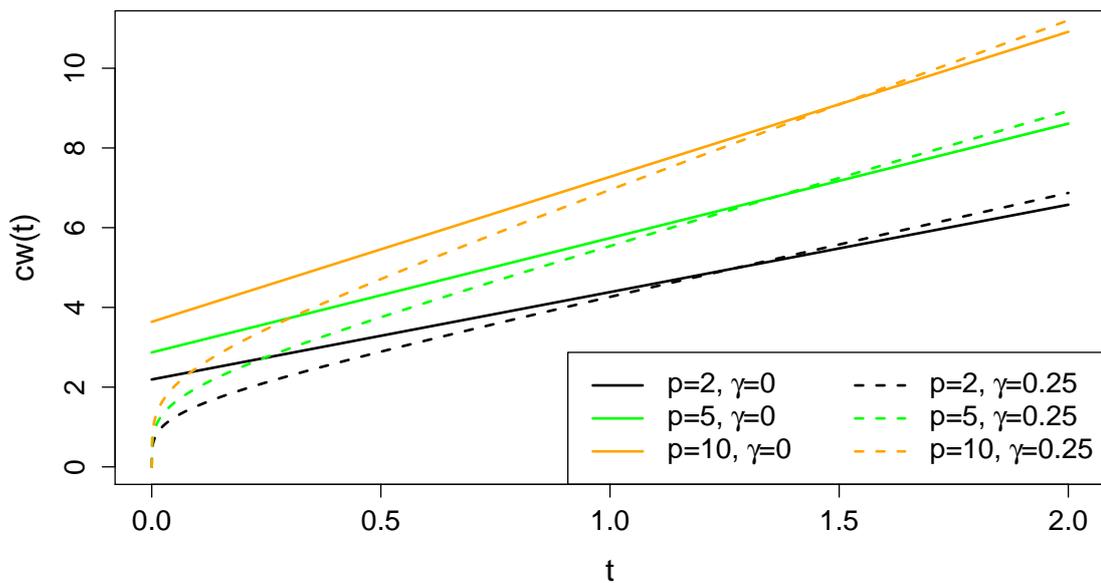


ABBILDUNG 3.1: Grenzfunktion  $cw(\cdot)$  für verschiedene Dimensionen und Werte des Justierungsparameters mit  $\tilde{\alpha} = 0.05$ .

### Der Detektor unter lokalen Alternativen

Die Menge der Alternativen zur Nullhypothese konstanter Varianzen ist ziemlich breit, da sie sowohl Szenarien mit einem einzigen oder mit multiplen Strukturbrüchen in einer oder mehreren Komponenten des Vektors der Varianzen umfasst, als auch Varianzänderungen von kleinerem und größerem Ausmaß. Eine derart umfassende Menge von Alternativen legt die Untersuchung der Güte des Verfahrens gegenüber lokalen Alternativen nahe. Ein solches Vorgehen wird beispielsweise in Kapitel 7.6.2. von Cameron und Trivedi (2005) beschrieben. Im dort behandelten Fall wird eine parametrische Funktion  $h(\cdot)$  betrachtet. Die Menge der Alternativen zur Nullhypothese  $H_0 : h(\theta) = 0$  ist unübersichtlich breit und beinhaltet eine Vielzahl an Szenarien. Um überhaupt Überlegungen zur Güte eines Verfahrens anstellen zu können, mit dem  $H_0$  überprüft werden kann, sollte deshalb eine feste Alternative  $h(\theta) = \delta$  für ein  $\delta \neq 0$  gewählt werden. Anschließend wird diese Konstante durch eine Nullfolge ersetzt. Mit Hilfe deren Folgenglieder kann wiederum eine Folge von festen Alternativen, die gegen die Nullhypothese strebt, definiert werden. Handelt es sich bei dem vom betrachteten Testverfahren verwendeten Schätzer für den Parameter oder Parametervektor  $\theta$  um einen  $\sqrt{T}$ -konsistenten Schätzer, so empfiehlt es sich, die  $\delta$ -Folge so zu wählen, dass sie ebenfalls mit der Rate  $\sqrt{T}$  gegen 0 strebt. Durch dieses Vorgehen wird die aus einem größeren Stichprobenumfang  $T$  resultierende Verbesserung des Parameterschätzers ausgeglichen, um eine nichtdegenerierte Grenzverteilung zu erhalten. Eine solche Folge lokaler Alternativen wäre beispielsweise gegeben durch

$$H_1^n : h(\theta) = \frac{\delta}{\sqrt{T}}.$$

Im hier betrachteten Fall kann die Annahme einer Folge von lokalen Alternativen wie folgt formuliert werden:

**Annahme 3.5** Für den Prozess  $\{X_t, t \in \mathbb{Z}\}$  mit  $X_t = (X_{1t}, \dots, X_{pt})'$  lassen sich die Varianzen der einzelnen Komponenten zerlegen in

$$\text{Var}(X_{jt}) = \bar{\sigma}_j^2 + \frac{1}{\sqrt{m}} g_j \left( \frac{t}{m} \right), \quad j = 1, \dots, p,$$

mit  $\bar{\sigma}_j^2$  und  $g_j(\cdot)$ ,  $j = 1, \dots, p$ , wie in (3.1).

**Theorem 3.2** Unter den Annahmen 3.1-3.4 sowie einer Folge von lokalen Alternativen wie in Annahme 3.5 gilt für jedes  $B > 0$

$$\lim_{m \rightarrow \infty} \mathbb{P}(\tau_m < \infty) = \lim_{m \rightarrow \infty} \mathbb{P} \left( \sup_{b \in [0, B]} \frac{\|V_{\lfloor m \cdot b \rfloor + 2}\|_2}{w(b)} > c \right) = \mathbb{P} \left( \sup_{b \in [0, B]} \frac{\|\mathcal{G}(b) + h(b)\|_2}{w(b)} > c \right),$$

mit  $\{\mathcal{G}(b), b \in [0, B]\}$  aus Definition 3.4. Die Vektorfunktion  $h(\cdot)$  ist gegeben durch  $h(\cdot) = H(h_1(\cdot), \dots, h_p(\cdot))'$ . Abgesehen von einer Konstanten entspricht  $H$  dem Grenzwert von  $\hat{D}_{p,m}$  unter der Nullhypothese. Außerdem gilt  $h_j(b) := \int_1^{b+1} g_j(u) du = 0$  für alle  $b \in [0, B]$  genau dann, wenn der  $j$ -te Komponentenprozess keiner Varianzänderung unterliegt.

## Güte des Überwachungsverfahrens

Theorem 3.2 sagt aus, dass auch kleinere Varianzänderungen in einzelnen Vektorkomponenten mit hoher Wahrscheinlichkeit nachgewiesen werden können, wenn der historische Zeitraum lang genug ist. Um allgemeinere Aussagen über die Güte des Tests treffen zu können, kann außerdem gezeigt werden, dass beliebige Varianzänderungen mit vorgegebener Wahrscheinlichkeit nachgewiesen werden können, wenn sie nur groß genug sind. Dies kann modelliert werden, indem eine der Funktionen  $g_j(\cdot)$ , die die strukturelle Stabilität der einzelnen Varianzen steuern, als durch den Faktor  $M$  skalierte Funktion  $g_*(\cdot)$  definiert wird. Es sei weiter angenommen, dass der Skalierungsfaktor gegen  $\infty$  strebt, was eine wachsende Differenz zwischen den Varianzen des betrachteten Komponentenprozesses vor und nach dem Strukturbruch impliziert.

**Annahme 3.6** Mindestens eine der Funktionen  $g_j(\cdot) \in \mathcal{F}_B$  mit  $\int_1^{b+1} |g_j(z)| dz > 0$  kann zerlegt werden in  $g_j(\cdot) = M \cdot g_*(\cdot)$  mit  $g_*(\cdot) \in \mathcal{F}_B$ .

**Definition 3.5** Unter der Alternative, dass mindestens eine Komponente des Vektors der Varianzen im Überwachungszeitraum von einem Strukturbruch betroffen ist, sei  $\mathbb{P}_{H_1}(M)$  die Wahrscheinlichkeit, dass der Detektor für ein gegebenes  $M$  während des Überwachungszeitraums die Grenzfunktion überschreitet.

**Theorem 3.3** Sei  $g_*(\cdot) \in \mathcal{F}_B$  beliebig, aber fest. Unter den Annahmen 3.1-3.4 und 3.6 existiert für jedes  $\varepsilon > 0$  ein  $M_\varepsilon$ , so dass für alle  $M > M_\varepsilon$  gilt

$$\lim_{m \rightarrow \infty} \mathbb{P}_{H_1}(M) > 1 - \varepsilon.$$

Theorem 3.3 belegt, dass eine ausreichend große Varianzänderung mit gegebener Wahrscheinlichkeit erkannt wird, wenn die Länge des historischen Zeitraums gegen  $\infty$  strebt. Dies gilt auch, wenn nur eine einzelne Komponente von einem Strukturbruch betroffen ist oder wenn verschiedene Komponenten entgegengesetzten Varianzänderungen ausgesetzt sind.

### Schätzung der Lage des Bruchpunkts

Wenn der Detektor tatsächlich im Laufe des Überwachungszeitraums die Grenzfunktion überschreitet, zeigt dies das Vorliegen eines Strukturbruchs an. In der Regel stimmt der Bruchpunkt allerdings nicht mit dem Stoppzeitpunkt  $\tau_m$  überein, da eine plötzliche Änderung in einer oder mehreren Komponenten des Varianzvektors etwas Zeit benötigt, um den Detektor stark genug zu beeinflussen, so dass dieser tatsächlich die Grenzfunktion überschreitet. Deshalb muss nach dem Anzeigen einer Varianzänderung noch die Lage des Bruchpunktes bestimmt werden. Eine Möglichkeit zur Lokation des Zeitpunktes der Varianzänderung liefert der folgende Schätzer, der eine multivariate Entsprechung der von Wied et al. (2012a) und Wied und Galeano (2013) verwendeten Schätzfunktion darstellt:

$$\hat{k} = \arg \max_{1 \leq j \leq \tau_m - 1} K_{j, \tau_m} \quad \text{mit} \quad K_{j, \tau_m} := \frac{j}{\sqrt{\tau_m}} \widehat{D}_{p, m}^{-1/2} \left\| \begin{pmatrix} [\overline{X}_1^2]_{m+1}^{m+j} - [\overline{X}_1^2]_{m+1}^{m+\tau_m-1} \\ \vdots \\ [\overline{X}_p^2]_{m+1}^{m+j} - [\overline{X}_p^2]_{m+1}^{m+\tau_m-1} \end{pmatrix} \right\|_2. \quad (3.6)$$

Durch den Vorfaktor werden die normierten Differenzen für kleine  $j$  herunterskaliert, da in diesem Fall erst wenige Beobachtungen aus dem Überwachungszeitraum vorliegen und deshalb die Mittelwerte, die als Subtrahenden verwendet werden, strukturbedingt noch stark schwanken können, ohne dass dies zwingend als Hinweis auf das Vorliegen eines Strukturbruchs verstanden werden muss. Für größere  $j$  hingegen nähern sich wegen  $\tau_m - 1 - j \rightarrow 0$  der Minuend und der Subtrahend immer stärker an, so dass schon kleinere Abweichungen auf das Vorliegen eines Strukturbruchs hinweisen können. Diesem Verhalten wird durch die Verwendung eines höheren Skalierungsfaktors Rechnung getragen.

Schätzer wie in (3.6) liefern im univariaten Fall wie beispielsweise in Wied et al. (2012a) und Wied und Galeano (2013) zufriedenstellende Ergebnisse und erscheinen deshalb auch im multivariaten Fall geeignet, um die Lage des Strukturbruchs zu schätzen. Gleichwohl überschreitet eine detaillierte Untersuchung der theoretischen Eigenschaften des Schätzers den Rahmen dieser Arbeit und kann als Ansatzpunkt für zukünftige Forschung dienen.

Die Leistungsfähigkeit des vorgeschlagenen Verfahrens und die empirischen Eigenschaften der Stoppzeit  $\tau_m$  und des geschätzten Bruchpunktes  $\hat{k}$  beziehungsweise der geschätzten relativen Position des Strukturbruchs im Überwachungszeitraum  $\hat{\lambda} = \frac{\hat{k}}{mB}$  werden im folgenden Abschnitt genauer untersucht.

### 3.3 Simulationen

Dieser Abschnitt behandelt das Verhalten des vorgeschlagenen Verfahrens in endlichen Stichproben. Jedes der betrachteten Szenarien wurde unter Einbeziehung verschiedener Parameter konstruiert. Die Dimension der betrachteten Zufallsvektoren beträgt  $p \in \{2, 5, 10\}$ . Da die Asymptotik in Kapitel 3.2 auf der Annahme beruht, dass die Länge des historischen Zeitraums gegen  $\infty$  strebt, werden mit  $m \in \{500, 1.000, 2.000\}$  relativ hohe Werte für diesen Parameter gewählt. Im Kontext von täglichen Daten entsprechen diese Werte Zeiträumen von etwa 2, 4 und 8 Jahren. Bei der Wahl von  $m$  ist zu beachten, dass in der Praxis kürzere Kalibrierungszeiträume häufig nichtinvertierbare Schätzungen für die Matrix  $\mathbf{D}_p$  liefern, insbesondere für höhere Dimensionen  $p$ . Andererseits verlangt Annahme 3.1, dass der historische Zeitraum frei von Strukturbrüchen ist. Da längere Zeiträume ohne augenscheinliche Varianzänderung in der Praxis kaum zu finden sind, werden größere Werte für  $m$  in dieser Arbeit nicht berücksichtigt. Des Weiteren erlaubt die Wahl von  $B \in \{0.5, 1, 2\}$  die Generierung von Zeitreihen, deren Überwachungszeitraum kürzer, gleich lang oder länger als die Kalibrierungsphase ist.

Abschließend müssen die Parameter, von denen die Grenzfunktion  $w(\cdot)$  in (3.5) abhängt, spezifiziert werden. Die Konstante  $\varepsilon$  wird in allen Simulationen als  $10^{-6}$  gewählt, während der Justierungsparameter die Werte  $\gamma \in \{0, 0.25\}$  annimmt. Diese Werte repräsentieren den Wunsch, Varianzänderungen aufzuspüren, die zu Beginn oder in der fortgeschrittenen Überwachungsphase auftreten. In allen Fällen beläuft sich das Signifikanzniveau auf  $\bar{\alpha} = 0.05$ . Für jeden Fall werden 10.000 Zeitreihen simuliert. Um kritische Werte zu bestimmen, werden jeweils 10.000 Brownsche Bewegungen auf einem Gitter von 10.000 äquidistanten Punkten simuliert.

#### 3.3.1 Zufallsfolgen unter der Nullhypothese

Zunächst werden Zeitreihen betrachtet, die weder im Längs- noch im Querschnitt, also weder über die Zeit noch zwischen den einzelnen Vektorkomponenten, Abhängigkeiten aufweisen. Dieses Vorgehen liefert Referenzwerte für das Verhalten in komplexeren Szenarien. Den einfachsten denkbaren Fall stellen dabei Zeitreihen dar, die sich als Realisation von Folgen von unabhängig identisch verteilten (u.i.v.) Zufallsvektoren ergeben. Konkret werden multivariat normal- oder  $t$ -verteilte Zufallsvektoren simuliert. Während die multivariate Normalverteilung ( $\mathcal{N}_p(\mu, \Sigma)$ ) etabliert und durch eine eindeutige Dichte charakterisiert ist, siehe beispielsweise Fahrmeir et al. (1996), wurden zur Erweiterung der  $t$ -Verteilung auf den multivariaten Fall verschiedene Varianten vorgeschlagen, von denen einige mit ihren Eigenschaften in Kotz und Nadarajah (2004) zu finden sind. Die Herleitung der intuitivsten Variante als Produkt eines normalverteilten Zufallsvektors und einer  $\chi^2$ - beziehungsweise  $\Gamma$ -verteilten Zufallsvariable kann Joe (2015), S.41f, entnommen werden. Gemeinsam mit Kotz und Nadarajah (2004) gilt dann:

**Definition 3.1 Multivariate  $t$ -Verteilung**

Sei  $\mathbf{X} = (X_1, \dots, X_p)'$  ein  $p$ -dimensionaler Zufallsvektor und  $f_{\mathbf{X}}(\cdot)$  die zugehörige Wahrscheinlichkeitsdichte. Der Zufallsvektor  $\mathbf{X}$  ist genau dann multivariat  $t$ -verteilt mit Mittelwertvektor  $\mu$ , Kovarianzmatrix  $\Sigma$  und  $\nu$  Freiheitsgraden, wenn gilt:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\gamma\left(\frac{\nu}{2}\right) [\pi\nu]^{p/2} [\det \Sigma]^{1/2}} \left[1 + \frac{1}{\nu} (\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu)\right]^{-\frac{\nu+p}{2}}.$$

Dies kann kurz geschrieben werden als:

$$\mathbf{X} \sim t_{p,\nu}(\mu, \Sigma).$$

**Definition 3.2 Multivariate Standardverteilungen**

Sei eine  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -Verteilung als  **$p$ -variante Standardnormalverteilung** bezeichnet, eine  $t_{p,\nu}(\mathbf{0}_p, \frac{\nu-2}{\nu}\mathbb{I}_p)$ -Verteilung als **multivariate Standard- $t$ -Verteilung** und eine  $t_{p,\nu}(\mathbf{0}_p, \frac{\nu-2}{\nu}\Sigma)$ -Verteilung als **standardisierte multivariate  $t$ -Verteilung**.

Dabei erlaubt die Standardisierung der Kovarianzmatrix der multivariaten  $t$ -Verteilung einen genaueren Vergleich der Simulationsergebnisse für die verschiedenen Verteilungen, da die Kovarianzmatrix einer nichtstandardisierten  $t_{p,\nu}(\mu, \Sigma)$ -Verteilung gerade  $\frac{\nu}{\nu-2}\Sigma$  beträgt, siehe Kotz und Nadarajah (2004). Somit sind sowohl die Varianzen der  $\mathcal{N}_p(\mathbf{0}_p, \Sigma)$ - und der  $t_{p,\nu}(\mathbf{0}_p, \frac{\nu-2}{\nu}\Sigma)$ -Verteilung als auch Änderungen der Diagonaleinträge von  $\Sigma$  gleich groß, was die Vergleichbarkeit der Simulationsergebnisse erhöht.

Für die in diesem Kapitel durchgeführten Simulationen wird die Anzahl der Freiheitsgrade bei der multivariaten  $t$ -Verteilung als  $\nu=8$  gewählt. Dieser Wert bewirkt moderat besetzte Ränder der Verteilung, die allerdings deutlich mehr Wahrscheinlichkeitsmasse beinhalten als im Falle der Normalverteilung.

Zur Festlegung der Kovarianzmatrizen wird die folgende Definition benötigt:

**Definition 3.3 Toeplitzmatrix** (Seber (2008), S.158)

Eine quadratische  $(n \times n)$ -Matrix  $A = [a_{ij}]_{i,j=1,\dots,n}$  wird dann als Toeplitzmatrix bezeichnet, wenn alle Einträge auf der Hauptdiagonalen gleich sind und außerdem die Einträge auf jeder der Nebendiagonalen jeweils gleich sind.

Das heißt, es gilt für alle  $i, j \in \{1, \dots, n\}$  und  $s \in \{-n+1, \dots, n-1\}$  mit  $|i+s|, |j+s| \in \{1, \dots, n\}$ :

$$a_{ij} = a_{i+s, j+s}.$$

Ist eine betrachtete Toeplitzmatrix zusätzlich noch symmetrisch, so ist ihre gesamte Struktur allein durch ihre erste Zeile festgelegt. Die zur Simulation verwendeten Kovarianzmatrizen sind zunächst Einheitsmatrizen und später symmetrische  $(p \times p)$ -Toeplitzmatrizen, deren erste Zeilen gegeben sind durch

$$\begin{aligned} \Sigma_{2,1\bullet} &= (1, 0.7) \\ \Sigma_{5,1\bullet} &= (1, 0.3, 0.4, 0.5, 0.6) \\ \Sigma_{10,1\bullet} &= (1, 0.1, 0.11, 0.12, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4). \end{aligned} \tag{3.7}$$

Die Wahl dieser Kovarianzmatrizen beruht darauf, dass die Matrizen zwar unterschiedliche Dimensionen aufweisen, sich aber in Struktur und Größe der Eigenwerte ähneln, was die Vergleichbarkeit der Simulationsergebnisse erhöhen sollte. Das empirische Niveau für Folgen von unabhängig identisch multivariat standardnormal- oder standard- $t$ -verteilten Zufallsvektoren findet sich in den Tabellen B.1 und B.2 in Anhang B sowie in Abbildung 3.2.

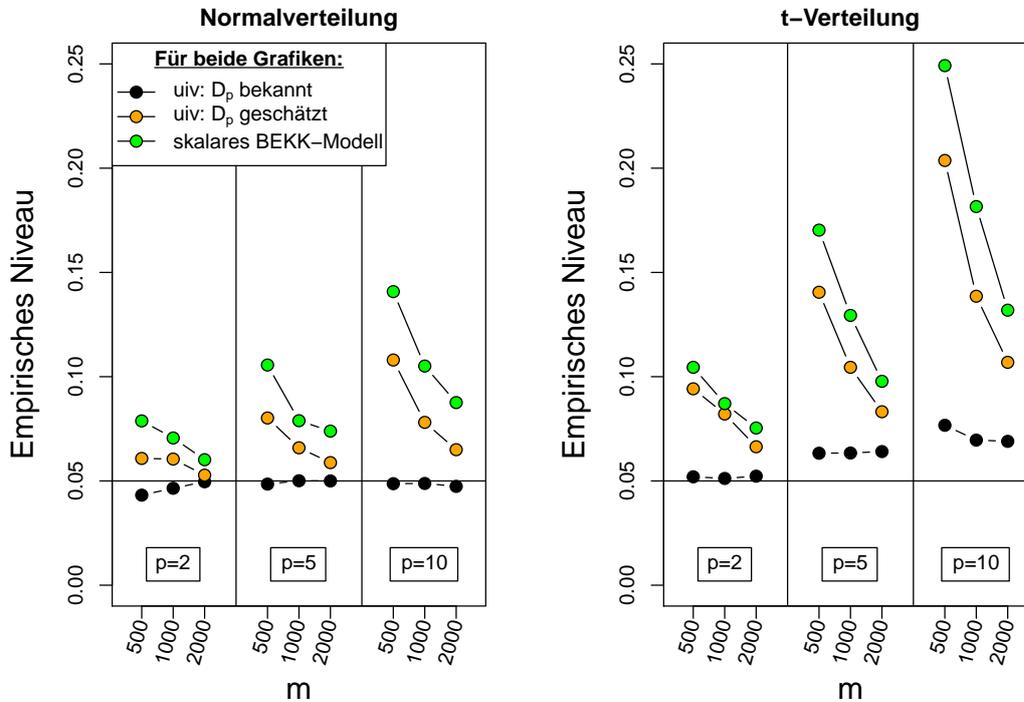


ABBILDUNG 3.2: Niveauvergleich: Anwendung auf u.i.v. Zufallsvektoren bei denen die Matrix  $D_p$  entweder bekannt ist oder geschätzt wurde und auf skalare BEKK-Zeitreihen.

Zugunsten einer übersichtlicheren Darstellung und aufgrund der Ähnlichkeit der Ergebnisse, zeigen die Abbildungen nur die Ergebnisse für  $\gamma = 0$  und den Fall, dass die Kovarianzmatrix  $\Sigma$  einer Einheitsmatrix passender Dimension  $p$  entspricht.

Die Tatsache, dass die Ergebnisse in den Tabellen B.1 und B.2 kaum von den Kovarianzmatrixeinträgen jenseits der Hauptdiagonalen abhängen, ist plausibel und zu erwarten, wenn man bedenkt, dass das Verfahren nur von Schätzungen der Hauptdiagonaleinträge, nicht aber von den übrigen Einträgen abhängt.

Im Allgemeinen steigt das empirische Niveau mit der Dimension  $p$ . Um den Ursprung dieses Phänomens zu klären, kann für einfache Szenarien die Matrix  $D_p$  selbst anstatt einer Schätzung verwendet werden. Für standardnormalverteilte Zufallsvektoren ergibt sich  $D_p = 2\mathbb{I}_p$  und für standard- $t$ -verteilte Zufallsvektoren  $D_p = \frac{2(\nu-1)}{\nu-4}\mathbb{I}_p$ . Die Ergebnisse für diese Simulationen können den Tabellen B.1 und B.2 in Anhang B entnommen werden und werden in Abbildung 3.2 illustriert. Sie zeigen, dass ein Großteil der Niveauperzerrung durch eine mangelhafte Schätzung der Variationsmatrix hervorgerufen wird. Einen weiteren Anstieg der Ablehnraten bewirken schwere Ränder der Innovationsverteilung. Auch eine mit dem Ziel einer Senkung des empirischen Niveaus durchgeführte Suche nach alternativen Schätzverfahren und Bandbreiten konnte keine Fortschritte zur Verbesserung des Verfahrens beitragen.

Da alle asymptotischen Aussagen auf der Annahme beruhen, dass die Länge des historischen Zeitraums gegen  $\infty$  strebt, nehmen auch die Niveauverzerrungen mit steigendem  $m$  ab. Während kein allgemeiner und gerichteter Einfluss von Änderungen des Parameters  $B$  auf das empirische Niveau festgestellt werden kann, steigt der Anteil der fälschlicherweise abgelehnten Nullhypothesen mit dem Wert des Justierungsparameters  $\gamma$ .

### 3.3.2 Zufallsfolgen unter der Alternativen

In diesem Abschnitt wird angenommen, dass die Kovarianzmatrix vor dem Strukturbruch in Abhängigkeit von der Dimension  $p$  einer der Matrizen  $\Sigma_p$  in (3.7) entspricht, deren Hauptdiagonaleinträge sich innerhalb des Überwachungszeitraums ändern. Im ersten Szenario steigen alle Varianzen von 1 auf 1.3 an, im zweiten springt nur eine der Varianzen von 1 auf 1.5. In beiden Szenarien ist von Interesse, wie zuverlässig ein früher oder späterer Strukturbruch entdeckt werden kann. Um die Simulationsergebnisse auch für verschiedene Längen des historischen Zeitraums vergleichen zu können, gibt der Parameter  $\lambda^* \in (0, 1)$  jeweils die Lage des Bruchpunktes innerhalb des Überwachungszeitraums an. Die Wahl  $\lambda^* \in \{0.05, 0.5\}$  signalisiert dabei eine Varianzänderung am Anfang ( $k^* = 0.05mB$ ) oder in der Mitte ( $k^* = 0.5mB$ ) des Überwachungszeitraums.

Die Ergebnisse im Fall einer kleineren Änderung, die alle Varianzen betrifft, finden sich in den Tabellen B.3 und B.4 in Anhang B und werden in Abbildung 3.3 grafisch dargestellt, während die Ergebnisse für den Fall einer einzelnen, aber stärkeren Varianzänderung den Tabellen B.5 und B.6 zu entnehmen und in Abbildung 3.4 dargestellt sind.

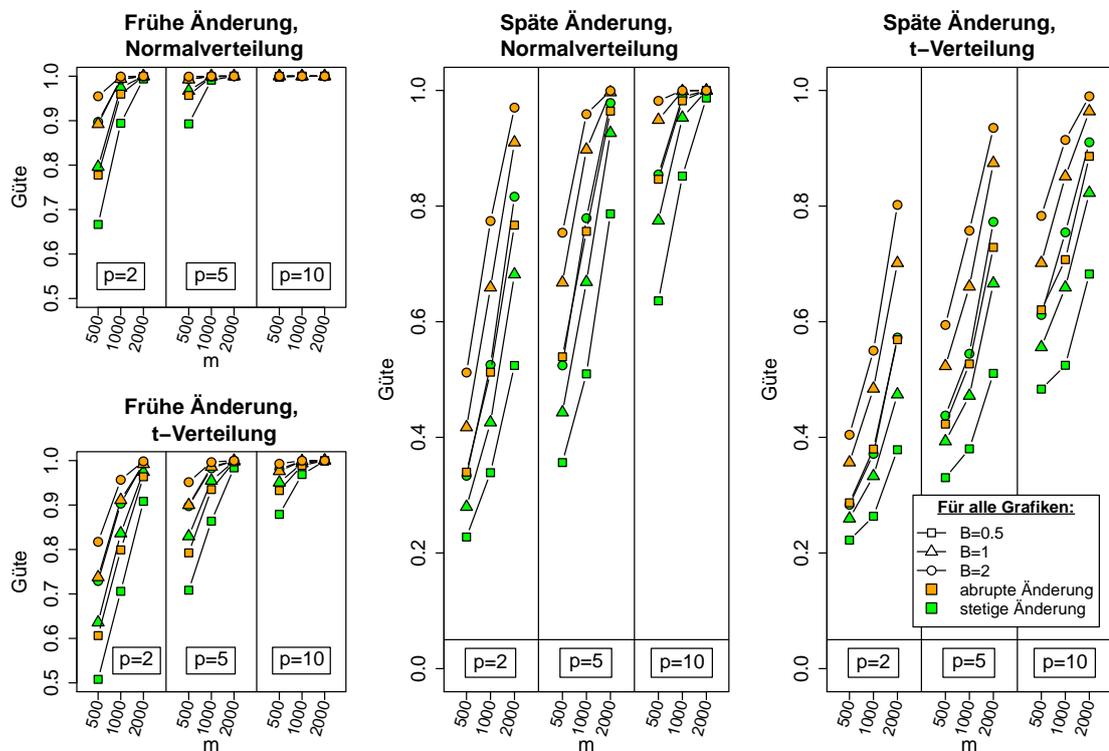


ABBILDUNG 3.3: Güte bei der Anwendung auf u.i.v. Zufallsvektoren, wenn sich alle Varianzen ändern.

Allgemein nimmt die Fähigkeit des Verfahrens, einen vorhandenen Strukturbruch nachzuweisen, mit wachsender Länge des historischen Zeitraums zu. Dieses Ergebnis steht in Einklang mit Theorem 3.3, das besagt, dass prinzipiell jede ausreichend große Varianzänderung mit einer beliebig hohen Wahrscheinlichkeit (kleiner Eins) erkannt werden kann, wenn der historische Zeitraum nur lang genug ist. Ändern sich alle Varianzen gemeinsam, nimmt die Güte mit steigender Dimension zu, wenn auch auf Kosten einer zunehmenden Niveauverzerrung, wie Abbildung 3.2 verdeutlicht. Ändert sich nur eine der Varianzen, sinkt die Güte mit steigender Dimension. Auch dieses Ergebnis ist plausibel, wenn berücksichtigt wird, dass der Anteil derjenigen Einzelzeitreihen, die nicht vom Strukturbruch betroffen sind, mit steigender Dimension immer größer wird. Entsprechend wird es immer

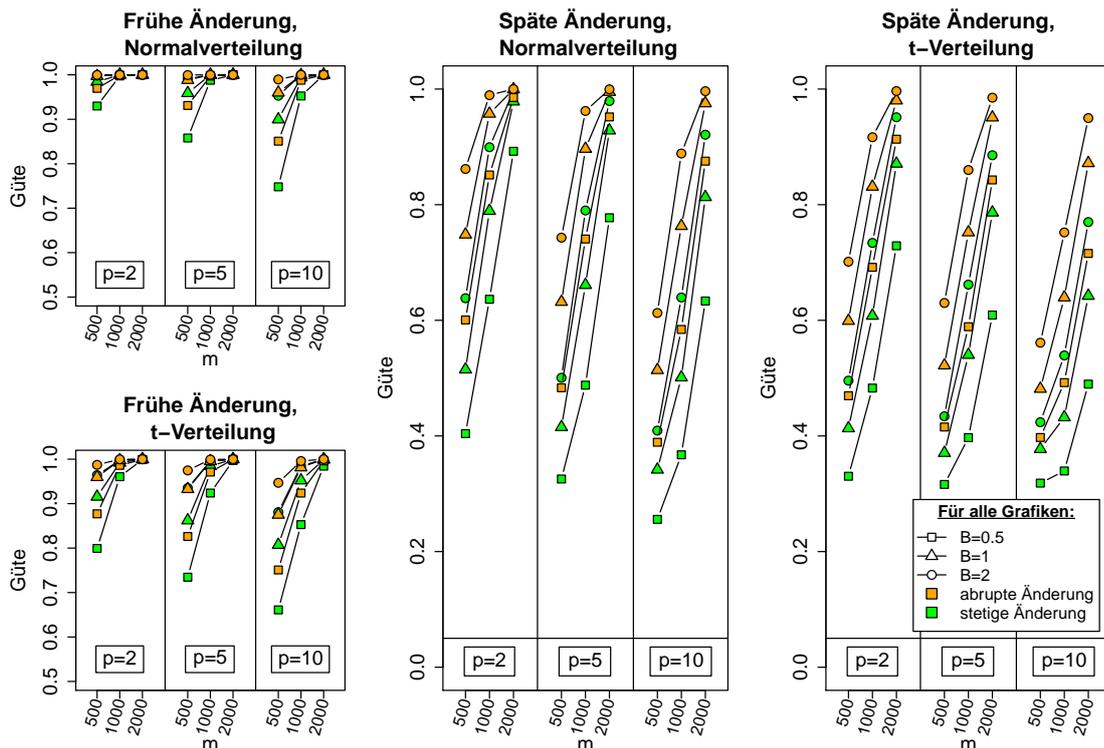


ABBILDUNG 3.4: Güte bei der Anwendung auf u.i.v. Zufallsvektoren, wenn sich nur eine der Varianzen ändert.

schwerer, den Strukturbruch nachzuweisen. Insgesamt sinkt die Güte sehr stark, wenn der Strukturbruch zu einem späteren Zeitpunkt eintritt, da in diesem Fall weniger Datenpunkte im verbleibenden Überwachungszeitraum übrig bleiben, anhand derer die Varianzänderung entdeckt werden könnte. Dieses Resultat tritt besonders deutlich für  $t$ -verteilte Zufallsvektoren zu Tage, wie die Ergebnisse in den Tabellen B.4 und B.6 in Anhang B belegen. Der direkte Vergleich der in den Abbildungen 3.3 und 3.4 visualisierten Ergebnisse zeigt, dass eine stärkere Änderung, die nur eine einzelne Komponente betrifft, nur dann zuverlässiger erkannt werden kann als eine kleinere Änderung aller Varianzen, wenn die Dimension nicht zu groß ist und der Strukturbruch nicht zu spät im Überwachungszeitraum auftritt. Wie schon unter der Nullhypothese beeinträchtigen auch hier schwere Ränder wie im Fall der  $t$ -Verteilung die Fähigkeit des Verfahrens, Änderungen im Vektor der Varianzen zu erkennen, deutlich. Dieser Einfluss klingt mit wachsender Länge des historischen Zeitraums ab. Außerdem scheint der Anteil der erkannten Strukturbrüche in beiden Szenarien für wachsende Werte des Justierungsparameters  $\gamma$  abzunehmen. Während dies für eine später eintretende Änderung ein plausibles Ergebnis darstellt, widerspricht es der Erwartung, dass

früh eintretende Änderungen zuverlässiger unter der Verwendung eines höheren  $\gamma$ -Wertes erkannt werden können. Eine Erklärung für dieses Verhalten könnte in der Struktur der skalierten Grenzfunktion zu finden sein. Diese weist im Falle eines höheren  $\gamma$ -Wertes eine stärkere Steigung auf. Da sich beide Grenzfunktionen schneiden, kann somit die Situation auftreten, dass die durch den Vorfaktor  $\frac{k}{\sqrt{m}}$  herunterskalierten Fluktuationsdifferenzen die zu  $\gamma = 0$  gehörige Grenzfunktion eher überschreiten als die zu  $\gamma = 0.25$  gehörige.

Bisher wurden Varianzänderungen unterstellt, die abrupt eintreten, also beispielsweise als Reaktion auf ein spontanes Ereignis wie eine Katastrophe oder einen Skandal. In der Praxis manifestieren sich Änderungen allerdings häufig erst mit der Zeit oder Schritt für Schritt, zum Beispiel im Falle einer langsam aufkommenden Krise. Deshalb wird nun die Güte des Verfahrens hinsichtlich der Erkennung von stetigen Varianzänderungen untersucht. Es wird dabei ein linearer Anstieg einer oder aller Varianzen entsprechend der zuvor betrachteten Szenarien angenommen, der bei  $\lambda_1^* = 0.05$  beziehungsweise bei 0.5 beginnt und bei  $\lambda_2^* = 0.3$  beziehungsweise bei 0.75 abgeschlossen ist. Auch die Ergebnisse für diese Alternativszenarien sind in den Tabellen B.3 bis B.6 verzeichnet. Sie entsprechen den blauen Werten. Um einen direkten Vergleich zu ermöglichen, werden sie gemeinsam mit der Güte im Fall einer abrupten Varianzänderung gleichen Ausmaßes in den Abbildungen 3.3 und 3.4 dargestellt. Im Vergleich zum Fall einer plötzlichen Änderung bleibt der Einfluss von Parameteränderungen auf die Güte in allen Fällen erhalten, allerdings sinkt die Güte insgesamt beträchtlich. Dieses Ergebnis ist wenig überraschend, da die Simulationen von abrupten Varianzänderungen gezeigt haben, dass ein Strukturbruch umso schlechter nachgewiesen werden kann, je später dieser eintritt und die stetigen Änderungen erst deutlich später abgeschlossen sind als die plötzlichen. Obwohl stetige Varianzänderungen, die erst in der fortgeschrittenen Überwachungsphase einsetzen, unter Verwendung eines kurzen historischen Zeitraums nur schwer zu erkennen sind, steigt die Güte schnell mit der Länge des historischen Zeitraums an, was die Abbildungen 3.3 und 3.4 verdeutlichen.

### 3.3.3 GARCH-Zeitreihen unter der Nullhypothese

Die im vorherigen Abschnitt betrachteten Zeitreihen weisen allenfalls eine Abhängigkeit zwischen den einzelnen Komponenten eines Vektors auf. In der Praxis hingegen ist es häufig notwendig, auch serielle Abhängigkeit zuzulassen, die in moderatem Ausmaß durch die Unterstellung eines *near-epoch dependenten* Prozesses wie in Annahme 3.4 erlaubt wird. Im Finanzbereich ist dabei die Überwachung der unbedingten Varianzen von Zufallsvektoren von besonderem Interesse, deren bedingte Varianzstruktur als zeitvariabel angenommen werden kann. Ein beliebter Ansatz, um ein solches Verhalten zu modellieren und gleichzeitig auch Varianzcluster zu berücksichtigen, wie sie in Finanzzeitreihen oft beobachtet werden, liegt in der Verwendung von multivariaten GARCH-Modellen, wie sie in Abschnitt 2.2 vorgestellt wurden.

#### Modellierung der GARCH-Zeitreihen

Da Bollerslev (1986) argumentiert, dass schon GARCH-Modelle von niedriger Ordnung sehr gut geeignet sind, um das typische Verhalten vieler Finanzzeitreihen zu erklären, werden in dieser Arbeit ausschließlich Modelle betrachtet, in denen die bedingte Kovarianzmatrix nur von den verzögerten Beobachtungen und bedingten Kovarianzen zum Lag 1 abhängt. Konkret wird zur Modellierung der bedingten Kovarianzmatrix in (2.2) das von Ding und Engle (2001) vorgeschlagene Zweiparametermodell verwendet. Dieses ergibt sich durch *Variance Targeting* wie in Engle und Mezrich (1996) aus einem skalaren Diagonalmodell, das

heißt aus einem GARCH-Modell, in dem alle einzelnen Varianzen und Kovarianzen nur von der eigenen Vergangenheit sowie den passenden Einträgen vergangener Beobachtungen abhängen und bei dem zusätzlich alle Einzelgleichungen durch die gleichen Parameter gesteuert werden. Im Zuge des *Variance Targeting* wird dabei die konstante Matrix in der Rekursionsgleichung durch die unbedingte Kovarianzmatrix  $H$  ersetzt. Diese Matrix wird zusätzlich noch mit einem Vorfaktor versehen, der dafür sorgt, dass sich die Vorfaktoren aller Summanden zu Eins addieren, was eine Stabilisierung der Rekursionsgleichung bewirkt. Die bedingte Kovarianzmatrix ist dann definiert als

$$H_t = (1 - \alpha - \beta)H + \alpha X_{t-1}X'_{t-1} + \beta H_{t-1}, \quad (3.8)$$

wobei  $\alpha$  und  $\beta$  positive Skalare sind, die die Stationaritätsbedingung  $|\alpha + \beta| < 1$  erfüllen. Sowohl das skalare Diagonal- als auch das durch (3.8) festgelegte Modell sind Spezialfälle des BEKK(1, 1, 1)-Modells nach Engle und Kroner (1995). Deshalb wird dieses Modell im Folgenden als **skalares BEKK-Modell** bezeichnet.

Um das vorgeschlagene Verfahren auf skalare BEKK-Zeitreihen anwenden zu können, muss sichergestellt werden, dass die Modellparameter so gewählt werden, dass die Annahmen 3.2-3.4 nicht verletzt werden. Während Annahme 3.2 als gültig angenommen werden muss, werden die verbleibenden beiden Annahmen genauer betrachtet. Hansen (1991) zeigt, dass univariate GARCH(1,1)-Prozesse mit u.i.v. Innovationen *near-epoch dependent* sind, wenn für die Modellparameter  $\alpha$  und  $\beta$  sowie für die Wölbung der Innovationsverteilung  $\kappa$  gilt:

$$\beta^2 + 2\alpha\beta + \alpha^2\kappa < 1. \quad (3.9)$$

Da (3.9) mit den in dieser Arbeit verwendeten Parametern unter Verwendung von normal- und  $t$ -verteilten Innovationen erfüllt wird, kann diese Eigenschaft auf jeden der einzelnen Komponentenprozesse in (3.8) übertragen werden. Im Allgemeinen ist die Rückführung des multivariaten auf den univariaten Fall nicht so einfach möglich. Da allerdings das sehr einfach strukturierte skalare BEKK-Modell keine Querverbindungen zwischen den Varianzgleichungen der einzelnen Komponenten zulässt, impliziert eine Gültigkeit für die einzelnen Komponentenprozesse in diesem konkreten Fall auch die  $\mathcal{L}_2$ -*near-epoch dependence* des gesamten Prozesses.

Zur Überprüfung der Gültigkeit von Annahme 3.3 kann das im Folgenden eingeführte Lemma verwendet werden. Dazu wird das Konzept einer sphärischen Verteilung benötigt. Diese ist mit Theorem 2.1 in Fang et al. (1990) wie folgt definiert:

**Definition 3.4 Sphärische Verteilung**

Sei  $\mathbf{X} = (X_1, \dots, X_p)'$  ein  $p$ -dimensionaler Zufallsvektor und  $f_{\mathbf{X}}(\cdot)$  die zugehörige Wahrscheinlichkeitsdichte. Der Zufallsvektor  $\mathbf{X}$  besitzt genau dann eine sphärische Verteilung, wenn eine Funktion  $g(\cdot)$  existiert, sodass gilt:

$$f_{\mathbf{X}}(\mathbf{x}) = g(\mathbf{x}'\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^p.$$

Es ist leicht zu sehen, dass sowohl die multivariate Standardnormal- als auch die Standard- $t$ -Verteilung, die zur Erzeugung der Innovationsvektoren genutzt werden, sphärische Verteilungen sind. Zum Nachweis der Existenz der vierten Momente definiert Hafner (2003) die folgende Matrix:

$$G_p := \frac{1}{3} \mathbb{E} \left( \epsilon_{t1}^4 \right) \left[ 2 \left( \mathbb{L}_p \otimes \left[ \mathbb{D}'_p \mathbb{D}_p \right]^{-1} \mathbb{D}'_p \right) \left( \mathbb{I}_p \otimes \mathbb{K}_{pp} \otimes \mathbb{I}_p \right) \left( \mathbb{D}_p \otimes \mathbb{D}_p \right) + \mathbb{I}_{[p]^2} \right].$$

Die Matrix  $G_p$  transformiert die zweiten Momente des Vektors der bedingten Varianzen  $h_t$  in die zweiten Momente der bedingten Varianzen in eine Vektor-ARMA-Darstellung, siehe Hafner (2003). Sei

$$\Gamma := \mathbb{E} \left( X_t^2 X_t^{2'} \right) \quad \text{mit} \quad X_t^2 = \text{vech} (X_t X_t')$$

die Matrix der unbedingten vierten Momente und Kreuzmomente des Zufallsvektors  $X_t$ . Dann gilt:

**Lemma 3.1** *Wenn die Innovationen  $\epsilon_t$  eine sphärische Verteilung mit endlichen vierten Momenten besitzen, existiert die Matrix  $\Gamma$  genau dann, wenn alle Eigenwerte von*

$$Z := \alpha^2 G_p + (2\alpha\beta + \beta^2) \mathbb{I}_{[p+]^2} \quad (3.10)$$

*im Betrag kleiner als 1 sind.*

Um einen nur mäßig starken Einfluss vergangener Beobachtungen und bedingter Varianzen auf die aktuellen bedingten Varianzen zuzulassen, werden die Modellparameter als  $(\alpha, \beta) = (0.03, 0.45)$  gewählt. Die Innovationsvektoren  $\epsilon_t$  sind unabhängig identisch multivariat standardnormal- oder standard- $t$ -verteilt mit 8 Freiheitsgraden. Die variablen Parameter werden wie in Abschnitt 3.3.1 gewählt. Entsprechend der Dimension  $p \in \{2, 5, 10\}$  wird die unbedingte Kovarianzmatrix  $H$  als eine der durch (3.7) definierten Toeplitzmatrizen festgelegt. Die Simulationsergebnisse lassen sich den Tabellen B.7 und B.8 in Anhang B entnehmen. Um den Vergleich mit dem Fall unabhängig identisch verteilter Zufallsvektoren zu erleichtern, sind die Anteile fälschlicherweise abgelehnter Nullhypothesen für  $\gamma = 0$  auch in Grafik 3.2 in Abschnitt 3.3.1 dargestellt. Obwohl sich das Niveau im Vergleich zu Zeitreihen ohne zeitliche Unabhängigkeit leicht erhöht, bleibt der Einfluss von Änderungen der Parameter ähnlich. Allerdings sinkt das empirische Niveau im Fall eines relativ langen Überwachungszeitraums im Vergleich zum historischen Zeitraum, wenn serielle Abhängigkeit in der Zeitreihe vorhanden ist.

### Univariate Überwachung von multivariaten Zeitreihen

Da jede  $p$ -variante Zeitreihe aus  $p$  univariaten und möglicherweise unkorrelierten Einzelzeitreihen besteht, könnte auch jede dieser einzelnen Zeitreihen mit einem univariaten Gegenstück der vorgeschlagenen Methode überwacht werden. Um zu untermauern, warum die vorgeschlagene multivariate Methode der Anwendung von mehreren univariaten Verfahren vorgezogen werden sollte, werden das empirische Niveau und die Güte beider Verfahren bei der Überwachung skalarer BEKK-Zeitreihen verglichen. Der normierte Detektor wird dann zu

$$|V_{j,k}| = \frac{k}{\sqrt{m}} \widehat{D}_{1,m}^{-1/2} \left| \left[ \hat{\sigma}_j^2 \right]_{m+1}^{m+k} - \left[ \hat{\sigma}_j^2 \right]_1^m \right|, \quad j = 1, \dots, p, \quad (3.11)$$

wobei  $\widehat{D}_{1,m}$  einen Skalar darstellt. Die globale Nullhypothese, dass in keiner der  $p$  Zeitreihen eine Varianzänderung auftritt, wird dann abgelehnt, falls einer der  $p$  Detektoren in (3.11) die Grenzfunktion im Laufe des Überwachungszeitraums überschreitet. Um sicherzustellen, dass die Wahrscheinlichkeit, die Nullhypothese fälschlicherweise abzulehnen, asymptotisch das Signifikanzniveau  $\tilde{\alpha} = 0.05$  nicht übersteigt, wird das Signifikanzniveau der univariaten Verfahren mit Hilfe der Bonferronikorrektur adjustiert und somit jeweils auf  $\tilde{\alpha}/p$  gesetzt. Die simulierten empirischen Niveauewerte finden sich gemeinsam mit den Ablehnraten des multivariaten Verfahrens in den Tabellen B.7 und B.8 in Anhang B. Zusätzlich illustriert

Abbildung 3.5 die Ergebnisse für  $\gamma = 0$  und  $B = 1$ , um den Vergleich des multivariaten mit den univariaten Verfahren zu erleichtern. Unter Verwendung der univariaten Verfahren wird etwas seltener ein Fehler erster Art begangen, allerdings nimmt der Unterschied mit zunehmender Länge des historischen Zeitraums ab. Auch unter Verwendung univariater Verfahren steigt das empirische Niveau an, wenn die Innovationsverteilung schwere Ränder aufweist.

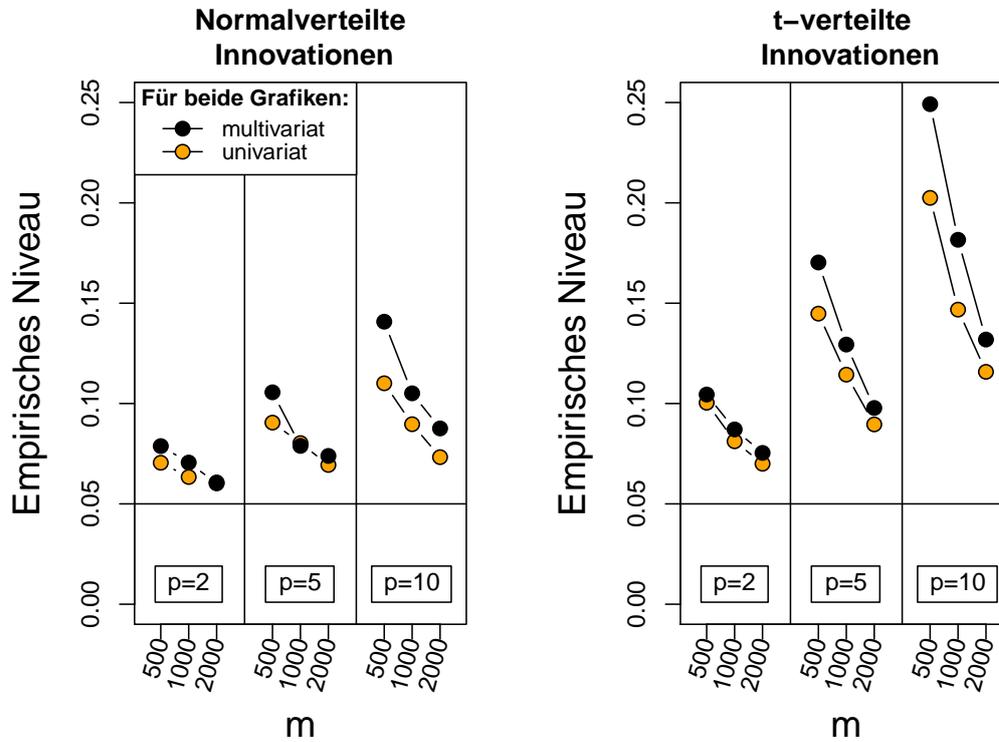


ABBILDUNG 3.5: Niveauevergleich: Überwachung von skalaren BEKK-Zeitreihen unter Verwendung von multi- und univariaten Methoden.

### 3.3.4 GARCH-Zeitreihen unter Alternativen

Nun werden sowohl das multivariate als auch die univariaten Verfahren mit verschiedenen Alternativszenarien konfrontiert. Diese entsprechen in ihrer Struktur denjenigen aus Abschnitt 3.3.2. Die Ergebnisse für die multivariate Methode finden sich in den Tabellen B.9-B.12 und diejenigen für die univariaten Verfahren in den Tabellen B.13-B.16 in Anhang B und werden außerdem in den Abbildungen 3.6 und 3.7 illustriert. Obwohl die Zeitreihen nun serielle Abhängigkeit aufweisen, ähneln die Ergebnisse denen für Folgen von u.i.v. Zufallsvektoren sehr stark. Die Fähigkeit, Strukturbrüche in der Varianz nachzuweisen sinkt etwas, wenn serielle Abhängigkeit zugelassen wird, allerdings bleibt der Einfluss von Parameteränderungen identisch im Vergleich zu Abschnitt 3.3.2.

Die Ablehnraten der univariaten Verfahren werden für den Fall normalverteilter Innovationen in Abbildung 3.8 gemeinsam mit der Güte des multivariaten Verfahrens grafisch dargestellt. Wie schon im Fall von u.i.v. Zufallsvektoren werden frühe Änderungen von beiden Methoden zuverlässig erkannt, während spätere Änderungen eher durch die multivariate Methode entdeckt werden. Letztere zeigt vor allem dann ihre Stärke, wenn alle Varianzen

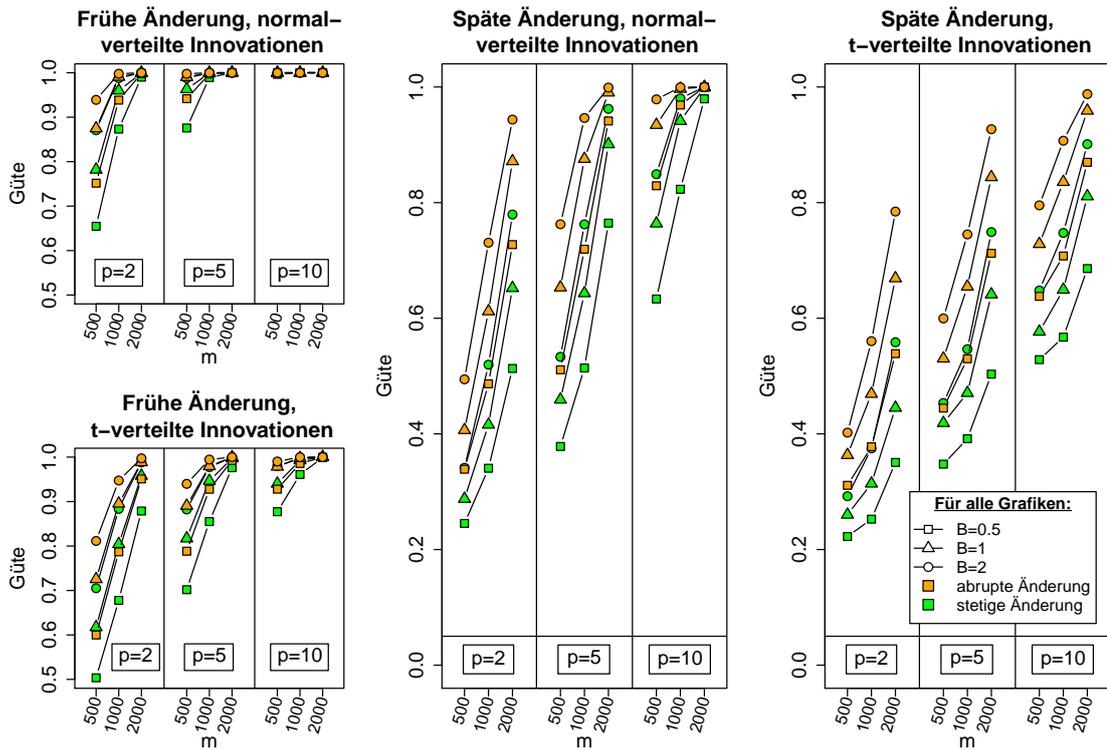


ABBILDUNG 3.6: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit der multivariaten Methode, wenn alle Varianzen von der Änderung betroffen sind.

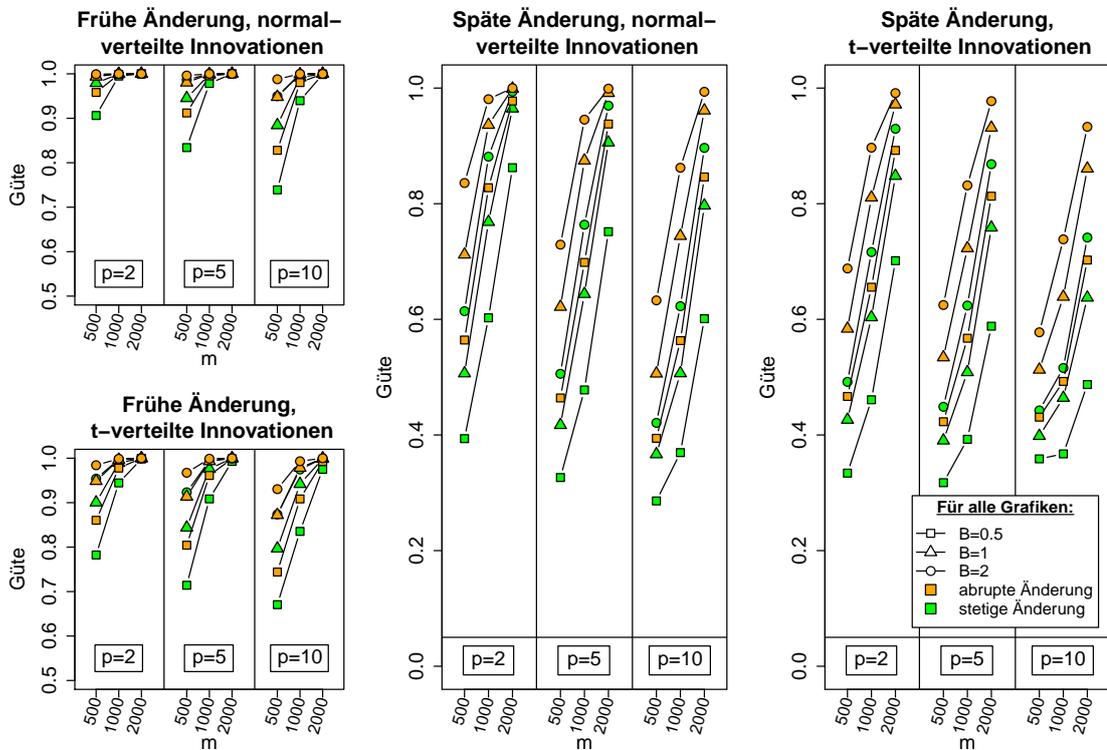


ABBILDUNG 3.7: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit der multivariaten Methode, wenn nur eine der Varianzen vom Strukturbruch betroffen ist.

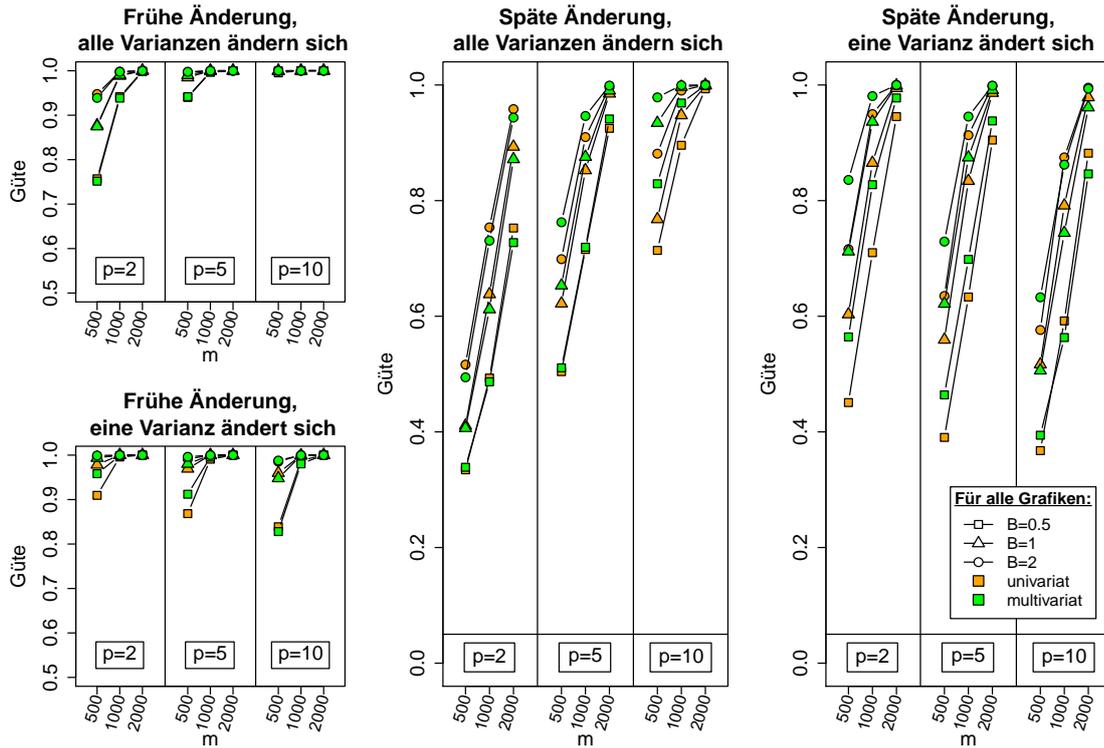


ABBILDUNG 3.8: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -verteilten Innovationen unter der Verwendung von univariaten Überwachungsverfahren.

eine kleinere Änderung erfahren oder wenn sich nur eine der Varianzen ändert, der historische Zeitraum aber kurz ist. Leider geht die höhere Güte der multivariaten Methode im Vergleich zur univariaten einher mit einem leicht erhöhten Niveau, wie in Abbildung 3.5 illustriert wird.

### Eigenschaften der Stoppzeiten und des Lokationsschätzers

Prinzipiell sollen Änderungen in den Varianzen nicht nur grundsätzlich entdeckt, sondern auch möglichst schnell nach ihrem Auftreten angezeigt werden. Deshalb werden auch die Reaktionszeit des Verfahrens sowie die Qualität der Lageschätzung für entdeckte Strukturbrüche untersucht. Während alle weiteren Parametereinstellungen wie zuvor gewählt werden, beschränkt sich die grafische Darstellung auf den Fall  $\gamma = 0$ . Um den Vergleich für verschiedene Zeitreihenlängen zu vereinfachen, illustrieren die Abbildungen 3.9 und 3.10 die standardisierten Anzeigeverzögerungen  $\frac{\tau_m - k^*}{mB}$  und die standardisierte Verzerrung des Lokationsschätzers  $\frac{\hat{k} - k^*}{mB}$ . Direkt unter den Boxplots werden außerdem die arithmetischen Mittel  $\pm$  der Standardabweichung dargestellt. Insgesamt nehmen sowohl die Verzögerung als auch die Verzerrung mit steigender Länge des historischen Zeitraums und Dimension ab, was eine schnellere Entdeckung und eine zuverlässigere Verortung erlaubt. Für kleinere Dimensionen bricht das Verfahren früher ab, wenn nur eine der Varianzen von einer stärkeren Änderung betroffen ist. Für höhere Dimensionen hingegen werden Änderungen, die mehrere Komponenten betreffen, schneller erkannt. Diese Ergebnisse stehen im Einklang mit den Güteeigenschaften, die im vorherigen Abschnitt diskutiert wurden. Da die Stoppzeiten  $\tau_m$  denjenigen Anteil des Datensatzes bestimmen, der zur Schätzung der Lage

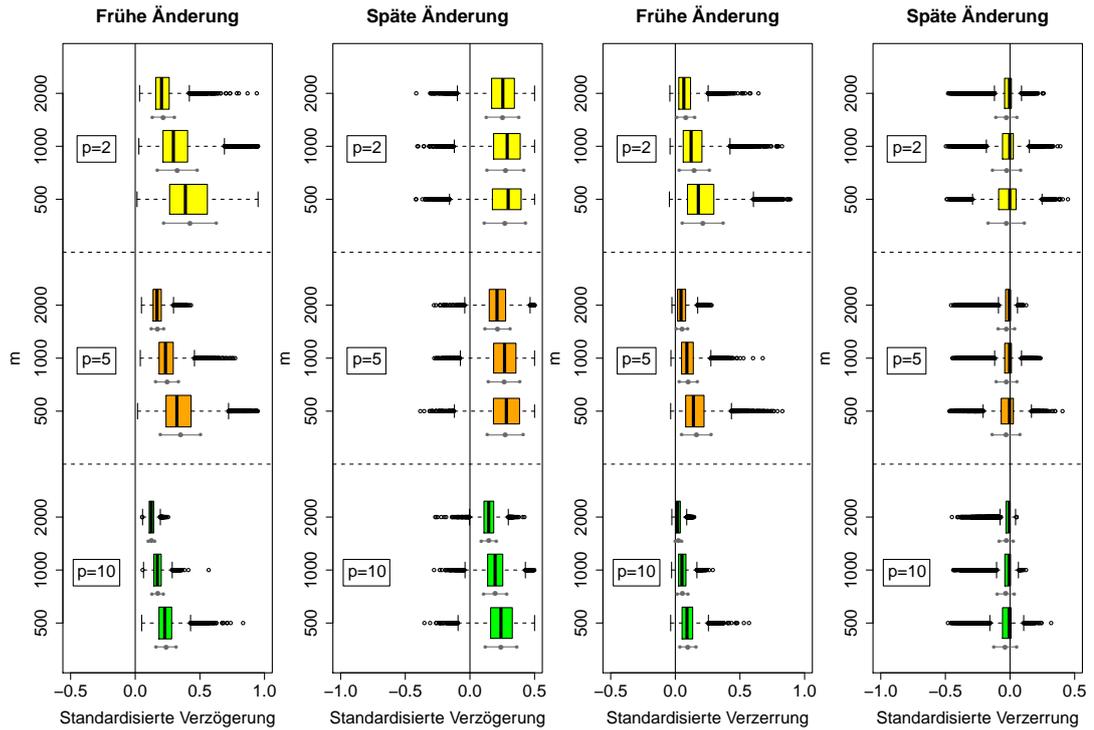


ABBILDUNG 3.9: Standardisierte Verzögerungszeiten und Verzerrung von  $\hat{\lambda}$ , wenn skalare BEKK-Zeitreihen mit  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -verteilten Innovationen überwacht werden und alle Varianzen von der Änderung betroffen sind.

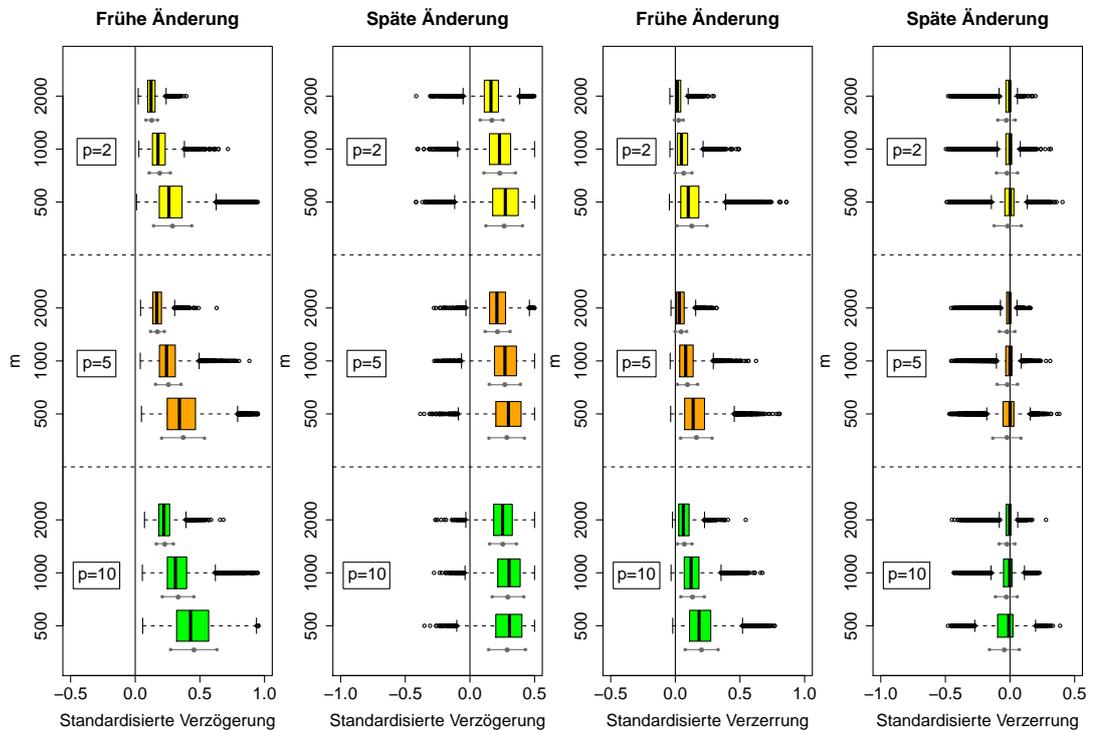


ABBILDUNG 3.10: Standardisierte Verzögerungszeiten und Verzerrung von  $\hat{\lambda}$ , wenn skalare BEKK-Zeitreihen mit  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -verteilten Innovationen überwacht werden und nur eine Varianz von der Änderung betroffen ist.

des Bruchpunktes verwendet wird, ist es nicht überraschend, dass die Verzögerung und die Verzerrung einen ähnlichen Verlauf aufweisen, wenn einzelne Parameter geändert werden.

### Änderung der Varianz der Innovationen

Bisher wurden Änderungen betrachtet, die die Einträge der unbedingten Kovarianzmatrix  $H$  direkt betreffen. In diesem Abschnitt hingegen steigen die Varianzen der Innovationsvektoren  $\epsilon_t$  von 1 auf 1.3, was auch einen Anstieg der Varianzen der resultierenden Vektoren  $y_t$  bewirkt. Die Ablehnraten können den Tabellen B.17 und B.18 in Anhang B entnommen werden und werden in Abbildung 3.11 für den Fall normalverteilter Innovationen gemeinsam mit den Ergebnissen für vergleichbare Varianzänderungen, die die unbedingte Kovarianzmatrix direkt betreffen, illustriert. Änderungen in den Varianzen der Innovationen können sogar noch etwas zuverlässiger erkannt werden als Änderungen der unbedingten Varianzen selbst. Dieses Ergebnis ist recht überraschend, da die Innovationen nur einen indirekten Einfluss auf die unbedingten Varianzen ausüben und somit im Hinblick auf die Modellgleichung (3.8) zu erwarten gewesen wäre, dass eine Änderung in den unbedingten Varianzen selbst zuverlässiger erkannt werden sollte.

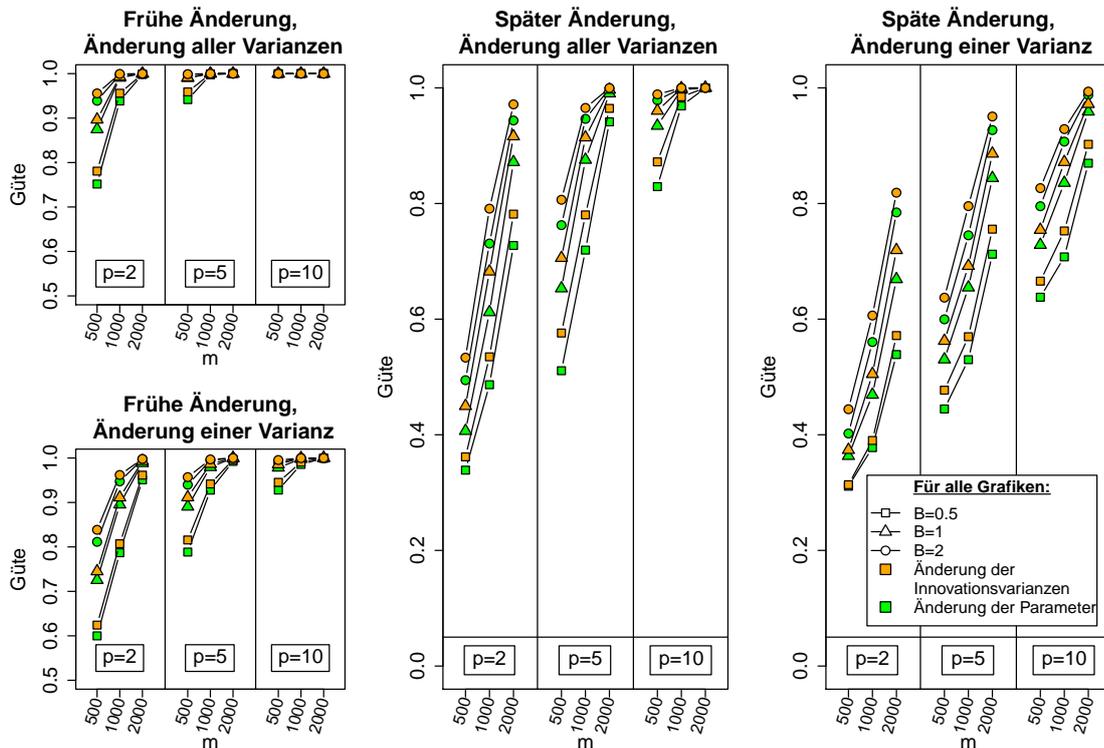


ABBILDUNG 3.11: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -verteilten Innovationen, wenn sich die Varianzen der Innovationen oder die bedingten Varianzen ändern.

### Varianzänderungen verschiedener Größe

Zum Abschluss der Simulationsstudie wird das Verhalten des Verfahrens bei unterschiedlich großen Varianzerhöhungen untersucht. Abbildung 3.12 illustriert die Ablehnraten der vorgestellten Methode, wenn sie mit einer früheren oder späteren Änderung aller Varianzen von 1 in  $1 + \Delta$  für  $\Delta \in \{-0.7, -0.6, \dots, 0.6, 0.7\}$  konfrontiert wird. Es werden

skalare BEKK-Zeitreihen mit multivariat standardnormal- oder standard- $t$ -verteilten Innovationen erzeugt, deren historischer Bereich und Überwachungszeitraum jeweils 1.000 Beobachtungen umfassen. Um die Invertierbarkeit in allen Fällen zu garantieren, wird die unbedingte Kovarianzmatrix  $H$  außerdem als Diagonalmatrix angenommen. Im Einklang mit den bisherigen Ergebnissen steigt die Güte mit dem Ausmaß der Varianzänderung und einer wachsenden Dimension der zugrundeliegenden Zufallsvektoren. Auffällig ist, dass für betragsmäßig kleine  $\Delta$ -Werte ein Anstieg der Varianzen häufiger erkannt wird als ein Abfall gleichen Ausmaßes.

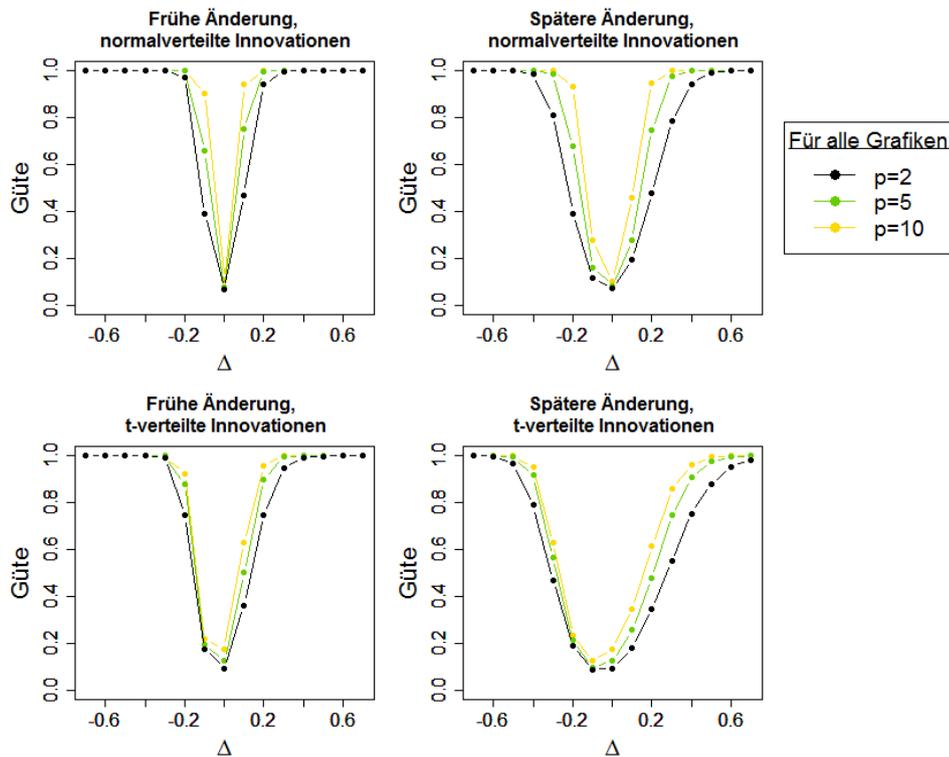


ABBILDUNG 3.12: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ - oder  $t_{p,\nu}(\mathbf{0}_p, \frac{\nu-2}{\nu}\mathbb{I}_p)$ -verteilten Innovationen und einer Varianzerhöhung um den Wert  $\Delta$ .

### 3.4 Anwendung auf Aktienrenditen

Abschließend wird die vorgestellte Methode genutzt, um die Varianzen einer Zeitreihe von Renditen verschiedener Aktien zu überwachen. Konkret werden Vektoren betrachtet, die die logarithmierten Eintagesrenditen der Zeitreihen der *DAX*-Unternehmen *Allianz*, *Bayer*, *Deutsche Bank*, *RWE* und *Siemens* von 1979 bis 2015 beinhalten. Da die Struktur der vorliegenden Daten mit verschiedenen multivariaten GARCH-Modellen, so auch mit dem skalaren BEKK-Modell, ansprechend modelliert werden kann, spricht nichts gegen eine Gültigkeit der Annahmen 3.3 und 3.4. Somit können die Varianzen der vorliegenden Zeitreihe mit der vorgeschlagenen Methode überwacht werden.

Weiterführende Simulationen, die nicht Teil dieser Arbeit sind, haben allerdings gezeigt, dass auch für relativ große Stichproben gravierende Niveauperzerrungen auftreten können, wenn die Summe der Parameter  $\alpha$  und  $\beta$  zu nahe bei Eins liegt, selbst wenn alle Bedingungen für die Anwendung des Verfahrens erfüllt sind. Je stärker sich die Parametersumme diesem Grenzwert annähert, desto eher neigt das sequenzielle Verfahren dazu, zu Beginn

des Überwachungszeitraums fälschlicherweise die Präsenz eines Strukturbruchs anzuzeigen. Eine Erklärung für dieses Verhalten liegt darin, dass eine Annäherung von  $\alpha + \beta$  an Eins bedeutet, dass auch der größte absolute Eigenwert der Matrix  $Z$  in (3.10) gegen Eins strebt. Dies kann als Annäherung an eine Verletzung von Annahme 3.3 interpretiert werden.

Passt man skalare BEKK-Modelle an den gesamten Datensatz oder kleinere Teildatensätze der log-Renditen an, legt dies nahe, dass die Summe von  $\alpha$  und  $\beta$  jeweils sehr nahe bei Eins liegt. Um zu verhindern, dass sich dies nachteilig auf das Niveau des Verfahrens auswirkt, wird die Methode deshalb anstelle der Originaldaten auf die Residuen angewendet, die sich bei der Anpassung eines skalaren BEKK-Modells an die vorliegende Zeitreihe ergeben. Dieses Vorgehen basiert auf der Idee, dass die Modellresiduen ein ähnliches Verhalten aufweisen wie die zugrundeliegenden Innovationen. Tatsächlich zeigten weiterführende Simulationen, die nicht Teil dieser Arbeit sind, dass ein Filtern von multivariaten GARCH-Zeitreihen und die Überwachung der resultierenden Residuenvektoren zu ähnlich hohen Ablehnraten führen wie im Fall, dass Prozesse von u.i.v. Zufallsvektoren mit gleicher Verteilung, wie sie für die Innovationsvektoren verwendet wurde, überwacht werden. Anhaltspunkte zur theoretischen Untermauerung dieser Annahme liefern Kulperger und Yu (2005), die zeigen, dass die auf GARCH-Residuen basierende Statistik des univariaten CUSUM-Tests das gleiche asymptotische Verhalten aufweist wie diejenige Teststatistik, die direkt auf den zugrundeliegenden Innovationen beruht. Falls außerdem die den überwachten Prozess antreibenden Innovationen aus einer der Normalverteilung ähnlichen Verteilung stammen, also beispielsweise einer Mischung aus Normalverteilungen oder einer  $t$ -Verteilung, so weisen diese ein Verhalten auf, das die Annahmen 3.2-3.4 erfüllt. Somit sollte auch der Detektor, der auf den Innovationen oder auf den den Innovationen ähnlichen Modellresiduen basiert, die gewünschten Eigenschaften bezüglich Niveau und Güte nicht erst für sehr große Stichproben, sondern auch schon für kleinere Datenumfänge aufweisen.

Wie Abschnitt 2.2 nahelegt, gibt es eine Vielzahl an multivariaten GARCH-Modellen, die an die vorliegenden Daten angepasst werden könnten. Damit die resultierenden Modellresiduen tatsächlich ein ähnliches Verhalten wie die zugrundeliegenden Innovationsvektoren aufweisen, sollte zur Modellanpassung sichergestellt werden, dass das Modell zur Struktur der Daten passt und dass außerdem konsistente Schätzer für die Parameter verwendet werden. Diese wurden für die meisten der gängigen Modelle wie das von Bollerslev et al. (1988) vorgestellte Vektor-GARCH-Modell oder die Modelle mit konstanter oder dynamischer Korrelation, siehe Bollerslev (1990) oder Jeantheau (1998) beziehungsweise Engle (2002) oder Tse und Tsui (2002), vorgeschlagen. Details zur konsistenten Parameterschätzung liefern beispielsweise Jeantheau (1998), Engle und Sheppard (2001) und Bauwens et al. (2006). Somit hätte die Methode auch auf Residuen, die bei der Anpassung eines anderen als des skalaren BEKK-Modells resultieren, angewandt werden können. Dennoch überschreitet ein detaillierter Nachweis, dass sich die asymptotische Verteilung des Detektors nicht ändert, den Rahmen dieser Arbeit und bleibt ein Ziel zukünftiger Forschung.

Im Folgenden werden die Parameter  $\alpha$  und  $\beta$  sowie die unbedingte Kovarianzmatrix  $H$  aus einer historischen Stichprobe vom Umfang  $m \in \{500, 1.000\}$  geschätzt. Dazu wird die in Pedersen und Rahbek (2014) vorgestellte zweistufige Quasi-Maximum-Likelihood-Methode verwendet. Da längere historische Zeiträume häufig nicht mehr als frei von Varianzänderungen angenommen werden können, wird  $m$  für die Anwendungen auf maximal 1.000 Beobachtungen beschränkt. Das verwendete Signifikanzniveau liegt bei  $\tilde{\alpha} = 0.05$ .

### Überprüfung der historischen Stichprobe auf Varianzänderungen

Da die Parameter zur Filterung der Zeitreihe aus der historischen Stichprobe geschätzt werden, muss diese Annahme 3.1 erfüllen, das heißt sie muss stabil bezüglich des Vektors der Varianzen sein. Um zu vermeiden, dass ein potenzieller Strukturbruch in der historischen Stichprobe übersehen wird, wird eine retrospektive Variante der vorgestellten Methode auf die Realisationen von  $X_1, \dots, X_m$  angewandt. Dieses Verfahren ähnelt der in Aue et al. (2009a) vorgestellten Methode und kann auch als multivariate Anlehnung an Wied et al. (2012a) oder Wied et al. (2012b) verstanden werden. Der verwendete Detektor lautet

$$Q_k = \frac{k}{\sqrt{m}} \widehat{D}_{p,m}^{-1/2} \left( [\widehat{\sigma}^2]_1^k - [\widehat{\sigma}^2]_1^m \right). \quad (3.12)$$

Dann gilt für das asymptotische Verhalten des Detektors

$$\lim_{m \rightarrow \infty} \mathbb{P} \left( \sup_{b \in [0,1]} \frac{\|Q_{\lfloor m \cdot b \rfloor + 2}\|_2}{w(b)} > c \right) = \mathbb{P} \left( \sup_{b \in [0,1]} \frac{\|B_p(b)\|_2}{w(b)} > c \right),$$

wobei  $B_p(\cdot)$  eine  $p$ -dimensionale Brownsche Brücke ist, deren Komponentenprozesse  $p$  ein-dimensionale unabhängige Brownsche Brücken sind. Wie in Aue et al. (2009a) kann nach dem Nachweis der bloßen Präsenz eines Strukturbruchs dessen Lage geschätzt werden durch

$$\hat{k}_r := \sup_{2 \leq k \leq m} Q_k.$$

Konkret wird bei der Anwendung auf echte Daten der folgende Algorithmus genutzt:

- (1) Wende die retrospektive Methode basierend auf dem Detektor in (3.12) auf die Teilstichprobe an, die aus den ersten  $m$  Datenpunkten besteht.
- (2a) Falls in diesem Teildatensatz ein Strukturbruch nachgewiesen werden kann, schätze die Lage des Bruchpunktes durch  $\hat{k}_r$ . Entferne dann alle Beobachtungen vor dem geschätzten Bruchpunkt und stocke die Teilstichprobe so weit mit nachfolgenden Datenpunkten auf, dass sie wieder  $m$  Werte umfasst. Kehre zurück zu Schritt (1). Falls es nicht ausreichend viele verbleibende Beobachtungen gibt, um die Teilstichprobe auf  $m$  Datenpunkte zu ergänzen, beende das Verfahren.
- (2b) Wenn in der Teilstichprobe kein Strukturbruch entdeckt wird, verwende diese Teilstichprobe als historischen Zeitraum und passe an diese Daten ein skalares BEKK-Modell an. Verwende die geschätzten Parameter, um Modellresiduen für den gesamten Überwachungszeitraum zu bestimmen und verwende die vorgestellte sequenzielle Methode, um den Vektor der Varianzen dieser Residuen zu überwachen.
- (3a) Kann mit der sequenziellen Methode ein Strukturbruch nachgewiesen werden, schätze die Lage des Bruchpunktes durch  $\hat{k}$ . Entferne alle Beobachtungen vor dem geschätzten Bruchpunkt und stocke die Teilstichprobe so weit mit nachfolgenden Datenpunkten auf, dass sie wieder  $m$  Werte umfasst. Kehre zurück zu Schritt (1). Falls es nicht ausreichend viele verbleibende Beobachtungen gibt, um die Teilstichprobe auf  $m$  Datenpunkte zu ergänzen, beende das Verfahren.
- (3b) Wenn mit Hilfe der sequenziellen Methode kein Strukturbruch im Überwachungszeitraum nachgewiesen werden kann, beende das Verfahren.

Allerdings führt die wiederholte Anwendung von retrospektiven Tests auf nur teilweise geänderte Daten augenscheinlich zu einer erhöhten Wahrscheinlichkeit für das Begehen eines Fehlers erster Art. Die Tatsache, dass der Anteil der ausgetauschten Datenpunkte von der Lage des Bruchpunkts in der historischen Stichprobe abhängt, gestaltet die Anwendung von Methoden für multiple Testprobleme schwierig. Deshalb wird dieses Problem an dieser Stelle vernachlässigt, um im Gegenzug die wichtige Eigenschaft eines konstanten Varianzvektors im historischen Zeitraum gewährleisten zu können. Wird  $\gamma = \{0, 0.25\}$  gewählt und  $B$  als die Anzahl der jeweils verbleibenden Datenpunkte dividiert durch  $m$ , liefert das sequenzielle Verfahren die Stoppzeiten  $\tau_m$  und die geschätzten Bruchpunkte  $\hat{k}$  beziehungsweise  $\hat{k}_r$ , die in Tabelle 3.2 angegeben sind.

$\gamma = 0$				$\gamma = 0.25$			
$\tau_m$	$\hat{k}, \hat{k}_r$	$\tau_m$	$\hat{k}, \hat{k}_r$	$\tau_m$	$\hat{k}, \hat{k}_r$	$\tau_m$	$\hat{k}, \hat{k}_r$
<b><math>m = 500</math></b>							
24.03.1983	12.05.1981		13.06.2002	24.06.1982	13.05.1981		13.06.2002
22.10.1984	13.08.1984		15.04.2003	22.10.1984	13.08.1984		15.04.2003
	16.09.1985		24.11.2003		16.09.1985		24.11.2003
12.03.1992	22.07.1988	22.05.2007	14.06.2006	10.02.1992	22.07.1988	15.06.2006	03.03.2006
07.03.1991	24.10.1990	19.09.2008	10.09.2008	09.10.1990	12.09.1990	18.03.2008	14.03.2008
08.11.1996	12.02.1996		02.04.2009	23.09.1997	11.06.1996	08.01.2013	31.08.2011
24.12.1999	15.03.1999	23.07.2012	05.12.2011	30.09.1998	13.07.1998	30.05.2014	18.11.2013
15.10.2001	07.08.2001	20.11.2014	17.09.2014	26.02.2003	07.08.2001		
<b><math>m = 1.000</math></b>							
12.04.1985	13.08.1984	25.11.2008	05.09.2008	30.10.1984	13.08.1984	22.12.2003	11.04.2003
26.03.1992	17.01.1991		19.05.2009	24.12.1991	17.10.1989		19.05.2004
27.01.1997	11.06.1996		08.08.2011		20.03.1991	10.10.2008	05.09.2008
30.06.2004	16.04.2003		19.11.2012	04.12.1996	11.06.1996		19.05.2009
	19.05.2004				13.07.1998		08.08.2011

TABELLE 3.2: Stoppzeitpunkte und geschätzte Bruchpunkte für die log-Renditen der Aktien von *Allianz*, *Bayer*, *Deutsche Bank*, *RWE* und *Siemens*.  
Blau: im historischen Datensatz entdeckte Bruchpunkte.

Gemeinsam mit den log-Renditen der betrachteten Kurswerte zeigt Abbildung 3.13 die entdeckten Bruchpunkte unter der Verwendung von  $\gamma = 0$  und  $m = 1.000$ . Alle Zeitreihen werden von dem Verfahren effektiv in Phasen mit höherer und niedrigerer Volatilität unterteilt.

Schätzzeitraum	Allianz	Bayer	Deutsche Bank	RWE	Siemens
02.01.1979 - 12.08.1984	0.0100	0.0103	0.0096	0.0084	0.0089
13.08.1984 - 16.01.1991	0.0214	0.0154	0.0167	0.0167	0.0156
17.01.1991 - 10.06.1996	0.0124	0.0116	0.0105	0.0102	0.0097
11.06.1996 - 12.07.1998	0.0193	0.0187	0.0173	0.0179	0.0173
13.07.1998 - 15.04.2003	0.0276	0.0255	0.0261	0.0218	0.0289
16.04.2003 - 18.05.2004	0.0218	0.0199	0.0171	0.0155	0.0170
19.05.2004 - 04.09.2008	0.0150	0.0148	0.0146	0.0130	0.0160
05.09.2008 - 18.05.2009	0.0498	0.0311	0.0627	0.0325	0.0431
19.05.2009 - 07.08.2011	0.0160	0.0157	0.0213	0.0133	0.0172
08.08.2011 - 31.12.2014	0.0174	0.0166	0.0236	0.0197	0.0137

TABELLE 3.3: Geschätzte Volatilitäten der log-Renditen zwischen aufeinanderfolgenden Bruchpunkten für  $m = 1.000$  basierend auf den Schätzwerten in Tabelle 3.2.

Basierend auf den Ergebnissen in Tabelle 3.2 können nun die Varianzen der log-Renditen zwischen jeweils zwei aufeinanderfolgenden Bruchpunkten geschätzt werden. Die resultierenden Standardabweichungen sind in Tabelle 3.3 angegeben und zum besseren Vergleich in Abbildung 3.14 illustriert.

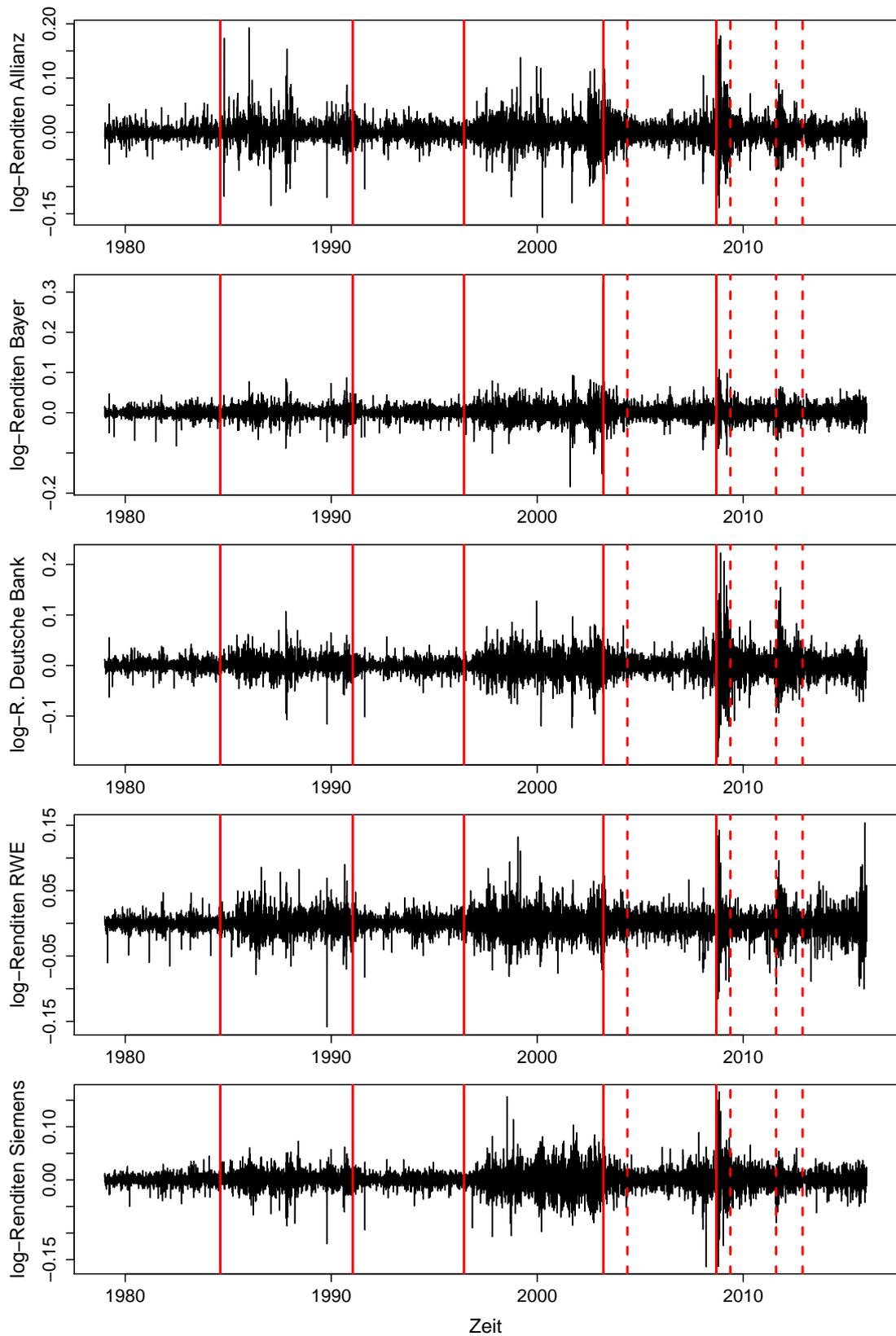


ABBILDUNG 3.13: log-Renditen mit den für  $\gamma = 0$  und  $m = 1.000$  ermittelten Bruchpunkten. — Bruchpunkt im Überwachungszeitraum; - - - Bruchpunkt im historischen Zeitraum.

Die Ergebnisse können mit auffälligen Ereignissen der jüngeren Geschichte in Zusammenhang gebracht werden. Die späten Achtziger Jahre wurden stark durch den Börsenkrach und die Tschernobyl-Katastrophe beeinflusst. Letztere ist hier von Bedeutung, da das betrachtete Portfolio unter anderem die Aktie des Energieerzeugers *RWE* beinhaltet, der seit den Siebzigern auch auf Kernenergie gesetzt hat. Mit dem Ende der Neunziger Jahre stiegen die Volatilitäten aller betrachteten Aktien im Zuge der Finanzkrise in Südostasien und der Währungskrise in Russland deutlich an. Dieser Trend wurde zur Jahrtausendwende nochmals verstärkt durch das Platzen der *dotcom*-Blase und den Beginn des Irakkrieges. Die darauf folgenden Jahre sinkender Volatilitäten wurden unterbrochen von der *Lehman Brothers*-Pleite, gefolgt von der globalen Finanzkrise sowie der Schuldenkrise, die große Teile Europas erfasste. Außerdem wurde wiederum der Kurs der *RWE*-Aktie stark durch die Nuklearkatastrophe in Fukushima im Jahr 2011 beeinflusst.

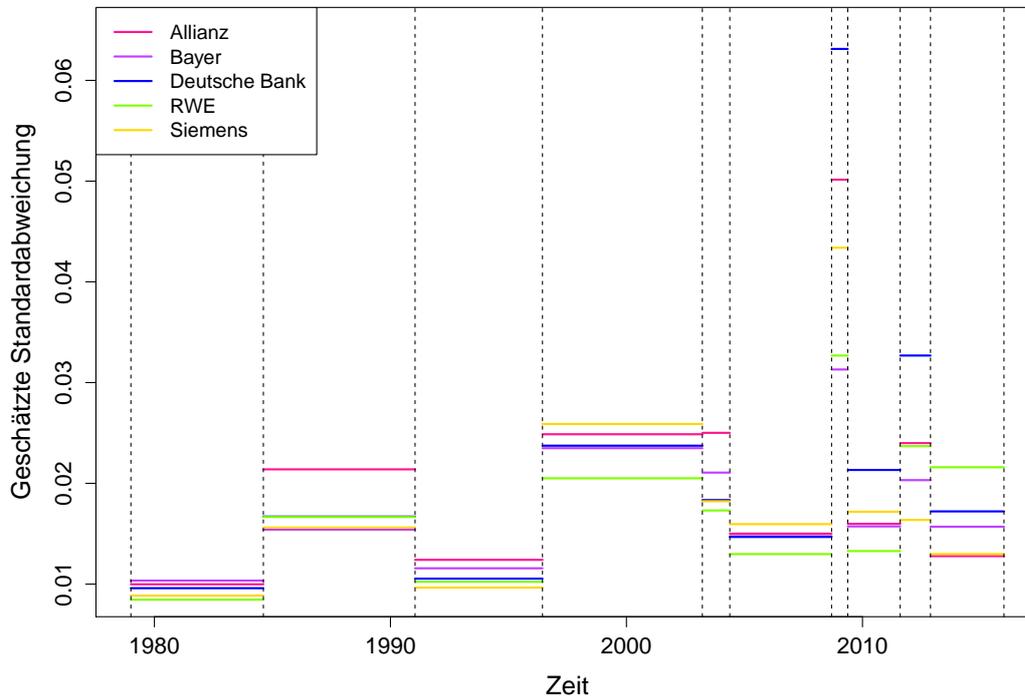


ABBILDUNG 3.14: Geschätzte Volatilitäten der log-Renditen zwischen aufeinanderfolgenden Bruchpunkten für  $m = 1.000$  basierend auf den Schätzwerten in Tabelle 3.2.

## Kapitel 4

# Das DCC-Modell nach Engle

Im Gegensatz zum im vorherigen Kapitel präsentierten nichtparametrischen Verfahren zur Überwachung der Varianzstruktur wird in diesem Kapitel eine parametrische Methode eingeführt, die die Überwachung derjenigen Parameter erlaubt, die neben der Varianz auch die Korrelationsstruktur steuern. Dazu muss zunächst das zugrundeliegende Modell genauer betrachtet werden.

### 4.1 Die Einführung des DCC-Modells

Sei  $\{y_t, t \in \mathbb{Z}\}$  eine Folge von  $p$ -dimensionalen Zufallsvektoren  $y_t = (y_{1t}, \dots, y_{pt})'$ , die einem multivariaten GARCH-Modell folgen. Wie in (2.2) in Abschnitt 2.2 erläutert, gilt somit:

$$y_t = H_t^{\frac{1}{2}} \epsilon_t \quad \text{mit} \quad H_t = \text{Cov}(y_t | \mathcal{F}_{t-1}).$$

Unter allen möglichen Spezifikationsformen für die bedingte Kovarianzmatrix  $H_t$  wird hier diejenige gewählt, die durch die dynamische Modellierung der bedingten Korrelation (*dynamic conditional correlation*) charakterisiert ist, das DCC-Modell. Dieses Modell, das eine flexible und praxisnahe multivariate Modellierung von Zeitreihen mit Varianzclustern erlaubt, wurde von Engle (2002) als Erweiterung des Modells mit konstanter bedingter Korrelation, oder kurz CCC-Modell, siehe Bollerslev (1990), vorgeschlagen. Es sei zu erwähnen, dass im gleichen Jahr, in dem Engle seine Erweiterung veröffentlichte, eine ähnliche, ebenfalls die dynamische Charakterisierung der bedingten Korrelation umfassende, Erweiterung des CCC-Modells durch Tse und Tsui (2002) vorgeschlagen wurde. Tse und Tsui (2002) bestimmen die für die bedingte Kovarianzmatrix benötigte bedingte Korrelationsmatrix direkt mit Hilfe einer GARCH-artigen Struktur auf Basis vergangener Korrelationen. Im Gegensatz dazu modelliert Engle (2002) eine Hilfsmatrix  $Q_t$  rekursiv und transformiert diese anschließend zu einer Korrelationsmatrix, das heißt zu einer Matrix, deren Hauptdiagonale mit Einsen besetzt ist und die unter gegebenen Annahmen positiv definit ist. Der Vorteil des Vorgehens von Engle (2002) wird durch die Betrachtung von „standardisierten“ Vektoren  $z_t$  deutlich, die weiter unten eingeführt werden und deren Kovarianz- und Korrelationsmatrix durch  $R_t$  gegeben ist. Engle (2002) verwendet diese „standardisierten“ Vektoren selbst, während Tse und Tsui (2002) auf deren empirische Korrelationsmatrix zurückgreifen, was den Einfluss einzelner Beobachtungen abmildert. Somit spiegelt der Ansatz von Engle (2002) stärker das Prinzip des klassischen univariaten GARCH-Modells von Bollerslev (1986) wider und wird deshalb im Folgenden vorgezogen.

Im Einzelnen wird dabei für die bedingte Kovarianzmatrix unterstellt, dass gilt:

$$H_t = D_t R_t D_t, \quad (4.1)$$

wobei  $D_t = \text{diag} \{h_{1t}^{1/2}, \dots, h_{pt}^{1/2}\}$  eine Diagonalmatrix ist, deren Hauptdiagonale mit den bedingten Standardabweichungen der einzelnen Komponenten von  $y_t$  besetzt ist. Dabei können die bedingten Varianzen  $h_{it}$ ,  $i = 1, \dots, p$ , beispielsweise als einfache univariate GARCH(1,1)-Prozesse modelliert werden:

$$h_{it} = \omega_i + \alpha_i y_{i,t-1}^2 + \beta_i h_{i,t-1}, \quad i = 1, \dots, p. \quad (4.2)$$

Es wäre auch möglich, die bedingten Varianzen unter Verwendung weiter zurückliegender Beobachtungen und bedingter Varianzen zu modellieren oder durch andere univariate GARCH-Modelle, die beispielsweise Asymmetrien oder andere Eigenschaften, die bei echten Aktienrenditen häufig beobachtet werden, stärker berücksichtigen. Die einfache Modellierung in (4.2) wird aus Gründen der Übersichtlichkeit vorgezogen, kann aber beliebig erweitert werden. Außerdem erlaubt bereits die Verwendung eines univariaten GARCH(1,1)-Modells eine zufriedenstellende Modellierung vieler Renditezeitreihen, wie Bollerslev (1986) argumentiert. Weiter bezeichnet  $R_t$  in (4.1) die bedingte Korrelationsmatrix von  $y_t$ , das heißt

$$R_t = \text{Cor}(y_t | \mathcal{F}_{t-1}).$$

Diese kann zerlegt werden in

$$R_t = Q_t^* Q_t Q_t^*, \quad (4.3)$$

wobei  $Q_t^*$  eine Normierungsmatrix ist, die gegeben ist durch:

$$Q_t^* := \text{diag} \{ [Q_t]_{11}^{-1/2}, \dots, [Q_t]_{pp}^{-1/2} \}. \quad (4.4)$$

Außerdem ist  $Q_t$  eine  $(p \times p)$ -Matrix ist, die sich rekursiv darstellen lässt als

$$Q_t = (1 - \alpha - \beta) \bar{Q} + \alpha z_{t-1} z_{t-1}' + \beta Q_{t-1} \quad (4.5)$$

mit  $z_t := D_t^{-1} y_t$ . Die Vektoren  $z_t$  werden im Folgenden als „standardisiert“ bezeichnet, da sie nur hinsichtlich der bedingten Varianz standardisiert sind, nicht aber bezüglich der bedingten Kovarianzen oder Korrelationen. Für diese Vektoren gilt:

$$\text{E}(z_t | \mathcal{F}_{t-1}) = \mathbf{0}_p \quad \text{und} \quad \text{E}(z_t z_t' | \mathcal{F}_{t-1}) = R_t.$$

Bei  $\alpha$  und  $\beta$  handelt es sich um nichtnegative skalare Parameter mit  $\alpha + \beta < 1$ . Außerdem ist  $\bar{Q} = [\bar{q}_{ij}]_{i,j=1,\dots,p}$  sowohl die unbedingte Kovarianz- und Korrelationsmatrix der „standardisierten“ Beobachtungen  $z_t$  als auch die unbedingte Korrelationsmatrix der  $y_t$ . Da diese Eigenschaft impliziert, dass die Hauptdiagonale mit Einsen besetzt ist und weil  $\bar{Q}$  per Definition eine symmetrische Matrix ist, sind die unbekannt Einträge alle im Vektor

$$\psi = \text{vecl}(\bar{Q}) = (\bar{q}_{21}, \dots, \bar{q}_{p,p-1})' \in [-1, 1]^{p^-}$$

enthalten. Zusammenfassend ergibt sich der Vektor der Parameter des DCC-Modells als

$$\theta = (\omega_1, \alpha_1, \beta_1, \dots, \omega_p, \alpha_p, \beta_p, \alpha, \beta, \bar{q}_{21}, \dots, \bar{q}_{p,p-1})',$$

was eine Gesamtzahl von  $d := 3p + 2 + p^- = \frac{1}{2}(p+1)(p+4)$  unbekannt Parameter im Modell ergibt.

Der Parametervektor lässt sich dabei zerlegen in einen Vektor der Varianzparameter

$$\boldsymbol{\theta}_1 = (\omega_1, \alpha_1, \beta_1, \dots, \omega_p, \alpha_p, \beta_p)' = (\phi'_1, \dots, \phi'_p)'$$

mit  $\phi_j := (\omega_j, \alpha_j, \beta_j)'$  den Parametern der  $j$ -ten Varianzgleichung für  $j = 1, \dots, p$ , sowie einen Vektor, der alle Parameter, die die Korrelationsstruktur charakterisieren, umfasst:

$$\boldsymbol{\theta}_2 = (\alpha, \beta, \bar{q}_{21}, \dots, \bar{q}_{p,p-1})' = (\alpha, \beta, \boldsymbol{\psi}')',$$

so dass insgesamt gilt:  $\boldsymbol{\theta} = (\boldsymbol{\theta}'_1, \boldsymbol{\theta}'_2)'$ .

### Bedingungen für die positive Definitheit der bedingten Kovarianzmatrix

Für viele theoretische Überlegungen, aber auch für Anwendungen, ist die positive Definitheit der bedingten Kovarianzmatrix unerlässlich. Proposition 2 in Engle und Sheppard (2001) liefert hinreichende Bedingungen unter denen diese Eigenschaft für alle  $t \in \mathbb{Z}$  erfüllt ist. Insbesondere ist die Matrix  $H_t$ , die durch (4.1)-(4.5) spezifiziert wird, für alle  $t \in \mathbb{Z}$  positiv definit, wenn die folgenden Annahmen erfüllt sind:

#### Annahme 4.1

1.  $\omega_i > 0, \forall i \in \{1, \dots, p\}$ .
2.  $\alpha_i > 0$  und  $\beta_i > 0$  mit  $\alpha_i + \beta_i < 1 \forall i \in \{1, \dots, p\}$ ,  
siehe auch Nelson und Cao (1992).
3.  $h_{i0} > 0, \forall i \in \{1, \dots, p\}$ .
4.  $\alpha > 0$  und  $\beta > 0$  mit  $\alpha + \beta < 1$ .
5. Es existiert eine Konstante  $\delta_1 > 0$  mit  $\lambda_{\min}(\bar{Q}) > \delta_1$ , wobei  $\lambda_{\min}(\cdot)$  der kleinste Eigenwert einer quadratischen Matrix ist.
6. Es existiert eine Konstante  $\delta_2 > 0$  mit  $\lambda_{\max}(Q_t) \stackrel{f.s.}{<} \delta_2, \forall t \in \mathbb{Z}$ , wobei  $\lambda_{\max}(\cdot)$  der größte Eigenwert einer quadratischen Matrix ist.

Dabei wird die Annahme 4.1(6.) nicht zwingend benötigt, um die positive Definitheit von  $R_t$  für alle  $t \in \mathbb{Z}$  nachzuweisen, aus der die positive Definitheit von  $H_t$  folgt. Diese Eigenschaft ergibt sich aus der positiven Definitheit von  $Q_t$  und Proposition 1 in Engle und Sheppard (2001), die auch als Proposition F.1 in Anhang F zu finden ist. Die Tatsache, dass  $Q_t$  für alle  $t \in \mathbb{Z}$  positiv definit ist, resultiert aus der Zerlegung

$$Q_t = \frac{1 - \alpha - \beta}{1 - \beta} \bar{Q} + \alpha \sum_{n=0}^{\infty} \beta^n z_{t-n-1} z'_{t-n-1}$$

und 6.70.(a) in Seber (2008):

$$\begin{aligned} \lambda_{\min}(Q_t) &\stackrel{f.s.}{\geq} \lambda_{\min}\left(\frac{1 - \alpha - \beta}{1 - \beta} \bar{Q}\right) + \lambda_{\min}\left(\alpha \sum_{n=0}^{\infty} \beta^n z_{t-n-1} z'_{t-n-1}\right) \\ &\stackrel{f.s.}{\geq} \frac{1 - \alpha - \beta}{1 - \beta} \lambda_{\min}(\bar{Q}) > \frac{1 - \alpha - \beta}{1 - \beta} \delta_1 > 0, \end{aligned}$$

da es sich bei  $\alpha \sum_{n=0}^{\infty} \beta^n z_{t-n-1} z'_{t-n-1}$  fast sicher um eine positiv definite Matrix handelt.

Allerdings kann Annahme 4.1(6.) genutzt werden, um feste Schranken für die positiven Eigenwerte von  $R_t$  zu bestimmen. Diese Schranken sind gültig für alle  $t \in \mathbb{Z}$  und erleichtern einige der theoretischen Überlegungen und Beweise deutlich.

Mit 6.17.(a) in Seber (2008) gilt:

$$\max_{1 \leq i \leq p} [Q_t]_{ii} < \sum_{i=1}^p [Q_t]_{ii} = \sum_{i=1}^p \lambda_i(Q_t) \leq p \lambda_{\max}(Q_t) < p \delta_2, \quad (4.6)$$

wobei  $\lambda_{\max}(Q_t) = \lambda_1(Q_t), \dots, \lambda_p(Q_t) = \lambda_{\min}(Q_t)$  die der Größe nach geordneten Eigenwerte von  $Q_t$  sind. Des Weiteren gilt für die positiven Einträge auf der Hauptdiagonalen:

$$\min_{1 \leq i \leq p} [Q_t]_{ii} \stackrel{f.s.}{>} \frac{1 - \alpha - \beta}{1 - \beta}. \quad (4.7)$$

Somit können mit Hilfe von (4.6), (4.7) und 6.95. in Seber (2008) die folgenden Schranken für die Eigenwerte von  $R_t$  hergeleitet werden:

$$\lambda_{\min}(R_t) \stackrel{f.s.}{\geq} \lambda_{\min}(Q_t^*)^2 \lambda_{\min}(Q_t) \stackrel{f.s.}{\geq} \left[ \max_{1 \leq i \leq p} [Q_t]_{ii} \right]^{-1} \frac{1 - \alpha - \beta}{1 - \beta} \delta_1 \stackrel{f.s.}{\geq} \frac{1 - \alpha - \beta}{1 - \beta} \frac{\delta_1}{p \delta_2} \quad (4.8)$$

$$\lambda_{\max}(R_t) \stackrel{f.s.}{\leq} \lambda_{\max}(Q_t^*)^2 \lambda_{\max}(Q_t) \stackrel{f.s.}{\leq} \left[ \min_{1 \leq i \leq p} [Q_t]_{ii} \right]^{-1} \delta_2 \stackrel{f.s.}{\leq} \frac{1 - \beta}{1 - \alpha - \beta} \delta_2. \quad (4.9)$$

Die Aussagen (4.8) und (4.9) werden in den folgenden Kapiteln und vor allem im Beweisteil in Anhang E ausgiebig angewandt.

## 4.2 Die Schätzung der Modellparameter

Ist eine beobachtete DCC-Zeitreihe  $y_1, \dots, y_T$  gegeben, wird der Quasi-Maximum-Likelihood-schätzer (QMLS) für den wahren Parametervektor  $\theta$  mit  $\hat{\theta}_T$  bezeichnet. Der Schätzer ergibt sich, indem die Gaußsche Quasi-Loglikelihoodfunktion (QLL-Funktion) maximiert wird. Diese ist gegeben durch:

$$L_T(\theta) = L_T(\theta|y_1, \dots, y_T) = \sum_{t=1}^T l_t(\theta|y_1, \dots, y_T), \quad (4.10)$$

mit den einzelnen QLL-Beiträgen

$$l_t(\theta|y_1, \dots, y_T) = -\frac{1}{2} \left( p \cdot \log 2\pi + \log \det(H_t) + y_t' H_t^{-1} y_t \right).$$

Sowohl der Einschnitt- als auch der Zweischritt-QMLS sind konsistent. Der Zweischritt-QMLS wurde von Engle und Sheppard (2001) vorgeschlagen, die auch Bedingungen für seine Konsistenz liefern. Bei diesem Schätzverfahren, dass zur Reduktion der Rechenzeit entwickelt wurde, werden in aufeinanderfolgenden Schritten zunächst die Varianz- und anschließend die Korrelationsparameter basierend auf den Varianzparameterschätzern aus dem ersten Schritt geschätzt. In weiterführenden Simulationen, die nicht Teil dieser Arbeit sind, zeigte sich, dass der Zweischrittschätzer für die im folgenden Kapitel in Betracht gezogenen Parameterkombinationen deutlich schlechtere Ergebnisse liefert als der rechenintensivere Einschnitt-QMLS, bei dem alle  $d$  Parameter in einem einzigen Schritt geschätzt werden. Deshalb wird letzterer Schätzer für die folgenden Simulationen und Anwendungen verwendet. Der Einschnitt-QMLS, der sich aus einer Zeitreihe der Länge  $T$  ergibt, wird

dabei mit  $\hat{\theta}_T$  bezeichnet. Außerdem ist  $L(\cdot)$  die Grenzfunktion der QLL-Funktion  $L_T(\cdot)$  für  $T \rightarrow \infty$  für die gelten muss:

**Annahme 4.2** Die Grenzfunktion  $L(\cdot)$  besitzt in  $\boldsymbol{\theta}$  ein striktes Maximum.

Die Tatsache, dass auch der Einschnitt-QMLS stark konsistent ist, folgt aus Annahme 4.2 und daraus, dass  $L_T(\cdot)$  auf dem eingeschränkten Parameterraum  $U$ , der im folgenden Kapitel ausführlich vorgestellt wird, gleichmäßig gegen  $L(\cdot)$  konvergiert. Die gleichmäßige Konvergenz kann unter Verwendung der Ergebnisse in Abschnitt E.2 analog zum Beweis der gleichmäßigen Konvergenz von  $D_m(\cdot)$  gegen  $D(\cdot)$  auf S.123 in dieser Arbeit gezeigt werden.

Um zusätzlich noch asymptotische Normalität nachweisen zu können, werden die folgenden Regularitätsbedingungen benötigt:

**Annahme 4.3** Sei  $\mathbf{A}(\boldsymbol{\theta}) := \lim_{T \rightarrow \infty} \mathbb{E} \left[ \frac{1}{T} \frac{\partial^2 L_T(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta=\boldsymbol{\theta}} \right]$  eine endliche und nichtsinguläre Matrix und  $\{\theta_T^*, T \in \mathbb{N}\}$  eine beliebige Folge mit  $\theta_T^* \xrightarrow{p} \boldsymbol{\theta}$  für  $T \rightarrow \infty$ . Dann gilt für  $T \rightarrow \infty$ :

$$\frac{1}{T} \frac{\partial^2 L_T(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta=\theta_T^*} \xrightarrow{p} \mathbf{A}(\boldsymbol{\theta}).$$

**Annahme 4.4** Sei  $\mathbf{B}(\boldsymbol{\theta}) := \lim_{T \rightarrow \infty} \mathbb{E} \left[ \frac{1}{T} \frac{\partial L_T(\theta)}{\partial \theta} \Big|_{\theta=\boldsymbol{\theta}} \frac{\partial L_T(\theta)}{\partial \theta'} \Big|_{\theta=\boldsymbol{\theta}} \right]$  eine endliche und nichtsinguläre Matrix. Dann gilt für  $T \rightarrow \infty$ :

$$\frac{1}{\sqrt{T}} \frac{\partial L_T(\theta)}{\partial \theta} \xrightarrow{d} \mathcal{N}_d(\mathbf{0}_d, \mathbf{B}(\boldsymbol{\theta})).$$

Da, wie in Abschnitt D.4 gezeigt wird, die Hessematrix von  $l_0(u)$  existiert und außerdem offensichtlich in allen Parametern in einer offenen und konvexen Umgebung von  $\boldsymbol{\theta}$  stetig ist, gilt mit Theorem 4.1.3 in Amemiya (1985), das sich als Theorem F.1 in Anhang F findet:

$$\sqrt{T} (\hat{\theta}_T - \boldsymbol{\theta}) \xrightarrow{d} \mathcal{N}_d(\mathbf{0}_d, \mathbf{A}(\boldsymbol{\theta})^{-1} \mathbf{B}(\boldsymbol{\theta}) \mathbf{A}(\boldsymbol{\theta})^{-1}). \quad (4.11)$$

### 4.3 Die Existenz einer stationären Lösung

Stationaritätseigenschaften sind von zentraler Bedeutung für eine verlässliche Parameterschätzung sowie eine sichere Strukturbruchererkennung. Hinreichende Bedingungen für die Existenz einer streng stationären Lösung des DCC-Modells werden in Fermanian und Malongo (2017) vorgestellt. Dazu werden zunächst die Modellgleichungen (2.2) und (4.1)-(4.5) in eine autoregressive Form umgeschrieben:

$$X_t = T_t X_{t-1} + \zeta_t \quad (4.12)$$

mit

$$X_t := \left( h_{1t}, \dots, h_{pt}, y_{1t}^2, \dots, y_{pt}^2, \text{vecl}(Q_t)', \text{vecl}(z_t z_t')' \right)', \quad (4.13)$$

$$\zeta_t := \left( \omega_1, \dots, \omega_p, \omega_1 z_{1t}^2, \dots, \omega_p z_{pt}^2, (1 - \alpha - \beta) \text{vecl}(\bar{Q})', \text{vecl}(z_t z_t')' \right)'$$

$$\text{und } T_t = \begin{bmatrix} \beta_1 & & 0 & \alpha_1 & & 0 & & & & \\ & \ddots & & & \ddots & & & & & \\ 0 & & \beta_p & 0 & & \alpha_p & & \mathbf{0}_{p \times p^-} & & \mathbf{0}_{p \times p^-} \\ \hline \beta_1 z_{1t}^2 & & 0 & \alpha_1 z_{1t}^2 & & 0 & & & & \\ & \ddots & & & \ddots & & & & & \\ 0 & & \beta_p z_{pt}^2 & 0 & & \alpha_p z_{pt}^2 & & \mathbf{0}_{p \times p^-} & & \mathbf{0}_{p \times p^-} \\ \hline & & \mathbf{0}_{p^- \times p} & & \mathbf{0}_{p^- \times p} & & \beta \mathbb{I}_{p^-} & & \alpha \mathbb{I}_{p^-} & \\ \hline & & \mathbf{0}_{p^- \times p} & & \mathbf{0}_{p^- \times p} & & \mathbf{0}_{p^- \times p^-} & & \mathbf{0}_{p^- \times p^-} & \end{bmatrix}.$$

Sei analog zu Fermanian und Malongo (2017)  $\mathcal{I} := \{\mathcal{I}_t, t \in \mathbb{Z}\}$  die an den Prozess  $\{X_t, t \in \mathbb{Z}\}$  in (4.13) adaptierte Filtration. Somit ist  $\mathcal{I}_t = \sigma(X_t, X_{t-1}, \dots)$  die Informationsmenge zum Zeitpunkt  $t$ . Außerdem sei im Folgenden der Prozess  $\{\eta_t, t \in \mathbb{Z}\}$  mit  $\eta_t := R_t^{-1/2} z_t$  betrachtet. Konstruktionsbedingt gilt für diese Vektoren:

$$\mathbb{E}(\eta_t | \mathcal{I}_{t-1}) = \mathbf{0}_p \quad \text{und} \quad \mathbb{E}(\eta_t \eta_t' | \mathcal{I}_{t-1}) = \mathbb{I}_p.$$

Bedingt auf die Informationsmenge zum Zeitpunkt  $t$  verhalten sich die  $\eta_t$  somit wie Innovationen und werden im Folgenden auch als solche bezeichnet.

Analog zu den Annahmen A0 und U0 in Fermanian und Malongo (2017) muss die folgende Annahme erfüllt sein:

**Annahme 4.5** Der Prozess  $\{\eta_t, t \in \mathbb{Z}\}$  besitzt die Markoveigenschaft bezüglich der Filtration  $\mathcal{I}$ . Dies bedeutet insbesondere, dass gilt:

$$\mathbb{E}(\eta_t | \mathcal{I}_{t-1}) = \mathbb{E}(\eta_t | X_{t-1}) \quad \forall t \in \mathbb{Z}.$$

Außerdem ist der Prozess  $\{\eta_t, t \in \mathbb{Z}\}$  stationär und ergodisch.

Mit Hilfe von Annahme 4.5 lässt sich (4.12) als nichtlineare Markovkette darstellen. Details zu diesem Vorgehen und zu Markovprozessen im Allgemeinen finden sich aus Gründen der Übersichtlichkeit in Anhang C. Für das hier betrachtete Modell, das durch die Modellgleichungen (4.1)-(4.5) gegeben ist, reduzieren sich die Annahmen U1-U3 in Fermanian und Malongo (2017) auf:

**Annahme 4.6**  $\max_{1 \leq i \leq p} \alpha_i + \max_{1 \leq i \leq p} \beta_i < 1$  und  $|\beta| < 1$ .

Fermanian und Malongo (2017) zeigen, dass die Bedingungen für Tweedies Theorem, das sich als Theorem F.3 in Anhang F findet, erfüllt sind, wenn Annahme 4.6 gilt. Somit existiert ein  $\sigma$ -endliches und endliches Maß  $\mu$  für die Übergangswahrscheinlichkeiten  $\mathbb{P}$  des Markovprozesses, das die Definition eines invarianten Maßes  $\pi = \frac{\mu}{\mu(E)}$  erlaubt. Da dieses invariante Maß auch die Übergangswahrscheinlichkeiten der nichtlinearen Markovkette steuert, folgt aus der Invarianz von  $\pi$  direkt die strikte Stationarität des Prozesses derjenigen  $X_t$ , die durch  $\mathbb{P}$  gesteuert werden. Dies impliziert die strikte Stationarität desjenigen Prozesses  $\{X_t, t \in \mathbb{Z}\}$  der die autoregressive Gleichung (4.12) löst. Somit ist auch der Prozess  $\{y_t, t \in \mathbb{Z}\}$ , der die Modellgleichungen (2.2) und (4.1)-(4.5) löst, strikt stationär.

### Die stationäre Lösung ist eindeutig

Fermanian und Malongo (2017) zeigen außerdem, dass unter Zuhilfenahme einiger weiterer Annahmen die strikt stationäre Lösung des DCC-Modells ebenfalls eindeutig ist. Diese Annahmen reduzieren sich im hier behandelten konkreten Fall auf:

**Annahme 4.7**

$$\mathbb{E} \left[ \ln \left( \beta^2 + \alpha^2 \frac{4(2p+1)\sqrt{p}}{\sqrt{C_\lambda} C_q} \|\eta_t\|_2^2 \right) \right] < 0$$

$$\text{mit Konstanten } C_\lambda = \frac{\lambda_{\min} \left( (1 - \alpha - \beta) \bar{Q} \right)}{1 - \beta^2} \quad \text{und} \quad C_q = \frac{(1 - \alpha - \beta) \min_{1 \leq i \leq p} \bar{q}_{ii}}{1 - \beta^2}.$$

Allgemein existiert nun unter den Annahmen 4.5 und 4.6 eine strikt stationäre Lösung der Modellgleichungen (2.2) und (4.1)-(4.5). Ist zusätzlich noch Annahme 4.7 erfüllt, so ist diese stationäre Lösung eindeutig.



## Kapitel 5

# Detektion von Parameteränderungen im DCC-Modell

### 5.1 Das Überwachungsszenario

Sei  $\theta_t \in \mathbb{R}^d$  der Parametervektor des DCC-Modells zum Zeitpunkt  $t$ . Wie bereits in Kapitel 3 wird angenommen, dass es einen historischen Zeitraum von Beobachtungen gibt, der nicht von einem Strukturbruch betroffen ist. Das heißt, es gilt:

**Annahme 5.1**  $\theta_1 = \dots = \theta_m$  für ein  $m \in \mathbb{N}$ .

Das Ziel ist nun, die Nullhypothese eines konstanten Parametervektors gegen die Alternative einer Änderung des Parametervektors zu einem unbekanntem Zeitpunkt im Überwachungszeitraum zu testen. Das Testproblem lässt sich somit formulieren als

$$H_0 : \theta_t = \boldsymbol{\theta}, \quad t = 1, \dots, m, m+1, \dots$$

gegen

$$H_1 : \theta_t = \begin{cases} \boldsymbol{\theta}, & t = 1, \dots, m, m+1, \dots, m+k^*-1 \\ \boldsymbol{\theta}^*, & t = m+k^*, m+k^*+1, \dots \end{cases}.$$

Dabei ist

$$\boldsymbol{\theta} = (\phi'_1, \dots, \phi'_p, \alpha, \beta, \psi')'$$

der Parametervektor vor dem Strukturbruch und

$$\boldsymbol{\theta}^* = (\phi^{*'}_1, \dots, \phi^{*'}_p, \alpha^*, \beta^*, \psi^{*'})'$$

der Parametervektor nach dem Strukturbruch. Es gilt:

$$\phi_i^* = (\omega_i^*, \alpha_i^*, \beta_i^*)', \quad i = 1, \dots, p, \quad \text{und} \quad \psi^* = \text{vecl}(\bar{Q}^*) \quad \text{mit} \quad \bar{Q}^* = [\bar{q}_{ij}^*]_{i,j=1,\dots,p}.$$

Somit tritt die Parameteränderung zum  $k^*$ -ten Zeitpunkt des Überwachungsbereiches ein. Dies ist gleichzeitig der  $(m+k^*)$ -te Punkt der gesamten Zeitreihe.

## Der Parameterraum

Als Grundlage für die Überlegungen und Verfahren der nächsten Kapitel wird analog zum Vorgehen von Berkes et al. (2003) zunächst der Raum aller möglichen Parameterkombinationen  $\Theta = \mathbb{R}^d$  auf eine durch konkrete Konstanten beschränkte Teilmenge eingegrenzt.

### Definition 5.1 Beschränkter Parameterraum

Es seien Konstanten  $0 < \underline{u} < \bar{u}$ ,  $0 < \rho < 1$  und  $0 < \delta_1 < \delta_2$  gegeben, so dass  $U \subset \Theta$  definiert ist als

$$\begin{aligned} U := \{u : \max \{t_1, \dots, t_p, b, a + b, |q_1|, \dots, |q_{p-}| \} &\leq \rho, \\ \lambda_{\min} (F_{\bar{Q}}(u)) &> \delta_1, \lambda_{\max} (F_{Q_t}(u)) < \delta_2 \quad \forall t \in \mathbb{Z} \\ \text{und } \underline{u} &< \min \{x_1, s_1, t_1, \dots, x_p, s_p, t_p, a, b\} \\ &\leq \max \{x_1, s_1, t_1, \dots, x_p, s_p, t_p, a, b\} \leq \bar{u}\}. \end{aligned}$$

Im Folgenden sei ein generisches Element aus dem eingeschränkten Parameterraum  $u \in U$  bezeichnet als

$$u = (x_1, s_1, t_1, \dots, x_p, s_p, t_p, a, b, q_1, \dots, q_{p-})'.$$

Die von diesem generischen Element abhängigen matrixwertigen Funktionen  $F_{\bar{Q}}(u)$  und  $F_{Q_t}(u)$  sind dabei wie folgt definiert:

**Definition 5.2** Seien für  $i \in \{1, \dots, p\}$ ,  $t \in \mathbb{Z}$  und  $u \in U$  die folgenden Terme definiert:

$$(i) \ w_{it}(u) := \frac{x_i}{1 - t_i} + s_i \sum_{k=1}^{\infty} t_i^{k-1} y_{i,t-k}^2 = \frac{x_i}{1 - t_i} + s_i \sum_{k=0}^{\infty} t_i^k y_{i,t-k-1}^2.$$

$$(ii) \ F_{D_t}(u) := \text{diag} \{w_{1t}(u)^{1/2}, \dots, w_{pt}(u)^{1/2}\}.$$

$$(iii) \ F_{R_t}(u) := F_{Q_t^*}(u) F_{Q_t}(u) F_{Q_t^*}(u)$$

$$\text{mit } F_{Q_t^*}(u) := \text{diag} \{[F_{Q_t}(u)]_{11}^{-1/2}, \dots, [F_{Q_t}(u)]_{pp}^{-1/2}\}.$$

$$(iv) \ F_{Q_t}(u) := \frac{1 - a - b}{1 - b} F_{\bar{Q}}(u) + a \sum_{k=1}^{\infty} b^{k-1} z_{t-k}(u) z'_{t-k}(u).$$

$$(v) \ F_{\bar{Q}}(u) := \begin{pmatrix} 1 & q_1 & q_2 & \dots & q_{p-1} \\ q_1 & 1 & q_p & \dots & \vdots \\ \vdots & q_p & \ddots & & q_{p-1} \\ \vdots & \vdots & & 1 & q_{p-} \\ q_{p-1} & \dots & q_{p-1} & q_{p-} & 1 \end{pmatrix}.$$

Dabei gilt  $z_t(u) = F_{D_t}(u)^{-1} y_t$ . Dies impliziert  $z_{it}(u) = \frac{y_{it}}{\sqrt{w_{it}(u)}}$  für alle  $i = 1, \dots, p$ .

Um eine konsistente Parameterschätzung überhaupt erst zu ermöglichen, wird angenommen, dass für den tatsächlichen Parametervektor  $\boldsymbol{\theta}$  gilt:

**Annahme 5.2**  $\boldsymbol{\theta} \in U$ .

Wird nun für die in 5.2 definierten Terme das generische Element  $u$  als der wahre Parametervektor  $\boldsymbol{\theta}$  gewählt, so gilt außerdem:

$$\begin{aligned} w_{it}(\boldsymbol{\theta}) &= h_{it}, \text{ für } i \in \{1, \dots, p\}, & F_{D_t}(\boldsymbol{\theta}) &= D_t, & F_{R_t}(\boldsymbol{\theta}) &= R_t, \\ F_{Q_t}(\boldsymbol{\theta}) &= Q_t, & F_{Q_t^*}(\boldsymbol{\theta}) &= Q_t^*, & F_{\bar{Q}}(\boldsymbol{\theta}) &= \bar{Q}, & \mathbf{z}_t(\boldsymbol{\theta}) &= z_t. \end{aligned}$$

Die QLL-Funktion in (4.10) lässt sich dann in Abhängigkeit von einem beliebigen, aber konkreten Element  $u \in U$  angeben als  $L_T(u) = \sum_{t=1}^T l_t(u)$  mit

$$l_t(u) = -\frac{1}{2} \left( p \cdot \log 2\pi + \sum_{i=1}^p \log w_{it}(u) + \log \det (F_{R_t}(u)) + \mathbf{z}'_t(u) F_{R_t}(u)^{-1} \mathbf{z}_t(u) \right). \quad (5.1)$$

Die Beiträge zur QLL-Funktion in (5.1) werden durch die Terme in Definition 5.2 gesteuert. Bis auf die unbedingte Kovarianz- beziehungsweise Korrelationsmatrix  $F_{\bar{Q}}(u)$  hängen diese von einer unendlichen Vergangenheit an Beobachtungen ab. Im Gegensatz dazu stehen in der Praxis nur endlich viele Werte zur Verfügung. Dies schafft die Notwendigkeit, Versionen der Terme aus Definition 5.2 mit endlicher Vergangenheit zu formulieren.

**Definition 5.3** Seien für  $i \in \{1, \dots, p\}$ ,  $t \in \mathbb{Z}$  und  $u \in U$  die folgenden Terme definiert:

$$(i) \hat{w}_{it}(u) := \frac{x_i}{1-t_i} + s_i \sum_{k=1}^{t-1} t_i^{k-1} y_{i,t-k}^2 = \frac{x_i}{1-t_i} + s_i \sum_{k=0}^{t-2} t_i^k y_{i,t-k-1}^2.$$

$$(ii) \hat{F}_{D_t}(u) := \text{diag} \left\{ \hat{w}_{1t}(u)^{1/2}, \dots, \hat{w}_{pt}(u)^{1/2} \right\}.$$

$$(iii) \hat{F}_{R_t}(u) := \hat{F}_{Q_t^*}(u) \hat{F}_{Q_t}(u) \hat{F}_{Q_t^*}(u)$$

$$\text{mit } \hat{F}_{Q_t^*}(u) := \text{diag} \left\{ \left[ \hat{F}_{Q_t}(u) \right]_{11}^{-\frac{1}{2}}, \dots, \left[ \hat{F}_{Q_t}(u) \right]_{pp}^{-\frac{1}{2}} \right\}.$$

$$(iv) \hat{F}_{Q_t}(u) := \frac{1-a-b}{1-b} F_{\bar{Q}}(u) + a \sum_{k=1}^{t-1} b^{k-1} \hat{\mathbf{z}}_{t-k}(u) \hat{\mathbf{z}}'_{t-k}(u) \text{ mit } \hat{\mathbf{z}}_t(u) = \hat{F}_{D_t}(u)^{-1} y_t.$$

$$(v) \hat{L}_T(u) = \sum_{t=1}^T \hat{l}_t(u), \text{ mit}$$

$$\hat{l}_t(u) = -\frac{1}{2} \left( p \cdot \log(2\pi) + \sum_{i=1}^p \log \hat{w}_{it}(u) + \log \det \left( \hat{F}_{R_t}(u) \right) + \hat{\mathbf{z}}'_t(u) \hat{F}_{R_t}(u)^{-1} \hat{\mathbf{z}}_t(u) \right).$$

## 5.2 Das Überwachungsverfahren

Das folgende Überwachungsverfahren stellt eine multivariate Erweiterung des Vorgehens von Berkes et al. (2004) dar. Es bezeichne  $l'_t(u)$  den Gradienten des Beitrags der Beobachtung zum Zeitpunkt  $t$  zur QLL-Funktion mit unendlicher Vergangenheit in (5.1) beziehungsweise  $\hat{l}'_t(u)$  den Gradienten im Fall einer endlichen Vergangenheit aus Definition 5.3. Die Gestalt der Gradienten ist etwas unübersichtlich und findet sich gemeinsam mit einer detaillierten Herleitung in Abschnitt D.3 in Anhang D. In allen Fällen, in denen die Transponierte eines Vektors mit der Ableitung verwechselt werden kann, wird die Transponierte durch  $^T$  markiert.

Wie in Kapitel 3 werden die asymptotischen Aspekte unter der Annahme, dass die Länge des historischen Zeitraums gegen  $\infty$  strebt, betrachtet. Deshalb werden auch in diesem Kapitel Charakteristika wie die Länge des Überwachungszeitraums und der Zeitpunkt des Strukturbruchs in Abhängigkeit von  $m$  und mit Hilfe der Parameter  $B$  und  $\lambda^*$  definiert.

### Detektor und Abbruchkriterium

Zunächst sei für jedes  $u \in U$  die Informationsmatrix

$$D(u) := \mathbb{E} \left[ l'_t(u) l'_t(u)^T \right]$$

gegeben, für die gilt:

**Annahme 5.3**  $D := D(\theta)$  ist eine endliche und nichtsinguläre Matrix.

Für eine endliche und beobachtete Stichprobe kann die unbekannte Informationsmatrix geschätzt werden durch:

$$\widehat{D}_m = \frac{1}{m} \sum_{t=1}^m \hat{l}'_t(\hat{\theta}_m) \hat{l}'_t(\hat{\theta}_m)^T.$$

Dann ergibt sich der Detektor

$$V_k = \sum_{t=m+1}^{m+k} \widehat{D}_m^{-\frac{1}{2}} \hat{l}'_t(\hat{\theta}_m)$$

mit dem Abbruchkriterium beziehungsweise der Stoppzeit

$$\tau_m = \min \left\{ k \leq mB : |V_k| > m^{\frac{1}{2}} \left( 1 + \frac{k}{m} \right) \mathfrak{b} \left( \frac{k}{m} \right) \right\}, \quad (5.2)$$

wobei  $\mathfrak{b}(\cdot)$  eine Grenzfunktion ist. Wie in Abschnitt 2.1 erläutert, signalisiert  $\tau_m < \infty$ , dass ein Strukturbruch im Vektor der Modellparameter entdeckt wurde. Analog zum Vorgehen in Berkes et al. (2004) werden moderate Annahmen an die Grenzfunktion  $\mathfrak{b}(\cdot)$  aufgestellt:

**Annahme 5.4**  $\mathfrak{b}(\cdot)$  ist stetig auf dem Intervall  $(0, \infty)$  und  $\inf_{0 < t < \infty} \mathfrak{b}(t) > 0$ .

Die Grenzfunktion  $\mathfrak{b}(\cdot)$  oder zumindest die variablen Teile der Funktion müssen nun so gewählt werden, dass gilt:

$$\lim_{m \rightarrow \infty} \mathbb{P}_{H_0} \{ \tau_m < \infty \} = \tilde{\alpha} \quad \text{und} \quad \lim_{m \rightarrow \infty} \mathbb{P}_{H_1} \{ \tau_m < \infty \} = 1.$$

Berkes et al. (2004) wählen die Grenzfunktion  $b(\cdot)$  als Konstante, deren Wert mit Hilfe von Monte-Carlo-Simulationen bestimmt werden kann. Testsimulationen, die vorab durchgeführt wurden, zeigen allerdings, dass für diese Wahl der Grenzfunktion das empirische Niveau des multivariaten Verfahrens sehr stark von der Wahl des Parameters  $B$  abhängt, wie es auch für den univariaten Fall in Berkes et al. (2004) beobachtet werden kann. Um diesen Effekt zumindest abzumildern, wird die Länge des Überwachungszeitraums bei der Bestimmung des Abbruchkriteriums (5.2) berücksichtigt. Außerdem wird eine potenziell kurvenförmige Grenzfunktion der strikt linearen Funktion vorgezogen, die sich ergibt, wenn  $b(\cdot)$  als konstant gewählt wird. Die Grenzfunktion aus (3.5) ergibt sich durch die Wahl

$$b(x) = \max \left\{ \left( \frac{x}{1+x} \right)^\gamma, \varepsilon \right\}$$

mit dem Justierungsparameter  $\gamma \in [0, 1/2)$  und der Konstanten  $\varepsilon > 0$ , die die Teilbarkeit durch die Werte der Grenzfunktion sicherstellt. Die Eigenschaften dieser Grenzfunktion und ihr Ursprung wurden bereits in Abschnitt 3.2 erläutert. Außerdem wird auch diese Grenzfunktion durch Multiplikation mit einer Konstanten  $c = c(\tilde{\alpha})$  skaliert, wobei diese Konstante so gewählt werden sollte, dass die Wahrscheinlichkeit für eine Überschreitung der Grenzfunktion durch die Detektorwerte innerhalb der Überwachungsphase gerade dem Signifikanzniveau  $\tilde{\alpha}$  entspricht.

### Asymptotische Eigenschaften des Detektors unter der Nullhypothese

Um Aussagen über das asymptotische Verhalten des Detektors unter der Nullhypothese oder unter der Alternativen treffen zu können, müssen die Zufallsvektoren  $y_t$  achte Momente und Kreuzmomente besitzen. Da diese Eigenschaft zu Teilen durch das Verhalten der zugrundeliegenden Innovationsvektoren gesteuert wird, wird zunächst angenommen, dass gilt:

**Annahme 5.5** Die Innovationsvektoren  $\epsilon_t$  besitzen absolute achte Momente und Kreuzmomente.

Die folgende Proposition liefert Bedingungen, die es erlauben, die Eigenschaft existenter achter Momente von den Innovationsvektoren auf die beobachtbaren Vektoren  $y_t$  zu übertragen. Wird der Vektor der bedingten Varianzen analog zu He und Teräsvirta (2004) als  $h_t = (h_{1t}, \dots, h_{pt})'$  bezeichnet und der Vektor der quadrierten Prozesswerte als  $y_t^{(2)} = (y_{1t}^2, \dots, y_{pt}^2)'$ , so ergibt sich eine Vektordarstellung der Varianzgleichungen in (4.2) als

$$h_t = \omega + \mathbf{A}y_{t-1}^{(2)} + \mathbf{B}h_{t-1} \quad (5.3)$$

mit  $\omega := (\omega_1, \dots, \omega_p)'$ ,  $\mathbf{A} := \text{diag} \{\alpha_1, \dots, \alpha_p\}$  und  $\mathbf{B} := \text{diag} \{\beta_1, \dots, \beta_p\}$ .

Sei außerdem  $Z_t := \text{diag} \{z_{1t}, \dots, z_{pt}\}$ . Daraus ergibt sich  $y_t^{(2)} = Z_t^2 h_t$ , was eine autoregressive Darstellung von (5.3) erlaubt:

$$h_t = \omega + \mathbf{C}_{t-1} h_{t-1}$$

mit  $\mathbf{C}_t := \mathbf{A}Z_t^2 + \mathbf{B}$ . Dabei ist die unbedingte Erwartung von  $\mathbf{C}_t$  unter den Stationaritätsbedingungen 5.5 und 5.6 zeitinvariant.

Es sei  $\otimes_k$  das  $k$ -fache Kroneckerprodukt identischer Matrizen und angenommen, dass gilt:

**Annahme 5.6**  $\lambda_{\max}\left(\mathbb{E}\left[\otimes_j \mathbf{C}_0\right]\right) < \infty, \quad \forall j \in \{1, \dots, 4\}.$

Dann gilt:

**Proposition 5.1** Sei  $\{y_t\}$  eine Folge von Zufallsvektoren, welche die Modellgleichungen (2.2) und (4.1)-(4.5) erfüllen. Dann gilt:

Unter den Annahmen 4.1-4.7, 5.5 und 5.6, besitzen die Zufallsvektoren  $y_t$  achte Momente und Kreuzmomente.

Alternativ zu Proposition 5.1 kann zum Nachweis der Existenz der Momente auch Theorem 1 in Fermanian und Malongo (2017) verwendet werden. Anstatt der Annahmen 5.5 und 5.6 müssen dann neben Annahme 4.5 auch E1 und E2 aus Fermanian und Malongo (2017) erfüllt sein. Während Annahme E1 in etwa unseren Annahmen entspricht, liefert Annahme E2 einige technische Annahmen an die strukturelle Gestalt der bedingten Innovationen. Diese Annahme wurde hier durch sehr grobes Abschätzen umgangen.

Basierend auf Proposition 5.1 kann nun eine weitere Proposition aufgestellt werden, die die Basis für die folgenden Theoreme bildet:

**Proposition 5.2** Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6, gilt

$$\widehat{\mathbf{D}}_m \xrightarrow{f.s.} \mathbf{D}.$$

**Theorem 5.1** Unter  $H_0$  und den Annahmen 4.1-4.7 und 5.1-5.6, gilt

$$\begin{aligned} \lim_{m \rightarrow \infty} \mathbb{P}_{H_0} \{\tau_m < \infty\} &= \lim_{m \rightarrow \infty} \mathbb{P}_{H_0} \left( \sup_{1 < k \leq mB} \frac{|V_k|}{m^{1/2} \left(1 + \frac{k}{m}\right) \mathbf{b}\left(\frac{k}{m}\right)} \geq c \right) \\ &= \mathbb{P}_{H_0} \left( \sup_{t \in (0, B]} \frac{|\mathcal{G}_d(t)|}{(1+t) \mathbf{b}(t)} \geq c \right) \end{aligned}$$

mit  $\{\mathcal{G}_d(b), b \in [0, B]\}$  aus Definition 3.4.

Analog zum Vorgehen in Galeano und Wied (2014) und Berkes et al. (2004) sowie in Abschnitt 3.2 gilt, dass  $|\mathcal{G}_d(t)|$  die gleiche Verteilung besitzt wie  $\max_{1 \leq i \leq d} \left| (1+t)W_i\left(\frac{t}{1+t}\right) \right|$  für alle  $t \in \mathbb{Z}$ . Dabei ist  $d$  die Gesamtzahl aller Parameter im DCC-Modell. Die Substitution von  $t = \frac{\tilde{s}}{1-\tilde{s}}$  liefert dann

$$\begin{aligned} \sup_{t \in (0, B]} \frac{|\mathcal{G}_d(t)|}{(1+t) \mathbf{b}(t)} &\stackrel{d}{=} \sup_{t \in (0, B]} \max_{1 \leq i \leq d} \frac{|W_i\left(\frac{t}{1+t}\right)|}{\mathbf{b}(t)} \stackrel{d}{=} \sup_{t \in (0, B]} \max_{1 \leq i \leq d} \frac{|W_i\left(\frac{t}{1+t}\right)|}{\max\left\{\left(\frac{t}{1+t}\right)^\gamma, \varepsilon\right\}} \\ &\stackrel{d}{=} \sup_{\tilde{s} \in \left(0, \frac{B}{1+B}\right]} \max_{1 \leq i \leq d} \frac{|W_i(\tilde{s})|}{\max\{\tilde{s}^\gamma, \varepsilon\}}. \end{aligned}$$

Des Weiteren ergibt sich durch die Wahl von  $\tilde{s} = \frac{s(1+B)}{B}$

$$\begin{aligned} \sup_{t \in (0, B]} \frac{|\mathcal{G}_d(t)|}{(1+t)\mathbf{b}(t)} &\stackrel{d}{=} \sup_{s \in (0, 1]} \max_{1 \leq i \leq d} \frac{\left| W_i \left( \frac{sB}{1+B} \right) \right|}{\max \left\{ \left( \frac{sB}{1+B} \right)^\gamma, \varepsilon \right\}} \\ &\stackrel{d}{=} \left( \frac{B}{1+B} \right)^{\frac{1}{2}-\gamma} \sup_{s \in (0, 1]} \max_{1 \leq i \leq d} \frac{|W_i(s)|}{\max \left\{ s^\gamma, \varepsilon \left( \frac{1+B}{B} \right)^\gamma \right\}}. \end{aligned}$$

Somit können Monte-Carlo-Simulationen genutzt werden, um kritische Werte  $c = c(\tilde{\alpha})$  in Abhängigkeit vom Signifikanzniveau  $\tilde{\alpha}$  zu bestimmen, indem die folgende Gleichheit ausgenutzt wird:

$$\begin{aligned} &\mathbb{P}_{H_0} \left( \left( \frac{B}{1+B} \right)^{1/2-\gamma} \sup_{s \in (0, 1]} \max_{1 \leq i \leq d} \frac{|W_i(s)|}{\max \left\{ s^\gamma, \varepsilon \left( \frac{1+B}{B} \right)^\gamma \right\}} \geq c(\tilde{\alpha}) \right) \\ &= 1 - \left[ \mathbb{P}_{H_0} \left( \sup_{s \in (0, 1]} \frac{|W_1(s)|}{\max \left\{ s^\gamma, \varepsilon \left( \frac{1+B}{B} \right)^\gamma \right\}} < \left( \frac{1+B}{B} \right)^{1/2-\gamma} c(\tilde{\alpha}) \right) \right]^d = \tilde{\alpha}. \end{aligned}$$

Alternativ kann auch die folgende Aussage genutzt werden, die sich durch einfache Umstellung der Terme ergibt:

$$\mathbb{P}_{H_0} \left( \sup_{s \in (0, 1]} \frac{|W_1(s)|}{\max \left\{ s^\gamma, \varepsilon \left( \frac{1+B}{B} \right)^\gamma \right\}} < \left( \frac{1+B}{B} \right)^{1/2-\gamma} c(\tilde{\alpha}) \right) = (1 - \tilde{\alpha})^{\frac{1}{d}}. \quad (5.4)$$

Für verschiedene Parametereinstellungen können die simulierten kritischen Werte Tabelle 5.1 entnommen werden. Dazu wurden 100.000 Brownsche Bewegungen auf einem Netz von 100.000 äquidistanten Punkten simuliert.

$B$	$\tilde{\alpha} = 0.01$			$\tilde{\alpha} = 0.05$			$\tilde{\alpha} = 0.1$		
	$\gamma = 0$	$0.2$	$0.4$	$\gamma = 0$	$0.2$	$0.4$	$\gamma = 0$	$0.2$	$0.4$
<b><math>p = 3</math></b>									
<b>0.1</b>	1.063	1.777	2.999	0.932	1.562	2.691	0.874	1.447	2.515
<b>0.2</b>	1.464	2.081	3.221	1.263	1.873	2.840	1.184	1.748	2.673
<b>0.3</b>	1.722	2.380	3.297	1.497	2.032	2.959	1.398	1.928	2.770
<b>0.4</b>	1.931	2.508	3.410	1.675	2.195	2.999	1.550	2.049	2.851
<b>0.5</b>	2.068	2.632	3.422	1.791	2.305	3.086	1.677	2.126	2.873
<b><math>p = 5</math></b>									
<b>0.1</b>	1.127	1.863	3.146	0.997	1.642	2.831	0.937	1.554	2.652
<b>0.2</b>	1.535	2.221	3.332	1.354	1.967	3.009	1.275	1.850	2.856
<b>0.3</b>	1.828	2.437	3.459	1.595	2.170	3.097	1.471	2.061	2.934
<b>0.4</b>	1.984	2.599	3.498	1.762	2.331	3.180	1.657	2.175	3.004
<b>0.5</b>	2.171	2.744	3.546	1.918	2.438	3.243	1.786	2.273	3.019

TABELLE 5.1: Simulierte kritische Werte  $c(\tilde{\alpha})$ .

### Kritische Werte unter Verwendung eines parametrischen Bootstrapverfahrens

Weiterführende Simulationen, die nicht Teil dieser Arbeit sind, haben gezeigt, dass kritische Werte, die mit Hilfe von Monte-Carlo-Simulationen basierend auf (5.4) bestimmt werden, untragbar hohe Niveauperzerrungen liefern. In der Konsequenz neigen die Detektorwerte dazu, die skalierte Grenzfunktion auch unter der Nullhypothese zügig nach dem Beginn des Überwachungszeitraums zu überschreiten. Diese Eigenschaft konnte auch unter Verwendung längerer historischer Zeiträume nicht auf ein praktikables Maß reduziert werden. Es werden historische Zeiträume mit Längen von deutlich mehr als 2.000 Datenpunkten benötigt. Da allerdings gemäß Annahme 5.1 sichergestellt werden muss, dass der historische Zeitraum frei von Strukturbrüchen ist und derartig lange Zeiträume ohne Parameteränderungen in der Praxis kaum gefunden werden können, müssen die kritischen Werte auf eine andere Art und Weise bestimmt werden. Das Ziel ist es hierbei, die Niveauperzerrungen so stark wie möglich zu reduzieren. Dazu bietet sich ein parametrisches und simulationsbasiertes Bootstrapverfahren an.

Es wird wie zuvor mit  $\hat{\theta}_m$  derjenige Schätzer für den Parametervektor verwendet, der sich aus der historischen Stichprobe ergibt. Es ist anzunehmen, dass der zugrundeliegende DCC-Prozess ein ähnliches Verhalten aufweist wie der durch die aus den historischen Daten geschätzten Parameter gesteuerte Prozess. Zumindest sollte diese Annahme zutreffen, wenn der historische Zeitraum ausreichend lang ist und die QML-Parameterschätzung somit ein zuverlässiges Ergebnis liefert. Deshalb werden  $b_{BT} = 199$  Realisationen von DCC-Prozessen generiert, deren stochastische Struktur durch den Parametervektor  $\hat{\theta}_m$  bestimmt wird. Diese Zeitreihen werden als  $Y^{*(i)} := \{y_1^{*(i)}, \dots, y_{m(B+1)}^{*(i)}\}$  bezeichnet, für  $i \in \{1, \dots, b_{BT}\}$ . Die intuitivste Herangehensweise wäre es nun, die Detektorwerte

$$|V_k^{*(i)}| = \left| \sum_{t=m+1}^{m+k} [\hat{D}_m^{*(i)}]^{-1/2} \hat{l}_t^{*(i)'} (\hat{\theta}_m) \right|$$

für jede Stichprobe  $Y^{*(i)}$  zu bestimmen. Dabei sind  $\hat{l}_t^{*(i)'} (\hat{\theta}_m)$  die Beiträge zur QLL-Funktion und

$$\hat{D}_m^{*(i)} = \frac{1}{m} \sum_{t=1}^m \hat{l}_t^{*(i)'} (\hat{\theta}_m) [\hat{l}_t^{*(i)'} (\hat{\theta}_m)]^T$$

der Schätzer für  $D_p$ , der sich aus den ersten  $m$  Beobachtungen der Zeitreihe  $Y^{*(i)}$  ergibt. Da die exakte Verteilung des Detektors allerdings nur von untergeordnetem Interesse ist und das Hauptaugenmerk auf der Bestimmung von kritischen Werten liegt, wird die Matrix  $D_p$  durch die Einheitsmatrix ersetzt. Durch dieses Vorgehen wird die zusätzliche Unsicherheit vermieden, die mit der Matrixschätzung einhergeht. Weiterführende Simulationen, die nicht Teil dieser Arbeit sind, haben gezeigt, dass dieser Ansatz im Vergleich zur Verwendung eines Schätzers für  $D_p$  deutlich geringere Niveauperzerrungen liefert. Es sei der unter Verwendung der Einheitsmatrix resultierende Detektor als

$$|\tilde{V}_k^{*(i)}| = \left| \sum_{t=m+1}^{m+k} \hat{l}_t^{*(i)'} (\hat{\theta}_m) \right|$$

bezeichnet und das Maximum der skalierten Detektorwerte, die sich aus  $Y^{*(i)}$  ergeben, als

$$T^{*(i)} := \max_{1 \leq k \leq [mB]} \frac{|V_k^{*(i)}|}{m^{1/2} \left(1 + \frac{k}{m}\right) \mathbf{b} \left(\frac{k}{m}\right)}, \quad \text{für } i \in \{1, \dots, b_{BT}\}.$$

Dann kann das  $(1 - \tilde{\alpha})$ -Quantil von  $\{T^{*(1)}, \dots, T^{*(b_{BT})}\}$  als kritischer Wert genutzt werden. In Anwendungen werden jeweils  $b_{BT} = 199$  Bootstrapwiederholungen durchgeführt. Diese auf den ersten Blick etwas überraschend anmutende Wahl der Iterationszahl entspricht einer allgemeinen Konvention im Bereich der Bootstrapmethoden, die in Davidson und MacKinnon (2000) näher erläutert wird. Diese beziehen sich zunächst auf Dufour und Kiviet (1998), die zeigen, dass für einen exakten Test zum Signifikanzniveau  $\tilde{\alpha}$  der Term  $\tilde{\alpha}(N + 1)$  ganzzahlig sein muss. Dieses Ergebnis übertragen Davidson und MacKinnon (2000) auf Bootstrapverfahren und argumentieren auch, dass somit für die gängigsten Wahlen des Signifikanzniveaus die Anzahl der Bootstrapwiederholungen um Eins reduziert werden muss, um einen exakten Test zu erhalten.

Obwohl eine detaillierte Analyse der resultierenden kritischen Werte und ihrer Eigenschaften jenseits des Rahmens dieser Arbeit liegt, zeigen diese ein zufriedenstellendes Verhalten in Simulationen und Anwendungen.

### Asymptotische Eigenschaften des Detektors unter Alternativen

Abschließend wird die asymptotische Verteilung des Detektors unter der Alternativen einer Parameteränderung von  $\theta$  zu  $\theta^*$  im  $k^*$ -ten Punkt des Überwachungszeitraums betrachtet. Dazu müssen zunächst Annahmen an die Parameter nach dem Strukturbruch aufgestellt werden. Diese entsprechen den Annahmen 5.2, 4.6, 4.7 und 5.6 für die Parameter vor dem Bruchpunkt.

**Annahme 5.7**  $\theta^* \in U$ .

**Annahme 5.8**  $\max_{1 \leq i \leq p} \alpha_i^* + \max_{1 \leq i \leq p} \beta_i^* < 1$  und  $|\beta^*| < 1$

**Annahme 5.9**

$$\mathbb{E} \left[ \ln \left( [\beta^*]^2 + [\alpha^*]^2 \frac{4(2p+1)\sqrt{p}}{\sqrt{C_\lambda^* C_q^*}} \|\eta_t\|_2^2 \right) \right] < 0$$

mit Konstanten  $C_\lambda^* = \frac{\lambda_{\min}((1 - \alpha^* - \beta^*)\bar{Q}^*)}{1 - [\beta^*]^2}$  und  $C_q^* = \frac{(1 - \alpha^* - \beta^*) \min_{1 \leq i \leq p} \bar{q}_{ii}^*}{1 - [\beta^*]^2}$ .

**Annahme 5.10**  $\lambda_{\max}(\mathbb{E}[\otimes_j C_0^*]) < \infty$ ,  $\forall j \in \{1, \dots, 4\}$ ,

wobei  $C_t^* := A^* Z_t^2 + B^*$  mit  $A^* := \text{diag}\{\alpha_1^*, \dots, \alpha_p^*\}$  und  $B^* := \text{diag}\{\beta_1^*, \dots, \beta_p^*\}$ .

**Theorem 5.2** *Unter der Alternativen eines Strukturbruchs und den Annahmen 4.1-4.7 und 5.1-5.10, gilt*

$$\lim_{m \rightarrow \infty} \mathbb{P}_{H_1} \{\tau_m < \infty\} = 1.$$

### Schätzung der Lage des Bruchpunkts

Sobald das Überwachungsverfahren die Präsenz eines Strukturbruchs angezeigt hat, kann die Lage des Bruchpunkts beispielsweise mit der folgenden Methode geschätzt werden:

$$\hat{k} := \arg \max_{1 \leq k \leq \tau_m - 1} \frac{k}{\sqrt{\tau_m}} \left| \frac{1}{\tau_m - 1} \sum_{t=m+1}^{m+\tau_m-1} \hat{\imath}_t(\hat{\theta}_m) - \frac{1}{k} \sum_{t=m+1}^{m+k} \hat{\imath}_t(\hat{\theta}_m) \right|. \quad (5.5)$$

Dieser Schätzer ist ähnlich aufgebaut wie die in Wied et al. (2012a), in Wied und Galeano (2013) oder in Abschnitt 3.2 vorgeschlagenen Schätzer. Zwar liegt auch hier eine genauere Analyse der Eigenschaften des Schätzers außerhalb des Rahmens dieser Arbeit. Da allerdings Schätzer dieser Art ein zufriedenstellendes Verhalten in Simulationen und Anwendungen zeigen, wird der Schätzer aus (5.5) in den nächsten Abschnitten zur Ermittlung der Lage eines nachgewiesenen Strukturbruchs verwendet.

## 5.3 Simulationen

Dieser Abschnitt befasst sich mit den Eigenschaften des vorgeschlagenen Verfahrens in endlichen Stichproben. Dazu müssen sowohl unter der Nullhypothese als auch unter der Alternativen verschiedene Parameter festgelegt werden. Zunächst wird die Länge des historischen Zeitraums als  $m \in \{500, 1.000, 2.000\}$  gewählt. Es wird angenommen, dass die Länge des Überwachungszeitraums deutlich kürzer ist als die des historischen Zeitraums und  $B \in \{0.1, 0.2, \dots, 0.5\}$  verwendet. Die Dimension der Zufallsvektoren beziehungsweise ihrer Realisationen beträgt  $p \in \{3, 5\}$ , während der Justierungsparameter  $\gamma$  Werte aus  $\{0, 0.2, 0.4\}$  annimmt. Diese Parametereinstellungen sollen jeweils die Fähigkeit des Verfahrens stärken, frühere oder spätere Strukturbrüche in einer vorliegenden Zeitreihe zu erkennen. Für jeden der betrachteten Fälle werden 1.000 DCC-Zeitreihen simuliert, auf die das Überwachungsverfahren angewandt werden kann. Diese niedrige Anzahl an Simulationsdurchläufen ist der hohen Rechenintensität geschuldet, die zu einem Großteil auf die Bestimmung der QML-Schätzwerte zurückgeführt werden kann. Für alle Testverfahren beläuft sich das Signifikanzniveau auf  $\tilde{\alpha} = 0.05$ .

### 5.3.1 Simulationen unter der Nullhypothese

Zunächst werden die Niveaueigenschaften des Überwachungsverfahrens unter der Nullhypothese, dass kein Strukturbruch im Parametervektor auftritt, untersucht. Für jede der Vektorkomponenten werden die Varianzparameter als  $\phi_i = (0.01, 0.05, 0.9)'$  beziehungsweise  $\phi_i = (0.01, 0.2, 0.7)'$ , für  $i \in \{1, \dots, p\}$ , gewählt. Dabei impliziert der zweite Fall einen stärkeren Einfluss einzelner Schocks auf die Volatilität folgender Beobachtungen. Die Korrelationsstruktur wird bestimmt durch  $(\alpha, \beta) = (0.05, 0.9)$ , während die konstante unbedingte Korrelationsmatrix  $\bar{Q}_p$  gegeben ist durch

$$\bar{Q}_3 = \begin{bmatrix} 1 & 0.5 & 0.1 \\ 0.5 & 1 & 0.5 \\ 0.1 & 0.5 & 1 \end{bmatrix} \quad \text{und} \quad \bar{Q}_5 = \begin{bmatrix} 1 & 0.5 & 0.3 & 0.2 & 0.1 \\ 0.5 & 1 & 0.5 & 0.3 & 0.2 \\ 0.3 & 0.5 & 1 & 0.5 & 0.3 \\ 0.2 & 0.3 & 0.5 & 1 & 0.5 \\ 0.1 & 0.2 & 0.3 & 0.5 & 1 \end{bmatrix}.$$

Die Ergebnisse der Simulationen sind in Tabelle 5.2 zu finden und belegen, dass das empirische Niveau mit  $B$  steigt. Dies ist plausibel, da größere Werte für diesen Parameter eine längere Überwachungsphase und somit ein zunehmendes Maß an Unsicherheit implizieren. Während hohe Werte für  $m$  und  $\gamma$  die Niveauverzerrung reduzieren, wächst die Wahrscheinlichkeit, einen Fehler erster Art zu begehen, mit der Dimension  $p$ . Insbesondere der Einfluss von Variationen der Länge des historischen Zeitraums und der Dimension verhalten sich so wie es im Voraus zu erwarten war. Des Weiteren ist das empirische Niveau deutlich höher, wenn die Varianzparameter als  $\phi_i = (0.01, 0.05, 0.9)'$  gewählt werden. Dieses Resultat ist wenig überraschend, wenn berücksichtigt wird, dass in diesem Fall die Summe  $\alpha_i + \beta_i$  näher an Eins ist und somit das Verhalten des Prozesses stärker dem eines Einheitswurzelprozesses ähnelt als im zweiten Szenario.

Die angegebene Richtung der beschriebenen Effekte kann nicht für jeden einzelnen der Fälle in Tabelle 5.2 beobachtet werden. Dies dürfte darauf zurückzuführen sein, dass nur eine relativ kleine Anzahl an Wiederholungen durchgeführt wurde.

		$\phi_i = (0.01, 0.05, 0.9)'$			$\phi_i = (0.01, 0.2, 0.7)'$			
$B$	$\gamma$	$m = 500$	$1.000$	$2.000$	$m = 500$	$1.000$	$2.000$	
$p = 3$	<b>0.1</b>	<b>0</b>	0.124	0.088	0.068	0.068	0.077	0.047
		<b>0.2</b>	0.133	0.084	0.074	0.067	0.070	0.052
		<b>0.4</b>	0.116	0.082	0.070	0.064	0.058	0.066
	<b>0.2</b>	<b>0</b>	0.150	0.101	0.094	0.069	0.057	0.088
		<b>0.2</b>	0.151	0.086	0.091	0.069	0.077	0.087
		<b>0.4</b>	0.118	0.080	0.083	0.065	0.060	0.060
	<b>0.3</b>	<b>0</b>	0.177	0.111	0.089	0.120	0.087	0.081
		<b>0.2</b>	0.151	0.094	0.073	0.095	0.075	0.083
		<b>0.4</b>	0.143	0.084	0.086	0.071	0.054	0.069
	<b>0.4</b>	<b>0</b>	0.213	0.113	0.118	0.106	0.109	0.104
		<b>0.2</b>	0.193	0.120	0.105	0.110	0.098	0.106
		<b>0.4</b>	0.147	0.094	0.072	0.077	0.066	0.063
<b>0.5</b>	<b>0</b>	0.197	0.118	0.141	0.129	0.110	0.122	
	<b>0.2</b>	0.179	0.135	0.097	0.100	0.109	0.115	
	<b>0.4</b>	0.166	0.111	0.090	0.086	0.073	0.073	
$p = 5$	<b>0.1</b>	<b>0</b>	0.139	0.093	0.079	0.080	0.085	0.062
		<b>0.2</b>	0.141	0.104	0.067	0.071	0.066	0.047
		<b>0.4</b>	0.117	0.093	0.072	0.070	0.048	0.060
	<b>0.2</b>	<b>0</b>	0.153	0.099	0.083	0.079	0.080	0.082
		<b>0.2</b>	0.161	0.112	0.085	0.088	0.081	0.068
		<b>0.4</b>	0.148	0.083	0.073	0.074	0.069	0.059
	<b>0.3</b>	<b>0</b>	0.181	0.109	0.087	0.102	0.118	0.116
		<b>0.2</b>	0.161	0.110	0.098	0.087	0.098	0.090
		<b>0.4</b>	0.148	0.087	0.085	0.074	0.064	0.071
	<b>0.4</b>	<b>0</b>	0.181	0.117	0.122	0.121	0.106	0.125
		<b>0.2</b>	0.199	0.109	0.111	0.102	0.108	0.101
		<b>0.4</b>	0.151	0.114	0.088	0.073	0.079	0.086
<b>0.5</b>	<b>0</b>	0.198	0.111	0.131	0.152	0.126	0.140	
	<b>0.2</b>	0.186	0.111	0.123	0.122	0.120	0.118	
	<b>0.4</b>	0.182	0.107	0.107	0.092	0.084	0.059	

TABELLE 5.2: Empirisches Niveau für verschiedene Parameterkombinationen bei einem Signifikanzniveau von  $\tilde{\alpha} = 0.05$ .

### 5.3.2 Simulationen unter verschiedenen Alternativszenarien

Im Folgenden wird die Fähigkeit des Überwachungsverfahrens, verschiedene Typen von Strukturbrüchen aufzuspüren, untersucht. Dazu werden zunächst Zeitreihen mit einem Strukturbruch in den Varianzparametern simuliert. Anschließend werden Zeitreihen, die eine Änderung in der unbedingten Korrelationsmatrix  $\bar{Q}$  aufweisen, betrachtet.

Da die Simulationsergebnisse unter der Nullhypothese eine starke Abhängigkeit des Niveaus beziehungsweise der Niveauverzerrung von der Länge des Überwachungszeitraums nahelegen, beschränkt sich dieser Abschnitt auf die Simulation von Alternativszenarien auf Zeitreihen, die einen Überwachungszeitraum der Länge  $0.2m$  aufweisen. Diese Wahl des Parameters  $B$  liefert für einen ausreichend langen historischen Zeitraum nur kleine Abweichungen des empirischen vom theoretischen Niveau, wie die Ergebnisse aus Tabelle 5.2 belegen. In Abhängigkeit von der Länge des historischen Zeitraums wird der Bruchpunkt  $k^*$  ähnlich dem Vorgehen in Kapitel 3 als  $k^* = \lceil mB\lambda^* \rceil$  mit  $\lambda^* \in \{0.05, 0.3, 0.5\}$  festgelegt. Diese Parameterwahl liefert Strukturbrüche, die entweder sehr früh oder erst später im Überwachungszeitraum auftreten.

### Änderungen in den Varianzparametern

In diesem Abschnitt werden zwei Szenarien von Änderungen in der Struktur der bedingten Varianzen betrachtet. Zunächst ändert sich für alle  $i \in \{1, \dots, p\}$  der Parametervektor  $\phi_i = (0.01, 0.05, 0.9)'$  in  $\phi_i^* = (0.005, 0.2, 0.7)'$ . Danach wird eine Änderung von  $\phi_i = (0.01, 0.2, 0.7)'$  in  $\phi_i^* = (0.05, 0.05, 0.9)'$  unterstellt. Diese beiden Situationen werden

$\gamma$	Güte	Stoppszeiten					Geschätzte Bruchpunkte				
		MW	SD	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$	MW	SD	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$
$\lambda^* = 0.05, m = 500 \Rightarrow k^* = 5$											
<b>0</b>	<b>0.973</b>	44.87	14.49	35	42	51	18.05	10.39	11	16	23
<b>0.2</b>	<b>0.971</b>	40.28	15.77	30	38	48	16.08	9.76	9	14	21
<b>0.4</b>	<b>0.971</b>	40.24	18.28	29	38	50	16.03	10.32	9	14	21
$\lambda^* = 0.05, m = 1.000 \Rightarrow k^* = 10$											
<b>0</b>	<b>1.000</b>	60.40	15.46	51	57	68	24.40	11.71	17	22	29
<b>0.2</b>	<b>0.999</b>	52.88	16.09	42	50	60	21.18	10.53	14	19	26
<b>0.4</b>	<b>1.000</b>	47.09	19.83	36	45	56	19.20	11.89	12	17	24
$\lambda^* = 0.3, m = 500 \Rightarrow k^* = 30$											
<b>0</b>	<b>0.916</b>	66.47	15.56	58	66	77	31.62	9.50	27	31	36
<b>0.2</b>	<b>0.891</b>	63.86	17.27	54	64	75	30.46	10.17	26	30	35
<b>0.4</b>	<b>0.849</b>	64.25	22.60	54	68	79	31	12.26	26	31	38
$\lambda^* = 0.3, m = 1.000 \Rightarrow k^* = 60$											
<b>0</b>	<b>0.998</b>	110.78	20.19	97	109	122	60.59	12.06	55	60	66
<b>0.2</b>	<b>0.998</b>	107.62	22.10	94	106	119	59.84	13.31	55	60	66
<b>0.4</b>	<b>0.994</b>	107.05	29.89	94	108	123	58.73	17.64	54	60	66
$\lambda^* = 0.5, m = 500 \Rightarrow k^* = 50$											
<b>0</b>	<b>0.724</b>	77.82	17.47	70.75	81	91	43.54	13.39	38	47	51
<b>0.2</b>	<b>0.703</b>	76.83	19.48	70	81	90	42.58	14.30	36	47	52
<b>0.4</b>	<b>0.558</b>	71.13	28.28	66	80	91	39.39	17.75	31	46	52
$\lambda^* = 0.5, m = 1.000 \Rightarrow k^* = 100$											
<b>0</b>	<b>0.965</b>	148.56	22.61	135	149	163	92.07	20.31	87	97	103
<b>0.2</b>	<b>0.972</b>	148.36	24.02	136	149	162	93.03	19.56	88	97	103
<b>0.4</b>	<b>0.931</b>	150.02	36.68	140	156	172	92.43	24.40	90	98	104

TABELLE 5.3: Güte gegenüber Parameteränderungen, die eine Reduktion der unbedingten Varianz bewirken (Szenario 1), mit den Eigenschaften der Stoppszeiten  $\tau_m$  und der geschätzten Bruchpunkte  $\hat{k}$  für  $p = 3$ .

im Folgenden als Szenario 1 und 2 bezeichnet. Der Strukturbruch im Vektor der Parameter bewirkt dabei nicht nur eine Änderung der bedingten Varianzstruktur, sondern in Szenario 1 auch eine Reduktion der unbedingten Varianzen sowie einen Varianzanstieg in Szenario 2. Die Ergebnisse für Szenario 1 können den Tabellen 5.3 sowie 5.4 entnommen werden und diejenigen für Szenario 2 den Tabellen 5.5 sowie 5.6. Um das Abschneiden der Überwachungsmethode besser beurteilen zu können, enthalten die Tabellen neben der Güte auch das arithmetische Mittel (MW), die Standardabweichung (SD) und die Quartile ( $q_p$ ) der Stoppzeiten sowie der geschätzten Bruchpunkte, die sich in denjenigen Fällen ergeben, in denen die Grenzfunktion tatsächlich im Laufe des Überwachungszeitraums überschritten wurde.

$\gamma$	Güte	Stoppzeiten					Geschätzte Bruchpunkte				
		MW	SD	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$	MW	SD	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$
$\lambda^* = 0.05, m = 500 \Rightarrow k^* = 5$											
<b>0</b>	<b>0.927</b>	44.46	13.82	36	42	50	17.23	9.07	11	15	22
<b>0.2</b>	<b>0.913</b>	39.51	15.13	30	37	45	15.31	8.99	9	13	19
<b>0.4</b>	<b>0.892</b>	38.88	16.83	29	37	48	15.46	10.01	9	13	20
$\lambda^* = 0.05, m = 1.000 \Rightarrow k^* = 10$											
<b>0</b>	<b>0.997</b>	60.62	14.83	51	58	67	23.61	11.11	16	21	28
<b>0.2</b>	<b>0.997</b>	51.60	14.42	42	49	59	20.66	10.31	14	18	25
<b>0.4</b>	<b>0.996</b>	47.04	18.90	37	45	57	18.85	11.01	12	17	24
$\lambda^* = 0.3, m = 500 \Rightarrow k^* = 30$											
<b>0</b>	<b>0.859</b>	66.82	15.25	58	67	77	32.15	9.05	27	31	37
<b>0.2</b>	<b>0.848</b>	64.89	16.99	56	65	76	31.01	10.60	26	30	37
<b>0.4</b>	<b>0.806</b>	64.09	22.07	55	66	78	30.23	11.81	25.25	30	37
$\lambda^* = 0.3, m = 1.000 \Rightarrow k^* = 60$											
<b>0</b>	<b>0.993</b>	108.95	17.86	97	107	119	59.93	12.89	54	60	65
<b>0.2</b>	<b>0.992</b>	106.35	20.53	94	106	118	58.72	13.69	54	59	64
<b>0.4</b>	<b>0.992</b>	107.68	28.43	97	109	123	58.45	16.38	54	60	66
$\lambda^* = 0.5, m = 500 \Rightarrow k^* = 50$											
<b>0</b>	<b>0.680</b>	79.04	16.92	73	82	91	43.53	12.98	37	46	51
<b>0.2</b>	<b>0.673</b>	76.54	20.58	70	82	91	42.48	14.36	37	46	51
<b>0.4</b>	<b>0.553</b>	70.67	30.36	65	82	93	38.90	18.00	31	46	51
$\lambda^* = 0.5, m = 1.000 \Rightarrow k^* = 100$											
<b>0</b>	<b>0.973</b>	147.58	22.32	136	148	161	91.81	19.18	86	96	102
<b>0.2</b>	<b>0.972</b>	145.77	25.43	135	148	161	90.26	21.93	84	97	102
<b>0.4</b>	<b>0.947</b>	148.30	38.44	139	155	169	90.00	25.49	87	97	103

TABELLE 5.4: Güte gegenüber Parameteränderungen, die eine Reduktion der unbedingten Varianz bewirken (Szenario 1), mit den Eigenschaften der Stoppzeiten  $\tau_m$  und der geschätzten Bruchpunkte  $\hat{k}$  für  $p = 5$ .

Die Güte hängt positiv von der Länge des historischen Zeitraums und negativ von der Dimension der zugrundeliegenden Zufallsvektoren ab. Während das erste Resultat zu erwarten war, dürfte der negative Einfluss von  $p$  auf die Güte dadurch zu erklären sein, dass der Anteil der  $3p$  Varianzparameter an der Gesamtzahl aller Modellparameter mit wachsender Dimension sinkt. Somit sind vergleichbare Varianzänderungen mit steigender Dimension schwerer nachzuweisen. Des Weiteren werden Änderungen, die früh eintreten, sehr viel häufiger entdeckt als Änderungen im fortgeschritteneren Überwachungszeitraum. Diese Eigenschaft kann typischerweise bei Verfahren beobachtet werden, die darauf beruhen, dass die in einem historischen Zeitraum gesammelte Information zur weiteren Überwachung eines Prozesses genutzt wird, siehe unter anderem Wied und Galeano (2013) und Abschnitt 3.3. Darüber hinaus können die Parameteränderungen, die eine Reduktion der unbedingten Varianz bewirken, deutlich zuverlässiger nachgewiesen werden als diejenigen Änderungen, die einen Varianzanstieg verursachen. Dieses Ergebnis steht in Einklang mit den Simulationsergebnissen aus dem vorherigen Abschnitt.

		Stoppzeiten					Geschätzte Bruchpunkte				
$\gamma$	Güte	MW	SD	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$	MW	SD	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$
$\lambda^* = 0.05, m = 500 \Rightarrow k^* = 5$											
<b>0</b>	<b>0.768</b>	64.97	19.92	51	66	80	24.62	16.68	12	20	32
<b>0.2</b>	<b>0.702</b>	56.90	23.55	39.25	57	75	21.57	15.39	10	17	29
<b>0.4</b>	<b>0.558</b>	49.34	29.99	21.25	48	75.75	20.44	17.00	7	15	29
$\lambda^* = 0.05, m = 1.000 \Rightarrow k^* = 10$											
<b>0</b>	<b>0.974</b>	116.05	32.61	95	114	138	39.92	30.82	16	29	55
<b>0.2</b>	<b>0.955</b>	107.65	38.44	80	107	133	36.77	30.67	14	26	49
<b>0.4</b>	<b>0.810</b>	107.81	49.94	72	109.5	146	39.06	35.14	14	25	53
$\lambda^* = 0.3, m = 500 \Rightarrow k^* = 30$											
<b>0</b>	<b>0.527</b>	76.44	15.92	65	79	90	32.30	11.46	27	32	37
<b>0.2</b>	<b>0.482</b>	73.28	18.45	61	76	88	31.38	12.84	25	31	36
<b>0.4</b>	<b>0.271</b>	62.76	27.40	46	68	85	29.00	14.92	20	30	36
$\lambda^* = 0.3, m = 1.000 \Rightarrow k^* = 60$											
<b>0</b>	<b>0.809</b>	151.87	28.70	130	155	176	58.77	15.35	52	60	65
<b>0.2</b>	<b>0.727</b>	145.37	33.39	121	149	172	55.86	16.11	48	59	65
<b>0.4</b>	<b>0.457</b>	142.46	47.80	122	153	179	54.78	22.12	46	59	65
$\lambda^* = 0.5, m = 500 \Rightarrow k^* = 50$											
<b>0</b>	<b>0.346</b>	81.76	14.91	72	84	95	44.04	15.42	36	48.5	54
<b>0.2</b>	<b>0.292</b>	76.91	19.49	70	81	92	40.72	16.24	29	46	52
<b>0.4</b>	<b>0.172</b>	58.26	35.67	18	72	87.25	32.97	23.10	6.75	39.5	52
$\lambda^* = 0.5, m = 1.000 \Rightarrow k^* = 100$											
<b>0</b>	<b>0.505</b>	167.32	24.59	153	172	187	86.40	22.86	74	94	102
<b>0.2</b>	<b>0.445</b>	161.34	32.89	146	168	184	83.36	25.98	71	91	101
<b>0.4</b>	<b>0.221</b>	138.81	61.53	122	164	182	74.26	37.55	52	90	101

TABELLE 5.5: Güte gegenüber Parameteränderungen, die einen Anstieg der unbedingten Varianz bewirken (Szenario 2), mit den Eigenschaften der Stoppzeiten  $\tau_m$  und der geschätzten Bruchpunkte  $\hat{k}$  für  $p = 3$ .

		Stoppzeiten					Geschätzte Bruchpunkte				
$\gamma$	Güte	MW	SD	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$	MW	SD	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$
$\lambda^* = 0.05, m = 500 \Rightarrow k^* = 5$											
<b>0</b>	<b>0.738</b>	67.80	19.61	54	70.50	83	25.32	17.10	12.25	20	34
<b>0.2</b>	<b>0.644</b>	60.38	23.72	43	62.5	80	23.15	17.21	11	18	31
<b>0.4</b>	<b>0.409</b>	46.89	30.46	15	49	72	18.67	16.67	6	13	26
$\lambda^* = 0.05, m = 1.000 \Rightarrow k^* = 10$											
<b>0</b>	<b>0.964</b>	127.63	32.68	105	127	151	43.43	31.97	18	33	62
<b>0.2</b>	<b>0.935</b>	117.04	39.51	90	116	145	38.39	31.36	15	27	54
<b>0.4</b>	<b>0.704</b>	112.02	55.44	73	120	157.25	41.77	37.69	14	27	59
$\lambda^* = 0.3, m = 500 \Rightarrow k^* = 30$											
<b>0</b>	<b>0.491</b>	76.75	16.17	66	80	89	30.91	11.62	25.5	31	36
<b>0.2</b>	<b>0.394</b>	72.05	20.39	59.25	73	89	30.59	12.72	24	31.5	36
<b>0.4</b>	<b>0.193</b>	56.33	34.87	30	64	88	25.09	17.89	8	29	34
$\lambda^* = 0.3, m = 500 \Rightarrow k^* = 60$											
<b>0</b>	<b>0.742</b>	155.38	29.41	137	159	178	58.46	15.89	53	60	65
<b>0.2</b>	<b>0.671</b>	149.38	33.91	128	154	176	57.44	15.32	51	60	65
<b>0.4</b>	<b>0.363</b>	135.31	58.76	114.5	151	181	53.02	26.20	43.5	59	65
$\lambda^* = 0.5, m = 500 \Rightarrow k^* = 50$											
<b>0</b>	<b>0.292</b>	79.91	14.59	70	82	92	41.55	15.36	31	46.5	53
<b>0.2</b>	<b>0.265</b>	76.12	19.48	65	81	91	39.81	15.82	29	46	52
<b>0.4</b>	<b>0.127</b>	46.23	36.34	4.5	56	78	25.44	21.42	2	24	48
$\lambda^* = 0.5, m = 1.000 \Rightarrow k^* = 100$											
<b>0</b>	<b>0.461</b>	166.82	23.85	150	171	186	83.79	21.66	71	91	100
<b>0.2</b>	<b>0.384</b>	160.36	33.85	144.75	169	185	82.73	24.10	68.75	90.5	101
<b>0.4</b>	<b>0.175</b>	125.46	71.73	64.5	156	182	65.53	39.96	31.5	80	99

TABELLE 5.6: Güte gegenüber Parameteränderungen, die einen Anstieg der unbedingten Varianz bewirken (Szenario 2), mit den Eigenschaften der Stoppzeiten  $\tau_m$  und der geschätzten Bruchpunkte  $\hat{k}$  für  $p = 5$ .

Diese belegen, dass der Detektor unter der Nullhypothese eine deutlich stärkere Tendenz aufweist, fälschlicherweise die Grenzfunktion zu überschreiten, wenn die Varianzparameter als  $\phi_i = (0.01, 0.05, 0.9)'$  gewählt werden, was der Ausgangssituation in Szenario 1 entspricht.

Abschließend fällt auf, dass wie unter der Nullhypothese die Güte mit einem steigenden Wert des Justierungsparameters  $\gamma$  sinkt. Dieses Ergebnis lässt sich sowohl für frühe als auch für spätere Strukturbrüche beobachten, was überraschend ist, da dieser Parameter die Grenzfunktion eigentlich hinsichtlich der Lage potenzieller Bruchpunkte sensibilisieren sollte. Wie schon in den Simulationen für das nichtparametrische Verfahren in Abschnitt 3.3, wird hier der Einfluss des Justierungsparameters scheinbar von anderen Effekten überlagert und entspricht somit in einigen der betrachteten Fälle nicht der Erwartung.

Die Kennzahlen der geschätzten Bruchpunkte in den Tabellen 5.3-5.6 legen nahe, dass die Lage des Strukturbruchs ohne größere Verzerrung geschätzt werden kann, wenn diese durch  $\lambda^* = 0.3$  bestimmt wird. Die Position von früheren und späteren Strukturbrüchen hingegen wird systematisch über- beziehungsweise unterschätzt. Die Eigenschaften der Stoppzeiten  $\tau_m$  und der geschätzten Bruchpunkte weisen starke Ähnlichkeiten auf. Die Ursache für dieses Verhalten ist in der Tatsache zu finden, dass die Stoppzeiten den Zeitraum definieren beziehungsweise einschränken, in dem überhaupt erst nach einem Strukturbruch gesucht werden kann.

### Änderungen in den Korrelationsparametern

Ein weiteres mögliches Alternativszenario betrifft Änderungen in der Korrelationsstruktur. Es wird angenommen, dass sowohl die Varianzparameter als auch  $\alpha$  und  $\beta$  über die Zeit hinweg konstant bleiben, während sich die unbedingte Kovarianzmatrix von  $\bar{Q}_p = \mathbb{I}_p$  in  $\bar{Q}_p^*$  ändert, wobei letztere Matrix durch Einsen auf der Hauptdiagonalen und den Wert  $\Delta$  in allen anderen Einträgen charakterisiert wird. Dabei wird  $\Delta$  aus  $\{0.1, \dots, 0.9\}$  gewählt. Die Wahl der Varianzparameter sowie  $\alpha$  und  $\beta$  erfolgt analog zu Abschnitt 5.3.1.

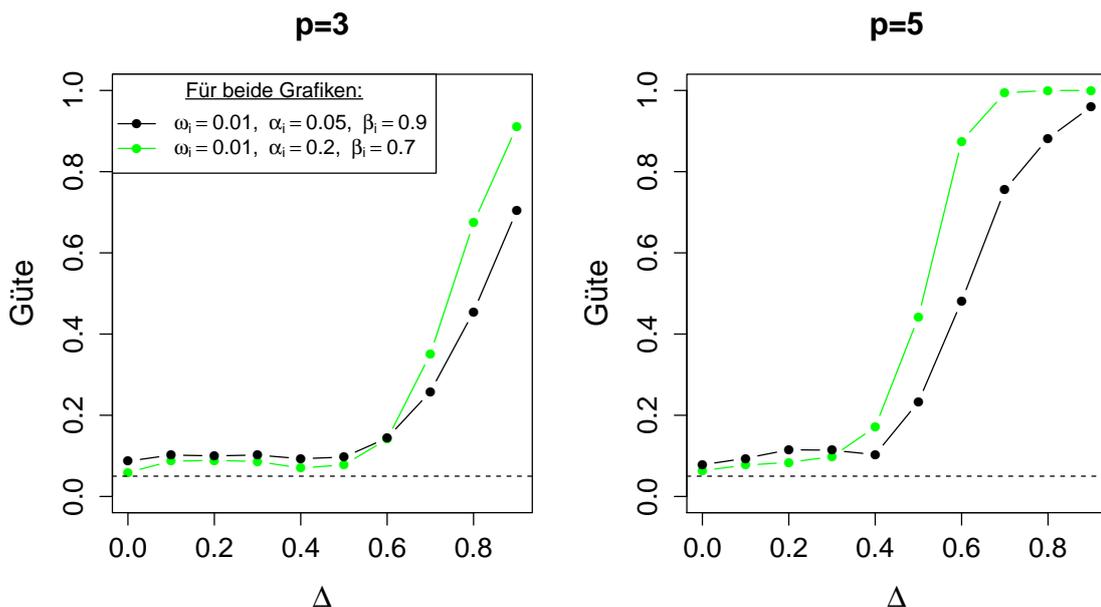


ABBILDUNG 5.1: Güte gegenüber Korrelationsänderungen.

Die simulierte Güte für Änderungen, deren Lage durch  $\lambda^* = 0.05$  bestimmt wird, ist in Abbildung 5.1 dargestellt. Dabei beträgt die Dimension der generierten Zeitreihen 3 oder 5, die Länge des historischen Zeitraums 1.000 Datenpunkte und der Justierungsparameter  $\gamma = 0.2$ . Die Ergebnisse enthüllen starke Probleme des Verfahrens, Korrelationsänderungen von geringerem bis mäßig hohem Ausmaß zu erkennen. Allerdings steigt die Gütefunktion in diesem Fall steil an und strebt gegen Eins. Während kleinere Änderungen in den Korrelationsparametern etwas zuverlässiger erkannt werden, wenn die Varianzparameter als  $\phi_i = (0.01, 0.05, 0.9)'$  gewählt werden, können größere Änderungen für  $\phi_i = (0.01, 0.2, 0.7)'$  häufiger nachgewiesen werden.

Die Tatsache, dass mehrere Gütewerte relativ niedrig sind, spricht dafür, dass die QLL-Funktion in einigen Bereichen ziemlich flach zu sein scheint, so dass sich manche Arten von Parameteränderungen nur marginal auf die Detektorwerte auswirken und somit nur schwer nachzuweisen sind.

## 5.4 Anwendung auf Aktienrenditen

Zur Untersuchung wie sich das Überwachungsverfahren in der Praxis bewährt, wird es auf die Renditen der Aktien mehrerer Versicherungsunternehmen angewandt, die an verschiedenen europäischen Börsenplätzen gehandelt werden. Im Einzelnen werden dabei die täglichen log-Renditen der Aktien von *Allianz* (abgekürzt durch *All*), *AXA*, *Generali* (*Gen*), *ING* und *Münchener Rück* (*MRu*) von Mai 1998 bis Oktober 2016 betrachtet, da eine gemeinsame Modellierung der Renditen von Aktien aus derselben Branche und demselben Währungsraum als grundsätzlich sinnvoll angesehen werden kann. Engle (2002) argumentiert, dass das DCC-Modell gut geeignet ist, um die typischen Auffälligkeiten in multivariaten Zeitreihen von Aktienrenditen zu modellieren. Außerdem erklärt Bollerslev (1986), dass schon GARCH-Modelle der Ordnung (1,1) sehr gut geeignet sind, um das charakteristische Verhalten von univariaten Renditezeitreihen zu modellieren. Deshalb werden vergleichsweise simple GARCH(1,1)-Modelle für die univariaten bedingten Varianzgleichungen in (4.2) verwendet, was auch dem Vorgehen in Kapitel 4.1 entspricht.

Wie die Simulationsergebnisse in Tabelle 5.2 nahelegen, steigt die Tendenz, eine gültige Nullhypothese abzulehnen, mit der Länge des Überwachungszeitraums beziehungsweise mit der Variablen  $B$ . Deshalb wird der vorliegende Datensatz nicht in einem Schritt, sondern etappenweise überwacht. Dazu werden in jedem Schritt die ersten  $m$  Beobachtungen als historische Stichprobe gewählt und die folgenden  $[mB]$  Datenpunkte als Überwachungszeitraum. Wenn in dieser Teilstichprobe ein Strukturbruch entdeckt wird, kann die Lage des Bruchpunkts durch  $\hat{k}$  in (5.5) geschätzt werden. Anschließend werden alle vorausgehenden Datenpunkte aus dem Datensatz entfernt. Die auf den geschätzten Bruchpunkt folgenden  $m$  und  $[mB]$  Datenpunkte werden dann als neue historische beziehungsweise zu überwachende Stichprobe gewählt und das Verfahren kann erneut angewandt werden. Wird in der betrachteten Teilstichprobe kein Strukturbruch nachgewiesen, so werden die ersten  $[mB]$  Beobachtungen des historischen Zeitraums entfernt und die Datenpunkte des vormaligen Überwachungszeitraums an deren Stelle in die historische Stichprobe einbezogen. Dadurch verschiebt sich das Datenfenster, das den historischen Zeitraum umfasst, um  $[mB]$  Zeitpunkte.

Die Ergebnisse für  $\gamma = 0.2$ ,  $B = 0.2$  und verschiedene Längen des historischen Zeitraums können Tabelle 5.7 entnommen werden. Außerdem finden sich die für  $m = 1.000$  ermittelten Strukturbrüche gemeinsam mit den überwachten Zeitreihen in Abbildung 5.2, die illustriert, dass die Zeitreihen durch die Bruchpunkte in ruhigere und turbulenter Phasen

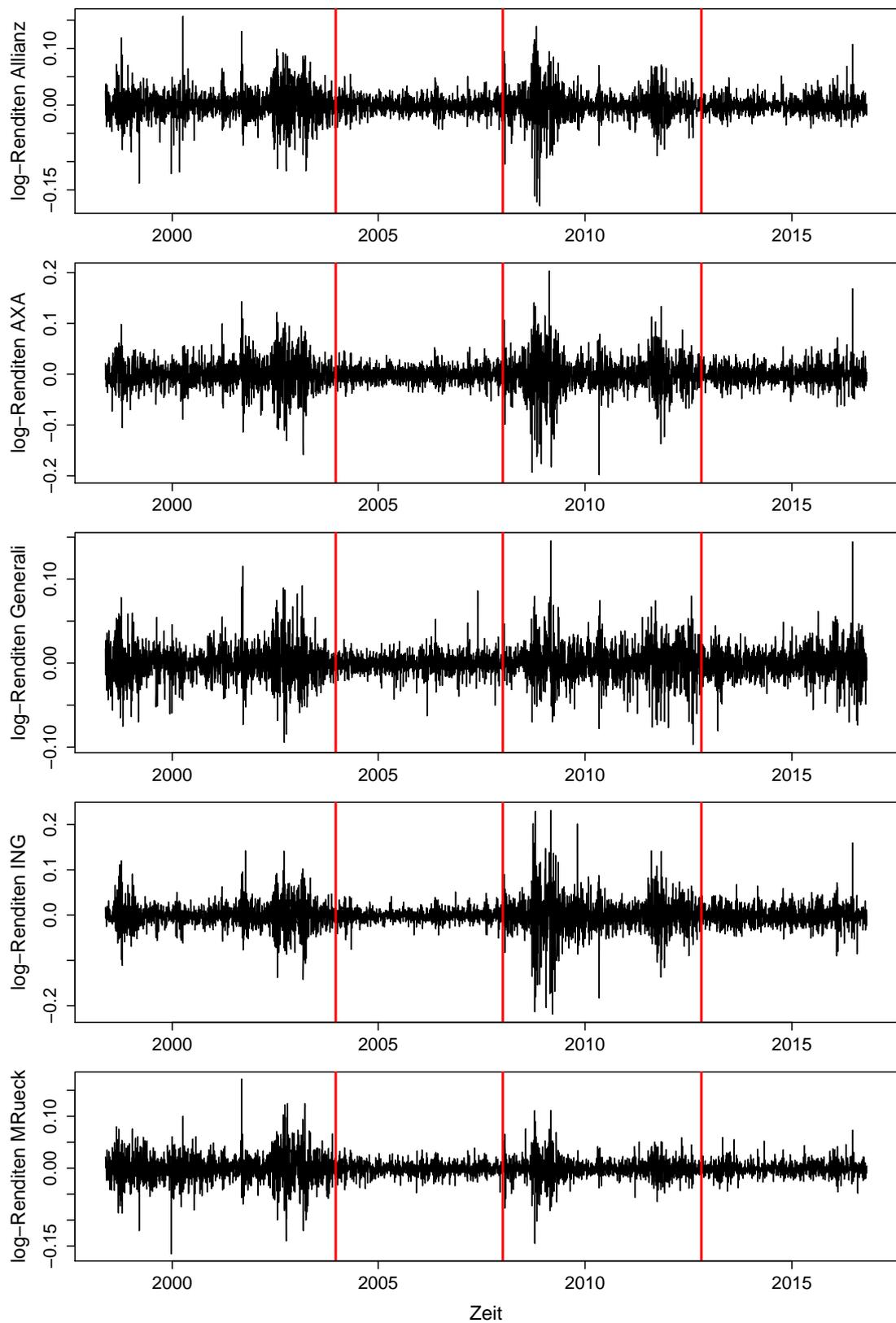


ABBILDUNG 5.2: log-Renditen der Aktienkurswerte mit den für  $m = 1.000$  ermittelten Bruchpunkten aus Tabelle 5.7 (rote Linien).

$\gamma = 0.2, B = 0.2, m = 500$		$\gamma = 0.2, B = 0.2, m = 1000$	
$\tau_m$	$\hat{k}$	$\tau_m$	$\hat{k}$
09.06.2000	22.05.2000	27.01.2004	22.12.2003
15.10.2002	10.10.2002	21.01.2008	04.01.2008
01.03.2006	05.12.2005	04.12.2012	15.10.2012
28.05.2008	22.05.2008		
04.08.2010	09.06.2010		
26.07.2012	05.07.2012		
09.07.2014	07.07.2014		

TABELLE 5.7: Abbruchzeiten und geschätzte Bruchpunkte bei der Anwendung des Überwachungsverfahrens auf die log-Renditen der Aktienkurswerte von *Allianz*, *AXA*, *Generali*, *ING* und *Münchener Rück*.

aufgespalten werden. Tabelle 5.7 belegt, dass unter Verwendung einer nur 500 Datenpunkte umfassenden historischen Stichprobe deutlich mehr Strukturbrüche nachgewiesen werden können als unter Verwendung eines historischen Zeitraums der Länge 1.000. Dies kann ein Verweis darauf sein, dass unter der Verwendung des längeren historischen Zeitraums Strukturbrüche übersehen wurden und somit möglicherweise die Annahme verletzt

	11.05.1998 -22.12.2003	23.12.2003 -04.01.2008	07.01.2008 -15.10.2012	16.10.2012 -25.10.2016
$\hat{\omega}_{All}$	0.00002	0.00002	0.00001	<0.00001
$\hat{\omega}_{AXA}$	0.00001	0.00002	<0.00001	<0.00001
$\hat{\omega}_{Gen}$	0.00002	0.00003	<0.00001	<0.00001
$\hat{\omega}_{ING}$	0.00002	0.00001	0.00002	0.00001
$\hat{\omega}_{MRe}$	0.00001	0.00001	0.00002	0.00009
$\hat{\alpha}_{All}$	0.099	0.072	0.083	0.034
$\hat{\alpha}_{AXA}$	0.085	0.040	0.095	0.045
$\hat{\alpha}_{Gen}$	0.096	0.097	0.058	0.028
$\hat{\alpha}_{ING}$	0.082	0.113	0.093	0.125
$\hat{\alpha}_{MRe}$	0.120	0.063	0.117	0.072
$\hat{\beta}_{All}$	0.879	0.811	0.905	0.944
$\hat{\beta}_{AXA}$	0.887	0.920	0.895	0.936
$\hat{\beta}_{Gen}$	0.877	0.786	0.941	0.967
$\hat{\beta}_{ING}$	0.900	0.806	0.900	0.667
$\hat{\beta}_{MRe}$	0.853	0.703	0.883	0.924
$\ \hat{\theta}_1\ _2$	0.182	0.171	0.167	0.140
$\hat{\alpha}$	0.011	0.022	0.032	0.013
$\hat{\beta}$	0.988	0.978	0.883	0.986
$\hat{Q}_{All,AXA}$	0.422	0.597	0.779	0.778
$\hat{Q}_{All,Gen}$	0.401	0.442	0.536	0.406
$\hat{Q}_{All,ING}$	0.519	0.523	0.774	0.697
$\hat{Q}_{All,MRe}$	0.651	0.505	0.745	0.704
$\hat{Q}_{AXA,Gen}$	0.424	0.487	0.579	0.538
$\hat{Q}_{AXA,ING}$	0.621	0.470	0.761	0.762
$\hat{Q}_{AXA,MRe}$	0.457	0.510	0.666	0.607
$\hat{Q}_{Gen,ING}$	0.468	0.245	0.511	0.442
$\hat{Q}_{Gen,MRe}$	0.366	0.366	0.439	0.256
$\hat{Q}_{ING,MRe}$	0.513	0.435	0.645	0.475
$\ \hat{\theta}_2\ _2$	2.698	2.532	3.027	2.908
$\ \hat{\theta}\ _2$	2.704	2.537	3.032	2.911

TABELLE 5.8: Jeweils aus den Datenpunkten zwischen zwei aufeinanderfolgenden Bruchpunkten aus Tabelle 5.7 geschätzte Modellparameter für  $m = 1.000$  mit der euklidischen Norm der geschätzten Parametervektoren.

ist, dass kein Strukturbruch in den ersten  $m$  Punkten vorliegt. Eine andere Erklärung ist, dass das Verfahren aufgrund der erheblichen Niveauperzerrung für  $m = 500$ , wie sie durch die Ergebnisse in Tabelle 5.2 belegt wird, dazu neigt, auch eine wahre Nullhypothese mit beträchtlich hoher Wahrscheinlichkeit abzulehnen. Ein Grund dafür liegt in der unzureichenden Schätzung der Modellparameter aus zu wenigen Datenpunkten. Deshalb erscheinen die Ergebnisse unter der Verwendung von  $m = 1.000$  vertrauenswürdiger. Die geschätzten Bruchpunktlokationen in Tabelle 5.7 können mit relevanten wirtschaftlichen Ereignissen der letzten zwei Dekaden in Verbindung gebracht werden. Der Strukturbruch im Jahr 2003 markiert das Ende der durch das Platzen der *dotcom*-Blase verursachten Krise und den Beginn einer ruhigeren Phase, die durch den Ausbruch der globalen Finanzkrise beendet wurde. Der letzte Bruchpunkt kann als Anfang einer Erholungsphase an den Börsen gesehen werden, die nach der Finanz- und Schuldenkrise langsam einsetzte.

Die QML-Schätzer für die Modellparameter zwischen jeweils zwei aufeinanderfolgenden entdeckten Bruchpunkten können Tabelle 5.8 entnommen werden. Um das Ausmaß der Parameteränderungen zwischen zwei aufeinanderfolgenden Strukturbrüchen leichter beurteilen zu können, gibt die Tabelle zusätzlich die euklidische Norm der aus den jeweiligen Teilstichproben geschätzten Parametervektoren sowie die Norm derjenigen Parametervektoren, die ausschließlich die Varianz- beziehungsweise Korrelationsparameter enthalten, an. Die stärkste Änderung im Sinne der euklidischen Norm kann dabei zwischen den Teilstichproben vor und nach dem Ausbruch der Finanzkrise beobachtet werden. Ein Großteil dieses Phänomens scheint dabei auf die Tatsache zurückzuführen zu sein, dass die Korrelation von Aktienrenditen in Krisenzeiten tendenziell zunimmt. Diese Besonderheit wird beispielsweise in Sandoval Jr. und De Paula Franca (2012) besprochen.



## Kapitel 6

# Fazit und Ausblick

Die beiden in den vergangenen Kapiteln vorgestellten sequenziellen Testverfahren können genutzt werden, um nach Änderungen in der Varianz- oder der zugrundeliegenden Parameterstruktur von multivariaten Finanzzeitreihen zu suchen. Während die in Kapitel 3 vorgestellte nichtparametrische Methode durch ihren modellunabhängigen Charakter und ihre im Finanzbereich vielseitige Anwendbarkeit besticht, erlaubt das parametrische Verfahren aus Kapitel 5 die Detektion einer Vielzahl verschiedener Arten von Änderungen im Vektor der Modellparameter. In dieser Arbeit werden bereits bekannte Methoden auf den multivariaten Fall und auf Zeitreihen, die als Realisationen von multivariaten GARCH-Prozessen angesehen werden können, erweitert. Da viele Zeitreihen aus dem finanzwirtschaftlichen Bereich diese Annahme erfüllen, wird durch dieses Vorgehen außerdem der Anwendungsbereich auf einige Fälle von hoher praktischer Relevanz erweitert.

Aufgrund der von Bissantz et al. (2011) belegten höheren Relevanz von Änderungen in den Varianzen als in den Kovarianzen in Bezug auf verschiedene Risikomaße, konzentriert sich die nichtparametrische Methode auf die Überwachung des Vektors der Varianzen. In einigen der simulierten Szenarien schlägt das Verfahren auch seinen univariaten Konkurrenten und weist insgesamt eine hohe Güte auf, wenn die Varianzänderung nicht zu spät im Überwachungszeitraum auftritt und wenn ein ausreichend langer historischer Zeitraum verwendet wird. Trotz auf den ersten Blick strenger theoretischer Annahmen an die betrachteten Prozesse, schränken diese die Anwendbarkeit auf Zeitreihen wie beispielsweise die hier betrachteten log-Renditen von Aktien nicht ein, da die aufgestellten Annahmen für solche Zeitreihen in der Regel als erfüllt betrachtet werden können.

Die Stärke des Verfahrens zur Überwachung der Parameter im DCC-Modell liegt in der Vielfalt der Alternativszenarien, die sich durch Änderungen im Vektor der Parameter gestalten lassen. Während sich viele konkurrierende Methoden auf die Überwachung der Varianz- oder Korrelationsstruktur beschränken, erlaubt die hier vorgestellte Methode eine gemeinsame Überwachung aller Modellparameter ohne eine vorherige Festlegung darauf, welche der Parameter von besonderem Interesse sein sollen. Des Weiteren ist es nicht notwendig, sich bei der Detektion im Voraus konkret auf Parameteränderungen zu konzentrieren, die ausschließlich Variationen in den bedingten oder in den unbedingten Momenten bewirken, da beide Strukturen gemeinsam überwacht werden können. Eine besonders hohe Güte erreicht das Verfahren, wenn der Strukturbruch eine Änderung in der Varianzstruktur des DCC-Modells bewirkt.

Die im DCC-Modell aufgestellten Annahmen 5.7-5.10 zur Gewährleistung einer eindeutigen stationären Lösung auch nach einem Strukturbruch sind dabei nicht zwingend notwendig. Unter analytischen Aspekten ist die Annahme eines solchen wohlgestalteten Verhaltens hilfreich. Liefert jedoch nur der Parametervektor vor dem Strukturbruch eine eindeutige stationäre Lösung, so sollte dies dafür sorgen, dass einige Änderungen sogar noch deutlicher

erkannt werden als wenn die Annahmen für die Existenz einer eindeutigen stationären Lösung auch nach dem Strukturbruch erfüllt sind. Beispielsweise könnte der Prozess nach dem Strukturbruch eine Einheitswurzel, explosives oder auf eine andere Art und Weise extremes Verhalten aufweisen. Dies würde eine derart starke Änderung des gesamten betrachteten Prozesses bewirken, dass der Strukturbruch noch leichter und schneller nachgewiesen werden können sollte.

Bei der Überwachung von DCC-Zeitreihen erweist sich die Annahme, dass der historische Zeitraum frei von Strukturbrüchen sein muss, als problematisch. Während im nichtparametrischen Fall einfache retrospektive Tests zur Überprüfung dieser Annahme genutzt werden können, bleibt offen, wie deren Gültigkeit für das DCC-Modell überprüft werden kann. Dies liegt in der Tatsache begründet, dass die Methode konstruktionsbedingt auf den Parameterschätzer auf Basis der Daten des historischen Zeitraums angewiesen ist. Für ein Äquivalent zur im nichtparametrischen Fall angewandten retrospektiven Methode müsste der Parametervektor aus teilweise sehr kleinen Teilstichproben geschätzt werden. Außerdem wäre die Rechenintensität unpraktikabel hoch, wenn für jeden Durchlauf der Parametervektor aus mehreren verschieden langen Teilstichproben geschätzt werden müsste. Da die Simulationen zeigen, dass schon die Verwendung von 500 Beobachtungen verbesserungswürdige Schätzergebnisse liefert, erscheint ein solcher Ansatz hier somit nicht sinnvoll.

Die vorgestellten Verfahren halten mit wachsender Dimension der betrachteten Zufallsvektoren das Signifikanzniveau immer weniger ein. Während beim nichtparametrischen Verfahren die Schätzung der Variationsmatrix  $D_p$  Niveauverzerrungen bewirkt, ist dieses Verhalten bei der parametrischen Methode auf die hohe Dimension des Parametervektors zurückzuführen. Diese ist schon für moderate Dimensionen der betrachteten Zufallsvektoren  $y_t$  relativ hoch. Eine Verbesserung der Parameterschätzung im DCC-Modell dürfte ohne eine Erhöhung des Umfangs der Teilstichprobe, aus der die Schätzung bestimmt wird, jedoch nur beschränkt möglich sein. Im Gegensatz dazu besteht für die Niveau- und Güteeigenschaften des nichtparametrischen Verfahrens durch eine optimierte Schätzung der Variationsmatrix  $D_p$  noch deutliches Verbesserungspotenzial.

Beide Verfahren weisen unter der Nullhypothese teilweise deutliche Niveauverzerrungen auf. Diesem Verhalten könnte durch Methoden zur Niveauadjustierung beziehungsweise der Bestimmung von korrigierten kritischen Werten, die eine Einhaltung des Niveaus unter der Nullhypothese erlauben, entgegengewirkt werden. Für das parametrische Verfahren würde ein solches Vorgehen immer vom konkret verwendeten Parametervektor abhängen. Somit müsste eine in diesem Fall zweifellos sehr rechenintensive Adjustierung der kritischen Werte für jede individuelle Parameterkombination einzeln durchgeführt werden. Für das nichtparametrische Verfahren jedoch könnten für die Simulationen unter verschiedenen Alternativen Methoden zur allgemeinen Niveauadjustierung herangezogen und deren Wirkung auf die Güte mit Hilfe von weiteren Simulationen näher untersucht werden.

Über den Rahmen dieser Arbeit hinaus würden sich einige Themengebiete zur weiteren Forschung anbieten. Zunächst weisen die in dieser Arbeit verwendeten Detektoren beide eine CUSUM-artige Struktur auf, da der Detektor für jeden Zeitpunkt im Überwachungszeitraum durch die Addition eines skalierten Terms ergänzt wird. Diese additive Struktur hat zur Folge, dass Änderungen, die erst spät eintreten, weniger zuverlässig nachgewiesen werden können, da nach dem Auftreten des Strukturbruchs zunächst einige neue Summanden in den Detektor eingehen müssen, bevor dieser in seinem Verhalten so stark beeinträchtigt wird, dass er die vorgegebene Grenzfunktion überschreitet. Um lange Verzögerungszeiten und eine teilweise niedrige Güte abzumildern, könnten MOSUM (*moving sum*)-artige Detektoren verwendet werden, wie sie von Bauer und Hackl (1978) zur Qualitätskontrolle

vorgeschlagen wurden. In diesem Fall würden nicht mehr alle skalierten Beobachtungsvektoren des Überwachungszeitraums betrachtet werden, sondern nur diejenigen in einem wandernden Zeitfenster gegebener Länge. Dies würde einen stärkeren Einfluss der Beobachtungen nach dem Strukturbruch bewirken und sollte sich sowohl positiv auf die bloße Erkennung vorliegender Strukturbrüche als auch auf die zugehörige Verzögerungszeit auswirken.

Des Weiteren sollte untersucht werden, ob die Verwendung alternativer Normen oder quadratischer Formen bei der Bestimmung der Detektorwerte die Simulationsergebnisse beeinflussen würden. Außerdem bieten die zur Schätzung der Lage eines entdeckten Bruchpunkts vorgeschlagenen Schätzer weiteres Untersuchungspotenzial. Interessant wäre hier vor allem die Effizienz der verwendeten Überwachungsverfahren, etwa in Bezug auf mittlere Verzögerungszeiten. Für die parametrische Methode könnte weiterhin die zur Schätzung der kritischen Werte verwendete parametrische Bootstrapmethode genauer analysiert und optimiert werden. Interessant wären hier beispielsweise asymptotische Resultate unter der Annahme eines wachsenden historischen Zeitraums.

Wie weiterführende Simulationen, die nicht Teil dieser Arbeit sind, zeigen, können unter Einbeziehung eines Schätzers für die Informationsmatrix  $\mathbf{D}$  in den aus den Bootstrapstichproben bestimmten Detektor einige Arten von Strukturbrüchen zuverlässiger erkannt werden als andere. Da dieses Resultat allerdings nicht für alle Alternativszenarien gilt und mit deutlich gesteigerter Wahrscheinlichkeit, einen Fehler erster Art zu begehen, einhergeht, liefert diese Erweiterung keine wirkliche Verbesserung des Verfahrens. Allerdings könnte untersucht werden, ob eine alternative Skalierung des Detektors das Verfahren bezüglich mancher sonst nur schwer nachzuweisender Parameteränderungen sensibilisieren kann. Sowohl allgemeingültig als auch in Bezug auf das nichtparametrische Verfahren, wäre außerdem eine ausführliche Untersuchung des Verhaltens der Residuen bei der Anpassung von multivariaten GARCH-Modellen von Interesse. Wenn gezeigt werden kann, dass sich einige Charakteristika von den zugrundeliegenden Innovationen auf die Residuen oder zumindest eine auf den Residuen basierende Teststatistik übertragen lassen, wie dies von Kulperger und Yu (2005) im univariaten Fall gezeigt wurde, könnte dies eine Vielzahl weiterer Anwendungsmöglichkeiten eröffnen.

Da sich das von Engle (2002) vorgeschlagene DCC-Modell großer Beliebtheit bei der Modellierung von Aktienrenditen erfreut, konzentriert sich die vorliegende Arbeit auf diesen Modelltyp. Allerdings lassen sich die hier präsentierten Ergebnisse auch leicht auf Modelle mit einer ähnlichen Struktur erweitern. Beispiele dafür sind unter anderem das Modell mit konstanter bedingter Korrelation (CCC-Modell) von Bollerslev (1990), das DCC-Modell in alternativer Darstellung von Tse und Tsui (2002) oder das asymmetrische generalisierte Modell mit dynamischer bedingter Korrelation (AG-DCC-Modell) von Capiello et al. (2006). Im Gegensatz dazu ist eine Erweiterung auf andere gebräuchliche multivariate GARCH-Modelle, deren Struktur sich zu stark vom DCC-Modell unterscheidet, deutlich schwieriger. Deshalb kann die vorgestellte Methode beispielsweise nicht ohne Weiteres auf das Vektor-GARCH-Modell von Bollerslev et al. (1988) oder auf das von Engle und Kroner (1995) vorgestellte BEKK-Modell übertragen werden.

Die Ergebnisse in Kapitel 5 legen auch nahe, im Voraus die Struktur der QLL-Funktion insbesondere in Hinblick auf die Präsenz sehr flacher Bereiche der Funktion zu untersuchen. Außerdem empfiehlt es sich, Modelle mit niedrigerer Parameterzahl auszuwählen oder zu spezifizieren. Zum einen erlaubt dieses Vorgehen eine zuverlässigere Parameterschätzung auf Basis einer historischen Stichprobe mit gegebener Länge. Zum anderen können sehr stark parametrisierte Modelle wie im Fall des klassischen Vektor-GARCH-Modells oder des unrestringierten BEKK-Modells Probleme wie eine mangelnde Identifizierbarkeit der

einzelnen Modellparameter mit sich bringen, siehe Scherrer und Ribarits (2007). Da jedoch das in dieser Arbeit betrachtete DCC-Modell vergleichsweise sparsam parametrisiert wurde, sollten alle Modellparameter identifizierbar sein.

## Anhang A

# Beweise der Aussagen aus Kap. 3

### A.1 Beweis von Theorem 3.1

Die Beweise der Theoreme 3.1-3.3 folgen sehr eng der Argumentation von Wied und Galeano (2013). Sei auch hier  $D[d_1, d_2]$  der Raum der càdlàg-Funktionen auf dem Intervall  $[d_1, d_2]$  ausgestattet mit der Supremumsnorm, die das Supremum der Absolutwerte einer Funktion angibt. Des Weiteren sei der Vektor der zeitinvarianten Varianzen gegeben als

$$\boldsymbol{\sigma}^2 = (\sigma_1^2, \dots, \sigma_p^2)' .$$

Außerdem sei der Prozess  $\{P_m(d), d \in [c, B]\}$  definiert durch

$$P_m(d) = \widehat{D}_{p,m}^{-1/2} \frac{[m \cdot d] - [m \cdot c]}{\sqrt{m}} \left( [\hat{\sigma}^2]_{[m \cdot c]}^{[m \cdot d]} - \sigma^2 \right)$$

$$\text{mit } \left( [\hat{\sigma}^2]_{[m \cdot c]}^{[m \cdot d]} - \sigma^2 \right) = \left( [\hat{\sigma}_1^2]_{[m \cdot c]}^{[m \cdot d]} - \sigma_1^2, \dots, [\hat{\sigma}_p^2]_{[m \cdot c]}^{[m \cdot d]} - \sigma_p^2 \right)' .$$

Um die Konvergenz des Prozesses  $\{P_m\}$  gegen eine Brownsche Bewegung auf dem Raum  $D[c, B]$  zu zeigen, kann das in Theorem 29.18 in Davidson (1994) formulierte Invarianzprinzip für abhängige Prozesse genutzt werden. Dieses findet sich als Theorem F.4 in Anhang F. Der hier verwendete Detektor weist eine ähnliche Grundstruktur auf wie der in Wied et al. (2012b) zur retrospektiven Detektion von Korrelationsänderungen verwendete. Somit kann analog zum Vorgehen im Beweis zu Lemma A.1 in Wied et al. (2012b) gezeigt werden, dass falls die Länge des historischen Zeitraums gegen  $\infty$  strebt, für festes  $c \geq 0$  gilt:

$$\{P_m(d), d \in [c, B]\} \xrightarrow{d} \{W_p(d) - W_p(c), d \in [c, B]\}. \quad (\text{A.1})$$

Dabei ist  $W_p(\cdot)$  eine  $p$ -dimensionale Standard-Brownsche Bewegung beziehungsweise ein  $p$ -variater Prozess dessen einzelne univariate Komponentenprozesse jeweils unabhängige univariate Standard-Brownsche Bewegungen sind. Wied et al. (2012b) argumentieren dabei, dass die Bedingungen für das Invarianzprinzip in Davidson (1994) erfüllt sind, wenn

$$c_{ni}^\lambda = \left( \frac{1}{n} \lambda' D_p \lambda \right)^{\frac{1}{2}} \quad \text{und} \quad \Omega = D_p$$

gewählt wird und außerdem die Annahmen 3.2-3.4 erfüllt sind. Des Weiteren wird die Funktion  $K_n$  so gewählt, dass jedem Anteil  $t$  ein fester Zeitpunkt  $[mt]$  im Überwachungszeitraum zugeordnet wird.

Für den sequenziellen Fall legen Wied und Galeano (2013) detailliert dar, dass das *Continuous Mapping Theorem* unter Verwendung eines passend gewählten stetigen Funktionals für  $b \in [0, B]$  die Konvergenzaussage

$$S_m(b) := \begin{pmatrix} \widehat{D}_{p,m}^{-1/2} \frac{[m \cdot b] + 2}{\sqrt{m}} \left( [\widehat{\sigma}^2]_{m+1}^{m+[m \cdot b] + 2} - \sigma^2 \right) \\ \widehat{D}_{p,m}^{-1/2} \sqrt{m} \left( [\widehat{\sigma}^2]_1^m - \sigma^2 \right) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} W_p(b+1) - W_p(1) \\ W_p(1) \end{pmatrix}$$

liefert. Als Konsequenz strebt auch der Prozess der Detektorwerte

$$V_{[m \cdot b] + 2} = \widehat{D}_{p,m}^{-1/2} \frac{[m \cdot b] + 2}{\sqrt{m}} \left( [\widehat{\sigma}^2]_{m+1}^{m+[m \cdot b] + 2} - \sigma^2 \right) - \widehat{D}_{p,m}^{-1/2} \frac{[m \cdot b] + 2}{\sqrt{m}} \left( [\widehat{\sigma}^2]_1^m - \sigma^2 \right)$$

in Verteilung gegen den Prozess  $\{W_p(b+1) - (b+1)W_p(1), b \in [0, B]\}$ . Eine weitere Anwendung des *Continuous Mapping Theorems* zur Skalierung durch die Werte der Grenzfunktion  $w(\cdot)$  liefert dann die Aussage von Theorem 3.1. Dabei kann die Kovarianzstruktur durch einfaches Nachrechnen der Eigenschaften des resultierenden Gaußprozesses ermittelt beziehungsweise nachgewiesen werden. ■

## A.2 Beweis von Theorem 3.2

Der Beweis nutzt die gleichen Argumente wie der Beweis von Theorem 3.1. Der Hauptunterschied besteht darin, dass die Aussage (A.1) unter Alternativszenarien nicht mehr gilt. Unter der Annahme einer Folge von lokalen Alternativen gilt, falls die Länge des historischen Zeitraums gegen  $\infty$  strebt, für festes  $c \geq 0$  und auf dem Raum  $D[c, B]$  die folgende Konvergenzaussage:

$$\{P_m(d), d \in [c, B]\} \xrightarrow{d} \left\{ W_p(d) - W_p(c) + H \begin{pmatrix} \int_c^d g_1(z) dz \\ \vdots \\ \int_c^d g_p(z) dz \end{pmatrix}, d \in [c, B] \right\} \quad (\text{A.2})$$

Wie in Wied und Galeano (2013) entspricht dabei die Konstante  $H$  bis auf eine Multiplikationskonstante dem Grenzwert der geschätzten Variationsmatrix  $\widehat{D}_{p,m}$  unter der Nullhypothese. Details können dem Beweis von Theorem 3.1 in Wied et al. (2012a) entnommen werden. Da die Folge der lokalen Alternativen dort sehr ähnlich gestaltet wurde, folgt der Beweis des in dieser Arbeit präsentierten Theorems sehr eng dem für Theorem 3.1 in Wied et al. (2012a) konstruierten Beweis. Im Einzelnen wird dort gezeigt, dass die einzelnen Komponenten von  $\{P_m(d), d \in [c, B]\}$  unter lokalen Alternativen nicht mehr gegen unabhängige Brownsche Bewegungen streben, sondern gegen durch einen additiven Term ergänzte Brownsche Bewegungen. Dieser additive Term hängt vom Grenzwert der geschätzten Variationsmatrix  $\widehat{D}_{p,m}$  und außerdem von den strukturdefinierenden Funktionen  $g_j(\cdot)$ , für  $j = 1, \dots, p$ , ab. Das *Continuous Mapping Theorem* kann anschließend genutzt werden, um die für die univariaten Einzelprozesse gültigen Aussagen auf den multivariaten Prozess zu übertragen. Außerdem können alle im Beweis von Theorem 3.1 verwendeten Transformationsschritte auch auf den Prozess (A.2) angewandt werden. Das Ergebnis bildet die in Theorem 3.2 präsentierte und durch einen Vektor von Konstanten additiv ergänzte Gaußprozess. ■

### A.3 Beweis von Theorem 3.3

O.B.d.A. wird angenommen, dass die erste Vektorkomponente von einer Varianzänderung betroffen ist, deren Umfang gegen  $\infty$  strebt. Das heißt, es existiert ein  $g^*(\cdot) \in \mathcal{F}_B$ , für das gilt:

$$g_1(\cdot) := Mg^*(\cdot).$$

Dann gilt für die Grenzverteilung des Detektors bezüglich  $m \rightarrow \infty$ :

$$\sup_{b \in [0, B]} \|V_{[mb]+2}\|_2 \xrightarrow{d} \sup_{b \in [0, B]} \left\| \begin{pmatrix} G_1(b) \\ \vdots \\ G_p(b) \end{pmatrix} + \mathbf{D}^{-\frac{1}{2}} \begin{pmatrix} M \int_1^{b+1} g^*(u) du \\ \int_1^{b+1} g_2(u) du \\ \vdots \\ \int_1^{b+1} g_p(u) du \end{pmatrix} \right\|_2. \quad (\text{A.3})$$

Im Folgenden wird die Bezeichnung  $\mathbf{D}^{-\frac{1}{2}} := (d_{ij})_{i,j=1,\dots,p}$ , verwendet. Außerdem werden in Abhängigkeit von den Funktionen, die die Struktur der Alternativen bestimmen, die Konstanten

$$c_1(b) := \int_1^{b+1} g^*(u) du \quad \text{und} \quad c_i(b) := \int_1^{b+1} g_i(u) du, \quad \text{für } i = 2, \dots, p,$$

definiert. Somit besitzt der Detektor auf der linken Seite von (A.3) asymptotisch die gleiche Verteilung wie

$$\sup_{b \in [0, B]} \sqrt{\sum_{j=1}^p \left( G_j(b) + Md_{j1}c_1(b) + \sum_{i=2}^p d_{ji}c_i(b) \right)^2}. \quad (\text{A.4})$$

Da  $\mathbf{D}_p$  eine positiv definite Matrix ist, existiert mindestens ein  $j \in \{1, \dots, p\}$  für das gilt:

$$d_{j1} \neq 0.$$

Unter der Annahme, dass  $M$  gegen  $\infty$  strebt, gilt auch:

$$|G_j(b) + Md_{j1}c_1(b)| \xrightarrow{P} \infty.$$

Dann strebt mit der Jensenungleichung auch die Quadratwurzel des Terms in (A.4) gegen  $\infty$ .

Die Tatsache, dass der Term in (A.4) für  $b = 0$  degeneriert, beeinträchtigt dieses Resultat nicht, da das zugehörige Ereignis Teil einer Nullmenge ist. Dies impliziert, dass der Term in (A.4) fast sicher jedes Quantil der asymptotischen Verteilung unter der Nullhypothese überschreitet, wenn  $M$  gegen  $\infty$  strebt. ■

## A.4 Beweis von Lemma 3.1

Hafner (2003) liefert Bedingungen für die Existenz der Matrix der unbedingten vierten Momente und Kreuzmomente für ein multivariates GARCH(1, 1)-Modell in vech-Darstellung:

$$\text{vech}(H_t) = C_0 + A_1 \text{vech}(X_{t-1} X'_{t-1}) + B_1 \text{vech}(H_{t-1}) \quad (\text{A.5})$$

mit  $C_0$  einem  $p^+$ -dimensionalen Parametervektor. Außerdem handelt es sich bei  $A_1$  und  $B_1$  um  $(p^+ \times p^+)$ -Parametermatrizen. Die mit der Variante (A.5) nah verwandte vec-Darstellung ergibt sich dann als

$$\text{vec}(H_t) = C_0^* + A_1^* \text{vec}(X_{t-1} X'_{t-1}) + B_1^* \text{vec}(H_{t-1}). \quad (\text{A.6})$$

Die Darstellung (A.6) enthält zahlreiche überflüssige Gleichungen, die zu stark aufgeblähten Parametermatrizen  $A_1^*$  und  $B_1^*$  der Dimension  $(p^2 \times p^2)$  sowie einem Parametervektor  $C_0^*$  der Länge  $p^2$  führt. Laut Engle und Kroner (1995) kann das Modell in (3.8) in vec-Darstellung angegeben werden, indem der Parametervektor und die Parametermatrizen in der Gleichung (A.6) gewählt werden als

$$C_0^* = (1 - \alpha - \beta) \left[ H^{\frac{1}{2}} \otimes H^{\frac{1}{2}} \right] \text{vec}(\mathbb{I}_p), \quad A_1^* = \alpha \mathbb{I}_{p^2} \quad \text{und} \quad B_1^* = \beta \mathbb{I}_{p^2}. \quad (\text{A.7})$$

Somit kann (3.8) in vech-Form dargestellt werden, indem die Terme zunächst in die vec- und dann in die vech-Darstellung überführt werden. Werden die Terme aus (A.7) in die Gleichung (A.6) eingesetzt und die Matrizen  $D_p$  und  $L_p$  multipliziert, ergibt sich für das Modell (A.5):

$$C_0 = (1 - \alpha - \beta) L_p \left[ H^{\frac{1}{2}} \otimes H^{\frac{1}{2}} \right] \text{vec}(\mathbb{I}_p), \quad A_1 = \alpha \mathbb{I}_{p^+} \quad \text{und} \quad B_1 = \beta \mathbb{I}_{p^+}. \quad (\text{A.8})$$

Außerdem gilt:

$$A_1 \otimes A_1 = \alpha^2 \mathbb{I}_{[p^+]^2}, \quad B_1 \otimes B_1 = \beta^2 \mathbb{I}_{[p^+]^2} \quad \text{und} \quad A_1 \otimes B_1 = B_1 \otimes A_1 = \alpha \beta \mathbb{I}_{[p^+]^2}.$$

Werden (A.8) und  $G_p$  aus Hafner (2003) genutzt, um die Matrix  $Z$  aufzustellen, so kann die letztere dazu genutzt werden, um die Existenz von  $\Gamma$  mit Hilfe von Theorem 3 in Hafner (2003) zu überprüfen, was den Beweis vervollständigt. ■

## Anhang B

# Simulationsergebnisse aus Kap. 3.3

### B.1 U.i.v. Zufallsvektoren unter der Nullhypothese

		$p = 2$			$p = 5$			$p = 10$				
		$\gamma$	$B$	$m = 500$	$1.000$	$2.000$	$m = 500$	$1.000$	$2.000$	$m = 500$	$1.000$	$2.000$
$D$ bekannt	$0$	$0.5$	$0.5$	0.048	0.043	0.045	0.049	0.052	0.054	0.050	0.050	0.052
		$1$	$1$	0.043	0.047	0.050	0.049	0.050	0.050	0.049	0.049	0.047
		$2$	$2$	0.046	0.044	0.051	0.051	0.052	0.054	0.053	0.053	0.049
	$0.25$	$0.5$	$0.5$	0.052	0.049	0.043	0.055	0.048	0.052	0.050	0.050	0.050
		$1$	$1$	0.049	0.047	0.050	0.047	0.047	0.046	0.047	0.049	0.049
		$2$	$2$	0.052	0.046	0.048	0.053	0.051	0.054	0.053	0.052	0.052
$\Sigma = I_p$	$0$	$0.5$	$0.5$	0.066	0.056	0.057	0.074	0.064	0.061	0.104	0.073	0.065
		$1$	$1$	0.061	0.061	0.053	0.080	0.066	0.059	0.108	0.078	0.065
		$2$	$2$	0.069	0.061	0.060	0.085	0.066	0.057	0.102	0.083	0.063
	$0.25$	$0.5$	$0.5$	0.064	0.061	0.055	0.084	0.068	0.061	0.103	0.075	0.063
		$1$	$1$	0.067	0.062	0.057	0.088	0.067	0.058	0.114	0.088	0.076
		$2$	$2$	0.070	0.056	0.054	0.082	0.075	0.061	0.113	0.081	0.067
$\Sigma = \Sigma_p$	$0$	$0.5$	$0.5$	0.064	0.059	0.057	0.079	0.065	0.061	0.101	0.074	0.065
		$1$	$1$	0.063	0.061	0.058	0.085	0.067	0.061	0.108	0.081	0.070
		$2$	$2$	0.066	0.059	0.059	0.084	0.074	0.064	0.102	0.079	0.064
	$0.25$	$0.5$	$0.5$	0.062	0.067	0.058	0.084	0.073	0.065	0.113	0.082	0.068
		$1$	$1$	0.065	0.057	0.052	0.090	0.080	0.065	0.113	0.085	0.076
		$2$	$2$	0.070	0.064	0.051	0.093	0.066	0.069	0.119	0.095	0.069

TABELLE B.1: Empirisches Niveau bei der Überwachung von Folgen von unabhängigen und identisch  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -verteilten Zufallsvektoren.

		$p = 2$			$p = 5$			$p = 10$				
		$\gamma$	$B$	$m = 500$	$1.000$	$2.000$	$m = 500$	$1.000$	$2.000$	$m = 500$	$1.000$	$2.000$
$D$ bekannt	$0$	$0.5$	$0.5$	0.049	0.051	0.052	0.071	0.061	0.061	0.078	0.072	0.070
		$1$	$1$	0.052	0.051	0.052	0.063	0.063	0.064	0.077	0.070	0.069
		$2$	$2$	0.053	0.053	0.056	0.063	0.063	0.064	0.075	0.077	0.072
	$0.25$	$0.5$	$0.5$	0.063	0.059	0.058	0.083	0.075	0.066	0.107	0.091	0.091
		$1$	$1$	0.064	0.060	0.059	0.081	0.076	0.071	0.095	0.091	0.081
		$2$	$2$	0.063	0.062	0.053	0.085	0.075	0.071	0.103	0.091	0.087
$\Sigma = I_p$	$0$	$0.5$	$0.5$	0.091	0.080	0.071	0.136	0.107	0.083	0.194	0.143	0.105
		$1$	$1$	0.094	0.082	0.066	0.141	0.105	0.083	0.204	0.139	0.107
		$2$	$2$	0.103	0.080	0.070	0.148	0.109	0.084	0.201	0.138	0.096
	$0.25$	$0.5$	$0.5$	0.110	0.091	0.078	0.163	0.124	0.099	0.233	0.164	0.117
		$1$	$1$	0.120	0.091	0.082	0.171	0.128	0.102	0.249	0.181	0.132
		$2$	$2$	0.108	0.097	0.074	0.174	0.126	0.096	0.249	0.167	0.121
$\Sigma = \Sigma_p$	$0$	$0.5$	$0.5$	0.097	0.082	0.072	0.138	0.107	0.090	0.201	0.138	0.101
		$1$	$1$	0.100	0.079	0.069	0.141	0.110	0.085	0.211	0.146	0.106
		$2$	$2$	0.098	0.079	0.075	0.141	0.111	0.086	0.196	0.141	0.101
	$0.25$	$0.5$	$0.5$	0.114	0.092	0.084	0.162	0.121	0.092	0.247	0.171	0.116
		$1$	$1$	0.114	0.088	0.071	0.185	0.132	0.103	0.252	0.166	0.122
		$2$	$2$	0.117	0.092	0.072	0.180	0.133	0.104	0.270	0.183	0.121

TABELLE B.2: Empirisches Niveau bei der Überwachung von Folgen von unabhängigen und identisch  $t_{p,\nu}(\mathbf{0}_p, \frac{\nu-2}{\nu}\mathbb{I}_p)$ -verteilten Zufallsvektoren mit  $\nu=8$  Freiheitsgraden.

## B.2 U.i.v. Zufallsvektoren unter verschiedenen Alternativen

Szenario 1: Alle Varianzen ändern sich

		$k^* = 0.05$						$k^* = 0.5$							
		$\gamma$	$B$	$m = 500$		1.000		2.000		$m = 500$		1.000		2.000	
$p = 2$	0	0.5	0.5	0.778	0.667	0.960	0.894	1.000	0.994	0.340	0.228	0.513	0.339	0.767	0.524
		1	0.892	0.796	0.994	0.974	1.000	1.000	0.417	0.280	0.659	0.426	0.910	0.682	
		2	0.955	0.897	0.999	0.993	1.000	1.000	0.512	0.334	0.774	0.525	0.970	0.816	
	0.25	0.5	0.766	0.648	0.955	0.882	1.000	0.993	0.314	0.207	0.485	0.312	0.748	0.499	
		1	0.888	0.785	0.992	0.964	1.000	1.000	0.393	0.260	0.610	0.375	0.889	0.634	
		2	0.943	0.869	0.999	0.990	1.000	1.000	0.459	0.289	0.736	0.477	0.960	0.784	
$p = 5$	0	0.5	0.957	0.893	0.999	0.991	1.000	1.000	0.539	0.356	0.756	0.510	0.964	0.787	
		1	0.993	0.967	1.000	1.000	1.000	1.000	0.668	0.443	0.898	0.669	0.997	0.926	
		2	0.999	0.993	1.000	1.000	1.000	1.000	0.754	0.525	0.959	0.779	0.999	0.978	
	0.25	0.5	0.955	0.885	0.999	0.990	1.000	1.000	0.516	0.334	0.743	0.493	0.955	0.750	
		1	0.992	0.963	1.000	1.000	1.000	1.000	0.641	0.414	0.875	0.629	0.995	0.904	
		2	0.999	0.992	1.000	1.000	1.000	1.000	0.729	0.489	0.951	0.751	0.999	0.973	
$p = 10$	0	0.5	1.000	0.998	1.000	1.000	1.000	1.000	0.846	0.636	0.982	0.852	1.000	0.987	
		1	1.000	1.000	1.000	1.000	1.000	1.000	0.949	0.775	0.999	0.953	1.000	0.999	
		2	1.000	1.000	1.000	1.000	1.000	1.000	0.982	0.855	1.000	0.989	1.000	1.000	
	0.25	0.5	1.000	0.997	1.000	1.000	1.000	1.000	0.827	0.608	0.974	0.817	1.000	0.984	
		1	1.000	1.000	1.000	1.000	1.000	1.000	0.935	0.743	0.999	0.943	1.000	0.999	
		2	1.000	1.000	1.000	1.000	1.000	1.000	0.977	0.850	1.000	0.987	1.000	1.000	

TABELLE B.3: Güte bei der Überwachung von Folgen von unabhängigen und identisch  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -verteilten Zufallsvektoren. Schwarz: sprunghafte Änderung; Blau: stetige Änderung.

		$k^* = 0.05$						$k^* = 0.5$							
		$\gamma$	$B$	$m = 500$		1.000		2.000		$m = 500$		1.000		2.000	
$p = 2$	0	0.5	0.606	0.508	0.799	0.706	0.964	0.908	0.287	0.222	0.380	0.264	0.569	0.379	
		1	0.738	0.636	0.911	0.836	0.992	0.975	0.357	0.260	0.484	0.333	0.702	0.474	
		2	0.817	0.729	0.957	0.903	0.998	0.993	0.405	0.283	0.550	0.372	0.802	0.573	
	0.25	0.5	0.615	0.507	0.788	0.683	0.955	0.891	0.282	0.222	0.358	0.245	0.525	0.340	
		1	0.734	0.622	0.898	0.810	0.990	0.967	0.348	0.256	0.437	0.297	0.665	0.434	
		2	0.803	0.705	0.951	0.890	0.998	0.992	0.378	0.268	0.518	0.345	0.773	0.530	
$p = 5$	0	0.5	0.792	0.708	0.935	0.864	0.997	0.984	0.423	0.330	0.527	0.380	0.729	0.511	
		1	0.900	0.830	0.986	0.955	1.000	0.998	0.523	0.393	0.661	0.472	0.875	0.666	
		2	0.952	0.898	0.997	0.983	1.000	1.000	0.594	0.438	0.758	0.545	0.935	0.773	
	0.25	0.5	0.806	0.712	0.936	0.860	0.996	0.980	0.425	0.336	0.515	0.370	0.701	0.480	
		1	0.900	0.825	0.984	0.950	1.000	0.998	0.509	0.390	0.641	0.454	0.850	0.634	
		2	0.949	0.888	0.996	0.981	1.000	1.000	0.576	0.420	0.741	0.529	0.923	0.748	
$p = 10$	0	0.5	0.933	0.879	0.989	0.969	1.000	0.999	0.621	0.483	0.707	0.525	0.886	0.682	
		1	0.976	0.951	0.999	0.994	1.000	1.000	0.702	0.556	0.851	0.659	0.964	0.823	
		2	0.993	0.981	1.000	0.999	1.000	1.000	0.783	0.612	0.914	0.755	0.990	0.910	
	0.25	0.5	0.936	0.878	0.988	0.965	1.000	0.999	0.612	0.480	0.683	0.498	0.877	0.665	
		1	0.978	0.951	0.999	0.992	1.000	1.000	0.696	0.556	0.829	0.629	0.958	0.803	
		2	0.995	0.977	1.000	0.999	1.000	1.000	0.768	0.614	0.897	0.718	0.991	0.898	

TABELLE B.4: Güte bei der Überwachung von Folgen von unabhängigen und identisch  $t_{p,\nu}(\mathbf{0}_p, \frac{\nu-2}{\nu}\mathbb{I}_p)$ -verteilten Zufallsvektoren mit  $\nu=8$  Freiheitsgraden überwacht werden. Schwarz: sprunghafte Änderung; Blau: stetige Änderung.

Szenario 2: Nur eine der Varianzen ändert sich

		$\gamma$	$B$	$k^* = 0.05$				$k^* = 0.5$						
				$m = 500$		1.000		2.000		$m = 500$		1.000		2.000
$p = 2$	0	0.5	0.970	0.930	0.999	0.998	1.000	1.000	0.601	0.404	0.851	0.636	0.985	0.892
		1	0.997	0.984	1.000	1.000	1.000	1.000	0.748	0.515	0.957	0.789	1.000	0.978
		2	1.000	0.998	1.000	1.000	1.000	1.000	0.862	0.638	0.989	0.899	1.000	0.997
	0.25	0.5	0.966	0.920	0.999	0.997	1.000	1.000	0.567	0.367	0.825	0.597	0.983	0.879
		1	0.997	0.981	1.000	1.000	1.000	1.000	0.717	0.482	0.952	0.773	0.999	0.971
		2	1.000	0.997	1.000	1.000	1.000	1.000	0.828	0.585	0.986	0.875	1.000	0.995
$p = 5$	0	0.5	0.931	0.858	0.997	0.988	1.000	1.000	0.483	0.326	0.741	0.488	0.952	0.777
		1	0.988	0.959	1.000	1.000	1.000	1.000	0.632	0.415	0.897	0.661	0.995	0.928
		2	1.000	0.991	1.000	1.000	1.000	1.000	0.743	0.501	0.962	0.790	1.000	0.979
	0.25	0.5	0.924	0.844	0.997	0.985	1.000	1.000	0.457	0.295	0.714	0.460	0.944	0.757
		1	0.987	0.953	1.000	1.000	1.000	1.000	0.614	0.395	0.871	0.615	0.994	0.914
		2	0.999	0.989	1.000	1.000	1.000	1.000	0.719	0.472	0.953	0.758	0.999	0.972
$p = 10$	0	0.5	0.851	0.748	0.988	0.952	1.000	1.000	0.389	0.256	0.584	0.367	0.875	0.633
		1	0.960	0.900	0.999	0.994	1.000	1.000	0.514	0.342	0.763	0.501	0.975	0.813
		2	0.990	0.953	1.000	1.000	1.000	1.000	0.613	0.409	0.889	0.640	0.996	0.921
	0.25	0.5	0.854	0.744	0.988	0.950	1.000	1.000	0.381	0.254	0.573	0.355	0.857	0.596
		1	0.958	0.889	0.999	0.993	1.000	1.000	0.490	0.324	0.745	0.477	0.970	0.790
		2	0.989	0.948	1.000	1.000	1.000	1.000	0.584	0.382	0.873	0.611	0.995	0.908

TABELLE B.5: Güte bei der Überwachung von Folgen von unabhängigen und identisch  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -verteilten Zufallsvektoren. Schwarz: sprunghafte Änderung; Blau: stetige Änderung.

		$\gamma$	$B$	$k^* = 0.05$				$k^* = 0.5$						
				$m = 500$		1.000		2.000		$m = 500$		1.000		2.000
$p = 2$	0	0.5	0.877	0.799	0.986	0.961	0.999	0.999	0.470	0.330	0.692	0.483	0.913	0.729
		1	0.960	0.916	0.998	0.992	1.000	1.000	0.599	0.413	0.831	0.608	0.980	0.871
		2	0.987	0.964	1.000	0.998	1.000	1.000	0.701	0.496	0.917	0.734	0.997	0.951
	0.25	0.5	0.872	0.788	0.984	0.954	0.999	0.999	0.452	0.321	0.660	0.454	0.897	0.693
		1	0.9564	0.906	0.997	0.990	1.000	1.000	0.578	0.397	0.812	0.577	0.973	0.843
		2	0.985	0.955	1.000	0.997	1.000	1.000	0.669	0.462	0.902	0.702	0.995	0.940
$p = 5$	0	0.5	0.826	0.735	0.971	0.924	1.000	0.997	0.416	0.316	0.589	0.397	0.843	0.609
		1	0.933	0.862	0.996	0.985	1.000	1.000	0.522	0.371	0.752	0.540	0.951	0.786
		2	0.975	0.935	0.999	0.996	1.000	1.000	0.630	0.434	0.860	0.662	0.985	0.886
	0.25	0.5	0.830	0.737	0.969	0.917	1.000	0.996	0.421	0.327	0.567	0.381	0.828	0.587
		1	0.931	0.855	0.995	0.980	1.000	1.000	0.510	0.369	0.720	0.502	0.942	0.765
		2	0.973	0.927	0.999	0.995	1.000	1.000	0.614	0.424	0.838	0.627	0.983	0.875
$p = 10$	0	0.5	0.751	0.661	0.924	0.853	0.996	0.984	0.397	0.318	0.492	0.339	0.716	0.490
		1	0.875	0.807	0.982	0.952	1.000	0.999	0.482	0.377	0.640	0.432	0.872	0.642
		2	0.947	0.881	0.996	0.985	1.000	0.999	0.561	0.424	0.752	0.539	0.950	0.770
	0.25	0.5	0.761	0.671	0.921	0.846	0.995	0.981	0.415	0.341	0.484	0.340	0.685	0.461
		1	0.876	0.801	0.982	0.951	1.000	0.998	0.483	0.388	0.635	0.434	0.855	0.618
		2	0.946	0.879	0.995	0.982	1.000	0.999	0.566	0.433	0.736	0.525	0.942	0.745

TABELLE B.6: Güte bei der Überwachung von Folgen von unabhängigen und identisch  $t_{p,\nu}(\mathbf{0}_p, \frac{\nu-2}{\nu}\mathbb{I}_p)$ -verteilten Zufallsvektoren mit  $\nu=8$  Freiheitsgraden. Schwarz: sprunghafte Änderung; Blau: stetige Änderung.

### B.3 GARCH-Prozesse unter der Nullhypothese

$\gamma$	B	$p = 2$			$p = 5$			$p = 10$		
		$m = 500$	1.000	2.000	$m = 500$	1.000	2.000	$m = 500$	1.000	2.000
Multivariates Verfahren										
0	0.5	0.073	0.069	0.061	0.098	0.078	0.066	0.133	0.103	0.083
	1	0.079	0.071	0.060	0.106	0.079	0.074	0.141	0.105	0.088
	2	0.076	0.072	0.069	0.106	0.086	0.072	0.133	0.097	0.079
0.25	0.5	0.075	0.070	0.058	0.110	0.082	0.073	0.136	0.099	0.080
	1	0.088	0.069	0.067	0.115	0.089	0.079	0.154	0.113	0.093
	2	0.084	0.069	0.060	0.115	0.081	0.073	0.148	0.111	0.091
Univariate Verfahren										
0	0.5	0.065	0.060	0.052	0.091	0.079	0.068	0.103	0.081	0.068
	1	0.071	0.063	0.061	0.091	0.080	0.070	0.110	0.090	0.073
	2	0.065	0.060	0.050	0.096	0.076	0.074	0.095	0.073	0.064
0.25	0.5	0.071	0.062	0.056	0.107	0.088	0.076	0.123	0.088	0.070
	1	0.075	0.070	0.057	0.100	0.077	0.069	0.131	0.099	0.085
	2	0.084	0.074	0.064	0.105	0.086	0.069	0.143	0.107	0.087

TABELLE B.7: Empirisches Niveau bei der Überwachung von skalaren BEKK-Zeitreihen mit Parametern  $\alpha = 0.03$ ,  $\beta = 0.45$  und  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -verteilten Innovationen.

$\gamma$	B	$p = 2$			$p = 5$			$p = 10$		
		$m = 500$	1.000	2.000	$m = 500$	1.000	2.000	$m = 500$	1.000	2.000
Multivariates Verfahren										
0	0.5	0.109	0.088	0.080	0.153	0.121	0.099	0.243	0.167	0.129
	1	0.105	0.087	0.075	0.170	0.129	0.098	0.249	0.182	0.132
	2	0.113	0.094	0.075	0.168	0.132	0.100	0.243	0.166	0.119
0.25	0.5	0.128	0.101	0.083	0.201	0.146	0.117	0.277	0.197	0.142
	1	0.134	0.108	0.089	0.210	0.157	0.109	0.311	0.214	0.157
	2	0.136	0.102	0.086	0.202	0.154	0.116	0.307	0.213	0.153
Univariate Verfahren										
0	0.5	0.091	0.071	0.066	0.144	0.116	0.092	0.183	0.137	0.106
	1	0.100	0.081	0.070	0.145	0.114	0.090	0.203	0.147	0.116
	2	0.099	0.085	0.066	0.151	0.122	0.090	0.194	0.137	0.108
0.25	0.5	0.107	0.092	0.079	0.176	0.148	0.111	0.279	0.210	0.159
	1	0.120	0.102	0.083	0.177	0.139	0.105	0.278	0.215	0.157
	2	0.124	0.102	0.084	0.193	0.148	0.116	0.271	0.196	0.142

TABELLE B.8: Empirisches Niveau bei der Überwachung von skalaren BEKK-Zeitreihen mit Parametern  $\alpha = 0.03$ ,  $\beta = 0.45$  und  $t_{p,\nu}(\mathbf{0}_p, \frac{\nu-2}{\nu}\mathbb{I}_p)$ -verteilten Innovationen mit  $\nu=8$  Freiheitsgraden.

## B.4 GARCH-Prozesse unter verschiedenen Alternativen

Szenario 1: Alle Varianzen ändern sich

		$k^* = 0.05$						$k^* = 0.5$							
		$\gamma$	$B$	$m = 500$		1.000		2.000		$m = 500$		1.000		2.000	
$p = 2$	0	0.5	0.752	0.655	0.938	0.873	0.998	0.990	0.339	0.245	0.487	0.341	0.727	0.513	
		1	0.875	0.782	0.991	0.960	1.000	0.999	0.407	0.288	0.612	0.416	0.872	0.652	
		2	0.939	0.871	0.998	0.988	1.000	1.000	0.494	0.341	0.731	0.520	0.944	0.779	
	0.25	0.5	0.750	0.641	0.931	0.860	0.997	0.987	0.317	0.234	0.458	0.312	0.695	0.476	
		1	0.871	0.761	0.989	0.952	1.000	0.999	0.383	0.264	0.576	0.382	0.852	0.615	
		2	0.927	0.852	0.998	0.985	1.000	1.000	0.459	0.311	0.694	0.480	0.932	0.743	
$p = 5$	0	0.5	0.942	0.876	0.997	0.989	1.000	1.000	0.511	0.378	0.719	0.514	0.941	0.765	
		1	0.990	0.963	1.000	0.999	1.000	1.000	0.653	0.459	0.876	0.643	0.990	0.901	
		2	0.998	0.988	1.000	1.000	1.000	1.000	0.763	0.533	0.946	0.763	0.999	0.962	
	0.25	0.5	0.940	0.865	0.997	0.987	1.000	1.000	0.484	0.354	0.683	0.480	0.925	0.730	
		1	0.990	0.959	1.000	0.999	1.000	1.000	0.625	0.436	0.861	0.615	0.988	0.884	
		2	0.997	0.985	1.000	1.000	1.000	1.000	0.728	0.501	0.931	0.731	0.999	0.954	
$p = 10$	0	0.5	1.000	0.996	1.000	1.000	1.000	1.000	0.829	0.633	0.969	0.823	1.000	0.980	
		1	1.000	1.000	1.000	1.000	1.000	1.000	0.934	0.764	0.997	0.941	1.000	0.999	
		2	1.000	1.000	1.000	1.000	1.000	1.000	0.979	0.849	1.000	0.980	1.000	1.000	
	0.25	0.5	1.000	0.996	1.000	1.000	1.000	1.000	0.811	0.609	0.962	0.803	1.000	0.974	
		1	1.000	1.000	1.000	1.000	1.000	1.000	0.923	0.742	0.996	0.929	1.000	0.998	
		2	1.000	1.000	1.000	1.000	1.000	1.000	0.973	0.832	1.000	0.977	1.000	1.000	

TABELLE B.9: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit  $\alpha = 0.03$ ,  $\beta = 0.45$  und  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -verteilten Innovationen. Schwarz: sprunghafte Änderung; Blau: stetige Änderung.

		$k^* = 0.05$						$k^* = 0.5$							
		$\gamma$	$B$	$m = 500$		1.000		2.000		$m = 500$		1.000		2.000	
$p = 2$	0	0.5	0.600	0.504	0.787	0.678	0.951	0.879	0.311	0.223	0.378	0.253	0.539	0.351	
		1	0.726	0.618	0.895	0.804	0.989	0.958	0.363	0.260	0.469	0.314	0.669	0.445	
		2	0.812	0.705	0.947	0.884	0.997	0.988	0.402	0.292	0.560	0.375	0.785	0.559	
	0.25	0.5	0.612	0.517	0.782	0.681	0.946	0.876	0.308	0.242	0.361	0.260	0.506	0.335	
		1	0.724	0.621	0.887	0.796	0.986	0.956	0.352	0.274	0.443	0.312	0.634	0.427	
		2	0.804	0.695	0.941	0.876	0.997	0.985	0.382	0.284	0.526	0.358	0.751	0.527	
$p = 5$	0	0.5	0.788	0.702	0.928	0.855	0.992	0.976	0.445	0.348	0.530	0.392	0.712	0.503	
		1	0.891	0.817	0.980	0.947	1.000	0.997	0.530	0.419	0.655	0.471	0.844	0.641	
		2	0.940	0.882	0.995	0.977	1.000	1.000	0.600	0.453	0.745	0.546	0.927	0.749	
	0.25	0.5	0.799	0.710	0.929	0.852	0.992	0.972	0.447	0.365	0.512	0.380	0.692	0.475	
		1	0.892	0.819	0.978	0.935	1.000	0.996	0.521	0.408	0.630	0.459	0.821	0.619	
		2	0.940	0.884	0.994	0.975	1.000	0.999	0.587	0.456	0.722	0.538	0.914	0.724	
$p = 10$	0	0.5	0.928	0.877	0.986	0.961	1.000	0.999	0.638	0.528	0.708	0.567	0.870	0.686	
		1	0.978	0.940	0.998	0.993	1.000	1.000	0.728	0.577	0.836	0.649	0.959	0.811	
		2	0.990	0.980	1.000	0.998	1.000	1.000	0.795	0.648	0.907	0.748	0.988	0.901	
	0.25	0.5	0.933	0.876	0.985	0.955	1.000	0.998	0.632	0.519	0.692	0.536	0.855	0.662	
		1	0.978	0.946	0.998	0.993	1.000	1.000	0.712	0.577	0.817	0.647	0.949	0.798	
		2	0.990	0.973	1.000	0.998	1.000	1.000	0.786	0.640	0.895	0.711	0.985	0.884	

TABELLE B.10: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit  $\alpha = 0.03$ ,  $\beta = 0.45$  und  $t_{p,\nu}(\mathbf{0}_p, \frac{\nu-2}{\nu}\mathbb{I}_p)$ -verteilten Innovationen mit  $\nu=8$  Freiheitsgraden. Schwarz: sprunghafte Änderung; Blau: stetige Änderung.

## Szenario 2: Nur eine der Varianzen ändert sich

		$\gamma$	$B$	$k^* = 0.05$			$k^* = 0.5$		
				$m = 500$	1.000	2.000	$m = 500$	1.000	2.000
$p = 2$	0	0.5	0.958 0.906	0.999 0.995	1.000 1.000	0.564 0.394	0.828 0.603	0.978 0.862	
		1	0.994 0.979	1.000 1.000	1.000 1.000	0.712 0.507	0.936 0.768	0.999 0.965	
		2	0.999 0.995	1.000 1.000	1.000 1.000	0.836 0.614	0.981 0.881	1.000 0.994	
	0.25	0.5	0.955 0.898	0.999 0.993	1.000 1.000	0.541 0.375	0.800 0.561	0.973 0.844	
		1	0.992 0.972	1.000 1.000	1.000 1.000	0.686 0.4710	0.925 0.732	0.999 0.958	
		2	0.999 0.995	1.000 1.000	1.000 1.000	0.810 0.586	0.975 0.848	1.000 0.991	
$p = 5$	0	0.5	0.912 0.834	0.995 0.978	1.000 1.000	0.464 0.326	0.698 0.478	0.938 0.752	
		1	0.980 0.945	1.000 0.999	1.000 1.000	0.622 0.417	0.875 0.644	0.991 0.906	
		2	0.996 0.986	1.000 1.000	1.000 1.000	0.729 0.506	0.945 0.764	0.999 0.970	
	0.25	0.5	0.903 0.826	0.995 0.976	1.000 1.000	0.437 0.302	0.665 0.446	0.923 0.730	
		1	0.978 0.936	1.000 0.999	1.000 1.000	0.596 0.393	0.859 0.608	0.989 0.887	
		2	0.996 0.981	1.000 1.000	1.000 1.000	0.700 0.487	0.933 0.733	0.999 0.961	
$p = 10$	0	0.5	0.828 0.739	0.980 0.940	1.000 0.999	0.394 0.286	0.563 0.370	0.846 0.601	
		1	0.948 0.884	0.999 0.992	1.000 1.000	0.506 0.367	0.744 0.507	0.961 0.797	
		2	0.988 0.950	1.000 1.000	1.000 1.000	0.633 0.421	0.862 0.623	0.994 0.897	
	0.25	0.5	0.827 0.723	0.977 0.930	1.000 0.998	0.379 0.278	0.538 0.355	0.823 0.562	
		1	0.946 0.873	0.999 0.988	1.000 1.000	0.491 0.349	0.721 0.489	0.953 0.767	
		2	0.986 0.948	1.000 0.998	1.000 1.000	0.610 0.395	0.840 0.589	0.990 0.876	

TABELLE B.11: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit  $\alpha = 0.03$ ,  $\beta = 0.45$  und  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -verteilten Innovationen. Schwarz: sprunghafte Änderung; Blau: stetige Änderung.

		$\gamma$	$B$	$k^* = 0.05$			$k^* = 0.5$		
				$m = 500$	1.000	2.000	$m = 500$	1.000	2.000
$p = 2$	0	0.5	0.860 0.782	0.978 0.944	1.000 0.998	0.467 0.334	0.656 0.461	0.892 0.701	
		1	0.949 0.900	0.997 0.990	1.000 1.000	0.584 0.426	0.811 0.604	0.972 0.848	
		2	0.984 0.954	0.999 0.997	1.000 1.000	0.688 0.492	0.897 0.716	0.991 0.930	
	0.25	0.5	0.857 0.773	0.976 0.937	1.000 0.998	0.453 0.319	0.630 0.446	0.875 0.672	
		1	0.946 0.884	0.996 0.987	1.000 1.000	0.565 0.400	0.789 0.574	0.964 0.821	
		2	0.982 0.945	0.999 0.997	1.000 1.000	0.666 0.483	0.881 0.691	0.989 0.921	
$p = 5$	0	0.5	0.804 0.714	0.961 0.909	0.999 0.993	0.423 0.318	0.567 0.393	0.813 0.588	
		1	0.914 0.844	0.993 0.976	1.000 1.000	0.534 0.390	0.723 0.509	0.931 0.759	
		2	0.967 0.923	0.999 0.995	1.000 1.000	0.625 0.449	0.832 0.624	0.977 0.868	
	0.25	0.5	0.805 0.717	0.960 0.898	0.999 0.994	0.427 0.330	0.555 0.386	0.795 0.562	
		1	0.911 0.851	0.992 0.977	1.000 0.999	0.532 0.400	0.703 0.514	0.920 0.738	
		2	0.966 0.921	0.999 0.993	1.000 1.000	0.614 0.446	0.818 0.595	0.974 0.846	
$p = 10$	0	0.5	0.744 0.671	0.908 0.835	0.993 0.975	0.431 0.359	0.493 0.367	0.703 0.487	
		1	0.872 0.797	0.978 0.942	1.000 0.998	0.513 0.398	0.639 0.464	0.861 0.638	
		2	0.930 0.874	0.993 0.975	1.000 0.999	0.578 0.443	0.738 0.516	0.933 0.742	
	0.25	0.5	0.751 0.674	0.903 0.824	0.992 0.970	0.448 0.376	0.486 0.365	0.679 0.461	
		1	0.873 0.794	0.975 0.937	1.000 0.996	0.522 0.432	0.621 0.453	0.838 0.615	
		2	0.933 0.869	0.992 0.974	1.000 0.999	0.586 0.466	0.727 0.519	0.924 0.724	

TABELLE B.12: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit  $\alpha = 0.03$ ,  $\beta = 0.45$  und  $t_{p,\nu}(\mathbf{0}_p, \frac{\nu-2}{\nu}\mathbb{I}_p)$ -verteilten Innovationen mit  $\nu=8$  Freiheitsgraden. Schwarz: sprunghafte Änderung; Blau: stetige Änderung.

## B.5 Vergleich der multivariaten Methode mit der Anwendung von mehreren univariaten Verfahren

Szenario 1: Alle Varianzen ändern sich

		$k^* = 0.05$			$k^* = 0.5$			
	$\gamma$	$B$	$m = 500$	$1.000$	$2.000$	$m = 500$	$1.000$	$2.000$
$p = 2$	<b>0</b>	<b>0.5</b>	0.757	0.942	0.999	0.334	0.493	0.753
		<b>1</b>	0.877	0.989	1.000	0.411	0.638	0.893
		<b>2</b>	0.948	0.998	1.000	0.516	0.754	0.958
	<b>0.25</b>	<b>0.5</b>	0.750	0.935	0.998	0.312	0.459	0.720
		<b>1</b>	0.872	0.987	1.000	0.380	0.599	0.869
		<b>2</b>	0.940	0.997	1.000	0.482	0.721	0.949
$p = 5$	<b>0</b>	<b>0.5</b>	0.940	0.997	1.000	0.504	0.716	0.925
		<b>1</b>	0.985	1.000	1.000	0.622	0.853	0.985
		<b>2</b>	0.996	1.000	1.000	0.699	0.910	0.996
	<b>0.25</b>	<b>0.5</b>	0.939	0.996	1.000	0.483	0.686	0.908
		<b>1</b>	0.983	1.000	1.000	0.585	0.820	0.980
		<b>2</b>	0.996	1.000	1.000	0.654	0.884	0.994
$p = 10$	<b>0</b>	<b>0.5</b>	0.995	1.000	1.000	0.714	0.896	0.993
		<b>1</b>	1.000	1.000	1.000	0.768	0.948	0.999
		<b>2</b>	1.000	1.000	1.000	0.881	0.990	1.000
	<b>0.25</b>	<b>0.5</b>	0.995	1.000	1.000	0.685	0.870	0.990
		<b>1</b>	0.999	1.000	1.000	0.735	0.929	0.999
		<b>2</b>	1.000	1.000	1.000	0.843	0.980	1.000

TABELLE B.13: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit  $\alpha = 0.03$ ,  $\beta = 0.45$  und  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -verteilten Innovationen durch univariate Verfahren.

		$k^* = 0.05$			$k^* = 0.5$			
	$\gamma$	$B$	$m = 500$	$1.000$	$2.000$	$m = 500$	$1.000$	$2.000$
$p = 2$	<b>0</b>	<b>0.5</b>	0.571	0.774	0.948	0.262	0.351	0.518
		<b>1</b>	0.716	0.895	0.989	0.348	0.468	0.681
		<b>2</b>	0.805	0.951	0.996	0.405	0.567	0.792
	<b>0.25</b>	<b>0.5</b>	0.585	0.777	0.947	0.269	0.346	0.503
		<b>1</b>	0.706	0.885	0.988	0.334	0.440	0.648
		<b>2</b>	0.784	0.941	0.995	0.375	0.524	0.754
$p = 5$	<b>0</b>	<b>0.5</b>	0.791	0.929	0.996	0.432	0.546	0.725
		<b>1</b>	0.891	0.984	0.999	0.514	0.656	0.866
		<b>2</b>	0.938	0.993	1.000	0.577	0.737	0.916
	<b>0.25</b>	<b>0.5</b>	0.802	0.929	0.995	0.433	0.528	0.704
		<b>1</b>	0.886	0.982	0.999	0.498	0.624	0.839
		<b>2</b>	0.933	0.993	1.000	0.551	0.705	0.895
$p = 10$	<b>0</b>	<b>0.5</b>	0.923	0.987	1.000	0.601	0.713	0.872
		<b>1</b>	0.960	0.998	1.000	0.637	0.781	0.935
		<b>2</b>	0.990	1.000	1.000	0.735	0.874	0.981
	<b>0.25</b>	<b>0.5</b>	0.922	0.987	1.000	0.596	0.691	0.848
		<b>1</b>	0.962	0.997	1.000	0.629	0.762	0.920
		<b>2</b>	0.989	1.000	1.000	0.705	0.845	0.973

TABELLE B.14: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit  $\alpha = 0.03$ ,  $\beta = 0.45$  und  $t_{p,\nu}(\mathbf{0}_p, \frac{\nu-2}{\nu}\mathbb{I}_p)$ -verteilten Innovationen mit  $\nu=8$  Freiheitsgraden durch univariate Verfahren.

## Szenario 2: Nur eine der Varianzen ändert sich

		$k^* = 0.05$			$k^* = 0.5$				
		$\gamma$	$B$	$m = 500$	$1.000$	$2.000$	$m = 500$	$1.000$	$2.000$
$p = 2$	$0$	$0.5$	$0.5$	0.910	0.996	1.000	0.451	0.710	0.945
		$1$	$1$	0.979	1.000	1.000	0.603	0.865	0.995
		$2$	$2$	0.996	1.000	1.000	0.716	0.950	1.000
	$0.25$	$0.5$	$0.5$	0.898	0.994	1.000	0.431	0.685	0.938
		$1$	$1$	0.975	1.000	1.000	0.559	0.839	0.990
		$2$	$2$	0.996	1.000	1.000	0.696	0.936	0.999
$p = 5$	$0$	$0.5$	$0.5$	0.868	0.991	1.000	0.390	0.633	0.905
		$1$	$1$	0.970	1.000	1.000	0.559	0.834	0.986
		$2$	$2$	0.991	1.000	1.000	0.635	0.913	0.998
	$0.25$	$0.5$	$0.5$	0.855	0.987	1.000	0.365	0.587	0.885
		$1$	$1$	0.959	0.999	1.000	0.510	0.795	0.980
		$2$	$2$	0.991	1.000	1.000	0.626	0.902	0.998
$p = 10$	$0$	$0.5$	$0.5$	0.839	0.988	1.000	0.368	0.592	0.882
		$1$	$1$	0.960	0.999	1.000	0.517	0.792	0.979
		$2$	$2$	0.985	1.000	1.000	0.576	0.875	0.995
	$0.25$	$0.5$	$0.5$	0.839	0.986	1.000	0.372	0.575	0.868
		$1$	$1$	0.958	0.999	1.000	0.517	0.779	0.976
		$2$	$2$	0.984	1.000	1.000	0.557	0.856	0.994

TABELLE B.15: Güte bei der Überwachung von skalaren BEKK- Zeitreihen mit  $\alpha = 0.03$ ,  $\beta = 0.45$  und  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -verteilten Innovationen durch univariate Verfahren.

		$k^* = 0.05$			$k^* = 0.5$				
		$\gamma$	$B$	$m = 500$	$1.000$	$2.000$	$m = 500$	$1.000$	$2.000$
$p = 2$	$0$	$0.5$	$0.5$	0.737	0.928	0.996	0.346	0.511	0.765
		$1$	$1$	0.871	0.981	1.000	0.447	0.662	0.895
		$2$	$2$	0.929	0.993	1.000	0.529	0.760	0.953
	$0.25$	$0.5$	$0.5$	0.730	0.920	0.995	0.337	0.488	0.736
		$1$	$1$	0.859	0.977	1.000	0.430	0.633	0.879
		$2$	$2$	0.922	0.992	1.000	0.508	0.735	0.946
$p = 5$	$0$	$0.5$	$0.5$	0.674	0.897	0.991	0.328	0.442	0.688
		$1$	$1$	0.834	0.972	0.998	0.432	0.616	0.866
		$2$	$2$	0.893	0.988	0.999	0.485	0.687	0.922
	$0.25$	$0.5$	$0.5$	0.677	0.892	0.990	0.339	0.432	0.668
		$1$	$1$	0.822	0.966	0.997	0.420	0.583	0.840
		$2$	$2$	0.890	0.987	0.999	0.486	0.671	0.913
$p = 10$	$0$	$0.5$	$0.5$	0.661	0.871	0.990	0.340	0.430	0.647
		$1$	$1$	0.817	0.961	0.998	0.434	0.586	0.823
		$2$	$2$	0.871	0.980	0.999	0.464	0.651	0.898
	$0.25$	$0.5$	$0.5$	0.682	0.872	0.990	0.385	0.446	0.635
		$1$	$1$	0.828	0.960	0.998	0.474	0.595	0.814
		$2$	$2$	0.870	0.978	0.999	0.480	0.642	0.885

TABELLE B.16: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit  $\alpha = 0.03$ ,  $\beta = 0.45$  und  $t_{p,\nu}(\mathbf{0}_p, \frac{\nu-2}{\nu}\mathbb{I}_p)$ -verteilten Innovationen mit  $\nu=8$  Freiheitsgraden durch univariate Verfahren.

## B.6 Alternative, dass sich die Innovationsvarianzen ändern

		$\gamma$	$B$	$k^* = 0.05$			$k^* = 0.5$		
				$m = 500$	$1.000$	$2.000$	$m = 500$	$1.000$	$2.000$
$p = 2$	$\mathbf{0}$	$\mathbf{0.5}$	0.781	0.956	0.999	0.362	0.535	0.782	
		$\mathbf{1}$	0.897	0.993	1.000	0.450	0.682	0.916	
		$\mathbf{2}$	0.956	0.999	1.000	0.534	0.791	0.972	
	$\mathbf{0.25}$	$\mathbf{0.5}$	0.777	0.952	0.999	0.341	0.504	0.755	
		$\mathbf{1}$	0.890	0.991	1.000	0.421	0.644	0.897	
		$\mathbf{2}$	0.950	0.999	1.000	0.497	0.761	0.965	
$p = 5$	$\mathbf{0}$	$\mathbf{0.5}$	0.959	0.999	1.000	0.576	0.780	0.965	
		$\mathbf{1}$	0.991	1.000	1.000	0.706	0.914	0.997	
		$\mathbf{2}$	0.999	1.000	1.000	0.807	0.965	1.000	
	$\mathbf{0.25}$	$\mathbf{0.5}$	0.957	0.999	1.000	0.551	0.754	0.956	
		$\mathbf{1}$	0.991	1.000	1.000	0.682	0.900	0.996	
		$\mathbf{2}$	0.999	1.000	1.000	0.787	0.958	1.000	
$p = 10$	$\mathbf{0}$	$\mathbf{0.5}$	1.000	1.000	1.000	0.872	0.985	1.000	
		$\mathbf{1}$	1.000	1.000	1.000	0.960	0.999	1.000	
		$\mathbf{2}$	1.000	1.000	1.000	0.989	1.000	1.000	
	$\mathbf{0.25}$	$\mathbf{0.5}$	1.000	1.000	1.000	0.858	0.981	1.000	
		$\mathbf{1}$	1.000	1.000	1.000	0.954	0.999	1.000	
		$\mathbf{2}$	1.000	1.000	1.000	0.985	1.000	1.000	

TABELLE B.17: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit  $\alpha = 0.03$ ,  $\beta = 0.45$  und  $\mathcal{N}_p(\mathbf{0}_p, \mathbb{I}_p)$ -verteilten Innovationen.

		$\gamma$	$B$	$k^* = 0.05$			$k^* = 0.5$		
				$m = 500$	$1.000$	$2.000$	$m = 500$	$1.000$	$2.000$
$p = 2$	$\mathbf{0}$	$\mathbf{0.5}$	0.624	0.807	0.961	0.314	0.390	0.572	
		$\mathbf{1}$	0.745	0.911	0.993	0.374	0.505	0.719	
		$\mathbf{2}$	0.839	0.962	0.998	0.444	0.606	0.819	
	$\mathbf{0.25}$	$\mathbf{0.5}$	0.628	0.800	0.956	0.313	0.370	0.542	
		$\mathbf{1}$	0.741	0.904	0.991	0.361	0.476	0.688	
		$\mathbf{2}$	0.831	0.957	0.998	0.423	0.574	0.789	
$p = 5$	$\mathbf{0}$	$\mathbf{0.5}$	0.816	0.942	0.996	0.477	0.570	0.756	
		$\mathbf{1}$	0.912	0.987	1.000	0.563	0.692	0.886	
		$\mathbf{2}$	0.957	0.997	1.000	0.637	0.796	0.951	
	$\mathbf{0.25}$	$\mathbf{0.5}$	0.821	0.940	0.996	0.474	0.549	0.730	
		$\mathbf{1}$	0.914	0.986	1.000	0.556	0.674	0.869	
		$\mathbf{2}$	0.957	0.997	1.000	0.623	0.775	0.940	
$p = 10$	$\mathbf{0}$	$\mathbf{0.5}$	0.945	0.991	1.000	0.666	0.753	0.902	
		$\mathbf{1}$	0.986	1.000	1.000	0.754	0.872	0.972	
		$\mathbf{2}$	0.996	1.000	1.000	0.827	0.929	0.994	
	$\mathbf{0.25}$	$\mathbf{0.5}$	0.949	0.990	1.000	0.669	0.740	0.892	
		$\mathbf{1}$	0.986	1.000	1.000	0.747	0.857	0.967	
		$\mathbf{2}$	0.996	1.000	1.000	0.819	0.919	0.992	

TABELLE B.18: Güte bei der Überwachung von skalaren BEKK-Zeitreihen mit  $\alpha = 0.03$ ,  $\beta = 0.45$  und  $t_{p,\nu}(\mathbf{0}_p, \frac{\nu-2}{\nu}\mathbb{I}_p)$ -verteilten Innovationen mit  $\nu=8$  Freiheitsgraden.



## Anhang C

# Darstellung der DCC-Modellgleichung als Markovkette

Ähnlich zum Beweis von Theorem 2.2 in Ling und McAleer (2003) und zum Vorgehen von Fermanian und Malongo (2017) kann (4.12) als nichtlineare Markovkette dargestellt werden. Um dieses Vorgehen zu veranschaulichen, müssen zunächst einige Terme definiert werden:

**Definition C.1 Total geordnete Menge** (Bartsch (2007), S.9ff)

Sei  $M$  eine Menge. Eine Relation  $R \subseteq M \times M$  heißt total geordnet, wenn sie die folgenden Bedingungen erfüllt:

**B1**  $\forall x \in M : (x, x) \in R$  (Reflexivität)

**B2**  $\forall x, y \in M : (x, y) \in R \wedge (y, x) \in R \Rightarrow x = y$  (Antisymmetrie)

**B3**  $\forall x, y, z \in M : (x, y) \in R \wedge (y, z) \in R \Rightarrow (x, z) \in R$  (Transitivität)

**B4**  $\forall x, y \in M : (x, y) \in R \vee (y, x) \in R$  (Linearität oder Totalität)

**Definition C.2 Lokal kompakter Raum** (Bartsch (2007), S.169f)

Ein metrischer Raum  $(E, \mathcal{B})$  ist lokal kompakt, wenn jedes  $x \in E$  eine kompakte Umgebung in  $E$  besitzt.

Sei im Folgenden  $S$  eine bezüglich der Relation  $\leq$  total geordnete Indexmenge. Sei außerdem  $E$  ein Zustandsraum mit  $\mathcal{B}$  einer Borel- $\sigma$ -Algebra auf  $E$ . Außerdem sei  $(E, \mathcal{B})$  ein lokal kompakter separabler metrischer Raum.

**Definition C.3 Markoveigenschaft** (Bauer (1991), S.365)

Sei  $\{Y_t, t \in \mathbb{Z}\}$  ein beliebiger Prozess und  $\mathbb{F}^Y = \{\mathcal{F}_t^Y, t \in \mathbb{Z}\}$  die an den Prozess adaptierte Filtration. Ein Prozess  $\{X_t, t \in \mathbb{Z}\}$  mit Zustandsraum  $(E, \mathcal{B})$  besitzt dann die Markoveigenschaft bezüglich der Filtration  $\mathbb{F}^Y$ , wenn fast sicher für alle  $i, j \in S$  mit  $i < j$  und alle  $B \in \mathcal{B}$  gilt:

$$\mathbb{P}(X_j \in B | \mathcal{F}_i^Y) = \mathbb{P}(X_j \in B | Y_i).$$

Analog zu Definition 42.1 in Bauer (1991) wird ein Prozess  $\mathbf{X} = \{X_t, t \in \mathbb{Z}\}$  dann als **Markovprozess** oder als **Markovkette** bezeichnet, wenn er die Markoveigenschaft bezüglich der an den Prozess  $\mathbf{X}$  selbst adaptierten Filtration  $\mathbb{F}^{\mathbf{X}}$  besitzt.

**Definition C.4 Übergangswahrscheinlichkeiten**

Für eine Markovkette  $\mathbf{X}$  mit Zustandsraum  $(E, \mathcal{B})$  wird die Einschnitt-Übergangswahrscheinlichkeit von  $x \in E$  in die Menge  $B \in \mathcal{B}$  bezeichnet als

$$P(x, B) = P(X_t \in B | X_{t-1} = x). \quad (\text{C.1})$$

**Definition C.5 Zeithomogene Markovkette** (Kijima (1997), S.16)

Eine Markovkette  $\mathbf{X}$  wird als zeithomogen (temporally homogeneous) bezeichnet, wenn die Übergangswahrscheinlichkeiten in (C.1) vom Zeitpunkt  $t$  unabhängig sind.

**Definition C.6 Halbstetigkeit von unten** (Meyn und Tweedie (2009), S.32)

Eine Funktion  $h(\cdot)$  von  $\mathbf{X}$  auf  $\mathbb{R}$  ist halbstetig von unten, wenn für alle  $x \in \mathbf{X}$  gilt:

$$\liminf_{y \rightarrow x} h(y) \geq h(x).$$

**Definition C.7 Fellerkette** (Meyn und Tweedie (2009), S.124)

Eine Markovkette  $\mathbf{X}$  mit Zustandsraum  $(E, \mathcal{B})$  heißt Fellerkette, wenn der Transitionoperator  $P(\cdot, O)$  für jede offene Menge  $O \in \mathcal{B}$  eine von unten halbstetige Funktion ist.

Die Aussage in Definition C.7 ist äquivalent dazu, dass für jede beschränkte stetige Funktion  $h(\cdot)$  auf  $S$  gilt, dass  $E[h(X_t) | X_{t-1} = x]$  stetig in  $x$  ist.

**Definition C.8 Nichtlineares Zustandsraummodell** (Meyn und Tweedie (2009), S.29)

Sei  $\mathbf{X}$  eine offene Teilmenge von  $\mathbb{R}^k$  und  $O_\omega$  eine offene Teilmenge von  $\mathbb{R}^p$ . Sei außerdem  $F : \mathbf{X} \times O_\omega \rightarrow \mathbf{X}$  eine glatte Funktion, das heißt, dass alle Ableitungen von  $F$  existieren und stetig sind.

Dann heißt ein Prozess  $\mathbf{X} = \{X_n, n \in \mathbb{N}_0\}$  nichtlineares Zustandsraummodell auf  $\mathbb{R}$ , das von  $F$  angetrieben wird mit Kontrollmenge  $O_\omega$ , wenn die folgenden Bedingungen erfüllt sind:

**B1** Für jedes  $n \geq 0$  sind  $X_n$  und  $Y_n$  Zufallsvektoren mit Werten in  $\mathbb{R}^k$  beziehungsweise in  $\mathbb{R}^p$ , für die induktiv gilt:

$$X_n = F(X_{n-1}, Y_n), \text{ für alle } n \geq 1.$$

**B2** Die Folge  $\{Y_n\}$  ist eine Sequenz von Störtermen in  $\mathbb{R}^p$ , deren Randverteilung  $L_Y$  eine Wahrscheinlichkeitsdichte  $f_\omega$  auf einer offenen Menge  $O_\omega$  besitzt.

Im Folgenden wird eine Markovkette  $\mathbf{X}$ , die einem nichtlinearen Zustandsraummodell genügt, kurz **nichtlineare Markovkette** genannt. Meyn und Tweedie (2009) belegen in Proposition 6.1.2, dass eine nichtlineare Markovkette immer auch eine Fellerkette ist.

Nun kann (4.12) als eine nichtlineare Markovkette dargestellt werden. Dazu argumentieren Fermanian und Malongo (2017), dass sowohl  $z_t$  als auch alle Funktionen von  $z_t$  also insbesondere auch  $\text{vecl}(z_t z_t')$  und  $z_{it}^2$ , für  $i = 1, \dots, p$ , Funktionen sind, die einzig von der

bedingten Korrelationsmatrix  $R_t$  und dem bedingten Innovationsvektor  $\eta_t$  abhängen. Gegeben  $X_{t-1}$  aus (4.13) ist  $R_t$  sogar deterministisch. Da sich die Vektoren  $\eta_t$  bedingt auf  $X_{t-1}$  wie gewöhnliche Innovationen verhalten, hängt die Verteilung von  $z_t$  bedingt auf  $\mathcal{I}_t$  nur von  $X_{t-1}$  ab. Fermanian und Malongo (2017) argumentieren, dass dies auch für die Matrix  $T_t$  sowie für die Vektoren  $\zeta_t$  und  $X_t$  gilt und zeigen, dass sich (4.12) so umstellen lässt, dass  $X_t$  eine Funktion von  $z_t$  und  $X_{n-t}$  ist. Da  $R_t$  messbar bezüglich  $\mathcal{I}_t$  beziehungsweise bezüglich  $X_{t-1}$  ist und außerdem  $z_t = R_t^{1/2} \eta_t$  gilt, ist  $X_t$  somit eine Funktion von  $X_{t-1}$  und  $\eta_t$ . Da die  $\eta_t$  mit Annahme 4.5 bedingt auf  $\mathcal{I}_t$  Innovationen sind, erfüllt (4.12) somit die Bedingungen in Definition C.8 und ist sowohl eine nichtlineare Markovkette als auch eine Fellerkette.



## Anhang D

# Ableitungen der Beiträge zur Quasi-Loglikelihoodfunktion

### D.1 Rechenregeln für Matrizen

Die folgenden Rechenregeln werden in diesem Kapitel benötigt:

**R1** Für die transformierten Duplikationsmatrizen gilt:

$$D_{p,-}^+ = \frac{1}{2} (D_{p,-})'.$$

**Beweis:** Für die Dimension der Matrizen gilt:

$$D_{p,-} \sim (p^2 \times p^-), \quad D_{p,-}^+ \sim (p^- \times p^2).$$

Dabei enthält die Matrix  $D_{p,-}$  nur die Einträge Null und Eins. Die Spaltensummen sind 2 und die Zeilensummen 0 oder 1. Genauer gesagt enthält jede Spalte zwei Einsen, allerdings befindet sich keine davon in derselben Zeile wie eine der Einsen aus einer anderen Spalte. Werden die Spalten von  $D_{p,-}$  mit  $\mathbf{d}_{\bullet,i}$ ,  $i \in \{1, \dots, p^-\}$ , bezeichnet, so gilt:

$$\mathbf{d}'_{\bullet,i} \mathbf{d}_{\bullet,j} = \begin{cases} 2, & i = j \\ 0, & i \neq j \end{cases} \Rightarrow D'_{p,-} D_{p,-} = 2 \cdot \mathbb{I}_{p^-}.$$

Somit besitzt die Matrix  $D'_{p,-} D_{p,-}$  den Eigenwert 2 beziehungsweise den Singulärwert  $\sqrt{2}$  mit Vielfachheit  $p^-$ . Mit Aussage (1) auf S.335 in Seber (2008) lautet die Singulärwertzerlegung von  $D_{p,-}$  in diesem Fall

$$D_{p,-} = U \begin{bmatrix} \sqrt{2} \cdot \mathbb{I}_{p^-} \\ \mathbf{0}_{p^+ \times p^-} \end{bmatrix} V',$$

wobei  $U \sim (p^2 \times p^2)$  und  $V \sim (p^- \times p^-)$  orthogonale Matrizen sind, die die orthonormalisierten Eigenvektoren von  $D_{p,-}$ ,  $D'_{p,-}$  beziehungsweise von  $D'_{p,-} D_{p,-}$  enthalten. Mit 5.5.1(9) in Lütkepohl (1996) lautet die Singulärwertzerlegung der Moore-Penrose-Inversen von  $D_{p,-}$ :

$$D_{p,-}^+ = V \begin{bmatrix} \frac{1}{\sqrt{2}} \cdot \mathbb{I}_{p^-} \\ \mathbf{0}_{p^+ \times p^-} \end{bmatrix} U' = \frac{1}{2} V \begin{bmatrix} \sqrt{2} \cdot \mathbb{I}_{p^-} \\ \mathbf{0}_{p^+ \times p^-} \end{bmatrix} U' = \frac{1}{2} D'_{p,-}.$$

■

**R2** Lütkepohl (1996), 10.4.2(1): Für eine symmetrische  $(m \times m)$ -Matrix  $X$  gilt:

$$\frac{\partial \text{vec}(X)}{\partial \text{vech}(X)'} = D_m \quad \Rightarrow \quad \frac{\partial \text{vec}(X)}{\partial \text{vecl}(X)'} = D_{m,-} \stackrel{R1}{=} 2 \left( D_{m,-}^+ \right)'.$$

**R3** Für symmetrische Matrizen  $X \sim (m, m)$  und  $Y \sim (n, n)$ , ergibt sich als direkte Konsequenz aus R2:

$$\frac{\partial \text{vech}(Y(X))}{\partial \text{vecl}(X)'} = \frac{\partial L_n \text{vec}(Y(X))}{\partial \text{vec}(X)'} \frac{\partial \text{vec}(X)}{\partial \text{vecl}(X)'} = L_n \frac{\partial \text{vec}(Y(X))}{\partial \text{vec}(X)'} D_{m,-}.$$

**R4** Für symmetrische Matrizen  $X \sim (n, n)$  und  $Y(X) \sim (n, n)$ , gilt

$$\frac{\partial \text{vec}(XYX)}{\partial \text{vec}(X)'} = (X \otimes X) \frac{\partial \text{vec}(Y)}{\partial \text{vec}(X)'} + (XY \otimes \mathbb{I}_n + \mathbb{I}_n \otimes XY).$$

**Beweis von R4:** Mit 17.30(h) in Seber (2008) gilt für eine  $(l \times m)$ -Matrix  $Z$  sowie für von den Einträgen in  $Z$  abhängige Matrizen  $U(Z) \sim (q \times r)$  und  $V(Z) \sim (r \times t)$ :

$$\frac{\partial \text{vec}(UV)}{\partial \text{vec}(Z)'} = (V \otimes \mathbb{I}_q)' \frac{\partial \text{vec}(U)}{\partial \text{vec}(Z)'} + (\mathbb{I}_t \otimes U) \frac{\partial \text{vec}(V)}{\partial \text{vec}(Z)'}. \quad (\text{D.1})$$

Unter der Annahme, dass es sich bei  $X$  und  $Y(X)$  um  $(n \times n)$ -Matrizen handelt, wird  $U := XY$  und  $V := X$  gewählt, was  $q = r = t = l = m = n$  impliziert. Dann gilt:

$$\frac{\partial \text{vec}(XY(X)X)}{\partial \text{vec}(X)'} = (X \otimes \mathbb{I}_n)' \frac{\partial \text{vec}(XY)}{\partial \text{vec}(X)'} + (\mathbb{I}_n \otimes XY) \frac{\partial \text{vec}(X)}{\partial \text{vec}(X)'}$$

Wird in einer zweiten Anwendung von (D.1)  $U := X$  und  $V := Y$  gewählt, so ergibt sich

$$\begin{aligned} \frac{\partial \text{vec}(XY(X)X)}{\partial \text{vec}(X)'} &= (X \otimes \mathbb{I}_n) \left[ (Y \otimes \mathbb{I}_n)' \frac{\partial \text{vec}(X)}{\partial \text{vec}(X)'} + (\mathbb{I}_n \otimes X) \frac{\partial \text{vec}(Y)}{\partial \text{vec}(X)'} \right] + (\mathbb{I}_n \otimes XY) \\ &= (X \otimes \mathbb{I}_n)(Y \otimes \mathbb{I}_n) + (X \otimes \mathbb{I}_n)(\mathbb{I}_p \otimes X) \frac{\partial \text{vec}(Y)}{\partial \text{vec}(X)'} + (\mathbb{I}_n \otimes XY) \\ &= (X \otimes X) \frac{\partial \text{vec}(Y)}{\partial \text{vec}(X)'} + (XY \otimes \mathbb{I}_n + \mathbb{I}_n \otimes XY). \end{aligned}$$

■

## D.2 Die ersten Ableitungen der Beiträge zur QLL-Funktion nach den Varianzparametern

Die dynamische Charakterisierung der bedingten Korrelationsmatrix  $R_t$  stellt den einzigen strukturellen Unterschied zum von Nakatani und Teräsvirta (2009) betrachteten CCC-Modell, siehe Bollerslev (1990), dar. Da  $R_t$  nicht von den Varianzparametern abhängt, ergeben sich die partiellen Ableitungen der Beiträge zur QLL-Funktion analog zu (3.2) in Nakatani und Teräsvirta (2009) als

$$\begin{aligned}\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1} &= -\frac{\partial \log \det D_t}{\partial \boldsymbol{\theta}_1} - \frac{1}{2} \frac{\partial y_t' H_t^{-1} y_t}{\partial \boldsymbol{\theta}_1} \\ &= -\frac{1}{2} \frac{\partial \text{vec}(D_t)'}{\partial \boldsymbol{\theta}_1} \text{vec} \left( 2D_t^{-1} - D_t^{-1} R_t^{-1} z_t z_t' - z_t z_t' R_t^{-1} D_t^{-1} \right).\end{aligned}$$

Die Definition von  $v_{it} := (1, y_{it}^2, h_{it})'$  erlaubt eine rekursive Darstellung der von Null verschiedenen partiellen Ableitungen der Hauptdiagonaleinträge von  $D_t$  nach den Varianzparametern  $\phi_i = (\omega_i, \alpha_i, \beta_i)'$  als

$$g_{it} := \frac{\partial h_{it}^{1/2}}{\partial \phi_i} = \frac{1}{2} h_{it}^{-1/2} \left( v_{i,t-1} + \beta_i \frac{\partial h_{i,t-1}}{\partial \phi_i} \right) \quad \text{und} \quad \frac{\partial h_{jt}^{1/2}}{\partial \phi_i} = \mathbf{0}_3, \quad \text{für } i \neq j.$$

Somit ergibt sich

$$\frac{\partial \text{vec}(D_t)'}{\partial \boldsymbol{\theta}_1} = \begin{bmatrix} g_{1t} & \mathbf{0}_{3 \times p} & \mathbf{0}_{3 \times p} & \mathbf{0}_3 & \dots & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{3 \times p} \\ \mathbf{0}_3 & \mathbf{0}_{3 \times p} & g_{2t} & \mathbf{0}_{3 \times p} & \dots & \mathbf{0}_3 & \mathbf{0}_{3 \times p} & \mathbf{0}_3 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \mathbf{0}_3 & \mathbf{0}_{3 \times p} & \mathbf{0}_3 & \mathbf{0}_{3 \times p} & \dots & g_{p-1,t} & \mathbf{0}_{3 \times p} & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_{3 \times p} & \mathbf{0}_3 & \mathbf{0}_{3 \times p} & \dots & \mathbf{0}_3 & \mathbf{0}_{3 \times p} & g_{pt} \end{bmatrix}.$$

Für theoretische Überlegungen werden Darstellungen der Ableitungen, die auf einer unendlichen Vergangenheit an Beobachtungen basieren, verwendet. Für Simulation und Schätzung hingegen werden rekursiv definierte Terme, die auf zuvor festgelegten Startwerten für  $t = 0$  basieren, herangezogen. Die bedingten Varianzen lassen sich dann für den Fall einer unendlichen Vergangenheit darstellen als:

$$\begin{aligned}\bullet \quad \frac{\partial h_{it}}{\partial \omega_i} &= 1 + \beta_i \frac{\partial h_{i,t-1}}{\partial \omega_i} = \frac{1}{1 - \beta_i} & \bullet \quad \frac{\partial h_{it}}{\partial \alpha_i} &= y_{i,t-1}^2 + \beta_i \frac{\partial h_{i,t-1}}{\partial \alpha_i} = \sum_{n=0}^{\infty} \beta_i^n y_{i,t-n-1}^2 \\ \bullet \quad \frac{\partial h_{it}}{\partial \beta_i} &= h_{i,t-1} + \beta_i \frac{\partial h_{i,t-1}}{\partial \beta_i} = \sum_{n=0}^{\infty} \beta_i^n h_{i,t-n-1} = \frac{\omega_i}{(1 - \beta_i)^3} + \alpha_i \sum_{n=0}^{\infty} \beta_i^n \sum_{k=0}^{\infty} \beta_i^k y_{i,t-n-k-2}^2.\end{aligned}$$

Startwerte für die bedingten Varianzen und ihre Ableitungen können beispielsweise wie im R-Paket `cggarch` gewählt werden als  $(s_1^2, \dots, s_p^2)'$  mit  $s_i^2 = \frac{1}{T} \sum_{t=1}^T y_{it}^2$ ,  $i = 1, \dots, p$ . Dieses Vorgehen steht in Einklang mit Annahme 4.1(3.), die positive Startwerte für alle bedingten Varianzgleichungen fordert. Als Konsequenz dieser Wahl der Startwerte gilt dann auch  $v_{i0} = (1, s_i^2, s_i^2)'$ ,  $i = 1, \dots, p$ , und  $\left( \frac{\partial h_{10}}{\partial \phi_1}, \dots, \frac{\partial h_{p0}}{\partial \phi_p} \right)' = \mathbf{0}_{p \times 3}$ .

### D.3 Die ersten Ableitungen der Beiträge zur QLL-Funktion nach den Korrelationsparametern

In diesem Abschnitt werden die partiellen Ableitungen der Beiträge zur QLL-Funktion nach denjenigen Parametern betrachtet, die die Korrelationsstruktur des DCC-Modells definieren. Dies sind die Einträge unter der Hauptdiagonalen der unbedingten Korrelationsmatrix  $\bar{Q}$  sowie die Parameter  $\alpha$  und  $\beta$  aus (4.5) in Kapitel 4.1. Ähnlich zu (3.3) in Nakatani und Teräsvirta (2009), ergeben sich die Ableitungen als:

$$\begin{aligned}
\frac{\partial l_t(\theta)}{\partial \theta_2} &= -\frac{1}{2} \frac{\partial \log \det(R_t)}{\partial \theta_2} - \frac{1}{2} \frac{\partial y_t' D_t^{-1} R_t^{-1} D_t^{-1} y_t}{\partial \theta_2} = -\frac{1}{2} \frac{\partial \text{vecl}(R_t)'}{\partial \theta_2} \frac{\partial \text{vec}(R_t)'}{\partial \text{vecl}(R_t)} \frac{\partial \log \det(R_t)}{\partial \text{vec}(R_t)} \\
&= -\frac{1}{2} \frac{\partial \text{vecl}(R_t)'}{\partial \theta_2} \frac{\partial \text{vec}(R_t)'}{\partial \text{vecl}(R_t)} \frac{\partial \text{vec}(R_t^{-1})'}{\partial \text{vec}(R_t)} \frac{\partial \text{vec}(D_t^{-1} R_t^{-1} D_t^{-1})'}{\partial \text{vec}(R_t^{-1})} \frac{\partial y_t' D_t^{-1} R_t^{-1} D_t^{-1} y_t}{\partial \text{vec}(D_t^{-1} R_t^{-1} D_t^{-1})} \\
&= -\frac{1}{2} \frac{\partial \text{vecl}(R_t)'}{\partial \theta_2} 2D_{p,-}^+ \text{vec}(R_t^{-1}) \\
&\quad - \frac{1}{2} \frac{\partial \text{vecl}(R_t)'}{\partial \theta_2} 2D_{p,-}^+ (-1) (R_t^{-1} \otimes R_t^{-1}) (D_t^{-1} \otimes D_t^{-1}) \text{vec}(y_t y_t') \\
&= -\frac{\partial \text{vecl}(R_t)'}{\partial \theta_2} D_{p,-}^+ \left[ \text{vec}(R_t^{-1}) - (R_t^{-1} \otimes R_t^{-1}) (D_t^{-1} \otimes D_t^{-1}) (y_t \otimes y_t) \right] \\
&= -\frac{\partial \text{vecl}(R_t)'}{\partial \theta_2} \text{vecl} \left( R_t^{-1} [\mathbb{I}_p - z_t z_t' R_t^{-1}] \right).
\end{aligned}$$

Wie in Nakatani und Teräsvirta (2009) ergibt sich die zweite Gleichheit durch eine wiederholte Anwendung der Kettenregel, siehe 10.2.1(6) in Lütkepohl (1996). Für die dritte Gleichheit werden 10.3.1(8), 10.3.3(10), 10.4.1(3), 10.4.2(1) und 10.6(1) aus Lütkepohl (1996) sowie R1 aus Abschnitt D.1 ausgenutzt. Die verbleibenden Gleichheiten basieren auf R2 aus Abschnitt D.1 und einfachen Matrixumformungen.

Im Folgenden werden die partiellen Ableitungen der Einträge unter der Hauptdiagonalen von  $R_t$  nach den Korrelationsparametern im Detail untersucht. Wie in Hafner und Herwatz (2008) gilt:

$$\frac{\partial \text{vecl}(R_t)}{\partial \theta_2'} = D_{p,-}^+ \frac{\partial \text{vec}(R_t)}{\partial \theta_2'} = D_{p,-}^+ \frac{\partial \text{vec}(R_t)}{\partial \text{vech}(Q_t)'} \frac{\partial \text{vech}(Q_t)}{\partial \theta_2'}, \quad (\text{D.2})$$

wobei

$$\frac{\partial \text{vec}(R_t)}{\partial \text{vech}(Q_t)'} = \frac{\partial \text{vec}(R_t)}{\partial \text{vec}(Q_t^*)'} \frac{\partial \text{vec}(Q_t^*)}{\partial \text{vech}(Q_t^*)'} \stackrel{\text{R2}}{=} \frac{\partial \text{vec}(Q_t^* Q_t Q_t^*)}{\partial \text{vec}(Q_t^*)'} D_p \frac{\partial \text{vech}(Q_t^*)}{\partial \text{vech}(Q_t)'} \quad (\text{D.3})$$

Ähnlich den Ausführungen in Abschnitt 4.3 in Hafner und Herwatz (2008) ergibt sich die Ableitung von  $\text{vec}(Q_t^* Q_t Q_t^*)$  nach  $\text{vec}(Q_t^*)$  als

$$\frac{\partial \text{vec}(Q_t^* Q_t Q_t^*)}{\partial \text{vec}(Q_t^*)'} \stackrel{\text{R4}}{=} (Q_t^* \otimes Q_t^*) \frac{\partial \text{vec}(Q_t)}{\partial \text{vec}(Q_t^*)'} + (Q_t^* Q_t \otimes \mathbb{I}_p + \mathbb{I}_p \otimes Q_t^* Q_t) \quad (\text{D.4})$$

und die Ableitung von  $\text{vech}(Q_t^*)$  nach  $\text{vech}(Q_t)$  als

$$\frac{\partial \text{vech}(Q_t^*)}{\partial \text{vech}(Q_t)'} = -\frac{1}{2} \text{diag} \left\{ \text{vech} \left( Q_t^{-\frac{3}{2}} \right) \right\} \cdot \text{diag} \{ \text{vech}(\mathbb{I}_p) \}$$

mit  $q_{ij,t} := [Q_t]_{ij}$  und  $Q_t^{-3/2} := [q_{ij,t}^{-3/2}]_{i,j=1,\dots,p}$ .

Somit implizieren (D.3) und (D.4), dass für Ableitungen der Einträge unter der Hauptdiagonalen von  $R_t$  nach den Korrelationsparametern gilt:

$$\begin{aligned} \frac{\partial \text{vecl}(R_t)}{\partial \boldsymbol{\theta}'_2} &= \mathbf{D}_{p,-}^+ \left[ (Q_t^* \otimes Q_t^*) \frac{\partial \text{vec}(Q_t)}{\partial \text{vec}(Q_t)'} + \left( Q_t^* Q_t \otimes \mathbb{I}_p + \mathbb{I}_p \otimes Q_t^* Q_t \right) \mathbf{D}_p \frac{\partial \text{vech}(Q_t^*)}{\partial \text{vech}(Q_t)'} \frac{\partial \text{vech}(Q_t)}{\partial \boldsymbol{\theta}'_2} \right] \\ &= \left[ \mathbf{D}_{p,-}^+ (Q_t^* \otimes Q_t^*) \mathbf{D}_p + \mathbf{D}_{p,-}^+ \left( Q_t^* Q_t \otimes \mathbb{I}_p + \mathbb{I}_p \otimes Q_t^* Q_t \right) \mathbf{D}_p \frac{\partial \text{vech}(Q_t^*)}{\partial \text{vech}(Q_t)'} \right] \frac{\partial \text{vech}(Q_t)}{\partial \boldsymbol{\theta}'_2}. \end{aligned}$$

Die Ableitung von  $\text{vech}(Q_t)$  nach  $\boldsymbol{\theta}_2$  in (D.2) wiederum kann ähnlich dem Vorgehen in Hafner und Herwatz (2008) aufgespalten werden in:

$$\begin{aligned} \frac{\partial \text{vech}(Q_t)}{\partial \alpha} &= -\text{vech}(\bar{Q}) + \text{vech}(z_{t-1} z'_{t-1}) + \beta \frac{\partial \text{vech}(Q_{t-1})}{\partial \alpha} \\ &= -\frac{1}{1-\beta} \text{vech}(\bar{Q}) + \sum_{n=0}^{\infty} \beta^n \text{vech}(z_{t-n-1} z'_{t-n-1}) \\ \frac{\partial \text{vech}(Q_t)}{\partial \beta} &= -\text{vech}(\bar{Q}) + \text{vech}(Q_{t-1}) + \beta \frac{\partial \text{vech}(Q_{t-1})}{\partial \beta} \\ &= -\frac{1}{1-\beta} \text{vech}(\bar{Q}) + \sum_{n=0}^{\infty} \beta^n \text{vech}(Q_{t-n-1}) \\ \frac{\partial \text{vech}(Q_t)}{\partial \text{vecl}(\bar{Q})} &= (1-\alpha-\beta) \frac{\partial \text{vech}(\bar{Q})}{\partial \text{vecl}(\bar{Q})'} + \alpha \frac{\partial \text{vech}(z_{t-1} z'_{t-1})}{\partial \text{vecl}(\bar{Q})'} + \beta \frac{\partial \text{vech}(Q_{t-1})}{\partial \text{vecl}(\bar{Q})'} \\ &\stackrel{\text{R3}}{=} \frac{1-\alpha-\beta}{1-\beta} \mathbf{L}_p \mathbf{D}_{p,-}. \end{aligned}$$

Für die rekursiven Darstellungen, die zur Simulation und Parameterschätzung verwendet werden, können Startwerte für  $t = 0$  wie im R-Paket `ccgarch` gewählt werden als  $\left( \frac{\partial \text{vech}(Q_0)}{\partial \alpha}, \frac{\partial \text{vech}(Q_0)}{\partial \beta} \right) = \mathbf{0}'_2$ .

## D.4 Die zweiten Ableitungen der Beiträge zur QLL-Funktion

Wie in Nakatani und Teräsvirta (2009) lässt sich die Hessematrix der Beiträge zur QLL-Funktion in verschiedene Blöcke aufspalten, die anschließend einzeln betrachtet werden können:

$$\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = \begin{bmatrix} \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'} & \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_2'} \\ \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_1'} & \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_2'} \end{bmatrix}.$$

### D.4.1 Die Berechnung von $\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'}$

Da die bedingte Korrelationsmatrix  $R_t$  nicht von den Varianzparametern abhängt, ergibt sich der Block links oben in der Hessematrix analog zu (3.6) in Nakatani und Teräsvirta (2009):

$$\begin{aligned} \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'} &= -\frac{1}{2} \frac{\partial}{\partial \boldsymbol{\theta}_1'} \left[ \frac{\partial \text{vec}(D_t)'}{\partial \boldsymbol{\theta}_1} \text{vec} \left( 2D_t^{-1} - D_t^{-1} R_t^{-1} z_t z_t' - z_t z_t' R_t^{-1} D_t^{-1} \right) \right] \\ &= -\frac{1}{2} \frac{\partial \text{vec}(D_t)'}{\partial \boldsymbol{\theta}_1} \left[ -2 \left( D_t^{-1} \otimes D_t^{-1} \right) + \left( z_t z_t' \otimes D_t^{-1} R_t^{-1} D_t^{-1} \right) \right. \\ &\quad + \left( D_t^{-1} R_t^{-1} D_t^{-1} \otimes z_t z_t' \right) + \left( D_t^{-1} \otimes D_t^{-1} R_t^{-1} z_t z_t' \right) + \left( D_t^{-1} R_t^{-1} z_t z_t' \otimes D_t^{-1} \right) \\ &\quad \left. + \left( z_t z_t' R_t^{-1} D_t^{-1} \otimes D_t^{-1} \right) + \left( D_t^{-1} \otimes z_t z_t' R_t^{-1} D_t^{-1} \right) \right] \frac{\partial \text{vec}(D_t)'}{\partial \boldsymbol{\theta}_1} \\ &\quad + \left[ \frac{1}{2} \left( \text{vec} \left( D_t^{-1} R_t^{-1} z_t z_t' \right) \otimes \mathbb{I}_{3p} \right) + \frac{1}{2} \left( \text{vec} \left( z_t z_t' R_t^{-1} D_t^{-1} \right) \otimes \mathbb{I}_{3p} \right) \right. \\ &\quad \left. - \left( \text{vec} \left( D_t^{-1} \right) \otimes \mathbb{I}_{3p} \right) \right] \frac{\partial^2 \text{vec}(D_t)'}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'}. \end{aligned}$$

Ein näherer Blick auf die einzelnen Bestandteile der Ableitung liefert:

$$\frac{\partial^2 \text{vec}(D_t)'}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'} = \frac{\partial}{\partial \boldsymbol{\theta}_1'} \text{vec} \left( \frac{\partial \text{vech}(D_t)'}{\partial \boldsymbol{\theta}_1} D_p' \right) = (D_p \otimes \mathbb{I}_{3p}) \frac{\partial}{\partial \boldsymbol{\theta}_1'} \text{vec} \left( \frac{\partial \text{vech}(D_t)'}{\partial \boldsymbol{\theta}_1} \right).$$

Außerdem werden diejenigen zweiten Ableitungen der Hauptdiagonaleinträge von  $D_t$ , die nicht Null sind, durch

$$g_{it}^{(2)} := \frac{\partial^2 h_{it}^{1/2}}{\partial \phi_i \partial \phi_i'}$$

bezeichnet. Somit ergeben sich die zweiten Ableitungen von  $\text{vech}(D_t)$  nach den Varianzparametern als

$$\frac{\partial}{\partial \boldsymbol{\theta}_1'} \text{vec} \left( \frac{\partial \text{vech}(D_t)'}{\partial \boldsymbol{\theta}_1} \right) = \begin{bmatrix} g_{1t}^{(2)} & \mathbf{0}_{3 \times 3p^2} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3p(p-1)} & \cdots & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 6p} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3p^2} & g_{2t}^{(2)} & \mathbf{0}_{3 \times 3p(p-1)} & \cdots & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 6p} & \mathbf{0}_{3 \times 3} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3p^2} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3p(p-1)} & \cdots & g_{p-1,t}^{(2)} & \mathbf{0}_{3 \times 6p} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3p^2} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3p(p-1)} & \cdots & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 6p} & g_{pt}^{(2)} \end{bmatrix}.$$

Mit Hilfe von  $V_{it} := \frac{\partial v_{it}}{\partial \phi'_i} = (\mathbf{0}_3, \mathbf{0}_3, \frac{\partial h_{i,t-1}}{\partial \phi_i})'$  lässt sich  $g_{it}^{(2)}$  auch darstellen als

$$g_{it}^{(2)} = -\frac{1}{4} \frac{1}{h_{it}^{3/2}} \frac{\partial h_{it}}{\partial \phi'_i} \left( v_{i,t-1} + \beta_i \frac{\partial h_{i,t-1}}{\partial \phi_i} \right) + \frac{1}{2} \frac{1}{h_{it}^{1/2}} \left( V_{i,t-1} + \beta_i \frac{\partial^2 h_{i,t-1}}{\partial \phi_i \partial \phi'_i} \right).$$

Dabei ergeben sich die partiellen Ableitungen der bedingten Varianzen als:

$$\begin{aligned} & \bullet \frac{\partial^2 h_{it}}{(\partial \omega_i)^2} = \beta_i \frac{\partial^2 h_{i,t-1}}{(\partial \omega_i)^2} = 0 & \bullet \frac{\partial^2 h_{it}}{\partial \omega_i \partial \alpha_i} = \beta_i \frac{\partial^2 h_{i,t-1}}{\partial \omega_i \partial \alpha_i} = 0 \\ & \bullet \frac{\partial^2 h_{it}}{\partial \omega_i \partial \beta_i} = \frac{\partial h_{it}}{\partial \omega_i} + \beta_i \frac{\partial^2 h_{i,t-1}}{\partial \omega_i \partial \beta_i} = \frac{1}{(1 - \beta_i)^2} \\ & \bullet \frac{\partial^2 h_{it}}{(\partial \alpha_i)^2} = \beta_i \frac{\partial^2 h_{i,t-1}}{(\partial \alpha_i)^2} = 0 & \bullet \frac{\partial^2 h_{it}}{\partial \alpha_i \partial \beta_i} = \frac{\partial h_{it}}{\partial \alpha_i} + \beta_i \frac{\partial^2 h_{i,t-1}}{\partial \alpha_i \partial \beta_i} = \sum_{n=0}^{\infty} n \beta_i^{n-1} y_{i,t-n-1}^2 \\ & \bullet \frac{\partial^2 h_{it}}{(\partial \beta_i)^2} = 2 \frac{\partial h_{i,t-1}}{\partial \beta_i} + \beta_i \frac{\partial^2 h_{i,t-1}}{(\partial \beta_i)^2} = 2 \sum_{n=0}^{\infty} \beta_i^n \frac{\partial h_{i,t-n-2}}{\partial \beta_i} \\ & \quad = \frac{2\omega_i}{(1 - \beta_i)^4} + 2\alpha_i \sum_{n=0}^{\infty} \beta_i^n \sum_{k=0}^{\infty} \beta_i^k \sum_{l=0}^{\infty} \beta_i^l y_{i,t-n-k-l-2}^2. \end{aligned}$$

Die Startwerte für  $t = 0$  können in den rekursiven Darstellungen wiederum auf Null gesetzt werden.

#### D.4.2 Die Berechnung von $\frac{\partial^2 l_t(\theta)}{\partial \theta_1 \partial \theta'_2}$ und $\frac{\partial^2 l_t(\theta)}{\partial \theta_2 \partial \theta'_1}$

Analog zu (3.9) und (3.10) in Nakatani und Teräsvirta (2009) können die Einträge links unten und rechts oben in der Hessematrix bestimmt werden. Diese beinhalten die Kreuzableitungen nach den Varianz- und Korrelationsparametern und ergeben sich unter der Beschränkung auf die Ableitungen der Einträge unter der Hauptdiagonalen von  $R_t$  nach den Korrelationsparametern als:

$$\begin{aligned} \frac{\partial^2 l_t(\theta)}{\partial \theta_1 \partial \theta'_2} &= -\frac{1}{2} \frac{\partial}{\partial \theta'_2} \left[ \frac{\partial \text{vec}(D_t)'}{\partial \theta_1} \text{vec} \left( 2D_t^{-1} - D_t^{-1} R_t^{-1} z_t z'_t - z_t z'_t R_t^{-1} D_t^{-1} \right) \right] \\ &= \frac{1}{2} \frac{\partial \text{vec}(D_t)'}{\partial \theta_1} \left[ \frac{\partial \text{vec} \left( D_t^{-1} R_t^{-1} z_t z'_t \right)}{\partial \theta'_2} + \frac{\partial \text{vec} \left( z_t z'_t R_t^{-1} D_t^{-1} \right)}{\partial \theta'_2} \right] \\ &= \frac{1}{2} \frac{\partial \text{vec}(D_t)'}{\partial \theta_1} \left[ \left( z_t z'_t \otimes D_t^{-1} \right) + \left( D_t^{-1} \otimes z_t z'_t \right) \right] \frac{\partial \text{vec} \left( R_t^{-1} \right)}{\partial \text{vec} \left( R_t \right)'} \frac{\partial \text{vec} \left( R_t \right)}{\partial \text{vecl} \left( R_t \right)'} \frac{\partial \text{vecl} \left( R_t \right)}{\partial \theta'_2} \\ &\stackrel{\text{R2}}{=} -\frac{\partial \text{vec}(D_t)'}{\partial \theta_1} \left[ \left( z_t z'_t R_t^{-1} \otimes D_t^{-1} R_t^{-1} \right) + \left( D_t^{-1} R_t^{-1} \otimes z_t z'_t R_t^{-1} \right) \right] \left( D_{p,-}^+ \right)' \frac{\partial \text{vecl} \left( R_t \right)}{\partial \theta'_2} \\ \frac{\partial^2 l_t(\theta)}{\partial \theta_2 \partial \theta'_1} &= \left[ \frac{\partial^2 l_t(\theta)}{\partial \theta_1 \partial \theta'_2} \right]' \\ &= -\frac{\partial \text{vecl} \left( R_t \right)'}{\partial \theta_2} D_{p,-}^+ \left[ \left( R_t^{-1} z_t z'_t \otimes R_t^{-1} D_t^{-1} \right) + \left( R_t^{-1} D_t^{-1} \otimes R_t^{-1} z_t z'_t \right) \right] \frac{\partial \text{vec} \left( D_t \right)}{\partial \theta'_1}. \end{aligned}$$

### D.4.3 Die Berechnung von $\frac{\partial^2 l_t(\theta)}{\partial \theta_2 \partial \theta_2'}$

Abschließend ist der Block rechts unten in der Hessematrix, der die zweiten Ableitungen der Beiträge der QLL-Funktion nach den Korrelationsparametern beinhaltet, gegeben als:

$$\begin{aligned}
\frac{\partial^2 l_t(\theta)}{\partial \theta_2 \partial \theta_2'} &= -\frac{\partial}{\partial \theta_2'} \left( \frac{\partial \text{vecl}(R_t)'}{\partial \theta_2} \text{vecl}(R_t^{-1}) \right) + \frac{\partial}{\partial \theta_2'} \left( \frac{\partial \text{vecl}(R_t)'}{\partial \theta_2} \text{vecl}(R_t^{-1} z_t z_t' R_t^{-1}) \right) \\
&= -\frac{\partial \text{vecl}(R_t)'}{\partial \theta_2} \left[ \frac{\partial \text{vecl}(R_t^{-1})}{\partial \theta_2'} - \frac{\partial \text{vecl}(R_t^{-1} z_t z_t' R_t^{-1})}{\partial \theta_2'} \right] \\
&\quad - \left[ \left( \text{vecl}(R_t^{-1})' \otimes \mathbb{I}_{p-+2} \right) - \left( \text{vecl}(R_t^{-1} z_t z_t' R_t^{-1})' \otimes \mathbb{I}_{p-+2} \right) \right] \frac{\partial}{\partial \theta_2'} \text{vec} \left( \frac{\partial \text{vecl}(R_t)'}{\partial \theta_2} \right) \\
&= \frac{\partial \text{vecl}(R_t)'}{\partial \theta_2} \mathbb{D}_{p,-}^+ \left[ \left( R_t^{-1} \otimes R_t^{-1} \right) - \left( R_t^{-1} z_t z_t' R_t^{-1} \otimes R_t^{-1} \right) \right. \\
&\quad \left. - \left( R_t^{-1} \otimes R_t^{-1} z_t z_t' R_t^{-1} \right) \right] \mathbb{D}_{p,-} \frac{\partial \text{vecl}(R_t)}{\partial \theta_2'} \\
&\quad + \left[ \left( \text{vecl}(R_t^{-1})' \otimes \mathbb{I}_{p-+2} \right) + \left( \text{vecl}(R_t^{-1} z_t z_t' R_t^{-1})' \otimes \mathbb{I}_{p-+2} \right) \right] \frac{\partial}{\partial \theta_2'} \text{vec} \left( \frac{\partial \text{vecl}(R_t)'}{\partial \theta_2} \right).
\end{aligned}$$

Die zweiten partiellen Ableitungen der Einträge unter der Hauptdiagonalen von  $R_t$  nach den Korrelationsparametern ergeben sich dann als:

$$\begin{aligned}
\frac{\partial}{\partial \theta_2'} \text{vec} \left( \frac{\partial \text{vecl}(R_t)'}{\partial \theta_2} \right) &= \frac{\partial}{\partial \theta_2'} \text{vec} \left( \frac{\partial \text{vech}(Q_t)'}{\partial \theta_2} \mathbb{D}_p' (Q_t^* \otimes Q_t^*) (\mathbb{D}_{p,-}^+)' \right) \\
&\quad + \frac{\partial}{\partial \theta_2'} \text{vec} \left( \frac{\partial \text{vech}(Q_t)'}{\partial \theta_2} \frac{\partial \text{vech}(Q_t^*)'}{\partial \text{vech}(Q_t)} \mathbb{D}_p' (Q_t Q_t^* \otimes \mathbb{I}_p + \mathbb{I}_p \otimes Q_t Q_t^*) (\mathbb{D}_{p,-}^+)' \right). \quad (\text{D.5})
\end{aligned}$$

Mit 10.5.5(4) in Lütkepohl (1996) entspricht der erste Summand in (D.5) gerade

$$\left[ \mathbb{D}_{p,-}^+ \otimes \frac{\partial \text{vech}(Q_t)'}{\partial \theta_2} \mathbb{D}_p' \right] \frac{\partial \text{vec}(Q_t^* \otimes Q_t^*)}{\partial \theta_2'} + \left[ \mathbb{D}_{p,-}^+ (Q_t^* \otimes Q_t^*) \mathbb{D}_p \otimes \mathbb{I}_{p-+2} \right] \frac{\partial}{\partial \theta_2'} \text{vec} \left( \frac{\partial \text{vech}(Q_t)'}{\partial \theta_2} \right).$$

Weiter gilt mit 10.5.5(7) in Lütkepohl (1996):

$$\begin{aligned}
\frac{\partial \text{vec}(Q_t^* \otimes Q_t^*)}{\partial \theta_2'} &= (\mathbb{I}_p \otimes \mathbb{K}_{pp} \otimes \mathbb{I}_p) \left[ \frac{\partial \text{vec}(Q_t^*)}{\partial \theta_2'} \otimes \text{vec}(Q_t^*) + \text{vec}(Q_t^*) \otimes \frac{\partial \text{vec}(Q_t^*)}{\partial \theta_2'} \right] \\
\text{mit} \quad \frac{\partial \text{vec}(Q_t^*)}{\partial \theta_2'} &= \mathbb{D}_p \frac{\partial \text{vech}(Q_t^*)}{\partial \text{vech}(Q_t)'} \frac{\partial \text{vech}(Q_t)}{\partial \theta_2'}.
\end{aligned}$$

Für die zweiten Ableitungen der Einträge von  $Q_t$  nach den Korrelationsparametern gilt:

$$\frac{\partial}{\partial \theta_2'} \text{vec} \left( \frac{\partial \text{vech}(Q_t)'}{\partial \theta_2} \right) = \mathbb{K}_{p^+, p-+2} \frac{\partial}{\partial \theta_2'} \text{vec} \left( \frac{\partial \text{vech}(Q_t)}{\partial \theta_2'} \right).$$

Dabei kann die Matrix der zweiten Ableitungen der Einträge von  $Q_t$  nach den Korrelationsparametern in einzelne Blockmatrizen unterschiedlicher Größe aufgespalten werden:

$$\begin{aligned}
& \begin{bmatrix} \frac{\partial^2 \text{vech}(Q_t)}{(\partial \alpha)^2} & \frac{\partial^2 \text{vech}(Q_t)}{\partial \alpha \partial \beta} & \frac{\partial^2 \text{vech}(Q_t)}{\partial \alpha \partial \text{vecl}(\bar{Q})'} \\ \frac{\partial^2 \text{vech}(Q_t)}{\partial \alpha \partial \beta} & \frac{\partial^2 \text{vech}(Q_t)}{(\partial \beta)^2} & \frac{\partial^2 \text{vech}(Q_t)}{\partial \beta \partial \text{vecl}(\bar{Q})'} \\ \text{vec} \left( \frac{\partial^2 \text{vech}(Q_t)}{\partial \alpha \partial \text{vecl}(\bar{Q})'} \right) & \text{vec} \left( \frac{\partial^2 \text{vech}(Q_t)'}{\partial \beta \partial \text{vecl}(\bar{Q})'} \right) & \frac{\partial}{\partial \text{vecl}(\bar{Q})'} \text{vec} \left( \frac{\partial \text{vech}(Q_t)'}{\partial \text{vecl}(\bar{Q})'} \right) \end{bmatrix} \\
& = \begin{bmatrix} \beta \frac{\partial^2 \text{vech}(Q_{t-1})}{(\partial \alpha)^2} & \frac{\partial \text{vech}(Q_{t-1})}{\partial \alpha} + \beta \frac{\partial^2 \text{vech}(Q_{t-1})}{\partial \alpha \partial \beta} & -\frac{1}{1-\beta} \mathbb{L}_p \mathbb{D}_{p,-} \\ \frac{\partial \text{vech}(Q_{t-1})}{\partial \alpha} + \beta \frac{\partial^2 \text{vech}(Q_{t-1})}{\partial \alpha \partial \beta} & 2 \frac{\partial \text{vech}(Q_{t-1})}{\partial \beta} + \beta \frac{\partial^2 \text{vech}(Q_{t-1})}{(\partial \beta)^2} & -\frac{2-\alpha-2\beta}{(1-\beta)^2} \mathbb{L}_p \mathbb{D}_{p,-} \\ -\frac{1}{1-\beta} \text{vec}(\mathbb{L}_p \mathbb{D}_{p,-}) & -\frac{2-\alpha-2\beta}{(1-\beta)^2} \text{vec}(\mathbb{L}_p \mathbb{D}_{p,-}) & \mathbf{0}_{p^+ p^- \times p^-} \end{bmatrix}.
\end{aligned}$$

Des Weiteren kann der zweite Summand in (D.5) unter Zuhilfenahme von 10.5.5(4) in Lütkepohl (1996) umgeformt werden zu

$$\begin{aligned}
& \left( \mathbb{D}_{p,-}^+ \otimes \frac{\partial \text{vech}(Q_t)'}{\partial \boldsymbol{\theta}_2} \frac{\partial \text{vech}(Q_t^*)'}{\partial \text{vech}(Q_t)} \mathbb{D}_p' \right) \frac{\partial \text{vec}(Q_t Q_t^* \otimes \mathbb{I}_p + \mathbb{I}_p \otimes Q_t Q_t^*)}{\partial \boldsymbol{\theta}_2'} \\
& + \left( \mathbb{D}_{p,-}^+ \left[ Q_t^* Q_t \otimes \mathbb{I}_p + \mathbb{I}_p \otimes Q_t^* Q_t \right] \mathbb{D}_p \otimes \mathbb{I}_{p^-+2} \right) \frac{\partial}{\partial \boldsymbol{\theta}_2'} \text{vec} \left( \frac{\partial \text{vech}(Q_t^*)'}{\partial \boldsymbol{\theta}_2} \right)
\end{aligned}$$

$$\begin{aligned}
\text{mit} \quad & \frac{\partial \text{vec}(Q_t Q_t^* \otimes \mathbb{I}_p + \mathbb{I}_p \otimes Q_t Q_t^*)}{\partial \boldsymbol{\theta}_2'} \\
& = \left( \mathbb{I}_p \otimes \mathbb{K}_{pp} \otimes \mathbb{I}_p \right) \left[ \frac{\partial \text{vec}(Q_t^* Q_t)}{\partial \boldsymbol{\theta}_2'} \otimes \text{vec}(\mathbb{I}_p) + \text{vec}(\mathbb{I}_p) \otimes \frac{\partial \text{vec}(Q_t^* Q_t)}{\partial \boldsymbol{\theta}_2'} \right] \\
\text{und} \quad & \frac{\partial \text{vec}(Q_t^* Q_t)}{\partial \boldsymbol{\theta}_2'} = (\mathbb{I}_p \otimes Q_t^*) \frac{\partial \text{vec}(Q_t)}{\partial \boldsymbol{\theta}_2'} + (Q_t \otimes \mathbb{I}_p) \frac{\partial \text{vec}(Q_t^*)}{\partial \boldsymbol{\theta}_2'} \\
& = \left[ (\mathbb{I}_p \otimes Q_t) \mathbb{D}_p \frac{\partial \text{vech}(Q_t^*)}{\partial \text{vech}(Q_t)'} + (Q_t^* \otimes \mathbb{I}_p) \mathbb{D}_p \right] \frac{\partial \text{vech}(Q_t)}{\partial \boldsymbol{\theta}_2'}.
\end{aligned}$$



## Anhang E

# Beweise der Aussagen aus Kap. 5

### E.1 Beweis von Proposition 5.1

Sei zunächst

$$\mathbf{y}_t := \text{vec} \left( y_t^{(2)} y_t^{(2)'} \right) \quad \text{mit} \quad y_t^{(2)} = \left( y_{1t}^2, \dots, y_{pt}^2 \right)'$$

definiert sowie

$$\mathbf{h}_t := \text{vec} \left( h_t h_t' \right) \quad \text{und} \quad \mathbf{Y}_t := \mathbb{E} \left[ \mathbf{y}_t \mathbf{y}_t' \right].$$

Somit existieren die achten Momente und Kreuzmomente von  $y_t$  genau dann, wenn  $\mathbf{Y}_t < \infty$  erfüllt ist. Da  $\otimes_k$  das  $k$ -fache Kroneckerprodukt identischer Matrizen ist, gilt mit Hilfe von 7.2(7) in Lütkepohl (1996):

$$\mathbf{y}_t \mathbf{y}_t' = \text{vec} \left( Z_t^2 h_t h_t' Z_t^2 \right) \text{vec} \left( Z_t^2 h_t h_t' Z_t^2 \right)' = \left[ \otimes_2 Z_t^2 \right] \mathbf{h}_t \mathbf{h}_t' \left[ \otimes_2 Z_t^2 \right]. \quad (\text{E.1})$$

Eine Vektorisierung beider Seiten von (E.1) bewirkt dann:

$$\text{vec} \left[ \mathbf{y}_t \mathbf{y}_t' \right] = \left[ \otimes_4 Z_t^2 \right] \text{vec} \left[ \mathbf{h}_t \mathbf{h}_t' \right]. \quad (\text{E.2})$$

Abschließend liefert eine Erwartungswertbildung auf beiden Seiten von (E.2):

$$\text{vec} \left( \mathbf{Y}_t \right) = \mathbb{E} \left( \left[ \otimes_4 Z_t^2 \right] \text{vec} \left[ \mathbf{h}_t \mathbf{h}_t' \right] \right). \quad (\text{E.3})$$

Im (erweiterten) CCC-Modell, das He und Teräsvirta (2004) betrachten, sind die Zufallsvektoren  $z_t$  identisch verteilt und unabhängig voneinander und von  $h_t$ . Diese Eigenschaft erlaubt eine direkte Aufspaltung von (E.3) in das Produkt der einzelnen Erwartungen der Potenzen von  $z_t$  und  $h_t$ . Leider ist dieses einfache Vorgehen im Modell mit dynamischer bedingter Korrelation nicht mehr möglich, da  $z_t$  und  $h_t$  messbar bezüglich  $\mathcal{F}_{t-1}$  sind und jeweils von der gesamten Prozessvergangenheit abhängen. Um eine Faktorisierung der Erwartung (E.3) zu ermöglichen, können die bedingten Innovationen  $\eta_t$  genutzt werden. Unter Verwendung der Bezeichnung

$$R_t^{1/2} := [\tilde{r}_{ij,t}]_{i,j=1,\dots,p},$$

gilt für den  $i$ -ten Eintrag des Vektors  $z_t$  und dessen Quadrat:

$$z_{it} = \sum_{j=1}^p \tilde{r}_{ij,t} \eta_{jt} \quad \text{und} \quad z_{it}^2 = \sum_{j=1}^p \sum_{k=1}^p \tilde{r}_{ij,t} \tilde{r}_{ik,t} \eta_{jt} \eta_{kt}.$$

Der  $i$ -te Eintrag der Hauptdiagonalen von  $R_t^{1/2} R_t^{1/2}$  ist gerade  $\sum_{j=1}^p \tilde{r}_{ij,t}^2$ . Da dieser Term auch dem  $i$ -ten Hauptdiagonaleintrag von  $R_t$  entspricht und somit Eins ist, sind alle Einträge von  $R_t^{1/2}$  betragsmäßig durch Eins beschränkt. Diese Eigenschaft impliziert

$$z_{it}^2 < \sum_{j=1}^p \sum_{k=1}^p |\eta_{jt} \eta_{kt}| =: \tilde{z}_t^2 \quad \text{für alle } i \in \{1, \dots, p\}.$$

Wird nun die Struktur der Zufallsvariablen  $\tilde{z}_t$  und  $h_t$  ausgenutzt, so ergibt sich für die Erwartung in (E.3)

$$\text{vec}(\mathbf{Y}_t) \leq \mathbf{E} \left( [\tilde{z}_t^2]^4 \right) \mathbf{E}(\text{vec}[\mathbf{h}_t \mathbf{h}_t']). \quad (\text{E.4})$$

Zunächst muss belegt werden, dass der erste Faktor in (E.4) endlich ist. Dazu liefert eine Substitution der einzelnen Terme

$$\eta_t = R_t^{-1/2} z_t = R_t^{-1/2} D_t^{-1} y_t = R_t^{-1/2} D_t^{-1} [D_t R_t D_t]^{1/2} \epsilon_t.$$

Somit sind die Einträge des Zufallsvektors  $\eta_t$  gewichtete Summen der ursprünglichen Innovationen  $\epsilon_{1t}, \dots, \epsilon_{pt}$ , deren achte Momente und Kreuzmomente gemäß Annahme 5.5 existieren. Somit ist  $\mathbf{E}([\tilde{z}_t^2]^4)$  endlich, falls diese Eigenschaft auch auf die Gewichte zutrifft. Dies ergibt sich direkt aus den Lemmas E.7 und E.14 und aus den Argumenten des Beweises von Lemma E.13 in Abschnitt E.2.

Die Endlichkeit des zweiten Faktors in (E.4) lässt sich wie in He und Teräsvirta (2004) zeigen. Werden zur Modellierung der bedingten Varianzen GARCH(1,1)-Modelle wie in (4.2) unterstellt, so ergibt sich analog zum Vorgehen im Beweis zu Theorem 2 in He und Teräsvirta (2004):

$$\bullet \mathbf{h}_t - [\otimes_2 \mathbf{C}_{t-1}] \mathbf{h}_{t-1} = \text{vec}(\boldsymbol{\omega} \boldsymbol{\omega}') + \sum_{k=1}^2 [E_{k,1} \otimes E_{k,2}] h_{t-1}, \quad (\text{E.5})$$

wobei  $\{(E_{k,1}, E_{k,2}), k \in \{1, 2\}\}$  die Menge aller Permutationen von  $(\boldsymbol{\omega}, \mathbf{C}_{t-1})$  ist.

$$\bullet \text{vec}[\mathbf{h}_t \mathbf{h}_t'] - [\otimes_3 \mathbf{C}_{t-1}] \text{vec}[\mathbf{h}_{t-1} \mathbf{h}_{t-1}'] \quad (\text{E.6})$$

$$= \text{vec}[\text{vec}(\boldsymbol{\omega} \boldsymbol{\omega}') \boldsymbol{\omega}'] + \sum_{k=1}^3 [E_{k,1}^{(1)} \otimes E_{k,2}^{(1)} \otimes E_{k,3}^{(1)}] h_{t-1} + \sum_{k=1}^3 [E_{k,1}^{(2)} \otimes E_{k,2}^{(2)} \otimes E_{k,3}^{(2)}] \mathbf{h}_{t-1},$$

wobei  $\{(E_{k,1}^{(1)}, \dots, E_{k,3}^{(1)}), k \in \{1, \dots, 3\}\}$  und  $\{(E_{k,1}^{(2)}, \dots, E_{k,3}^{(2)}), k \in \{1, \dots, 3\}\}$  die Mengen aller Permutationen von  $(\boldsymbol{\omega}, \boldsymbol{\omega}, \mathbf{C}_{t-1})$  und  $(\boldsymbol{\omega}, \mathbf{C}_{t-1}, \mathbf{C}_{t-1})$  sind.

$$\bullet \text{vec}[\mathbf{h}_t \mathbf{h}_t'] - [\otimes_4 \mathbf{C}_{t-1}] \text{vec}[\mathbf{h}_{t-1} \mathbf{h}_{t-1}'] \quad (\text{E.7})$$

$$= \text{vec}[\text{vec}(\boldsymbol{\omega} \boldsymbol{\omega}') \text{vec}(\boldsymbol{\omega} \boldsymbol{\omega}')'] + \sum_{k=1}^4 [E_{k,1}^{(1)} \otimes E_{k,2}^{(1)} \otimes E_{k,3}^{(1)} \otimes E_{k,4}^{(1)}] h_{t-1}$$

$$+ \sum_{k=1}^6 [E_{k,1}^{(2)} \otimes E_{k,2}^{(2)} \otimes E_{k,3}^{(2)} \otimes E_{k,4}^{(2)}] \mathbf{h}_{t-1} + ([\mathbf{C}_{t-1} \otimes \boldsymbol{\omega} + \boldsymbol{\omega} \otimes \mathbf{C}_{t-1}] \otimes \otimes_2 \mathbf{C}_{t-1}$$

$$+ [\otimes_2 \mathbf{C}_{t-1} \otimes (\mathbf{C}_{t-1} \otimes \boldsymbol{\omega} + \boldsymbol{\omega} \otimes \mathbf{C}_{t-1})] \mathbf{K}_{p^2 p}) \text{vec}(\mathbf{h}_{t-1} \mathbf{h}_{t-1}'),$$

wobei  $\{(E_{k,1}^{(1)}, \dots, E_{k,4}^{(1)}), k \in \{1, \dots, 4\}\}$  und  $\{(E_{k,1}^{(2)}, \dots, E_{k,4}^{(2)}), k \in \{1, \dots, 6\}\}$   
die Mengen aller Permutationen von  $(\omega, \omega, \omega, \mathbf{C}_{t-1})$  und  $(\omega, \omega, \mathbf{C}_{t-1}, \mathbf{C}_{t-1})$  sind.

Im Folgenden werden Matrixpolynome  $\Psi^{(j)}(L)$  im Lagoperator  $L$  betrachtet, die definiert sind als

$$\Psi^{(j)}(L) = \sum_{i=0}^{\infty} \psi_{i,t-1}^{(j)} L^i \quad \text{für } j \in \{1, 2, 3\},$$

$$\begin{aligned} \text{wobei } \psi_{0,t-1}^{(j)} &= \mathbb{I}_{p^{j+1}} \\ \psi_{1,t-1}^{(j)} &= -\otimes_{j+1} \mathbf{C}_{t-1} \\ \psi_{l,t-1}^{(j)} &= 0, \quad \text{für alle } l \geq 2 \text{ und } j \in \{1, 2, 3\}. \end{aligned}$$

Wie in He und Teräsvirta (2004) existiert die Inverse von  $\Psi^{(j)}(L)$  für  $j \in \{1, 2, 3\}$  genau dann, wenn gilt:

$$\lambda_{\max}(\mathbf{E}[\otimes_k \mathbf{C}_0]) < \infty \quad \forall k \in \{1, \dots, j+1\}. \quad (\text{E.8})$$

Werden nun die entsprechenden Inversen von links an die Terme (E.5)-(E.7) multipliziert, erlaubt die Eigenschaft (E.8) jeweils eine Filterdarstellung von  $\mathbf{h}_t$ ,  $\text{vec}(\mathbf{h}_t \mathbf{h}_t')$  und  $\text{vec}(\mathbf{h}_t \mathbf{h}_t')$  basierend auf dem Prozess  $\{h_t\}$ . Da  $h_t < \infty$  fast sicher für alle  $t \in \mathbb{Z}$  gilt, impliziert (E.8) die Endlichkeit und somit die Existenz der rechten Seite von Gleichung (E.4). Daraus folgt die Existenz der achten Momente und Kreuzmomente von  $y_t$ , was den Beweis vervollständigt. ■

## E.2 Beweis von Proposition 5.2

Der Beweis von Proposition 5.2 ist recht umfangreich und wie folgt gegliedert:

Die Aussage  $\widehat{\mathbf{D}}_m \xrightarrow{f.s.} \mathbf{D}$  gilt, wenn

$$\left| \widehat{\mathbf{D}}_m(\hat{\theta}_m) - \mathbf{D}(\theta) \right| \xrightarrow{f.s.} 0 \quad (\text{E.9})$$

erfüllt ist. Die linke Seite von (E.9) lässt sich dabei abschätzen durch:

$$\left| \widehat{\mathbf{D}}_m(\hat{\theta}_m) - \mathbf{D}(\theta) \right| \leq \left| \widehat{\mathbf{D}}_m(\hat{\theta}_m) - \mathbf{D}_m(\hat{\theta}_m) \right| + \left| \mathbf{D}_m(\hat{\theta}_m) - \mathbf{D}_m(\theta) \right| + \left| \mathbf{D}_m(\theta) - \mathbf{D}(\theta) \right|. \quad (\text{E.10})$$

Der Beweis von Proposition 5.2 beruht darauf, dass die fast sichere Konvergenz aller Summanden auf der rechten Seite von (E.10) gegen Null gezeigt wird. Für den ersten Summanden bedeutet dies, dass gezeigt wird, dass die Variationsmatrix, die aus einer endlichen Vergangenheit von Beobachtungen bestimmt wird, ein geeigneter Ersatz für die Matrix mit unendlicher Vergangenheit ist. Es wird somit gezeigt, dass für  $m \rightarrow \infty$  gilt:

$$\sup_{u \in U} \left| \widehat{\mathbf{D}}_m(u) - \mathbf{D}_m(u) \right| \xrightarrow{f.s.} 0 \quad (\text{E.11})$$

mit

$$\widehat{\mathbf{D}}_m(u) = \frac{1}{m} \sum_{t=1}^m \hat{l}_t'(u) \hat{l}_t'(u)^T.$$

Da  $\hat{\theta}_m$  unter den Annahmen 4.2-4.4 ein konsistenter Schätzer ist, strebt der zweite Summand in (E.10) mit dem *Continuous Mapping Theorem* gegen Null. Für den dritten Summanden gilt dies, wenn die Grenzmatrix

$$\mathbf{D}(u) = \mathbf{E} \left[ l'_0(u) l'_0(u)^T \right]$$

tatsächlich für alle  $u \in U$  existiert und außerdem  $\mathbf{D}_m(\cdot)$  in  $U$  gleichmäßig gegen  $\mathbf{D}(\cdot)$  konvergiert, wenn also gilt:

$$\sup_{u \in U} |\mathbf{D}_m(u) - \mathbf{D}(u)|.$$

Dazu muss gezeigt werden, dass gilt:

$$\mathbf{E} \left[ \sup_{u \in U} |l'_0(u) l'_0(u)^T| \right] < \infty. \quad (\text{E.12})$$

Dann folgt die Aussage der Proposition aus der fast sicheren Konvergenz der rechten Seite von (E.10) gegen Null beziehungsweise der Gültigkeit von (E.11) und (E.12).

**Beweis von**  $\sup_{u \in U} |\hat{\mathbf{D}}_m(u) - \mathbf{D}_m(u)| \xrightarrow{f.s.} 0$

Wie in Berkes et al. (2003) wird in diesem Kapitel eine multivariate Version des Mittelwertsatzes verwendet. Diese lässt sich wie folgt formulieren:

**Lemma E.1** Sei  $f : \mathbb{R}^l \rightarrow \mathbb{R}$  eine  $l$ -variante Funktion, die in allen Argumenten stetig ist. Seien weiter  $\mathbf{x}$  und  $\mathbf{y}$  beliebige  $l$ -dimensionale reellwertige Vektoren. Dann existiert ein Vektor  $\xi \in \mathbb{R}^l$  mit  $|\xi - \mathbf{x}| \leq |\mathbf{x} - \mathbf{y}|$  und  $|\xi - \mathbf{y}| \leq |\mathbf{x} - \mathbf{y}|$ , so dass gilt

$$|f(\mathbf{x}) - f(\mathbf{y})| = \left| \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\xi} \right| |\mathbf{x} - \mathbf{y}|. \quad (\text{E.13})$$

**Beweis:** Die beliebigen Vektoren aus  $\mathbb{R}^l$  werden als

$$\mathbf{x} := (x_1, \dots, x_l)' \quad \text{und} \quad \mathbf{y} := (y_1, \dots, y_l)'$$

gewählt. Außerdem wird ein Hilfsvektor

$$\mathbf{y}_i := (x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_l)', \quad \text{für } i \in \{1, \dots, l\},$$

definiert. Dann gilt konstruktionsbedingt:

$$|\mathbf{x} - \mathbf{y}_i| = |x_i - y_i|, \quad \text{für } i \in \{1, \dots, l\}.$$

Eine komponentenweise Anwendung des univariaten Mittelwertsatzes impliziert, dass für jedes  $i \in \{1, \dots, l\}$  ein  $\xi_i$  und ein  $\xi_i$  existieren mit

$$\xi_i \in [x_i, y_i] \quad \text{und} \quad \xi_i := (x_1, \dots, x_{i-1}, \xi_i, x_{i+1}, \dots, x_l)',$$

für die gilt:

$$|f(\mathbf{x}) - f(\mathbf{y}_i)| = \left| \frac{\partial f(\mathbf{x})}{\partial x_i} \Big|_{\mathbf{x}=\xi_i} \right| |x_i - y_i|.$$

Die Wahl  $\xi := (\xi_1, \dots, \xi_l)'$  liefert dann für alle  $i \in \{1, \dots, p\}$ :

$$|\xi_i - \mathbf{x}| = |\xi_i - x_i| \leq |x_i - y_i| \quad \text{und} \quad |\xi_i - \mathbf{y}| \leq |x_i - y_i|$$

und somit auch die Aussage (E.13). ■

Mit Lemma E.1 ergibt sich dann für die linke Seite von (E.11):

$$\begin{aligned} \sup_{u \in U} \left| \widehat{\mathbf{D}}_m(u) - \mathbf{D}_m(u) \right| &= \frac{1}{m} \sup_{u \in U} \left| \sum_{t=1}^m \left[ \hat{l}'_t(u) \hat{l}'_t(u)^T - l'_t(u) l'_t(u)^T \right] \right| \\ &\leq \frac{2d}{m} \sup_{u \in U, t \in \mathbb{Z}} |v_t(u)| \sup_{u \in U} \left| \sum_{t=1}^m \left[ \hat{l}'_t(u) - l'_t(u) \right] \right|, \end{aligned} \quad (\text{E.14})$$

wobei der Vektor  $v_t(\cdot) \in \mathbb{R}^d$  so gestaltet ist, dass gilt:

$$|v_t(u) - l'_t(u)| \leq \left| \hat{l}'_t(u) - l'_t(u) \right| \quad \text{und} \quad |v_t(u) - \hat{l}'_t(u)| \leq \left| \hat{l}'_t(u) - l'_t(u) \right|.$$

Für die Summe in (E.14) gilt außerdem:

$$\begin{aligned} \sup_{u \in U} \left| \sum_{t=1}^m \left[ \hat{l}'_t(u) - l'_t(u) \right] \right| \\ = \max \left\{ \max_{1 \leq j \leq p} \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \hat{l}_t(u)}{\partial r_j} - \frac{\partial l_t(u)}{\partial r_j} \right) \right|, \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \hat{l}_t(u)}{\partial u_2} - \frac{\partial l_t(u)}{\partial u_2} \right) \right| \right\}, \end{aligned}$$

wobei  $u_1 := (r'_1, \dots, r'_p)'$  mit  $r_j := (x_j, s_j, t_j)'$ ,  $j = 1, \dots, p$ , die Varianzparameter und  $u_2 := (a, b, q_1, \dots, q_p)'$  die Korrelationsparameter enthält.

**(I) Der Beweis von**  $\sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \hat{l}_t(u)}{\partial u_1} - \frac{\partial l_t(u)}{\partial u_1} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1)$ .

In diesem Kapitel wird die folgende Aussage wiederholt angewandt. Es handelt sich um eine Verallgemeinerung von Lemma 2.2 in Berkes et al. (2003), das auch als Lemma F.1 in Anhang F zu finden ist. Unter Verwendung der Notation in Berkes et al. (2003), sei

$$\log^+ x := \begin{cases} \log x, & x > 1 \\ 0, & \text{sonst} \end{cases}.$$

**Lemma E.2** Sei  $\{X_t, t \in \mathbb{N}_0\}$  eine Folge von identisch verteilten, aber nicht notwendigerweise unabhängigen Zufallsvariablen, für die gilt:

$$\mathbb{E} \left[ \log^+ |X_0| \right] < \infty. \quad (\text{E.15})$$

Dann konvergiert die Summe  $\sum_{k=0}^{\infty} k^j a^k X_k$  fast sicher für jede Konstante  $a \in \mathbb{R}$  mit  $|a| < 1$  und für jedes feste  $j \in \mathbb{N}_0$ .

**Beweis:** Der Beweis folgt der Struktur des Beweises von Lemma 2.2 in Berkes et al. (2003) und erweitert diesen nur geringfügig, um ihn der hier gegebenen Situation anzupassen. Berkes et al. (2003) nutzen das Borel-Cantelli-Lemma, das als Lemma F.2 in Anhang F zu finden ist, um die fast sichere Konvergenz der betrachteten Reihe zu zeigen. Für eine gegebene Konstante  $\varepsilon > 0$  wird dazu eine Folge von Ereignissen  $A_k$  definiert durch

$$A_k := \left\{ \xi_k a^k > \varepsilon \right\}.$$

Somit wird das Eintreten von  $A_k$  impliziert durch das Eintreten von  $\tilde{A}_k$  mit

$$\tilde{A}_k := \left\{ \xi_k > [a^k]^{-1} \varepsilon^k \right\} \supset \left\{ \xi_k > [a^k]^{-1} \varepsilon^k \varepsilon^{\frac{1}{k-1}} \right\} = A_k.$$

Wenn nun für eine Konstante  $\zeta := \frac{\varepsilon}{a}$  gilt:

$$\sum_{k=0}^{\infty} \mathbf{P} \left( \xi_k > \zeta^k \right) < \infty, \quad (\text{E.16})$$

so folgt mit dem Borel-Cantelli-Lemma, dass nur endlich viele der Ereignisse  $\left\{ \xi_k > \zeta^k \right\}$  eintreten können. Gilt diese Eigenschaft für jedes  $\zeta > 1$  und somit auch für jedes  $\varepsilon > 0$ , so konvergiert die Reihe  $\sum_{k=0}^{\infty} a^k \xi_k$  fast sicher.

Im hier betrachteten Fall muss dieses Prinzip von  $a^k \xi_k$  erweitert werden auf Summanden der Form  $k^j a^k \xi_k$ . Die Funktion

$$f : \mathbb{R}_+ \rightarrow \mathbb{R}_+ \quad \text{mit} \quad f(x) = \frac{\zeta^x}{x^j}$$

besitzt ein Minimum in  $x_{\min} = -\frac{j}{\log \zeta}$  und ist auf  $[x_{\min}, \infty)$  streng monoton steigend in  $x$ . Somit existiert für jedes  $\zeta$  eine Konstante  $\zeta_0$  aus dem Intervall  $(1, \zeta)$ , für die gilt:

$$\frac{\zeta^k}{k^j} > \zeta_0^k \quad \forall k \geq k_0.$$

Dabei ist  $k_0$  die kleinste ganze Zahl größer als  $x_{\min}$ . Für das in dieser Arbeit benötigte Gegenstück von (E.16) beziehungsweise von (2.5) in Berkes et al. (2003) ergibt sich dann:

$$\begin{aligned} \sum_{k=0}^{\infty} \mathbf{P} \left( |X_k| > \frac{\zeta^k}{k^j} \right) &\leq \sum_{k=0}^{k_0-1} \mathbf{P} \left( |X_k| > \frac{\zeta^k}{k^j} \right) + \sum_{k=k_0}^{\infty} \mathbf{P} \left( |X_k| > \frac{\zeta^k}{k^j} \right) \\ &\leq k_0 + \sum_{k=0}^{\infty} \mathbf{P} \left( |X_k| > \zeta_0^k \right). \end{aligned} \quad (\text{E.17})$$

Entsprechend der Argumentation in Berkes et al. (2003) ist die rechte Seite von (E.17) endlich, wenn (E.15) erfüllt ist. Somit liefert das Borel-Cantelli-Lemma die fast sichere Konvergenz der Summe für jede nichtnegative ganze Zahl  $j$ , was den Beweis vervollständigt. ■

**Lemma E.3** Sei  $T_{jt} := \sum_{k=t-1}^{\infty} \rho^k y_{j,t-k-1}^2$ , für  $j = 1, \dots, p$  und  $t = 1, \dots, m$ .  
 Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :

$$\sum_{t=1}^m T_{jt} \stackrel{f.s.}{=} \mathcal{O}(1) \quad \text{und} \quad \left| \sum_{t=1}^m T_{it} y_{it} y_{jt} \right| \stackrel{f.s.}{=} \mathcal{O}(1) \quad \forall i, j \in \{1, \dots, p\}.$$

**Beweis:** Mit den Annahmen 4.1 und 4.5-4.7 gilt, dass  $\{y_{jt}^2, t \in \mathbb{Z}\}$  und  $\{y_{it} y_{jt}, t \in \mathbb{Z}\}$  Folgen von wenn auch nicht bedingt, so doch unbedingt identisch verteilten Zufallsvariablen sind. Unter den Annahmen 5.5 und 5.6 gilt für alle  $i, j \in \{1, \dots, p\}$ :

$$\begin{aligned} \mathbb{E} \left[ \log^+ (y_{jt}^2) \right] &\leq \mathbb{E} \left[ y_{jt}^2 \right] < \infty \\ \text{und} \quad \mathbb{E} \left[ \log^+ (y_{it} y_{jt}) \right] &\leq \mathbb{E} \left[ \log^+ (y_{it}^2 y_{jt}^2) \right] \leq \mathbb{E} \left[ y_{it}^2 y_{jt}^2 \right] < \infty. \end{aligned}$$

Somit gilt mit Lemma E.2 zunächst  $T_{jt} \stackrel{f.s.}{=} \mathcal{O}(1)$ , für alle  $j \in \{1, \dots, p\}$  und  $t \in \mathbb{Z}$ .  
 Des Weiteren ergibt sich für alle  $i, j \in \{1, \dots, p\}$  und für  $m \rightarrow \infty$ :

$$\begin{aligned} \sum_{t=1}^m T_{jt} &= \sum_{t=1}^m \sum_{k=t-1}^{\infty} \rho^k y_{j,t-k-1}^2 = \sum_{t=0}^m \sum_{l=0}^{\infty} \rho^{l+t} y_{jl}^2 = \sum_{t=0}^{m-1} \rho^t \sum_{l=0}^{\infty} \rho^l y_{jl}^2 = \frac{1-\rho^m}{1-\rho} \sum_{l=0}^{\infty} \rho^l y_{jl}^2 \stackrel{a.s.}{=} \mathcal{O}(1) \\ \text{und} \quad \left| \sum_{t=1}^m T_{it} y_{it} y_{jt} \right| &= \left| \sum_{t=1}^m y_{it} y_{jt} \sum_{l=0}^{\infty} \rho^{l+t-1} y_{il}^2 \right| \leq \left| \sum_{t=0}^{m-1} \rho^t y_{i,t+1} y_{j,t+1} \right| \cdot \left| \sum_{l=0}^{\infty} \rho^l y_{il}^2 \right| \stackrel{a.s.}{=} \mathcal{O}(1), \end{aligned}$$

wobei die jeweils letzte Gleichheit in beiden Fällen durch Lemma E.2 impliziert wird. ■

**Lemma E.4** Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für alle  $j \in \{1, \dots, p\}$  und für  $m \rightarrow \infty$ :

$$\begin{aligned} \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{w}_{jt}(u)^{\frac{1}{2}} - w_{jt}(u)^{\frac{1}{2}} \right) \right| &\stackrel{f.s.}{=} \mathcal{O}(1), \quad \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{w}_{jt}(u)^{-\frac{1}{2}} - w_{jt}(u)^{-\frac{1}{2}} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1) \\ \text{und} \quad \sup_{u \in U} \left| \sum_{t=1}^m \left[ \left( \widehat{w}_{it}(u) \widehat{w}_{jt}(u) \right)^{-\frac{1}{2}} - \left( w_{it}(u) w_{jt}(u) \right)^{-\frac{1}{2}} \right] \right| &\stackrel{f.s.}{=} \mathcal{O}(1). \end{aligned}$$

**Beweis:** Analog zu den Argumenten im Beweis von Lemma 5.4 in Berkes et al. (2003), gibt es positive Konstanten  $C_1 := \frac{u}{1-u}$  und  $C_2 := \frac{u}{1-\rho}$  mit

$$0 < C_1 \leq \widehat{w}_{jt}(u) \leq w_{jt}(u) \leq C_2 \left( 1 + \sum_{k=0}^{\infty} \rho^k y_{j,t-k-1}^2 \right) \stackrel{f.s.}{=} \mathcal{O}(1) \quad \forall j \in \{1, \dots, p\}. \quad (\text{E.18})$$

Wegen  $\widehat{w}_{jt}(u) = \min \{ \widehat{w}_{jt}(u), w_{jt}(u) \}$  gilt mit dem Mittelwertsatz für jedes  $j \in \{1, \dots, p\}$ :

$$\sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{w}_{jt}(u)^{\frac{1}{2}} - w_{jt}(u)^{\frac{1}{2}} \right) \right| \leq \sup_{u \in U} \sum_{t=1}^m \frac{1}{2\widehat{w}_{jt}(u)} \left( w_{jt}(u) - \widehat{w}_{jt}(u) \right) \leq \frac{1}{2C_1} \sum_{t=1}^m T_{jt} \stackrel{f.s.}{=} \mathcal{O}(1). \quad (\text{E.19})$$

Die Gültigkeit der zweiten Aussage folgt durch ein Herausziehen von  $\widehat{w}_{jt}(u)^{-1}$  beziehungsweise von  $w_{jt}(u)^{-1}$  mit  $w_{jt}(u)^{-1} \leq \widehat{w}_{jt}(u)^{-1}$  aus (E.18) und (E.19):

$$\sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{w}_{jt}(u)^{-\frac{1}{2}} - w_{jt}(u)^{-\frac{1}{2}} \right) \right| \leq \sup_{u \in U} \sum_{t=1}^m \frac{1}{\widehat{w}_{jt}(u)} \left( w_{jt}(u)^{\frac{1}{2}} - \widehat{w}_{jt}(u)^{\frac{1}{2}} \right) \stackrel{f.s.}{=} \mathcal{O}(1).$$

Unter der Berücksichtigung von  $\widehat{w}_{jt}(u) = \min \{ \widehat{w}_{jt}(u), w_{jt}(u) \}$  gilt mit dem Mittelwertsatz und (E.18) für jedes  $i, j \in \{1, \dots, p\}$ :

$$\begin{aligned} & \sup_{u \in U} \sum_{t=1}^m \left[ \left( \widehat{w}_{it}(u) \widehat{w}_{jt}(u) \right)^{-\frac{1}{2}} - \left( w_{it}(u) w_{jt}(u) \right)^{-\frac{1}{2}} \right] \\ &= \sup_{u \in U} \sum_{t=1}^m \left( \widehat{w}_{it}(u) \widehat{w}_{jt}(u) w_{it}(u) w_{jt}(u) \right)^{-\frac{1}{2}} \left[ \left( w_{it}(u) w_{jt}(u) \right)^{\frac{1}{2}} - \left( \widehat{w}_{it}(u) \widehat{w}_{jt}(u) \right)^{\frac{1}{2}} \right] \\ &\leq \frac{1}{C_1^2} \sup_{u \in U} \sum_{t=1}^m \frac{1}{2} \left( \widehat{w}_{it}(u) \widehat{w}_{jt}(u) \right)^{-\frac{1}{2}} \left| w_{it}(u) w_{jt}(u) - \widehat{w}_{it}(u) \widehat{w}_{jt}(u) \right| \\ &\leq \frac{1}{2C_1^3} \sup_{u \in U} \sum_{t=1}^m \left( \left| w_{it}(u) - \widehat{w}_{it}(u) \right| \widehat{w}_{jt}(u) + \left| w_{jt}(u) - \widehat{w}_{jt}(u) \right| w_{it}(u) \right) \\ &\leq \frac{\rho}{2C_1^3} \sup_{u \in U} \sum_{t=1}^m \left( T_{it} \widehat{w}_{jt}(u) + T_{jt} w_{it}(u) \right) \leq \frac{\rho}{C_1^3} \max_{1 \leq j \leq p} \sup_{u \in U, t \in \mathbb{Z}} w_{jt}(u) \sum_{t=1}^m T_{it} \stackrel{f.s.}{=} \mathcal{O}(1). \end{aligned}$$

■

**Lemma E.5** Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für alle  $j \in \{1, \dots, p\}$  und für  $m \rightarrow \infty$ :

$$\begin{aligned} & \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vec}(\mathbf{F}_{D_t}(u))'}{\partial r_j} \right| \stackrel{f.s.}{=} \mathcal{O}(1) \\ \text{und} \quad & \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vec}(\widehat{\mathbf{F}}_{D_t}(u))'}{\partial r_j} - \frac{\partial \text{vec}(\mathbf{F}_{D_t}(u))'}{\partial r_j} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1). \end{aligned}$$

**Beweis:** Unter Verwendung von Lemma E.2 und (E.18) gilt für alle  $j \in \{1, \dots, p\}$ :

$$\bullet \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vec}(\mathbf{F}_{D_t}(u))'}{\partial x_j} \right| \leq \frac{1}{2} \frac{1}{1 - \rho} \sup_{u \in U, t \in \mathbb{Z}} w_{jt}(u)^{-\frac{1}{2}} \stackrel{f.s.}{=} \mathcal{O}(1) \quad (\text{E.20})$$

$$\bullet \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vec}(\mathbf{F}_{D_t}(u))'}{\partial s_j} \right| \leq \frac{1}{2} \sup_{u \in U, t \in \mathbb{Z}} w_{jt}(u)^{-\frac{1}{2}} \sup_{u \in U, t \in \mathbb{Z}} \sum_{k=0}^{\infty} \rho^k y_{j,t-k-1}^2 \stackrel{f.s.}{=} \mathcal{O}(1) \quad (\text{E.21})$$

$$\bullet \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vec}(\mathbf{F}_{D_t}(u))'}{\partial t_j} \right| \leq \frac{1}{2} \frac{\rho}{\underline{u}} \sup_{u \in U, t \in \mathbb{Z}} w_{jt}(u)^{-\frac{1}{2}} \sup_{u \in U, t \in \mathbb{Z}} \sum_{k=0}^{\infty} k \rho^k y_{j,t-k-1}^2 \stackrel{f.s.}{=} \mathcal{O}(1). \quad (\text{E.22})$$

Die erste Aussage des Lemmas ergibt sich als direkte Konsequenz von (E.20)-(E.22).

Die zweite Aussage wird durch die folgenden Teilaussagen (E.23)-(E.25) impliziert. Mit Lemma E.4 gilt zunächst:

$$\begin{aligned} \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vec}(\widehat{F}_{D_t}(u))'}{\partial x_j} - \frac{\partial \text{vec}(F_{D_t}(u))'}{\partial x_j} \right) \right| \\ \leq \frac{1}{2(1-\rho)} \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{w}_{jt}(u)^{-\frac{1}{2}} - w_{jt}(u)^{-\frac{1}{2}} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1). \end{aligned} \quad (\text{E.23})$$

Außerdem liefern (E.18) und die Lemmas E.3 und E.4:

$$\begin{aligned} \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vec}(\widehat{F}_{D_t}(u))'}{\partial s_j} - \frac{\partial \text{vec}(F_{D_t}(u))'}{\partial s_j} \right) \right| \\ \leq \frac{1}{2} \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{1}{\widehat{w}_{jt}(u)^{\frac{1}{2}}} - \frac{1}{w_{jt}(u)^{\frac{1}{2}}} \right) \right| \sup_{u \in U, t \in \mathbb{Z}} \sum_{k=0}^{t-2} t_j^k y_{j,t-k-1}^2 + \frac{1}{2C_1^{\frac{1}{2}}} \sum_{t=1}^m T_{jt} \stackrel{f.s.}{=} \mathcal{O}(1). \end{aligned} \quad (\text{E.24})$$

Abschließend gilt:

$$\begin{aligned} \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vec}(\widehat{F}_{D_t}(u))'}{\partial t_j} - \frac{\partial \text{vec}(F_{D_t}(u))'}{\partial t_j} \right) \right| \\ \leq \frac{1}{2C_1^{\frac{3}{2}}} \frac{\rho}{u} \sum_{t=1}^m \left[ \sum_{k=t-1}^{\infty} k \rho^k y_{j,t-k-1}^2 \right] \left[ \sum_{k=0}^{t-2} k \rho^k y_{j,t-k-1}^2 \right] + \frac{1}{2C_1^{\frac{1}{2}}} \frac{\rho}{u} \sum_{t=1}^m \sum_{k=t-1}^{\infty} k \rho^k y_{j,t-k-1}^2 \end{aligned} \quad (\text{E.25})$$

mit

$$\sum_{t=1}^m \sum_{k=t-1}^{\infty} k \rho^k y_{j,t-k-1}^2 = \sum_{t=0}^m \sum_{l=0}^{\infty} (l+t) \rho^{l+t} y_{jl}^2 = \sum_{t=0}^m \rho^t \sum_{l=0}^{\infty} l \rho^l y_{jl}^2 + \sum_{t=0}^m t \rho^t \sum_{l=0}^{\infty} \rho^l y_{jl}^2. \quad (\text{E.26})$$

Da (E.26) mit Lemma E.2 fast sicher  $\mathcal{O}(1)$  ist, gilt dies auch für (E.25). ■

**Lemma E.6** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\sup_{u \in U} \left| \sum_{t=1}^m (\widehat{F}_{D_t}(u) - F_{D_t}(u)) \right| \stackrel{f.s.}{=} \mathcal{O}(1) \quad \text{und} \quad \sup_{u \in U} \left| \sum_{t=1}^m (\widehat{F}_{Q_t}(u) - F_{Q_t}(u)) \right| \stackrel{f.s.}{=} \mathcal{O}(1).$$

**Beweis:** Für die erste Aussage impliziert Lemma E.4

$$\sup_{u \in U} \left| \sum_{t=1}^m (\widehat{F}_{D_t}(u) - F_{D_t}(u)) \right| = \max_{1 \leq j \leq p} \left| \sum_{t=1}^m (\widehat{w}_{jt}(u)^{\frac{1}{2}} - w_{jt}(u)^{\frac{1}{2}}) \right| \stackrel{f.s.}{=} \mathcal{O}(1).$$

Für die zweite Aussage liefern (E.18) und die Lemmas E.3 und E.4:

$$\begin{aligned}
& \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{F}_{Q_t}(u) - F_{Q_t}(u) \right) \right| = \sup_{u \in U} a \left| \sum_{t=1}^m \sum_{k=1}^{\infty} b^{k-1} \left( \widehat{z}_t(u) \widehat{z}'_t(u) - z_t(u) z'_t(u) \right) \right| \\
& \leq \sup_{u \in U} \max_{1 \leq i, j \leq p} \rho \left| \sum_{t=1}^m \sum_{k=1}^{\infty} \rho^{k-1} y_{i,t-k} y_{j,t-k} \right. \\
& \quad \times \left. \left[ \left( \widehat{w}_{i,t-k}(u) \widehat{w}_{j,t-k}(u) \right)^{-1/2} - \left( w_{i,t-k}(u) w_{j,t-k}(u) \right)^{-1/2} \right] \right| \\
& \leq \sup_{t \in \mathbb{Z}} \max_{1 \leq i, j \leq p} \left| \sum_{k=1}^{\infty} \rho^k y_{i,t-k} y_{j,t-k} \right| \\
& \quad \times \sup_{u \in U} \left| \sum_{t=1}^m \left[ \left( \widehat{w}_{i,t-k}(u) \widehat{w}_{j,t-k}(u) \right)^{-1/2} - \left( w_{i,t-k}(u) w_{j,t-k}(u) \right)^{-1/2} \right] \right| \stackrel{f.s.}{=} \mathcal{O}(1). \quad \blacksquare
\end{aligned}$$

Im Folgenden verlaufen die Beweise unter der Annahme einer endlichen beziehungsweise unendlichen Vergangenheit analog und werden für jeweils einen dieser beiden Fälle ausgelassen.

**Lemma E.7** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\begin{aligned}
& \sup_{u \in U, t \in \mathbb{Z}} |F_{D_t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1) \quad \text{und} \quad \sup_{u \in U, t \in \mathbb{Z}} |\widehat{F}_{D_t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1); \\
& \sup_{u \in U, t \in \mathbb{Z}} |F_{Q_t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1) \quad \text{und} \quad \sup_{u \in U, t \in \mathbb{Z}} |\widehat{F}_{Q_t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1); \\
& \sup_{u \in U, t \in \mathbb{Z}} |F_{R_t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1) \quad \text{und} \quad \sup_{u \in U, t \in \mathbb{Z}} |\widehat{F}_{R_t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1).
\end{aligned}$$

**Beweis:** Mit Hilfe von (E.18) und Lemma 2.2 in Berkes et al. (2003) gilt:

$$\begin{aligned}
& \sup_{u \in U, t \in \mathbb{Z}} |F_{D_t}(u)| = \sup_{u \in U, t \in \mathbb{Z}} \max_{1 \leq i \leq p} w_{it}(u)^{1/2} \stackrel{f.s.}{=} \mathcal{O}(1) \\
& \text{und} \quad \sup_{u \in U, t \in \mathbb{Z}} |F_{Q_t}(u)| \leq \sup_{u \in U, t \in \mathbb{Z}} \max_{1 \leq i \leq p} \frac{1-a-b}{1-b} q_i + a \sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{k=0}^{\infty} b^k z_{t-k-1}(u) z'_{t-k-1}(u) \right| \\
& \quad \leq \frac{1-2u}{1-\rho} + \rho \sup_{u \in U, t \in \mathbb{Z}} \max_{1 \leq i, j \leq p} \left| \sum_{k=0}^{\infty} \rho^k \frac{y_{i,t-k-1}}{w_{i,t-k-1}(u)^{1/2}} \frac{y_{j,t-k-1}}{w_{j,t-k-1}(u)^{1/2}} \right| \stackrel{f.s.}{=} \mathcal{O}(1).
\end{aligned}$$

Die letzte Gleichheit folgt dabei mit Hilfe von Lemma E.2 und den Annahmen 5.5 und 5.6 aus

$$\begin{aligned}
\mathbb{E} \left[ \log^+ \left( \frac{y_{it}}{w_{it}(u)^{1/2}} \frac{y_{jt}}{w_{jt}(u)^{1/2}} \right) \right] & \leq \mathbb{E} \left[ \log^+ \left( \frac{y_{it}^2}{w_{it}(u)^{1/2}} \frac{y_{jt}^2}{w_{jt}(u)^{1/2}} \right) \right] \\
& \leq \mathbb{E} \left[ \frac{y_{it}^2}{w_{it}(u)^{1/2}} \frac{y_{jt}^2}{w_{jt}(u)^{1/2}} \right] \leq C_1^{-1} \mathbb{E} \left[ y_{it}^2 y_{jt}^2 \right] < \infty.
\end{aligned}$$

Abschließend folgt aus der Eigenschaft, dass  $F_{R_t}(u)$  für jedes  $t \in \mathbb{Z}$  fast sicher eine Korrelationsmatrix ist, dass alle Absoluteinträge der Matrix fast sicher durch 1 beschränkt sind. \blacksquare

In den folgenden Aussagen werden die Determinanten von Matrizen betrachtet. Dazu werden zunächst einige Bezeichnungen eingeführt:

**Definition E.1 Permutationen** (Bosch (2006), S.134ff)

Sei  $S_p$  die Menge aller  $p!$  Permutationen von  $(1, \dots, p)$ . Sei außerdem  $\pi \in S_p$  mit  $\pi = (\pi(1), \dots, \pi(p))'$  eine konkrete Permutation und  $\text{sgn}(\pi) = (-1)^s$  das Vorzeichen der Permutation  $\pi$ , wobei  $s$  die Anzahl der Fehlstände in  $\pi$  ist. Das heißt,  $s$  entspricht der Anzahl aller Paare  $(\pi(i), \pi(j))$ ,  $1 \leq i < j \leq p$ , mit  $\pi(i) > \pi(j)$  und gibt somit an, ob eine gerade oder eine ungerade Anzahl an paarweisen Vertauschungen benachbarter Einträge in  $(1, \dots, p)$  notwendig ist, um die Permutation  $\pi$  zu erhalten.

**Definition E.2 Determinante** (Seber (2008), S.57)

Sei  $A := [a_{ij}]_{i,j=1,\dots,p}$  eine quadratische Matrix. Dann ist die Determinante der Matrix  $A$  gegeben durch

$$\det A := \sum_{\pi \in S_p} \text{sgn}(\pi) \prod_{i=1}^p a_{i\pi(i)}.$$

**Lemma E.8** Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :

$$\sup_{u \in U, t \in \mathbb{Z}} |\det F_{Q_t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1), \quad \sup_{u \in U, t \in \mathbb{Z}} |\det \hat{F}_{Q_t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1)$$

$$\text{und} \quad \sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{t=1}^m (\det \hat{F}_{Q_t}(u) - \det F_{Q_t}(u)) \right| \stackrel{f.s.}{=} \mathcal{O}(1).$$

**Beweis:** Mit Annahme 4.1(6.) gilt für die Determinante der Matrix  $F_{Q_t}(u)$ :

$$\sup_{u \in U, t \in \mathbb{Z}} |\det F_{Q_t}(u)| \leq \left[ \sup_{u \in U, t \in \mathbb{Z}} \lambda_{\max}(F_{Q_t}(u)) \right]^p = \delta_2^p.$$

Für die zweite Aussage liefern die Lemmas E.6 und E.7 die folgende Abschätzung:

$$\begin{aligned} & \sup_{u \in U} \left| \sum_{t=1}^m (\det \hat{F}_{Q_t}(u) - \det F_{Q_t}(u)) \right| \\ &= \sup_{u \in U} \left| \sum_{t=1}^m \sum_{\pi \in S_p} \left( [\hat{F}_{Q_t}(u)]_{1\pi(1)} \cdots [\hat{F}_{Q_t}(u)]_{p\pi(p)} - [F_{Q_t}(u)]_{1\pi(1)} \cdots [F_{Q_t}(u)]_{p\pi(p)} \right) \right| \\ &= \sup_{u \in U} \left| \sum_{t=1}^m \sum_{\pi \in S_p} \sum_{j=1}^{p-1} \left( [F_{Q_t}(u)]_{j\pi(j)} - [\hat{F}_{Q_t}(u)]_{j\pi(j)} \right) \prod_{i=1}^{j-1} [F_{Q_t}(u)]_{i\pi(i)} \prod_{k=j+1}^p [\hat{F}_{Q_t}(u)]_{k\pi(k)} \right| \\ &\leq p!(p-1) \sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{t=1}^m (F_{Q_t}(u) - \hat{F}_{Q_t}(u)) \right| \max_{1 \leq j \leq p} \left[ \sup_{u \in U, t \in \mathbb{Z}} |F_{Q_t}(u)| \right]^{j-1} \\ &\quad \times \max_{1 \leq j \leq p} \left[ \sup_{u \in U, t \in \mathbb{Z}} |\hat{F}_{Q_t}(u)| \right]^{p-j} \stackrel{f.s.}{=} \mathcal{O}(1). \quad \blacksquare \end{aligned}$$

**Lemma E.9** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\sup_{u \in U, t \in \mathbb{Z}} \det F_{Q_t^*}(u) \stackrel{f.s.}{=} \mathcal{O}(1), \quad \sup_{u \in U, t \in \mathbb{Z}} \det \widehat{F}_{Q_t^*}(u) \stackrel{f.s.}{=} \mathcal{O}(1)$$

$$\text{und } \sup_{u \in U} \left| \sum_{t=1}^m \left[ \left( \det \widehat{F}_{Q_t^*}(u) \right)^2 - \left( \det F_{Q_t^*}(u) \right)^2 \right] \right| \stackrel{f.s.}{=} \mathcal{O}(1).$$

**Beweis:** Für die erste Aussage gilt mit Hilfe von (4.7):

$$\sup_{u \in U, t \in \mathbb{Z}} \det F_{Q_t^*}(u) = \sup_{u \in U, t \in \mathbb{Z}} \prod_{i=1}^p [F_{Q_t}(u)]_{ii}^{-\frac{1}{2}} \stackrel{f.s.}{\leq} \sup_{u \in U} \left( \frac{1-a-b}{1-b} \right)^{-\frac{p}{2}} \leq \left( \frac{1-u}{1-\rho} \right)^{\frac{p}{2}} = \mathcal{O}(1).$$

Um die zweite Aussage des Lemmas beweisen zu können, wird das Intervall

$$I_{Q_t^*}(u) := \left[ \xi_t^*(u), \bar{\xi}_t^*(u) \right]$$

mit  $\xi_t^*(u) := \min \left\{ \det \widehat{F}_{Q_t^*}(u), \det F_{Q_t^*}(u) \right\}$  und  $\bar{\xi}_t^*(u) := \max \left\{ \det \widehat{F}_{Q_t^*}(u), \det F_{Q_t^*}(u) \right\}$  betrachtet. Dann existiert gemäß Mittelwertsatz (MWS) für jedes  $t \in \mathbb{Z}$  und  $u \in U$ , ein  $\xi_t^*(u) \in I_{Q_t^*}(u)$ , für das unter Verwendung der Lemmas E.6, E.9 und (4.7) gilt:

$$\begin{aligned} & \sup_{u \in U} \left| \sum_{t=1}^m \left[ \left( \det \widehat{F}_{Q_t^*}(u) \right)^2 - \left( \det F_{Q_t^*}(u) \right)^2 \right] \right| \stackrel{MWS}{=} \sup_{u \in U} \left| \sum_{t=1}^m 2\xi_t^*(u) \left[ \det \widehat{F}_{Q_t^*}(u) - \det F_{Q_t^*}(u) \right] \right| \\ & \leq 2 \sup_{u \in U, t \in \mathbb{Z}} \bar{\xi}_t^*(u) \sup_{u \in U} \left| \sum_{t=1}^m \left( \prod_{i=1}^p [\widehat{F}_{Q_t}(u)]_{ii}^{-\frac{1}{2}} - \prod_{i=1}^p [F_{Q_t}(u)]_{ii}^{-\frac{1}{2}} \right) \right| \\ & = 2 \sup_{u \in U, t \in \mathbb{Z}} \bar{\xi}_t^*(u) \sup_{u \in U} \left| \sum_{t=1}^m \sum_{j=1}^p \left( [\widehat{F}_{Q_t}(u)]_{jj}^{-\frac{1}{2}} - [F_{Q_t}(u)]_{jj}^{-\frac{1}{2}} \right) \prod_{i=1}^{j-1} [F_{Q_t}(u)]_{ii}^{-\frac{1}{2}} \prod_{k=j+1}^p [\widehat{F}_{Q_t}(u)]_{kk}^{-\frac{1}{2}} \right| \\ & \stackrel{MWS}{\leq} 2 \sup_{u \in U, t \in \mathbb{Z}} \bar{\xi}_t^*(u) \sup_{u \in U} \sum_{j=1}^p \left| \sum_{t=1}^m \frac{1}{2} [F_{Q_t}(u)]_{jj}^{-\frac{3}{2}} \left( [\widehat{F}_{Q_t}(u)]_{jj} - [F_{Q_t}(u)]_{jj} \right) \right| \sup_{u \in U} \left( \frac{1-a-b}{1-b} \right)^{-\frac{p-1}{2}} \\ & \leq \left( \frac{1-u}{1-\rho} \right)^{\frac{p-1}{2}} \sup_{u \in U, t \in \mathbb{Z}} \bar{\xi}_t^*(u) p \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{F}_{Q_t}(u) - F_{Q_t}(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1). \quad \blacksquare \end{aligned}$$

**Lemma E.10** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\sup_{u \in U, t \in \mathbb{Z}} \frac{1}{\det F_{R_t}(u)} \stackrel{f.s.}{=} \mathcal{O}(1), \quad \sup_{u \in U, t \in \mathbb{Z}} \frac{1}{\det \widehat{F}_{R_t}(u)} \stackrel{f.s.}{=} \mathcal{O}(1)$$

$$\text{und } \sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{t=1}^m \left( \frac{1}{\det \widehat{F}_{R_t}(u)} - \frac{1}{\det F_{R_t}(u)} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).$$

**Beweis:** Für die inverse Determinante gilt mit Hilfe von (4.8):

$$\sup_{u \in U, t \in \mathbb{Z}} \det F_{R_t}(u) \geq \left( \sup_{u \in U, t \in \mathbb{Z}} \lambda_{\min}(F_{R_t}(u)) \right)^p \stackrel{f.s.}{>} \left[ \frac{1-a-b}{1-b} \frac{\delta_1}{\delta_2^p} \right]^p. \quad (\text{E.27})$$

Aus (E.27) folgt dann:

$$\sup_{u \in U, t \in \mathbb{Z}} \frac{1}{\det \mathbb{F}_{R_t}(u)} \stackrel{f.s.}{<} \left[ \frac{1-b}{1-a-b} \frac{\delta_2^p}{\delta_1} \right]^p = \mathcal{O}(1).$$

Für die letzte Aussage kann das Intervall

$$I_{R_t}(u) := [\underline{\xi}_t(u), \bar{\xi}_t(u)]$$

mit  $\underline{\xi}_t(u) := \min \left\{ \det \widehat{\mathbb{F}}_{R_t}(u), \det \mathbb{F}_{R_t}(u) \right\}$  und  $\bar{\xi}_t(u) := \max \left\{ \det \widehat{\mathbb{F}}_{R_t}(u), \det \mathbb{F}_{R_t}(u) \right\}$  betrachtet werden. Mit dem Mittelwertsatz existiert dann für jedes  $t \in \mathbb{Z}$  und  $u \in U$  ein  $\xi_t(u) \in I_{R_t}(u)$  für das mit Hilfe der Lemmas E.8 und E.9 gilt:

$$\begin{aligned} \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{1}{\det \widehat{\mathbb{F}}_{R_t}(u)} - \frac{1}{\det \mathbb{F}_{R_t}(u)} \right) \right| &\leq \sup_{u \in U, t \in \mathbb{Z}} \left| -\frac{1}{\xi_t^2(u)} \right| \sup_{u \in U} \left| \sum_{t=1}^m \left( \det \widehat{\mathbb{F}}_{R_t}(u) - \det \mathbb{F}_{R_t}(u) \right) \right| \\ &\leq \sup_{u \in U, t \in \mathbb{Z}} \frac{1}{\xi_t^2(u)} \sup_{u \in U} \left| \sum_{t=1}^m \left( [\det \widehat{\mathbb{F}}_{Q_t^*}(u)]^2 \det \widehat{\mathbb{F}}_{Q_t}(u) - [\det \mathbb{F}_{Q_t^*}(u)]^2 \det \mathbb{F}_{Q_t}(u) \right) \right| \\ &\leq \sup_{u \in U, t \in \mathbb{Z}} \frac{1}{\xi_t^2(u)} \sup_{u \in U, t \in \mathbb{Z}} \det \widehat{\mathbb{F}}_{Q_t}(u) \sup_{u \in U} \left| \sum_{t=1}^m \left( [\det \widehat{\mathbb{F}}_{Q_t^*}(u)]^2 - [\det \mathbb{F}_{Q_t^*}(u)]^2 \right) \right| \\ &\quad + \sup_{u \in U, t \in \mathbb{Z}} \frac{1}{\xi_t^2(u)} \left[ \sup_{u \in U, t \in \mathbb{Z}} \det \mathbb{F}_{Q_t^*}(u) \right]^2 \sup_{u \in U} \left| \sum_{t=1}^m \left( \det \widehat{\mathbb{F}}_{Q_t}(u) - \det \mathbb{F}_{Q_t}(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1) \quad \blacksquare \end{aligned}$$

**Definition E.3** *Streichungsmatrix* (Knabner und Barth (2013), S.289)

Sei  $X$  eine beliebige  $(p \times p)$ -Matrix. Dann sei  $X^{(i,j)}$  die Streichungsmatrix zur Stelle  $(i, j)$ , also diejenige  $(p-1 \times p-1)$ -Matrix, die durch Streichung der  $i$ -ten Zeile und  $j$ -ten Spalte aus  $X$  entsteht.

**Lemma E.11** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\sup_{u \in U, t \in \mathbb{Z}} \det \mathbb{F}_{R_t}(u)^{(i,j)} \stackrel{f.s.}{=} \mathcal{O}(1), \quad \sup_{u \in U, t \in \mathbb{Z}} \det \widehat{\mathbb{F}}_{R_t}(u)^{(i,j)} \stackrel{f.s.}{=} \mathcal{O}(1)$$

$$\text{und } \sup_{u \in U} \left| \sum_{t=1}^m \left( \det \widehat{\mathbb{F}}_{R_t}(u)^{(i,j)} - \det \mathbb{F}_{R_t}(u)^{(i,j)} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).$$

**Beweis:** Der Beweis von Lemma E.11 funktioniert analog zum Beweis von Lemma E.8 und nutzt zusätzlich einige der Argumente aus dem Beweis von Lemma E.10. Aufgrund der hohen Ähnlichkeit mit den genannten Beweisen wird der Beweis hier nicht im Detail dargelegt.

**Lemma E.12** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{t=1}^m \left( \widehat{F}_{D_t}(u)^{-1} - F_{D_t}(u)^{-1} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1)$$

und

$$\sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{t=1}^m \left( \widehat{F}_{R_t}(u)^{-1} - F_{R_t}(u)^{-1} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).$$

**Beweis:** Mit Hilfe der Lemmas E.4, E.10 und E.11 gilt:

$$\sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{F}_{D_t}(u)^{-1} - F_{D_t}(u)^{-1} \right) \right| = \sup_{u \in U} \max_{1 \leq j \leq p} \left| \sum_{t=1}^m \left( \widehat{w}_{jt}(u)^{-\frac{1}{2}} - w_{jt}(u)^{-\frac{1}{2}} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).$$

Außerdem gilt:

$$\begin{aligned} & \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{F}_{R_t}(u)^{-1} - F_{R_t}(u)^{-1} \right) \right| \\ &= \sup_{u \in U} \max_{1 \leq i, j \leq p} \left| \sum_{t=1}^m \left( \frac{1}{\det \widehat{F}_{R_t}(u)} \det \widehat{F}_{R_t}(u)^{(i,j)} - \frac{1}{\det F_{R_t}(u)} \det F_{R_t}(u)^{(i,j)} \right) \right| \\ &\leq \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{1}{\det \widehat{F}_{R_t}(u)} - \frac{1}{\det F_{R_t}(u)} \right) \right| \max_{1 \leq i, j \leq p} \sup_{u \in U, t \in \mathbb{Z}} \left| \det \widehat{F}_{R_t}(u)^{(i,j)} \right| \\ &\quad + \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{1}{\det F_{R_t}(u)} \right| \max_{1 \leq i, j \leq p} \sup_{u \in U} \left| \sum_{t=1}^m \left( \det \widehat{F}_{R_t}(u)^{(i,j)} - \det F_{R_t}(u)^{(i,j)} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1). \quad \blacksquare \end{aligned}$$

**Lemma E.13** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\begin{aligned} & \sup_{u \in U, t \in \mathbb{Z}} \left| F_{D_t}(u)^{1/2} \right| \stackrel{f.s.}{=} \mathcal{O}(1) \quad \text{und} \quad \sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{F}_{D_t}(u)^{1/2} \right| \stackrel{f.s.}{=} \mathcal{O}(1); \\ & \sup_{u \in U, t \in \mathbb{Z}} \left| F_{Q_t}(u)^{1/2} \right| \stackrel{f.s.}{=} \mathcal{O}(1) \quad \text{und} \quad \sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{F}_{Q_t}(u)^{1/2} \right| \stackrel{f.s.}{=} \mathcal{O}(1); \\ & \sup_{u \in U, t \in \mathbb{Z}} \left| F_{Q_t^*}(u)^{1/2} \right| \stackrel{f.s.}{=} \mathcal{O}(1) \quad \text{und} \quad \sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{F}_{Q_t^*}(u)^{1/2} \right| \stackrel{f.s.}{=} \mathcal{O}(1). \end{aligned}$$

**Beweis:** Zunächst gilt mit (E.18):

$$\sup_{u \in U, t \in \mathbb{Z}} \left| F_{D_t}(u)^{1/2} \right| = \sup_{u \in U, t \in \mathbb{Z}} \max_{1 \leq i \leq p} w_{it}(u)^{1/4} \stackrel{f.s.}{=} \mathcal{O}(1).$$

Sei  $U_t(u)$  eine Matrix, deren Spalten den orthonormalisierten Eigenvektoren der Matrix  $F_{Q_t}(u)$  entsprechen. Sei außerdem  $\Lambda_t(u)$  eine Diagonalmatrix, deren Hauptdiagonale mit den Eigenwerten der Matrix  $F_{Q_t}(u)$  besetzt ist. Dann ergibt sich eine Spektralzerlegung der Matrix  $F_{Q_t}(u)$  als

$$F_{Q_t}(u) = U_t(u) \Lambda_t(u) U_t(u)'$$

Aufgrund der Konstruktion der Matrix  $U_t(u)$  aus normierten Eigenvektoren gilt:

$$\sup_{u \in U, t \in \mathbb{Z}} |U_t(u)| \leq 1 \quad (\text{E.28})$$

und als Konsequenz aus Lemma E.7:

$$\lambda_{\max}(\mathbf{F}_{Q_t}(u)) \stackrel{f.s.}{=} \mathcal{O}(1). \quad (\text{E.29})$$

Aus (E.28) und (E.29) folgt dann:

$$\begin{aligned} \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{Q_t}(u)^{\frac{1}{2}} \right| &= \sup_{u \in U, t \in \mathbb{Z}} \left| U_t(u) \Lambda_t^{\frac{1}{2}}(u) U_t(u)' \right| \\ &\stackrel{f.s.}{\leq} 2p^2 \sup_{u \in U, t \in \mathbb{Z}} |U_t(u)| \sup_{u \in U, t \in \mathbb{Z}} \lambda_{\max}(\mathbf{F}_{Q_t}(u))^{\frac{1}{2}} \stackrel{f.s.}{=} \mathcal{O}(1) \\ \text{und } \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{Q_t^*}(u)^{\frac{1}{2}} \right| &\stackrel{f.s.}{\leq} \sup_{u \in U} \left| \frac{1-a-b}{1-b} \min_{1 \leq i \leq p} \left[ \mathbf{F}_{\bar{Q}}(u) \right]_{ii} \right|^{-\frac{1}{4}} \stackrel{f.s.}{\leq} \left( \frac{1-u}{1-2\rho} \right)^{\frac{1}{4}} = \mathcal{O}(1). \quad \blacksquare \end{aligned}$$

Für den Beweis des folgenden Lemmas werden zusätzliche Kennzahlen von Matrizen benötigt.

**Definition E.4 Inverse Matrix** (Knabner und Barth (2013), S.290ff, und Horn und Johnson (1985), S.17)

Sei  $A$  eine beliebige  $(p \times p)$ -Matrix. Bezeichne

$$M_{ij} := \det A^{(i,j)}, \quad \text{für } 1 \leq i, j \leq p,$$

als Minor der Matrix  $A$  zur Stelle  $(i, j)$  und

$$\mathbf{a}_{ij} := (-1)^{i+j} M_{ij}$$

als den zugehörigen Kofaktor. Bezeichne außerdem

$$A^{adj} := [\mathbf{a}_{ji}]_{i,j=1,\dots,p}$$

als Adjunkte der Matrix  $A$ . Dann kann die Inverse von  $A$  angegeben werden als

$$A^{-1} = \frac{1}{\det A} A^{adj}.$$

**Lemma E.14** Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :

$$\begin{aligned} \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{D_t}(u)^{-1} \right| &\stackrel{f.s.}{=} \mathcal{O}(1) \quad \text{und} \quad \sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{\mathbf{F}}_{D_t}(u)^{-1} \right| \stackrel{f.s.}{=} \mathcal{O}(1); \\ \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{R_t}(u)^{-1} \right| &\stackrel{f.s.}{=} \mathcal{O}(1) \quad \text{und} \quad \sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{\mathbf{F}}_{R_t}(u)^{-1} \right| \stackrel{f.s.}{=} \mathcal{O}(1). \end{aligned}$$

**Beweis:** Die Aussage (E.18) impliziert

$$\sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{D_t}(u)^{-1} \right| = \sup_{u \in U, t \in \mathbb{Z}} \max_{1 \leq j \leq p} w_{jt}(u)^{-1/2} \leq C_1^{-1/2} \stackrel{f.s.}{=} \mathcal{O}(1).$$

Um zu beweisen, dass  $\sup_{u \in U, t \in \mathbb{Z}} |\mathbf{F}_{R_t}(u)^{-1}| \stackrel{f.s.}{=} \mathcal{O}(1)$  gilt, werden die Einträge der Matrix  $\mathbf{F}_{R_t}(u)^{-1}$  im Detail betrachtet. Sei

$$M_{ij,t}(u) := \det \mathbf{F}_{R_t}(u)^{(i,j)}$$

der Minor von  $\mathbf{F}_{R_t}(u)$  zur Stelle  $(i, j)$ . Da die Einträge von  $\mathbf{F}_{R_t}(u)$  im Betrag kleiner als 1 sind, ist der Betrag von  $M_{ij,t}(u)$  beschränkt durch

$$\sup_{u \in U, t \in \mathbb{Z}} \max_{1 \leq i, j \leq p} |M_{ij,t}(u)| \leq \sup_{u \in U, t \in \mathbb{Z}} \max_{1 \leq i, j \leq p} \sum_{\pi \in S_{p-1}} \prod_{k=1}^{p-1} \left| \left[ \mathbf{F}_{R_t}(u)^{(i,j)} \right]_{k\pi(k)} \right| < (p-1)! \quad (\text{E.30})$$

Somit liefern (4.8) und (E.30) analog zur Argumentation im Beweis von Lemma E.10:

$$\begin{aligned} \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{R_t}(u)^{-1} \right| &\leq \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{1}{\det \mathbf{F}_{R_t}(u)} \right| \max_{1 \leq i, j \leq p} \sup_{u \in U, t \in \mathbb{Z}} \left| \det \mathbf{F}_{R_t}(u)^{(i,j)} \right| \\ &\stackrel{f.s.}{<} \left[ \frac{1-b}{1-a-b} \frac{\delta_2^p}{\delta_1} \right]^p (p-1)! = \mathcal{O}(1). \end{aligned}$$

■

**Lemma E.15** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\sup_{u \in U, t \in \mathbb{Z}} |z_t(u)z'_t(u)| \stackrel{f.s.}{=} \mathcal{O}(1), \quad \sup_{u \in U, t \in \mathbb{Z}} |\hat{z}_t(u)\hat{z}'_t(u)| \stackrel{f.s.}{=} \mathcal{O}(1)$$

$$\text{und } \sup_{u \in U} \left| \sum_{t=1}^m \left( \hat{z}_t(u)\hat{z}'_t(u) - z_t(u)z'_t(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).$$

**Beweis:** Unter Verwendung von

$$\mathbf{F}_{H_t}(u) := \mathbf{F}_{D_t}(u)\mathbf{F}_{R_t}(u)\mathbf{F}_{D_t}(u)$$

ergibt sich mit Hilfe von Annahme 5.5 und den Lemmas E.13 und E.14:

$$\begin{aligned} \sup_{u \in U, t \in \mathbb{Z}} |z_t(u)z'_t(u)| &\leq p^2 \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{D_t}(u)^{-1} \right| \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{H_t}(u)^{1/2} \epsilon_t \epsilon'_t \mathbf{F}_{H_t}(u)^{1/2} \right| \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{D_t}(u)^{-1} \right| \\ &\leq p^4 \left( \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{D_t}(u)^{-1} \right| \right)^2 \left( \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{D_t}(u)^{1/2} \right| \right)^4 \left( \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{R_t}(u)^{1/2} \right| \right)^2 \sup_{t \in \mathbb{Z}} |\epsilon_t \epsilon'_t| \\ &\leq p^6 \left( \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{D_t}(u)^{-1} \right| \right)^2 \left( \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{D_t}(u)^{1/2} \right| \right)^4 \left( \sup_{u \in U} \left| \mathbf{F}_{Q_t^*}(u)^{1/2} \right| \right)^4 \left( \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{Q_t}(u)^{1/2} \right| \right)^2 \\ &\quad \times \sup_{t \in \mathbb{Z}} |\epsilon_t \epsilon'_t| \stackrel{f.s.}{=} \mathcal{O}(1). \end{aligned}$$

Für den zweiten Teil der Aussage gilt mit Hilfe von (E.18) und Lemma E.3:

$$\begin{aligned}
& \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{\mathbf{z}}_t(u) \widehat{\mathbf{z}}_t'(u) - \mathbf{z}_t(u) \mathbf{z}_t'(u) \right) \right| \\
&= \sup_{u \in U} \max_{1 \leq i, j \leq p} \left| \sum_{t=1}^m y_{it} y_{jt} \left[ \left( \widehat{\mathbf{w}}_{it}(u) \widehat{\mathbf{w}}_{jt}(u) \right)^{-\frac{1}{2}} - \left( \mathbf{w}_{it}(u) \mathbf{w}_{jt}(u) \right)^{-\frac{1}{2}} \right] \right| \\
&\leq \frac{\rho}{2C_1^3} \sup_{u \in U} \max_{1 \leq i, j \leq p} \left| \sum_{t=1}^m \left( T_{it} \widehat{\mathbf{w}}_{jt}(u) + T_{jt} \mathbf{w}_{it}(u) \right) y_{it} y_{jt} \right| \\
&\leq \frac{\rho}{2C_1^3} \max_{1 \leq i \leq p} \sup_{u \in U, t \in \mathbb{Z}} \mathbf{w}_{it}(u) \max_{1 \leq i, j \leq p} \left( \left| \sum_{t=1}^m T_{it} y_{it} y_{jt} \right| + \left| \sum_{t=1}^m T_{jt} y_{it} y_{jt} \right| \right) \stackrel{f.s.}{=} \mathcal{O}(1).
\end{aligned}$$

Im Folgenden wird die Bezeichnung

$$K_{1t}(u) := \mathbf{F}_{D_t}(u)^{-1} + \mathbf{F}_{D_t}(u)^{-1} \mathbf{F}_{R_t}(u)^{-1} \mathbf{z}_t(u) \mathbf{z}_t'(u)$$

und analog  $\widehat{K}_{1t}(u)$  in Abhängigkeit von einer endlichen Vergangenheit verwendet.

**Lemma E.16** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\sup_{u \in U, t \in \mathbb{Z}} |K_{1t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1), \quad \sup_{u \in U, t \in \mathbb{Z}} |\widehat{K}_{1t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1)$$

$$\text{und } \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{K}_{1t}(u) - K_{1t}(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).$$

**Beweis:** Die Lemmas E.12, E.14 und E.15 implizieren, dass gilt:

$$\begin{aligned}
& \sup_{u \in U, t \in \mathbb{Z}} |K_{1t}(u)| \leq \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{D_t}(u)^{-1} \right| \\
& \quad + p^2 \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{D_t}(u)^{-1} \right| \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{R_t}(u)^{-1} \right| \sup_{u \in U, t \in \mathbb{Z}} |\mathbf{z}_t(u) \mathbf{z}_t'(u)| \stackrel{f.s.}{=} \mathcal{O}(1) \\
& \text{und } \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{K}_{1t}(u) - K_{1t}(u) \right) \right| \\
& \leq p^2 \sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{\mathbf{F}}_{R_t}(u)^{-1} \right| \sup_{u \in U, t \in \mathbb{Z}} |\widehat{\mathbf{z}}_t(u) \widehat{\mathbf{z}}_t'(u)| \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{\mathbf{F}}_{D_t}(u)^{-1} - \mathbf{F}_{D_t}(u)^{-1} \right) \right| \\
& \quad + p^2 \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{D_t}(u)^{-1} \right| \sup_{u \in U, t \in \mathbb{Z}} |\widehat{\mathbf{z}}_t(u) \widehat{\mathbf{z}}_t'(u)| \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{\mathbf{F}}_{R_t}(u)^{-1} - \mathbf{F}_{R_t}(u)^{-1} \right) \right| \\
& \quad + p^2 \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{D_t}(u)^{-1} \right| \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{R_t}(u)^{-1} \right| \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{\mathbf{z}}_t(u) \widehat{\mathbf{z}}_t'(u) - \mathbf{z}_t(u) \mathbf{z}_t'(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).
\end{aligned}$$

**Lemma E.17** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial l_t(u)}{\partial r_j} \right| \stackrel{f.s.}{=} \mathcal{O}(1) \quad \text{und} \quad \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \hat{l}_t(u)}{\partial r_j} - \frac{\partial l_t(u)}{\partial r_j} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).$$

**Beweis:** Mit Hilfe der Lemmas E.5, E.12 und E.16 gilt:

$$\begin{aligned} \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial l_t(u)}{\partial r_j} \right| &\leq \sup_{u \in U, t \in \mathbb{Z}} \frac{\partial \text{vec}(\mathbf{F}_{D_t}(u))'}{\partial r_j} \sup_{u \in U, t \in \mathbb{Z}} K_{1t}(u) \stackrel{f.s.}{=} \mathcal{O}(1) \\ \text{und} \quad \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \hat{l}_t(u)}{\partial r_j} - \frac{\partial l_t(u)}{\partial r_j} \right) \right| \\ &= \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vec}(\widehat{\mathbf{F}}_{D_t}(u))'}{\partial r_j} \widehat{K}_{1t}(u) - \frac{\partial \text{vec}(\mathbf{F}_{D_t}(u))'}{\partial r_j} K_{1t}(u) \right) \right| \\ &\leq p \sup_{u \in U, t \in \mathbb{Z}} |\widehat{K}_{1t}(u)| \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vec}(\widehat{\mathbf{F}}_{D_t}(u))'}{\partial r_j} - \frac{\partial \text{vec}(\mathbf{F}_{D_t}(u))'}{\partial r_j} \right) \right| \\ &\quad + p \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vec}(\mathbf{F}_{D_t}(u))'}{\partial r_j} \right| \sup_{u \in U} \left| \sum_{t=1}^m (\widehat{K}_{1t}(u) - K_{1t}(u)) \right| \\ &\quad + p \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vec}(\mathbf{F}_{D_t}(u))'}{\partial r_j} \right| \sup_{u \in U} \left| \sum_{t=1}^m (\widehat{\mathbf{F}}_{D_t}(u)^{-1} - \mathbf{F}_{D_t}(u)^{-1}) \right| \stackrel{f.s.}{=} \mathcal{O}(1). \quad \blacksquare \end{aligned}$$

**(II) Der Beweis von**  $\sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \hat{l}_t(u)}{\partial u_2} - \frac{\partial l_t(u)}{\partial u_2} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1)$ .

Im Folgenden wird die Bezeichnung

$$K_{2t}(u) := \mathbf{F}_{R_t}(u)^{-1} - \mathbf{F}_{R_t}(u)^{-1} \mathbf{z}_t(u) \mathbf{z}'_t(u) \mathbf{F}_{R_t}(u)^{-1}$$

und analog  $\widehat{K}_{2t}(u)$  in Abhängigkeit von einer endlichen Vergangenheit verwendet.

**Lemma E.18** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\begin{aligned} \sup_{u \in U, t \in \mathbb{Z}} |K_{2t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1), \quad \sup_{u \in U, t \in \mathbb{Z}} |\widehat{K}_{2t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1) \\ \text{und} \quad \sup_{u \in U} \left| \sum_{t=1}^m (\widehat{K}_{2t}(u) - K_{2t}(u)) \right| \stackrel{f.s.}{=} \mathcal{O}(1). \end{aligned}$$

**Beweis:** Mit den Lemmas E.12, E.14 und E.15, gilt:

$$\begin{aligned} \sup_{u \in U, t \in \mathbb{Z}} |K_{2t}(u)| \\ \leq \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{R_t}(u)^{-1} \right| - p^2 \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{R_t}(u)^{-1} \right| \sup_{u \in U, t \in \mathbb{Z}} |\mathbf{z}_t(u) \mathbf{z}'_t(u)| \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{R_t}(u)^{-1} \right| \stackrel{f.s.}{=} \mathcal{O}(1) \end{aligned}$$

$$\begin{aligned}
& \text{und } \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{K}_{2t}(u) - K_{2t}(u) \right) \right| \\
& \leq \sup_{u \in U, t \in \mathbb{Z}} |\widehat{z}_t(u) \widehat{z}'_t(u)| \left( \sup_{u \in U, t \in \mathbb{Z}} |\widehat{F}_{R_t}(u)^{-1}| + \sup_{u \in U, t \in \mathbb{Z}} |F_{R_t}(u)^{-1}| \right) \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{F}_{R_t}(u)^{-1} - F_{R_t}(u)^{-1} \right) \right| \\
& \quad + \sup_{u \in U, t \in \mathbb{Z}} |\widehat{F}_{R_t}(u)^{-1}| \sup_{u \in U, t \in \mathbb{Z}} |F_{R_t}(u)^{-1}| \sup_{u \in U} \left| \sum_{t=1}^m \left( \widehat{z}_t(u) \widehat{z}'_t(u) - z_t(u) z'_t(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).
\end{aligned}$$

■

**Lemma E.19** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\begin{aligned}
& \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vech}(F_{Q_t}(u))'}{\partial u_2} \right| \stackrel{f.s.}{=} \mathcal{O}(1) \\
& \text{und } \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vech}(F_{Q_t}(u))'}{\partial u_2} - \frac{\partial \text{vech}(\widehat{F}_{Q_t}(u))'}{\partial u_2} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).
\end{aligned}$$

**Beweis:** Für den ersten Teil der Aussage gilt mit Hilfe von  $\sup_{u \in U} |F_{\bar{Q}}(u)| \stackrel{f.s.}{\leq} 1$ :

$$\bullet \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vech}(F_{Q_t}(u))'}{\partial \text{vecl}(F_{\bar{Q}}(u))} \right| = \frac{1-a-b}{1-b} = \mathcal{O}(1) \quad (\text{E.31})$$

$$\bullet \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vech}(F_{Q_t}(u))'}{\partial a} \right| \leq \frac{1}{1-\rho} + \sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{n=0}^{\infty} \rho^n z_{t-n-1}(u) z'_{t-n-1}(u) \right| \stackrel{f.s.}{=} \mathcal{O}(1) \quad (\text{E.32})$$

$$\bullet \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vech}(F_{Q_t}(u))'}{\partial b} \right| \leq \frac{\bar{u}}{(1-\rho)^2} + \sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{n=0}^{\infty} n \rho^{n-1} z_{t-n-1}(u) z'_{t-n-1}(u) \right| \stackrel{f.s.}{=} \mathcal{O}(1). \quad (\text{E.33})$$

Dabei werden (E.32) und (E.33) durch die Lemmas E.2 und E.24 impliziert. Das letztere Lemma wurde dabei aus Gründen der Übersichtlichkeit auf S.124 in Abschnitt E.2.1 verschoben. Die Aussage (E.31) wiederum gilt mit den Lemmas E.7 und E.15. Des Weiteren gilt mit Hilfe der Lemmas E.6 und E.15:

$$\begin{aligned}
& \bullet \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vech}(F_{Q_t}(u))'}{\partial \text{vecl}(F_{\bar{Q}}(u))} - \frac{\partial \text{vech}(\widehat{F}_{Q_t}(u))'}{\partial \text{vecl}(F_{\bar{Q}}(u))} \right) \right| = 0 \\
& \bullet \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vech}(F_{Q_t}(u))'}{\partial a} - \frac{\partial \text{vech}(\widehat{F}_{Q_t}(u))'}{\partial a} \right) \right| \\
& \quad = \sup_{u \in U} \left| \sum_{t=1}^m \sum_{k=0}^{\infty} b^k \left( z_{t-k}(u) z'_{t-k}(u) - \widehat{z}_{t-k}(u) \widehat{z}'_{t-k}(u) \right) \right| \\
& \quad \leq \frac{1}{1-b} \sup_{u \in U} \left| \sum_{t=1}^m \left( z_{t-k}(u) z'_{t-k}(u) - \widehat{z}_{t-k}(u) \widehat{z}'_{t-k}(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1)
\end{aligned}$$

$$\begin{aligned}
& \bullet \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vech}(\mathbf{F}_{Q_t}(u))'}{\partial b} - \frac{\partial \text{vech}(\widehat{\mathbf{F}}_{Q_t}(u))'}{\partial b} \right) \right| \\
&= \sup_{u \in U} \left| \sum_{t=1}^m \sum_{k=0}^{\infty} b^k \left( \text{vech}(\mathbf{F}_{Q_t}(u)) - \text{vech}(\widehat{\mathbf{F}}_{Q_t}(u)) \right) \right| \\
&\leq \frac{1}{1-b} \sup_{u \in U} \left| \sum_{t=1}^m \left( \mathbf{F}_{Q_t}(u) - \widehat{\mathbf{F}}_{Q_t}(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).
\end{aligned}$$

**Lemma E.20** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\begin{aligned}
& \sup_{u \in U} \left| \sum_{t=1}^m \left( \mathbf{F}_{Q_t^*}(u) \otimes \mathbf{F}_{Q_t^*}(u) - \widehat{\mathbf{F}}_{Q_t^*}(u) \otimes \widehat{\mathbf{F}}_{Q_t^*}(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1) \\
& \text{und} \quad \sup_{u \in U} \left| \sum_{t=1}^m \left( \mathbf{F}_{Q_t^*}(u) \mathbf{F}_{Q_t}(u) - \widehat{\mathbf{F}}_{Q_t^*}(u) \widehat{\mathbf{F}}_{Q_t}(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).
\end{aligned}$$

**Beweis:** Mit Hilfe der Lemmas E.6 und E.7 sowie dem Mittelwertsatz gilt:

$$\begin{aligned}
& \sup_{u \in U} \left| \sum_{t=1}^m \left( \mathbf{F}_{Q_t^*}(u) \otimes \mathbf{F}_{Q_t^*}(u) - \widehat{\mathbf{F}}_{Q_t^*}(u) \otimes \widehat{\mathbf{F}}_{Q_t^*}(u) \right) \right| \\
&= \sup_{u \in U} \max_{1 \leq i, j \leq p} \left| \sum_{t=1}^m \left( [\mathbf{F}_{Q_t}(u)]_{ii}^{-1/2} [\mathbf{F}_{Q_t}(u)]_{jj}^{-1/2} - [\widehat{\mathbf{F}}_{Q_t}(u)]_{ii}^{-1/2} [\widehat{\mathbf{F}}_{Q_t}(u)]_{jj}^{-1/2} \right) \right| \\
&\leq 2 \sup_{u \in U} \max_{1 \leq j \leq p} \left| [\mathbf{F}_{Q_t}(u)]_{jj}^{-1/2} \right| \sup_{u \in U} \max_{1 \leq i \leq p} \left| \sum_{t=1}^m \left( [\mathbf{F}_{Q_t}(u)]_{ii}^{-1/2} - [\widehat{\mathbf{F}}_{Q_t}(u)]_{ii}^{-1/2} \right) \right| \\
&\leq 2 \sup_{u \in U} \left| \min_{1 \leq j \leq p} [\mathbf{F}_{Q_t}(u)]_{jj} \right|^{-1/2} \frac{1}{2} \sup_{u \in U} \left| \min_{1 \leq j \leq p} [\mathbf{F}_{Q_t}(u)]_{jj} \right|^{-3/2} \sup_{u \in U} \left| \sum_{t=1}^m \left( \mathbf{F}_{Q_t}(u) - \widehat{\mathbf{F}}_{Q_t}(u) \right) \right| \\
&\leq \left( \frac{1-u}{1-\rho} \right)^2 \sup_{u \in U} \left| \sum_{t=1}^m \left( \mathbf{F}_{Q_t}(u) - \widehat{\mathbf{F}}_{Q_t}(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1)
\end{aligned}$$

$$\begin{aligned}
& \text{und} \quad \sup_{u \in U} \left| \sum_{t=1}^m \left( \mathbf{F}_{Q_t^*}(u) \mathbf{F}_{Q_t}(u) - \widehat{\mathbf{F}}_{Q_t^*}(u) \widehat{\mathbf{F}}_{Q_t}(u) \right) \right| \\
&\leq \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{Q_t}(u) \right| \sup_{u \in U} \left| \sum_{t=1}^m \left( \mathbf{F}_{Q_t^*}(u) - \widehat{\mathbf{F}}_{Q_t^*}(u) \right) \right| \\
&\quad + \sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{\mathbf{F}}_{Q_t^*}(u) \right| \sup_{u \in U} \left| \sum_{t=1}^m \left( \mathbf{F}_{Q_t}(u) - \widehat{\mathbf{F}}_{Q_t}(u) \right) \right| \\
&= \left( \sup_{u \in U, t \in \mathbb{Z}} \left| \mathbf{F}_{Q_t}(u) \right| + \sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{\mathbf{F}}_{Q_t^*}(u) \right| \right) \sup_{u \in U} \left| \sum_{t=1}^m \left( \mathbf{F}_{Q_t}(u) - \widehat{\mathbf{F}}_{Q_t}(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).
\end{aligned}$$

Im Folgenden wird die Bezeichnung

$$K_{3t}(u) := D_{p,-}^+ \left( F_{Q_t^*}(u) \otimes F_{Q_t^*}(u) \right) D_p \\ + D_{p,-}^+ \left( F_{Q_t^*}(u) F_{Q_t}(u) \otimes \mathbb{I}_p + \mathbb{I}_p \otimes F_{Q_t^*}(u) F_{Q_t}(u) \right) D_p \frac{\partial \text{vech} \left( F_{Q_t^*}(u) \right)}{\partial \text{vech} \left( F_{Q_t}(u) \right)'}$$

und analog  $\widehat{K}_{3t}(u)$  in Abhängigkeit von einer endlichen Vergangenheit verwendet.

**Lemma E.21** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\sup_{u \in U, t \in \mathbb{Z}} |K_{3t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1), \quad \sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{K}_{3t}(u) \right| \stackrel{f.s.}{=} \mathcal{O}(1) \\ \text{und} \quad \sup_{u \in U} \left| \sum_{t=1}^m \left( K_{3t}(u) - \widehat{K}_{3t}(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).$$

**Beweis:** Zunächst gilt:

$$\sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vech} \left( F_{Q_t^*}(u) \right)}{\partial \text{vech} \left( F_{Q_t}(u) \right)'} \right| \leq \frac{1}{2} \sup_{u \in U, t \in \mathbb{Z}} \left( \min_{1 \leq j \leq p} \left| [F_{Q_t}(u)]_{jj} \right| \right)^{-\frac{3}{2}} \stackrel{f.s.}{=} \left( \frac{1-u}{1-\rho} \right)^{\frac{3}{2}} = \mathcal{O}(1). \quad (\text{E.34})$$

Des Weiteren gilt mit dem Mittelwertsatz und Lemma E.6:

$$\sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vech} \left( F_{Q_t^*}(u) \right)}{\partial \text{vech} \left( F_{Q_t}(u) \right)'} - \frac{\partial \text{vech} \left( \widehat{F}_{Q_t^*}(u) \right)}{\partial \text{vech} \left( F_{Q_t}(u) \right)'} \right) \right| \\ = \frac{1}{2} \max_{1 \leq j \leq p} \sup_{u \in U} \left| \sum_{t=1}^m \left( [F_{Q_t}(u)]_{jj}^{-\frac{3}{2}} - [\widehat{F}_{Q_t}(u)]_{jj}^{-\frac{3}{2}} \right) \right| \\ \leq \frac{3}{4} \sup_{u \in U, t \in \mathbb{Z}} \left( \min_{1 \leq j \leq p} [F_{Q_t}(u)]_{jj} \right)^{-\frac{5}{2}} \sup_{u \in U} \left| \sum_{t=1}^m \left( F_{Q_t}(u) - \widehat{F}_{Q_t}(u) \right) \right| \\ \leq \frac{3}{4} \left( \frac{1-u}{1-\rho} \right)^{\frac{5}{2}} \sup_{u \in U} \left| \sum_{t=1}^m \left( F_{Q_t}(u) - \widehat{F}_{Q_t}(u) \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1). \quad (\text{E.35})$$

In Kombination mit den Lemmas E.7 und E.20 implizieren die Aussagen (E.34) und (E.35), dass gilt:

$$\sup_{u \in U, t \in \mathbb{Z}} |K_{3t}(u)| \\ \leq \sup_{u \in U, t \in \mathbb{Z}} \left| F_{Q_t^*}(u) \otimes F_{Q_t^*}(u) \right| + \sup_{u \in U, t \in \mathbb{Z}} \left| F_{Q_t^*}(u) F_{Q_t}(u) \right| \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vech} \left( F_{Q_t^*}(u) \right)}{\partial \text{vech} \left( F_{Q_t}(u) \right)'} \right| \\ \leq \left( \sup_{u \in U, t \in \mathbb{Z}} \left| F_{Q_t^*}(u) \right| \right)^2 + p \sup_{u \in U, t \in \mathbb{Z}} \left| F_{Q_t^*}(u) \right| \sup_{u \in U, t \in \mathbb{Z}} \left| F_{Q_t}(u) \right| \left( \frac{1-u}{1-\rho} \right)^{3/2} \stackrel{f.s.}{=} \mathcal{O}(1)$$

$$\begin{aligned}
& \text{und } \sup_{u \in U} \left| \sum_{t=1}^m (K_{3t}(u) - \widehat{K}_{3t}(u)) \right| \leq \sup_{u \in U} \left| \sum_{t=1}^m (F_{Q_t^*}(u) \otimes F_{Q_t}(u) - \widehat{F}_{Q_t^*}(u) \otimes \widehat{F}_{Q_t}(u)) \right| \\
& + 2 \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vech}(\widehat{F}_{Q_t^*}(u))}{\partial \text{vech}(F_{Q_t}(u))'} \right| \sup_{u \in U} \left| \sum_{t=1}^m (F_{Q_t^*}(u) F_{Q_t}(u) - \widehat{F}_{Q_t^*}(u) \widehat{F}_{Q_t}(u)) \right| \\
& + 2 \sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{F}_{Q_t^*}(u) \widehat{F}_{Q_t}(u) \right| \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vech}(F_{Q_t^*}(u))}{\partial \text{vech}(F_{Q_t}(u))'} - \frac{\partial \text{vech}(\widehat{F}_{Q_t^*}(u))}{\partial \text{vech}(F_{Q_t}(u))'} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1). \blacksquare
\end{aligned}$$

**Lemma E.22** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\begin{aligned}
& \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vecl}(F_{R_t}(u))'}{\partial u_2} \right| \stackrel{f.s.}{=} \mathcal{O}(1), \quad \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vecl}(\widehat{F}_{R_t}(u))'}{\partial u_2} \right| \stackrel{f.s.}{=} \mathcal{O}(1) \\
& \text{und } \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vecl}(F_{R_t}(u))'}{\partial u_2} - \frac{\partial \text{vecl}(\widehat{F}_{R_t}(u))'}{\partial u_2} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).
\end{aligned}$$

**Beweis:** Mit Hilfe der Lemmas E.19 und E.21, gilt:

$$\begin{aligned}
& \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vecl}(F_{R_t}(u))'}{\partial u_2} \right| \leq \sup_{u \in U, t \in \mathbb{Z}} |K_{3t}(u)| \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vech}(F_{Q_t}(u))'}{\partial u_2} \right| \stackrel{f.s.}{=} \mathcal{O}(1) \\
& \text{und } \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vecl}(F_{R_t}(u))'}{\partial u_2} - \frac{\partial \text{vecl}(\widehat{F}_{R_t}(u))'}{\partial u_2} \right) \right| \\
& \leq \sup_{u \in U, t \in \mathbb{Z}} |K_{3t}(u)| \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vech}(F_{Q_t}(u))'}{\partial u_2} - \frac{\partial \text{vech}(F_{Q_t}(u))'}{\partial u_2} \right) \right| \\
& + \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vech}(F_{Q_t}(u))'}{\partial u_2} \right| \sup_{u \in U} \left| \sum_{t=1}^m (K_{3t}(u) - \widehat{K}_{3t}(u)) \right| \stackrel{f.s.}{=} \mathcal{O}(1). \blacksquare
\end{aligned}$$

**Lemma E.23** *Unter den Annahmen 4.1-4.7, 5.1-5.3, 5.5 und 5.6 gilt für  $m \rightarrow \infty$ :*

$$\sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial l_t(u)}{\partial u_2} \right| \stackrel{f.s.}{=} \mathcal{O}(1) \quad \text{und} \quad \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \hat{l}_t(u)}{\partial u_2} - \frac{\partial l_t(u)}{\partial u_2} \right) \right| \stackrel{f.s.}{=} \mathcal{O}(1).$$

**Beweis:** Mit Hilfe der Lemmas E.18 und E.19, gilt:

$$\sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial l_t(u)}{\partial u_2} \right| \leq \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vecl}(F_{R_t}(u))'}{\partial u_2} \right| \sup_{u \in U, t \in \mathbb{Z}} |K_{2t}(u)| \stackrel{f.s.}{=} \mathcal{O}(1)$$

$$\begin{aligned}
& \text{und } \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \hat{l}_t(u)}{\partial u_2} - \frac{\partial l_t(u)}{\partial u_2} \right) \right| \\
& \leq \sup_{u \in U, t \in \mathbb{Z}} |K_{2t}(u)| \sup_{u \in U} \left| \sum_{t=1}^m \left( \frac{\partial \text{vecl}(F_{R_t}(u))'}{\partial u_2} - \frac{\partial \text{vecl}(\hat{F}_{R_t}(u))'}{\partial u_2} \right) \right| \\
& \quad + \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vecl}(\hat{F}_{R_t}(u))'}{\partial u_2} \right| \sup_{u \in U} \left| \sum_{t=1}^m (F_{R_t}(u)^{-1} - \hat{F}_{R_t}(u)^{-1}) \right| \\
& \quad + \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{\partial \text{vecl}(\hat{F}_{R_t}(u))'}{\partial u_2} \right| \sup_{u \in U} \left| \sum_{t=1}^m (\hat{K}_{2t}(u) - K_{2t}(u)) \right| \stackrel{f.s.}{=} \mathcal{O}(1).
\end{aligned}$$

Aus den Lemmas E.17 und E.23 folgt dann:

$$\sup_{u \in U} \left| \hat{D}_m(u) - D_m(u) \right| \stackrel{f.s.}{=} \mathcal{O}\left(\frac{1}{m}\right).$$

Dies vervollständigt den Beweis von Aussage (E.11). ■

**Beweis von**  $\mathbb{E} \left[ \sup_{u \in U} |l'_0(u)l'_0(u)^T| \right] < \infty$

Analog zum Vorgehen in Berkes et al. (2003), gilt mit Hilfe der Lemmas E.17 und E.23:

$$\sup_{u \in U} |l'_0(u)l'_0(u)^T| \leq \left( \sup_{u \in U} \left| \frac{\partial l_0(u)}{\partial u} \right| \right)^2 \stackrel{f.s.}{=} \mathcal{O}(1).$$

Dies impliziert

$$\mathbb{E} \left[ \sup_{u \in U} |l'_0(u)l'_0(u)^T| \right] < \infty.$$

### Beweis der gleichmäßigen Konvergenz von $D_m(\cdot)$ gegen $D(\cdot)$

Die gleichmäßige Konvergenz kann mit Hilfe von Theorem A.2.2 in White (1994) nachgewiesen werden, das sich als Theorem F.2 in Anhang F findet. Die Bedingungen des Theorems sind erfüllt, da  $U$  eine kompakte Menge und  $l'_t(u)l'_t(u)^T$  ergodisch sowie stetig in  $u$  für alle  $y_t$  und messbar in  $y_t$  für alle  $u \in U$  ist. Wird nun die dominierende Funktion als  $\sup_{u \in U} |l'_t(u)l'_t(u)^T|$  gewählt, so folgt die Existenz der Erwartung aus (E.12). Dies impliziert die gleichmäßige Konvergenz von  $D_m(\cdot)$  gegen  $D(\cdot)$ . ■

### Beweis von Proposition 5.2

Da  $\hat{\theta}_m$  ein stark konsistenter Schätzer für den Parametervektor ist und  $D$  mit Annahme 5.3 eine nichtsinguläre Matrix ist, folgt  $\hat{D}_m \xrightarrow{f.s.} D$  direkt aus diesen beiden Eigenschaften gemeinsam mit der gleichmäßigen Konvergenz von  $D_m(\cdot)$  gegen  $D(\cdot)$ . ■

### E.2.1 Beweis, dass die zweiten partiellen Ableitungen von endlicher Erwartung sind

#### Grundlegendes zur Notation

Für  $\varphi \in (0, 1)$  und  $i, j \in \{1, \dots, p\}$  wird in diesem Abschnitt die folgende Schreibweise verwendet:

##### Definition E.5

$$G_0^y(i, j, \varphi) := \sum_{n=0}^{\infty} \varphi^n y_{i, -n-1} y_{j, -n-1} \quad (\text{E.36})$$

$$G_1^y(i, j, \varphi) := \sum_{n=0}^{\infty} n \varphi^{n-1} y_{i, -n-1} y_{j, -n-1} \quad (\text{E.37})$$

$$G_2^y(i, j, \varphi) := \sum_{n=0}^{\infty} n(n-1) \varphi^{n-2} y_{i, -n-1} y_{j, -n-1} \quad (\text{E.38})$$

$$G_0^z(i, j, u) := \sum_{n=0}^{\infty} b^n z_{i, -n-1}(u) z_{j, -n-1}(u) \quad (\text{E.39})$$

$$G_1^z(i, j, u) := \sum_{n=0}^{\infty} n b^{n-1} z_{i, -n-1}(u) z_{j, -n-1}(u) \quad (\text{E.40})$$

$$G_2^z(i, j, u) := \sum_{n=0}^{\infty} n(n-1) b^{n-2} z_{i, -n-1}(u) z_{j, -n-1}(u) \quad (\text{E.41})$$

$$Z(u) := \left( \sum_{s=1}^p z_{s0}(u) \right)^2 = \sum_{s=1}^p \sum_{t=1}^s z_{t0}(u) z_{s-t,0}(u). \quad (\text{E.42})$$

Mit (E.18) gilt für alle  $u \in U$ ,  $k \in \{1, \dots, 3\}$  und  $i, j \in \{1, \dots, p\}$ :

$$G_k^z(i, j, u) \stackrel{f.s.}{\leq} C_1^{-1} G_k^y(i, j, b).$$

Außerdem belegt die folgende Aussage die Existenz der Erwartungen einiger Produkte der in Definition E.5 eingeführten Terme:

**Lemma E.24** *Seien  $G_1, G_2, G_3$  und  $G_4$  beliebige Terme aus (E.36)-(E.42) für beliebige Indizes  $i, j \in \{1, \dots, p\}$  und Parameterwerte  $\varphi \in (0, 1)$  beziehungsweise  $u \in U$ . Dann gilt:*

$$\mathbb{E} \left( \prod_{i=1}^k G_i \right) < \infty \quad \text{für } k = 1, \dots, 4 \quad (\text{E.43})$$

**Beweis:** Der Beweis kann anhand des Produkts von vier Termen der Art (E.36) demonstriert und beliebig erweitert werden. Dass (E.43) auch dann Gültigkeit behält, wenn einzelne oder alle Faktoren durch Terme der Art (E.37)-(E.42) ersetzt werden, folgt mit analoger Argumentation.

Sei nun  $G_k = G_0^y(i_k, j_k, \varphi_k)$  für  $k = 1, \dots, 4$  und  $\bar{\varphi} := \max\{\varphi_1, \dots, \varphi_4\}$ .

Zunächst gilt für  $\varphi_1, \varphi_2 \in (0, 1)$  und alle Kombinationen von  $i, j, q, r \in \{1, \dots, p\}$  und unter Verwendung der Cauchyproduktregel:

$$\begin{aligned} G_0^y(i, j, \varphi_1) G_0^y(q, r, \varphi_2) &= \sum_{m=0}^{\infty} \sum_{n=0}^m \varphi_1^n \varphi_2^{m-n} y_{i, -n-1} y_{j, -n-1} y_{q, -m+n-1} y_{r, -m+n-1} \\ &\leq \sum_{m=0}^{\infty} \left(\bar{\varphi}^{\frac{1}{2}}\right)^m \sum_{n=0}^m \left(\bar{\varphi}^{\frac{1}{2}}\right)^n y_{i, -n-1} y_{j, -n-1} y_{q, -m+n-1} y_{r, -m+n-1}. \end{aligned} \quad (\text{E.44})$$

Unter den Annahmen 5.5 und 5.6 ist die Doppelsumme in (E.44) stochastisch beschränkt, da mit der Cauchy-Schwarz-Ungleichung für alle  $s, t \in \mathbb{Z}$  und alle  $i, j, k, l \in \{1, \dots, p\}$  gilt:

$$\mathbb{E}|y_{is} y_{js} y_{kt} y_{lt}| \leq \left[\mathbb{E}\left(y_{is}^2 y_{js}^2\right)\right]^{1/2} \left[\mathbb{E}\left(y_{kt}^2 y_{lt}^2\right)\right]^{1/2} < \infty.$$

Analog sind auch alle Produkte  $G_{m_1}^y(i, j, \varphi_1) G_{m_2}^y(k, l, \varphi_2)$  oder  $G_{m_1}^z(i, j, u) G_{m_2}^z(k, l, u)$  mit  $\varphi_1, \varphi_2 \in (0, 1)$  und  $m_1, m_2 \in \{0, 1, 2\}$  von endlicher Erwartung. Diese Eigenschaft kann auf Produkte von bis zu vier verschiedenen Termen wie in (E.36)-(E.42) erweitert werden. Dies folgt direkt aus den Annahmen 5.5 und 5.6 sowie einer wiederholten Anwendung der Cauchyproduktregel sowie der verallgemeinerten Hölder-Ungleichung auf die resultierenden Summanden. Die verallgemeinerte Hölder-Ungleichung findet sich als Lemma F.3 in Anhang F. Als Parameter werden in diesem Fall  $m = 8$ ,  $q = 1$  und  $p_i = 8$ , für  $i = 1, \dots, 8$ , gewählt. ■

Allgemein gilt mit Lemma E.14 für die inverse Matrix  $F_{R_0}(u)^{-1} := \left[r_{ij,0}^-(u)\right]_{i,j=1,\dots,p}$ :

$$\sup_{u \in U, t \in \mathbb{Z}} \left| r_{ij,t}^-(u) \right| \stackrel{f.s.}{<} \left[ \frac{1-b}{1-a-b} \frac{\delta_2^p}{\delta_1} \right]^p (p-1)! =: \delta_* \quad \forall i, j \in \{1, \dots, p\}. \quad (\text{E.45})$$

Diese Aussage wird in den nächsten Abschnitten der Arbeit wiederholt genutzt.

### Die partiellen Ableitungen von $w_{i0}(u)$

Die folgenden Aussagen über die partiellen Ableitungen von  $w_{i0}(u)$  nach verschiedenen Parametern werden in den folgenden Abschnitten benötigt:

$$\left(\frac{\partial w_{i0}(u)}{\partial x_i}\right)^2 = \frac{1}{4} \frac{1}{(1-t_i)^2} \quad \text{und} \quad \left(\frac{\partial w_{i0}(u)}{\partial s_i}\right)^2 = \frac{1}{4} G_0^y(i, i, t_i)^2 \quad (\text{E.46})$$

$$\left(\frac{\partial w_{i0}(u)}{\partial t_i}\right)^2 = \frac{1}{4} \frac{x_i^2}{(1-t_i)^4} + \frac{1}{2} \frac{x_i}{(1-t_i)^2} G_1^y(i, i, t_i) + \frac{1}{4} G_1^y(i, i, t_i)^2 \quad (\text{E.47})$$

$$\begin{aligned} \frac{\partial w_{i0}(u)}{\partial x_i} \frac{\partial w_{i0}(u)}{\partial s_i} &= \frac{1}{4} \frac{1}{1-t_i} G_0^y(i, i, t_i) \\ \frac{\partial w_{i0}(u)}{\partial x_i} \frac{\partial w_{i0}(u)}{\partial t_i} &= \frac{1}{4} \frac{x_i}{(1-t_i)^3} + \frac{1}{4} \frac{1}{1-t_i} G_1^y(i, i, t_i) \end{aligned} \quad (\text{E.48})$$

$$\frac{\partial w_{i0}(u)}{\partial s_i} \frac{\partial w_{i0}(u)}{\partial t_i} = \frac{1}{4} \frac{x_i}{(1-t_i)^2} G_0^y(i, i, t_i) + \frac{1}{4} G_0^y(i, i, t_i) G_1^y(i, i, t_i). \quad (\text{E.49})$$

Außerdem gilt für die partiellen Ableitungen der Quadratwurzel von  $w_{j0}(u)$ :

$$\frac{\partial w_{j0}(u)^{\frac{1}{2}}}{\partial x_j} \leq \frac{1}{2C_1^{\frac{1}{2}}(1-t_j)} \quad \text{und} \quad \frac{\partial w_{j0}(u)^{\frac{1}{2}}}{\partial s_j} \leq \frac{G_0^y(j, j, t_j)}{2C_1^{\frac{1}{2}}} \quad (\text{E.50})$$

$$\frac{\partial w_{j0}(u)^{\frac{1}{2}}}{\partial t_j} \leq \frac{G_1^y(j, j, t_j)}{2C_1^{\frac{1}{2}}} \quad (\text{E.51})$$

$$\left| \frac{\partial^2 w_{j0}(u)^{\frac{1}{2}}}{(\partial x_j)^2} \right| \leq \frac{1}{4C_1^{\frac{3}{2}}(1-t_j)^2} \quad \text{und} \quad \left| \frac{\partial^2 w_{j0}(u)^{\frac{1}{2}}}{(\partial s_j)^2} \right| \leq \frac{G_0^y(j, j, t_j)^2}{4C_1^{\frac{3}{2}}} \quad (\text{E.52})$$

$$\left| \frac{\partial^2 w_{j0}(u)^{\frac{1}{2}}}{(\partial t_j)^2} \right| \leq \frac{G_1^y(j, j, t_j)}{4C_1^{\frac{3}{2}}} + \frac{G_2^y(j, j, t_j)}{2C_1^{\frac{1}{2}}} \quad (\text{E.53})$$

$$\left| \frac{\partial^2 w_{j0}(u)^{\frac{1}{2}}}{\partial x_j \partial s_j} \right| \leq \frac{G_0^y(j, j, t_j)}{4C_1^{\frac{3}{2}}(1-t_j)} \quad \text{und} \quad \left| \frac{\partial^2 w_{j0}(u)^{\frac{1}{2}}}{\partial x_j \partial t_j} \right| \leq \frac{1}{2C_1^{\frac{1}{2}}(1-t_j)^2} + \frac{G_1^y(j, j, t_j)}{4C_1^{\frac{3}{2}}(1-t_j)} \quad (\text{E.53})$$

$$\left| \frac{\partial^2 w_{j0}(u)^{\frac{1}{2}}}{\partial s_j \partial t_j} \right| \leq \frac{G_0^y(j, j, t_j)^2}{4C_1^{\frac{3}{2}}} + \frac{G_1^y(j, j, t_j)}{2C_1^{\frac{1}{2}}}. \quad (\text{E.54})$$

### Die partiellen Ableitungen von $F_{Q_0}(u)$ und $F_{Q_0^*}(u)$

Für die Ableitungen der einzelnen Einträge der Matrix  $F_{Q_0}(u)$  gilt:

$$\frac{\partial [F_{Q_0}(u)]_{ij}}{\partial a} = G_0^z(i, j, u) - \frac{1}{1-b} \quad \text{und} \quad \frac{\partial [F_{Q_0}(u)]_{ij}}{\partial b} = aG_1^z(i, j, u) - \frac{a}{(1-b)^2} \quad (\text{E.55})$$

$$\frac{\partial^2 [F_{Q_0}(u)]_{ij}}{(\partial a)^2} = 0 \quad \text{und} \quad \frac{\partial^2 [F_{Q_0}(u)]_{ij}}{(\partial b)^2} = aG_2^z(i, j, u) - \frac{2a}{(1-b)^3} \quad (\text{E.56})$$

$$\frac{\partial^2 [F_{Q_0}(u)]_{ij}}{\partial a \partial b} = G_1^z(i, j, u) - \frac{1}{(1-b)^2}. \quad (\text{E.57})$$

Die Einträge der Matrix  $F_{Q_0^*}(u)$  hingegen sind entweder 0 oder  $[F_{Q_0}(u)]_{ii}^{-\frac{1}{2}}$  für  $i \in \{1, \dots, p\}$ . Deshalb sind im Folgenden nur die Ableitungen der von Null verschiedenen Einträge von Interesse.

$$\frac{\partial^2 [F_{Q_0^*}(u)]_{ii}}{(\partial a)^2} = \frac{3}{4} [F_{Q_0}(u)]_{ii}^{-\frac{5}{2}} \left( \frac{\partial [F_{Q_0}(u)]_{ii}}{\partial a} \right)^2 < \frac{3a^2}{4} \left( \frac{1-b}{1-a-b} \right)^{\frac{5}{2}} G_0^z(i, i, u)^2 \quad (\text{E.58})$$

$$\begin{aligned} \frac{\partial^2 [F_{Q_0^*}(u)]_{ii}}{(\partial b)^2} &= \frac{3}{4} [F_{Q_0}(u)]_{ii}^{-\frac{5}{2}} \left( \frac{\partial [F_{Q_0}(u)]_{ii}}{\partial b} \right)^2 - \frac{1}{2} [F_{Q_0}(u)]_{ii}^{-\frac{3}{2}} \frac{\partial^2 [F_{Q_0}(u)]_{ii}}{(\partial b)^2} \\ &< \frac{3a^2}{4} \left( \frac{1-b}{1-a-b} \right)^{\frac{5}{2}} G_1^z(i, i, u)^2 + \frac{1}{2} \left( \frac{1-b}{1-a-b} \right)^{\frac{3}{2}} \frac{2}{(1-\rho)^3} \end{aligned} \quad (\text{E.59})$$

$$\begin{aligned} \frac{\partial^2 [F_{Q_0^*}(u)]_{ii}}{\partial a \partial b} &= \frac{3}{4} [F_{Q_0}(u)]_{ii}^{-\frac{5}{2}} \frac{\partial [F_{Q_0}(u)]_{ii}}{\partial a} \frac{\partial [F_{Q_0}(u)]_{ii}}{\partial b} - \frac{1}{2} [F_{Q_0}(u)]_{kk}^{-\frac{3}{2}} \frac{\partial^2 [F_{Q_0}(u)]_{ii}}{\partial a \partial b} \\ &< \frac{3a^2}{4} \left( \frac{1-b}{1-a-b} \right)^{\frac{5}{2}} G_0^z(i, i, u) G_1^z(i, i, u) + \frac{1}{2} \left( \frac{1-b}{1-a-b} \right)^{\frac{3}{2}} \frac{1}{(1-\rho)^2} \end{aligned} \quad (\text{E.60})$$

Mit Lemma E.24 sind alle Terme (E.46)-(E.60) stochastisch beschränkt.

**Beweis, dass die Erwartung von  $\frac{\partial^2 l_0(u)}{\partial u_1 \partial u_1'}$  endlich ist**

Für die Erwartung der partiellen Ableitungen zweiter Ordnung nach den Varianzparametern gilt:

$$\mathbb{E} \left( \frac{\partial^2 l_0(u)}{\partial u_1 \partial u_1'} \right) = -\frac{1}{2} \mathbb{E} \left( \frac{\partial \text{vec}(\mathbf{F}_{D_0}(u))'}{\partial u_1} \left[ 2 \left( \mathbf{F}_{D_0}(u)^{-1} \otimes \mathbf{F}_{D_0}(u)^{-1} \right) \right. \right. \quad (\text{E.61})$$

$$\left. + \left( \mathbf{z}_0(u) \mathbf{z}_0'(u) \otimes \mathbf{F}_{D_0}(u)^{-1} \mathbf{F}_{R_0}(u)^{-1} \mathbf{F}_{D_0}(u)^{-1} \right) \right] \quad (\text{E.62})$$

$$\left. + \left( \mathbf{F}_{D_0}(u)^{-1} \mathbf{F}_{R_0}(u)^{-1} \mathbf{F}_{D_0}(u)^{-1} \otimes \mathbf{z}_0(u) \mathbf{z}_0'(u) \right) \right] \quad (\text{E.63})$$

$$\left. + \left( \mathbf{F}_{D_0}(u)^{-1} \otimes \mathbf{F}_{D_0}(u)^{-1} \mathbf{F}_{R_0}(u)^{-1} \mathbf{z}_0(u) \mathbf{z}_0'(u) \right) \right] \quad (\text{E.64})$$

$$\left. + \left( \mathbf{F}_{D_0}(u)^{-1} \mathbf{F}_{R_0}(u)^{-1} \mathbf{z}_0(u) \mathbf{z}_0'(u) \otimes \mathbf{F}_{D_0}(u)^{-1} \right) \right] \quad (\text{E.65})$$

$$\left. + \left( \mathbf{z}_0(u) \mathbf{z}_0'(u) \mathbf{F}_{R_0}(u)^{-1} \mathbf{F}_{D_0}(u)^{-1} \otimes \mathbf{F}_{D_0}(u)^{-1} \right) \right] \quad (\text{E.66})$$

$$\left. + \left( \mathbf{F}_{D_0}(u)^{-1} \otimes \mathbf{z}_0(u) \mathbf{z}_0'(u) \mathbf{F}_{R_0}(u)^{-1} \mathbf{F}_{D_0}(u)^{-1} \right) \right] \left) \frac{\partial \text{vec}(\mathbf{F}_{D_0}(u))}{\partial u_1'} \quad (\text{E.67})$$

$$+ \mathbb{E} \left( \left[ \frac{1}{2} \left( \text{vec} \left( \mathbf{F}_{D_0}(u)^{-1} \mathbf{F}_{R_0}(u)^{-1} \mathbf{z}_0(u) \mathbf{z}_0'(u) \right) \otimes \mathbb{I}_{3p} \right) \right] \right) \quad (\text{E.68})$$

$$+ \frac{1}{2} \left( \text{vec} \left( \mathbf{z}_0(u) \mathbf{z}_0'(u) \mathbf{F}_{R_0}(u)^{-1} \mathbf{F}_{D_0}(u)^{-1} \right) \otimes \mathbb{I}_{3p} \right) \quad (\text{E.69})$$

$$- \left( \text{vec} \left( \mathbf{F}_{D_0}(u)^{-1} \right) \otimes \mathbb{I}_{3p} \right) \left] \frac{\partial^2 \text{vec}(\mathbf{F}_{D_0}(u))'}{\partial u_1 \partial u_1'} \right). \quad (\text{E.70})$$

Die Erwartungen der Kroneckerprodukte in (E.61)-(E.67) besitzen alle eine ähnliche Struktur. Durch eine geeignete Wahl von Matrizen

$$\mathbf{V}^{(l)} := \left[ \mathbf{v}_{ij}^{(l)} \right]_{i,j=1,\dots,p}, \quad l \in \{1, 2\},$$

können diese Terme jeweils angegeben werden als

$$\frac{\partial \text{vec}(\mathbf{F}_{D_0}(u))'}{\partial u_1} \left( \mathbf{V}^{(1)} \otimes \mathbf{V}^{(2)} \right) \frac{\partial \text{vec}(\mathbf{F}_{D_0}(u))}{\partial u_1'}. \quad (\text{E.71})$$

Dabei ist (E.71) jeweils eine Blockdiagonalmatrix mit  $(3 \times 3)$ -Blöcken der Form

$$\frac{\mathbf{v}_{kk}^{(1)} \mathbf{v}_{kk}^{(2)}}{\mathbf{w}_{k0}(u)} \frac{\partial \mathbf{w}_{k0}(u)}{\partial r_k} \frac{\partial \mathbf{w}_{k0}(u)}{\partial r_k'}, \quad k \in \{1, \dots, p\}, \quad (\text{E.72})$$

auf der Hauptdiagonalen.

Für den Term (E.61) wird (E.72) zu

$$\frac{\mathbf{v}_{ii}^{(1)} \mathbf{v}_{ii}^{(2)}}{\mathbf{w}_{i0}(u)} = \mathbf{w}_{i0}(u)^{-2} \leq C_1^{-2} \quad \forall i \in \{1, \dots, p\}.$$

Somit ist die Erwartung dieses Blocks endlich.

Für (E.62) und (E.63) wird (E.72) zu

$$\frac{\mathbf{v}_{ii}^{(1)} \mathbf{v}_{ii}^{(2)}}{\mathbf{w}_{i0}(u)} \leq C_1^{-3} y_{i0}^2 \left| r_{ii,0}^-(u) \right| \leq \delta_* C_1^{-3} y_{i0}^2 \quad (\text{E.73})$$

und für (E.64)-(E.67) zu

$$\frac{\mathbf{v}_{ii}^{(1)} \mathbf{v}_{ii}^{(2)}}{\mathbf{w}_{i0}(u)} = \frac{y_{i0}}{\mathbf{w}_{i0}(u)^2} \sum_{k=1}^p \frac{y_{k0}}{\mathbf{w}_{k0}(u)} \left| r_{ik,0}^-(u) \right| \leq \frac{\delta_*}{C_1^3} \sum_{k=1}^p y_{i0} y_{k0}. \quad (\text{E.74})$$

Unter Verwendung von Lemma E.24 sind dann Produkte von einem der Terme (E.46)-(E.49) mit (E.73) oder (E.74) von endlicher Erwartung. Außerdem sind die Summanden (E.68)-(E.70) von endlicher Erwartung, falls die folgenden Terme für alle  $u \in U$  sowie  $i_1, \dots, i_4 \in \{1, \dots, p\}$  und  $x, y \in \{x_{i_1}, s_{i_1}, t_{i_1}, x_{i_4}, s_{i_4}, t_{i_4}\}$  endlich sind:

$$\mathbb{E} \left( w_{i_1 0}(u)^{-1/2} \left| r_{i_2 i_3, 0}^-(u) \right| Z(u) \frac{\partial^2 w_{i_4 0}(u)^{1/2}}{\partial x \partial y} \right) \leq \frac{\delta_*}{C_1^{1/2}} \mathbb{E} \left( Z(u) \frac{\partial^2 w_{i_4 0}(u)^{1/2}}{\partial x \partial y} \right) \quad (\text{E.75})$$

$$\mathbb{E} \left( w_{i_1 0}(u)^{-1/2} \frac{\partial^2 w_{i_1 0}(u)^{1/2}}{\partial x \partial y} \right) \leq \frac{1}{C_1^{1/2}} \mathbb{E} \left( \frac{\partial^2 w_{i_1 0}(u)^{1/2}}{\partial x \partial y} \right). \quad (\text{E.76})$$

Die Endlichkeit von (E.75) und (E.76) ist dann eine direkte Konsequenz von (E.45), (E.51)-(E.54) und Lemma E.24. Gemeinsam belegen diese Resultate, dass gilt:

$$\mathbb{E} \left( \frac{\partial^2 l_0(u)}{\partial u_1 \partial u_1'} \right) < \infty. \quad \blacksquare$$

**Beweis, dass die Erwartung von  $\frac{\partial^2 l_0(u)}{\partial u_1 \partial u_2'}$  und  $\frac{\partial^2 l_0(u)}{\partial u_2 \partial u_1'}$  endlich ist**

Für die Erwartung der Kreuzableitungen nach den Varianz- und Korrelationsparametern gilt:

$$\begin{aligned} \mathbb{E} \left( \frac{\partial^2 l_0(u)}{\partial u_1 \partial u_2'} \right) &= -\frac{1}{4} \mathbb{E} \left( \frac{\partial \text{vec}(\mathbf{F}_{D_0}(u))'}{\partial u_1} \left[ \left( \mathbf{z}_0(u) \mathbf{z}_0'(u) \mathbf{F}_{R_0}(u)^{-1} \otimes \mathbf{F}_{D_0}(u)^{-1} \mathbf{F}_{R_0}(u)^{-1} \right) \right. \right. \\ &\quad \left. \left. + \left( \mathbf{F}_{D_0}(u)^{-1} \mathbf{F}_{R_0}(u)^{-1} \otimes \mathbf{z}_0(u) \mathbf{z}_0'(u) \mathbf{F}_{R_0}(u)^{-1} \right) \right] 2 \left( \mathbf{D}_{p,-}^+ \right)' \frac{\partial \text{vecl}(\mathbf{F}_{R_0}(u))}{\partial u_2'} \right). \quad (\text{E.77}) \end{aligned}$$

Diese Erwartung ist endlich, falls dies auch für den folgenden Term und für alle  $u \in U$ ,  $i_1, \dots, i_5 \in \{1, \dots, p\}$ ,  $x \in \{x_{i_1}, s_{i_1}, t_{i_1}\}$  und  $y \in \{a, b, q_1, \dots, q_{p-}\}$  gilt:

$$\mathbb{E} \left( \frac{\partial w_{i_1 0}^{\frac{1}{2}}}{\partial x} Z(u) \left| r_{i_2 i_3}^-(u) \right|^2 w_{i_4 0}^{-\frac{1}{2}} \frac{\partial \left[ \mathbf{F}_{Q_0^*}(u) \right]_{i_5 i_5}}{\partial y} \right) \leq \frac{\delta_*^2}{C_1^{\frac{1}{2}}} \mathbb{E} \left( \frac{\partial w_{i_1 0}^{\frac{1}{2}}}{\partial x} Z(u) \frac{\partial \left[ \mathbf{F}_{Q_0^*}(u) \right]_{i_5 i_5}}{\partial y} \right) \quad (\text{E.78})$$

Die Existenz der rechten Seite von (E.78) folgt dabei aus (E.50), (E.51), (E.55) und Lemma E.24. ■

**Beweis, dass die Erwartung von  $\frac{\partial^2 l_0(u)}{\partial u_2 \partial u'_2}$  endlich ist**

Zunächst muss gezeigt werden, dass die folgenden Erwartungen existieren:

$$\mathbb{E} \left[ \frac{\partial \text{vecl}(\mathbf{F}_{R_0}(u))'}{\partial u_2} \mathbf{D}_{p,-}^+ \left( \mathbf{F}_{R_0}(u)^{-1} \otimes \mathbf{F}_{R_0}(u)^{-1} \right) \mathbf{D}_{p,-} \frac{\partial \text{vecl}(\mathbf{F}_{R_0}(u))}{\partial u'_2} \right] \quad (\text{E.79})$$

$$\mathbb{E} \left[ \frac{\partial \text{vecl}(\mathbf{F}_{R_0}(u))'}{\partial u_2} \mathbf{D}_{p,-}^+ \left( \mathbf{F}_{R_0}(u)^{-1} \mathbf{z}_t(u) \mathbf{z}'_t(u) \mathbf{F}_{R_0}(u)^{-1} \otimes \mathbf{F}_{R_0}(u)^{-1} \right) \mathbf{D}_{p,-} \frac{\partial \text{vecl}(\mathbf{F}_{R_0}(u))}{\partial u'_2} \right] \quad (\text{E.80})$$

$$\mathbb{E} \left[ \frac{\partial \text{vecl}(\mathbf{F}_{R_0}(u))'}{\partial u_2} \mathbf{D}_{p,-}^+ \left( \mathbf{F}_{R_0}(u)^{-1} \otimes \mathbf{F}_{R_0}(u)^{-1} \mathbf{z}_t(u) \mathbf{z}'_t(u) \mathbf{F}_{R_0}(u)^{-1} \right) \mathbf{D}_{p,-} \frac{\partial \text{vecl}(\mathbf{F}_{R_0}(u))}{\partial u'_2} \right] \quad (\text{E.81})$$

$$\mathbb{E} \left[ \left( \text{vecl} \left( \mathbf{F}_{R_0}(u)^{-1} \left[ \mathbb{I}_p - \mathbf{z}_t(u) \mathbf{z}'_t(u) \mathbf{F}_{R_0}(u)^{-1} \right] \right)' \mathbf{D}_{p,-}^+ \otimes \frac{\partial \text{vech}(\mathbf{F}_{Q_0}(u))'}{\partial u_2} \mathbf{D}'_p \right) \times \frac{\partial \text{vecl}(\mathbf{F}_{Q_0^*}(u) \otimes \mathbf{F}_{Q_0^*}(u))}{\partial u'_2} \right] \quad (\text{E.82})$$

$$\mathbb{E} \left[ \left( \text{vecl} \left( \mathbf{F}_{R_0}(u)^{-1} \left[ \mathbb{I}_p - \mathbf{z}_t(u) \mathbf{z}'_t(u) \mathbf{F}_{R_0}(u)^{-1} \right] \right)' \mathbf{D}_{p,-}^+ \otimes \frac{\partial \text{vech}(\mathbf{F}_{Q_0^*}(u))'}{\partial u_2} \mathbf{D}'_p \right) \times \left( \mathbb{I}_p \otimes \mathbf{K}_{pp} \otimes \mathbb{I}_p \right) \left( \frac{\partial \text{vec}(\mathbf{F}_{Q_0^*}(u) \mathbf{F}_{Q_0}(u))}{\partial u'_2} \otimes \text{vec}(\mathbb{I}_p) + \text{vec}(\mathbb{I}_p) \otimes \frac{\partial \text{vec}(\mathbf{F}_{Q_0^*}(u) \mathbf{F}_{Q_0}(u))}{\partial u'_2} \right) \right] \quad (\text{E.83})$$

$$\mathbb{E} \left[ \left( \text{vecl} \left( \mathbf{F}_{R_0}(u)^{-1} \left[ \mathbb{I}_p - \mathbf{z}_t(u) \mathbf{z}'_t(u) \mathbf{F}_{R_0}(u)^{-1} \right] \right)' \mathbf{D}_{p,-}^+ \left( \mathbf{F}_{Q_0}(u)^* \otimes \mathbf{F}_{Q_0}(u)^* \right) \mathbf{D}_p \otimes \mathbb{I}_{p-+2} \right) \times \frac{\partial}{\partial u'_2} \text{vec} \left( \frac{\partial \text{vech}(\mathbf{F}_{Q_0}(u))'}{\partial u_2} \right) \right] \quad (\text{E.84})$$

$$\mathbb{E} \left[ \left( \text{vecl} \left( \mathbf{F}_{R_0}(u)^{-1} \left[ \mathbb{I}_p - \mathbf{z}_t(u) \mathbf{z}'_t(u) \mathbf{F}_{R_0}(u)^{-1} \right] \right)' \mathbf{D}_{p,-}^+ \times \left[ \mathbf{F}_{Q_0^*}(u) \mathbf{F}_{Q_0}(u) \otimes \mathbb{I}_p + \mathbb{I}_p \otimes \mathbf{F}_{Q_0^*}(u) \mathbf{F}_{Q_0}(u) \right] \mathbf{D}_p \otimes \mathbb{I}_{p-+2} \right) \frac{\partial}{\partial u'_2} \text{vec} \left( \frac{\partial \text{vech}(\mathbf{F}_{Q_0^*}(u))'}{\partial u_2} \right) \right] \quad (\text{E.85})$$

Zum Nachweis der Endlichkeit von (E.79)-(E.85) können die folgenden Aussagen genutzt werden:

- $\mathbf{d}(i, j) := \left( [\mathbf{F}_{Q_0}(u)]_{ii} [\mathbf{F}_{Q_0}(u)]_{jj} \right)^{-1/2} \stackrel{f.s.}{<} \frac{1-b}{1-a-b}$
- $[\mathbf{F}_{Q_0}(u)]_{ij} \mathbf{d}(i, j) \leq 1 \Rightarrow \frac{[\mathbf{F}_{Q_0}(u)]_{ij}}{[\mathbf{F}_{Q_0}(u)]_{ii}^{1/2}} \leq [\mathbf{F}_{Q_0}(u)]_{jj}^{1/2} < [1 + a\mathbf{G}_0^z(j, j, u)]^{1/2}$ .
- $\frac{\mathbf{G}_0^z(k, k, u)}{[\mathbf{F}_{Q_0}(u)]_{kk}} \stackrel{f.s.}{\leq} \frac{\mathbf{G}_0^z(k, k, u)}{a\mathbf{G}_0^z(k, k, u)} = \frac{1}{a}$ .

Die Existenz der Erwartungen (E.80) und (E.81) ist eine direkte Konsequenz der Endlichkeit der folgenden Ausdrücke für alle  $i_1, \dots, i_5 \in \{1, \dots, p\}$ ,  $x \in \{a, b, q_1, \dots, q_{p-}\}$  und  $u \in U$ :

$$\begin{aligned} \mathbb{E} \left( \left[ \frac{\partial [\mathbb{F}_{Q_t^*}(u)]_{i_1 i_1}}{\partial x} \right]^2 \left[ \frac{\partial [\mathbb{F}_{Q_t}(u)]_{i_2 i_3}}{\partial [\mathbb{F}_{Q_t}(u)]_{i_2 i_2}^{1/2}} \right]^2 |r_{i_4 i_5, 0}^-(u)|^3 Z(u) \right) \\ \leq \delta_*^3 \mathbb{E} \left( \left[ \frac{\partial [\mathbb{F}_{Q_t^*}(u)]_{i_1 i_1}}{\partial x} \right]^2 [1 + aG_0^z(i_3, i_3, u)] Z(u) \right) \end{aligned} \quad (\text{E.86})$$

$$\begin{aligned} \mathbb{E} \left( \left[ \frac{\partial [\mathbb{F}_{Q_t^*}(u)]_{i_1 i_1}}{\partial x} \right]^2 [\mathbb{F}_{Q_0^*}(u)]^4 |r_{i_4 i_5, 0}^-(u)|^3 Z(u) \right) \\ \leq \delta_*^3 \mathbb{E} \left( \left[ \frac{\partial [\mathbb{F}_{Q_t^*}(u)]_{i_1 i_1}}{\partial x} \right]^2 \left( \frac{1-b}{1-a-b} \right)^2 Z(u) \right). \end{aligned} \quad (\text{E.87})$$

Die Endlichkeit der rechten Seite von (E.86) und (E.87) folgt aus (E.45), (E.55) sowie Lemma E.24. Analog ergibt sich die Existenz von (E.79), indem  $|r_{i_4 i_5, 0}^-(u)|^3 Z(u)$  durch  $|r_{i_4 i_5, 0}^-(u)|^2$  ersetzt wird.

Des Weiteren ist die Existenz von (E.82) und (E.83) eine direkte Konsequenz der Endlichkeit der folgenden Terme für alle  $u \in U$ ,  $i_1, \dots, i_5 \in \{1, \dots, p\}$  und  $x, y \in \{a, b, q_1, \dots, q_{p-}\}$ :

$$\begin{aligned} \mathbb{E} \left( |r_{i_1 i_2, 0}^-(u)|^2 Z(u) \frac{\partial [\mathbb{F}_{Q_0}(u)]_{i_3 i_4}}{\partial x} \frac{\partial [\mathbb{F}_{Q_0^*}(u)]_{i_5 i_5}}{\partial y} [\mathbb{F}_{Q_t}(u)]_{i_6 i_6}^{-1/2} \right) \\ \leq \delta_*^2 \left( \frac{1-b}{1-a-b} \right)^{1/2} \mathbb{E} \left( Z(u) \frac{\partial [\mathbb{F}_{Q_0}(u)]_{i_3 i_4}}{\partial x} \frac{\partial [\mathbb{F}_{Q_0^*}(u)]_{i_5 i_5}}{\partial y} \right) \end{aligned} \quad (\text{E.88})$$

$$\begin{aligned} \mathbb{E} \left( |r_{i_1 i_2, 0}^-(u)|^2 Z(u) \left[ \frac{\partial [\mathbb{F}_{Q_0^*}(u)]_{i_3 i_3}}{\partial x} \right]^2 [\mathbb{F}_{Q_0}(u)]_{i_3 i_3} \right) \leq \delta_*^2 \mathbb{E} \left( Z(u) \left[ \frac{\partial [\mathbb{F}_{Q_0}(u)]_{i_3 i_3}}{\partial x} \right]^2 \right) \end{aligned} \quad (\text{E.89})$$

$$\begin{aligned} \mathbb{E} \left( |r_{i_1 i_2, 0}^-(u)|^2 Z(u) \frac{\partial [\mathbb{F}_{Q_0^*}(u)]_{i_3 i_3}}{\partial x} \frac{\partial [\mathbb{F}_{Q_0^*}(u)]_{i_3 i_3}}{\partial y} [\mathbb{F}_{Q_0^*}(u)]_{i_3 i_3} \right) \\ \leq \delta_*^2 \left( \frac{1-b}{1-a-b} \right)^{1/2} \mathbb{E} \left( Z(u) \left[ \frac{\partial [\mathbb{F}_{Q_0^*}(u)]_{i_3 i_3}}{\partial x} \right]^2 \right). \end{aligned} \quad (\text{E.90})$$

Die Endlichkeit der jeweiligen Terme auf der rechten Seite von (E.88)-(E.90) ergibt sich aus (E.45) und Lemma E.24.

Abschließend sind die Terme (E.84) und (E.85) unter Berücksichtigung von

$$\frac{\partial}{\partial u'_2} \text{vec} \left( \frac{\partial \text{vech} (F_{Q_0^*}(u))}{\partial u'_2} \right) = \begin{bmatrix} \frac{\partial^2 \text{vech} (F_{Q_0^*}(u))}{(\partial a)^2} & \frac{\partial^2 \text{vech} (F_{Q_0^*}(u))}{\partial a \partial b} & \mathbf{0}_{p^+ \times p^-} \\ \frac{\partial^2 \text{vech} (F_{Q_0^*}(u))}{\partial a \partial b} & \frac{\partial^2 \text{vech} (F_{Q_0^*}(u))}{(\partial b)^2} & \mathbf{0}_{p^+ \times p^-} \\ \mathbf{0}_{p^+ p^-} & \mathbf{0}_{p^+ p^-} & \mathbf{0}_{p^+ p^- \times p^-} \end{bmatrix}$$

endlich, falls dies auch für die folgenden Erwartungen und für alle  $u \in U$  sowie für  $i_1, \dots, i_6 \in \{1, \dots, p\}$  und  $x, y \in \{a, b, q_1, \dots, q_{p^-}\}$  gilt:

$$\mathbb{E} \left( \left| r_{i_1 i_2, 0}^- (u) \right|^2 Z(u) [F_{Q_0}(u)]_{i_3 i_3}^{-1} \frac{\partial^2 [F_{Q_0}(u)]_{i_4 i_5}}{\partial x \partial y} \right) \leq \delta_*^2 \frac{1-b}{1-a-b} \mathbb{E} \left( Z(u) \frac{\partial^2 [F_{Q_0}(u)]_{i_4 i_5}}{\partial x \partial y} \right) \quad (\text{E.91})$$

$$\begin{aligned} \mathbb{E} \left( \left| r_{i_1 i_2, 0}^- (u) \right|^2 Z(u) \frac{[F_{Q_0}(u)]_{i_3 i_4} \partial^2 [F_{Q_0}(u)]_{i_5 i_6}}{[F_{Q_0}(u)]_{i_3 i_3}^{1/2} \partial x \partial y} \right) \\ \leq \delta_*^2 \mathbb{E} \left( Z(u) [1 + a G_0^z(i_4, i_4, u)]^{1/2} \frac{\partial^2 [F_{Q_0}(u)]_{i_5 i_6}}{\partial x \partial y} \right). \end{aligned} \quad (\text{E.92})$$

Mit (E.45), (E.55) sowie Lemma E.24 sind die rechtseitigen Terme in (E.91) und (E.92) endlich. Dies vervollständigt den Beweis und liefert:

$$\mathbb{E} \left( \frac{\partial^2 l_0(u)}{\partial u_2 \partial u'_2} \right) < \infty. \quad \blacksquare$$

### E.3 Beweis von Theorem 5.1

Der Beweis folgt sehr eng den Argumenten des Beweises von Theorem 3.1 in Berkes et al. (2004). Das Vorgehen wird im Folgenden skizziert.

Zunächst gilt für  $m \rightarrow \infty$  die dem Lemma 6.3 in Berkes et al. (2004) entsprechende Aussage

$$\sup_{1 \leq k < \infty} \frac{\sum_{i=m+1}^{m+k} l'_i(\hat{\theta}_m) - \left[ \sum_{i=m+1}^{m+k} l'_i(\theta) + (\hat{\theta}_m - \theta) k \mathbb{E}(l''_0(\theta)) \right]}{\sqrt{m} \left( 1 + \frac{k}{m} \right) \mathbf{b} \left( \frac{k}{m} \right)} = o_P(1)$$

unter Berücksichtigung der Argumente im Beweis von Proposition 5.2 und der gleichmäßigen Konvergenz von  $l''_t(\cdot)$  gegen  $\mathbb{E}[l''_0(\cdot)]$ . Diese folgt mit Hilfe von Theorem A.2.2 in White (1994), das als Theorem F.2 in Anhang F zu finden ist. Da  $U$  eine kompakte Menge ist und  $l''_t(u)$  sowohl stetig in  $u$  für alle  $y_t$  und messbar in  $y_t$  für alle  $u \in U$ , bleibt nur die Dominanzbedingung nachzuweisen. Wird die dominierende Funktion als  $\sup_{u \in U} |l''_t(u)|$  gewählt, folgt mit den Resultaten aus Abschnitt E.2.1:

$$\mathbb{E} \left[ \sup_{u \in U} |l''_t(u)| \right] < \infty.$$

Mit Theorem A.2.2 in White (1994) gilt somit:

$$\sup_{u \in U} \left| \frac{1}{m} \sum_{i=1}^m l_i''(u) - \mathbb{E} [l_0''(u)] \right| = \frac{1}{m} \sup_{u \in U} \left| \sum_{i=1}^m (l_i''(u) - \mathbb{E} [l_0''(u)]) \right| \xrightarrow{f.s.} 0.$$

Mit den Annahmen 4.2-4.4 gilt außerdem analog zu Lemma 6.4 in Berkes et al. (2004) für  $m \rightarrow \infty$ :

$$(\hat{\theta}_m - \theta) = \frac{1}{m} \sum_{i=1}^m l_i'(\theta) \left[ \mathbb{E} (l_0''(\theta)) \right]^{-1} [1 + o(1)].$$

Des Weiteren gilt mit den Annahmen 4.2-4.4 und 5.4 sowie Lemma 6.5 in Berkes et al. (2004):

$$\sup_{1 \leq k < \infty} \frac{\left| \sum_{i=m+1}^{m+k} l_i'(\hat{\theta}_m) - \left[ \sum_{i=m+1}^{m+k} l_i'(\theta) - \frac{k}{m} \sum_{i=1}^m l_i'(\theta) \right] \right|}{\sqrt{m} \left(1 + \frac{k}{m}\right) \mathbf{b} \left(\frac{k}{m}\right)} = o_P(1).$$

Mit Hilfe von Lemma 6.6 in Berkes et al. (2004) gilt außerdem:

$$\sup_{1 \leq k < \infty} \frac{\left| \sum_{i=m+1}^{m+k} l_i'(\theta) - \frac{k}{m} \sum_{i=1}^m l_i'(\theta) \right|}{\sqrt{m} \left(1 + \frac{k}{m}\right) \mathbf{b} \left(\frac{k}{m}\right)} \xrightarrow{d} \sup_{0 < t < \infty} \frac{|\mathbf{W}_D(1+t) - (1+t)\mathbf{W}_D(1)|}{(1+t)\mathbf{b}(t)}. \quad (\text{E.93})$$

Dabei ist der Grenzprozess  $\{\mathbf{W}_D(t), t \in [0, B]\}$  ein  $d$ -variater Gaußprozess mit Mittelwertfunktion

$$\mathbb{E} [\mathbf{W}_D(s)] = \mathbf{0}_d, \quad \forall s \in [0, B],$$

und Kovarianzkern

$$\mathbb{E} [\mathbf{W}_D^T(k) \mathbf{W}_D(l)] = \min \{k, l\} \mathbf{D}, \quad \forall k, l \in [0, B].$$

Die Lemmas E.17 und E.23 implizieren außerdem:

$$\sup_{u \in U} \left| \sum_{i=1}^n (\hat{l}_i'(u) - l_i'(u)) \right| \stackrel{f.s.}{=} \mathcal{O}(1). \quad (\text{E.94})$$

Aus (E.93) und (E.94) folgt:

$$\sup_{1 \leq k < mB} \frac{\left| \sum_{i=m+1}^{m+k} \hat{D}_m^{-\frac{1}{2}} \hat{l}_i'(\hat{\theta}_m) \right|}{m^{\frac{1}{2}} \left(1 + \frac{k}{m}\right) \mathbf{b} \left(\frac{k}{m}\right)} \xrightarrow{d} \sup_{0 < t \leq B} \frac{\left| \mathbf{D}^{-\frac{1}{2}} [\mathbf{W}_D(1+t) - (1+t)\mathbf{W}_D(1)] \right|}{(1+t)\mathbf{b}(t)}.$$

Ein einfaches Nachrechnen der Eigenschaften des Prozesses zeigt, dass

$$\left\{ \mathbf{D}^{-\frac{1}{2}} [\mathbf{W}_D(1+t) - (1+t)\mathbf{W}_D(1)], t \in [0, \infty) \right\}$$

und  $\{\mathcal{G}(t), t \in [0, \infty)\}$  aus Definition 3.4 die gleiche Verteilung besitzen. Dies vervollständigt den Beweis des Theorems. ■

## E.4 Beweis von Theorem 5.2

Wie in Berkes et al. (2004) bietet sich zur Untersuchung des Detektors unter Alternativen die folgende Zerlegung an:

$$\frac{\sum_{i=m+1}^{m+k} l'_i(\hat{\theta}_m)}{\sqrt{m} \left(1 + \frac{k}{m}\right) \mathbf{b} \left(\frac{k}{m}\right)} = \frac{\sum_{i=m+1}^{m+k^*-1} l'_i(\hat{\theta}_m)}{\sqrt{m} \left(1 + \frac{k}{m}\right) \mathbf{b} \left(\frac{k}{m}\right)} + \frac{\sum_{i=m+k^*}^{m+k} l'_i(\hat{\theta}_m)}{\sqrt{m} \left(1 + \frac{k}{m}\right) \mathbf{b} \left(\frac{k}{m}\right)}. \quad (\text{E.95})$$

Der erste Summand auf der rechten Seite von (E.95), der die Gradientenbeiträge der Beobachtungen vor dem Strukturbruch enthält, kann analog zum Beweis von Theorem 5.1 oder zum Beweis von Theorem 3.1. in Berkes et al. (2004) behandelt werden. Für  $m \rightarrow \infty$ , was auch  $k^* \rightarrow \infty$  impliziert, gilt dann:

$$\sup_{k^* \leq k < \infty} \frac{\left| \sum_{i=m+1}^{m+k^*-1} l'_i(\hat{\theta}_m) \right|}{\sqrt{m} \left(1 + \frac{k}{m}\right) \mathbf{b} \left(\frac{k}{m}\right)} \xrightarrow{d} \sup_{t \in [\lambda^* B, \infty)} \frac{|\mathbf{W}_D (1 + \lambda^* B) - (1 + \lambda^* B) \mathbf{W}_D(1)|}{(1+t) \mathbf{b}(t)}. \quad (\text{E.96})$$

Der zweite Summand auf der rechten Seite von (E.95) enthält die Gradientenbeiträge der Beobachtungen nach dem Strukturbruch. Diese werden durch den Parametervektor  $\boldsymbol{\theta}^*$  gesteuert. Somit liefert eine Taylorreihenentwicklung von  $l'_i(\hat{\theta}_m)$  im Punkt  $\boldsymbol{\theta}^*$  für  $m \rightarrow \infty$ :

$$\sup_{k^* \leq k < \infty} \frac{\left| \sum_{i=m+k^*}^{m+k} l'_i(\hat{\theta}_m) - \left[ \sum_{i=m+k^*}^{m+k} l'_i(\boldsymbol{\theta}^*) + (\hat{\theta}_m - \boldsymbol{\theta}^*)' \sum_{i=m+k^*}^{m+k} l''_i(\boldsymbol{\theta}^*) \right] \right|}{\sqrt{m} \left(1 + \frac{k}{m}\right) \mathbf{b} \left(\frac{k}{m}\right)} = o_P(1).$$

Analog zum Beweis von Theorem 4.4 in Berkes et al. (2003) gilt mit Hilfe von (4.11):

$$(\hat{\theta}_m - \boldsymbol{\theta}^*) = (\hat{\theta}_m - \boldsymbol{\theta}) + (\boldsymbol{\theta} - \boldsymbol{\theta}^*) = -\frac{1}{m} \sum_{i=1}^m l'_i(\boldsymbol{\theta}) \left[ \mathbf{E}(l''_0(\boldsymbol{\theta})) \right]^{-1} \left(1 + o_P(1)\right) + (\boldsymbol{\theta} - \boldsymbol{\theta}^*).$$

Mit Annahme 5.7 existiert außerdem eine Umgebung  $U_2$  von  $\boldsymbol{\theta}^*$ , in der die Funktion  $\frac{1}{m} \sum_{i=1}^m l''_i(u)$  mit Theorem A.2.2 in White (1994), das als Theorem F.2 in Anhang F zu finden ist, gleichmäßig gegen ihr theoretisches Gegenstück konvergiert. Dies gilt analog zum Vorgehen im Beweis von Theorem 5.1. Die gleichmäßige Konvergenz impliziert auch die Konvergenz in Wahrscheinlichkeit von

$$\sup_{k^* \leq k < \infty} \frac{\left| (\hat{\theta}_m - \boldsymbol{\theta}^*)' \sum_{i=m+k^*}^{m+k} l''_i(\boldsymbol{\theta}^*) \right|}{\sqrt{m} \left(1 + \frac{k}{m}\right) \mathbf{b} \left(\frac{k}{m}\right)} \frac{\left| \left[ -\frac{k-k^*+1}{m} \left[ \mathbf{E}(l''_0(\boldsymbol{\theta})) \right]^{-1} \mathbf{E}(l''_0(\boldsymbol{\theta}^*)) \sum_{i=1}^m l'_i(\boldsymbol{\theta}) + (k-k^*+1) (\boldsymbol{\theta} - \boldsymbol{\theta}^*)' \mathbf{E}(l''_0(\boldsymbol{\theta}^*)) \right] \right|}{\sqrt{m} \left(1 + \frac{k}{m}\right) \mathbf{b} \left(\frac{k}{m}\right)}$$

gegen 0, siehe Lemma 2.2 in Hayashi (2000).

Außerdem gilt mit der Dreiecksungleichung:

$$\begin{aligned} & \sup_{k^* \leq k < \infty} \left| \frac{\left[ \sum_{i=m+k^*}^{m+k} l'_i(\boldsymbol{\theta}^*) - \frac{k-k^*+1}{m} \left[ \mathbb{E}(l''_0(\boldsymbol{\theta})) \right]^{-1} \mathbb{E}(l''_0(\boldsymbol{\theta}^*)) \sum_{i=1}^m l'_i(\boldsymbol{\theta}) \right]}{\sqrt{m} \left(1 + \frac{k}{m}\right) \mathbf{b}\left(\frac{k}{m}\right)} \right. \\ & \quad \left. - \frac{(k-k^*+1) (\boldsymbol{\theta} - \boldsymbol{\theta}^*)' \mathbb{E}(l''_0(\boldsymbol{\theta}^*))}{\sqrt{m} \left(1 + \frac{k}{m}\right) \mathbf{b}\left(\frac{k}{m}\right)} \right| \\ & \geq \left| \sup_{k^* \leq k < \infty} \frac{\left| \sum_{i=m+k^*}^{m+k} l'_i(\boldsymbol{\theta}^*) - \frac{k-k^*+1}{m} \left[ \mathbb{E}(l''_0(\boldsymbol{\theta})) \right]^{-1} \mathbb{E}(l''_0(\boldsymbol{\theta}^*)) \sum_{i=1}^m l'_i(\boldsymbol{\theta}) \right|}{\sqrt{m} \left(1 + \frac{k}{m}\right) \mathbf{b}\left(\frac{k}{m}\right)} \right. \end{aligned} \quad (\text{E.97})$$

$$\left. - \sup_{k^* \leq k < \infty} \frac{\left| (k-k^*+1) (\boldsymbol{\theta} - \boldsymbol{\theta}^*)' \mathbb{E}(l''_0(\boldsymbol{\theta}^*)) \right|}{\sqrt{m} \left(1 + \frac{k}{m}\right) \mathbf{b}\left(\frac{k}{m}\right)} \right|. \quad (\text{E.98})$$

Dabei konvergiert der Minuend in (E.97) für  $m \rightarrow \infty$  und wegen  $k^* = \lambda^* m B$  auch für  $k^* \rightarrow \infty$  in Verteilung gegen

$$\sup_{t \in (\lambda^*, \infty)} \frac{\left| \mathbf{W}_{D^*}(1+t) - \mathbf{W}_{D^*}(1+\lambda^*) - (t-\lambda^*) \left[ \mathbb{E}(l''_0(\boldsymbol{\theta})) \right]^{-1} \mathbb{E}(l''_0(\boldsymbol{\theta}^*)) \mathbf{W}_D(1) \right|}{(1+t) \mathbf{b}(t)}, \quad (\text{E.99})$$

wobei  $D^* := \text{Cov}[l'_0(\boldsymbol{\theta}^*)]$  und  $\{\mathbf{W}_{D^*}(t), t \in [0, \infty)\}$  ein  $d$ -variater Gaußprozess mit Mittelwertfunktion

$$\mathbb{E}[\mathbf{W}_{D^*}(s)] = \mathbf{0}_d, \quad \forall s \in [0, B),$$

und Kovarianzkern

$$\mathbb{E}[\mathbf{W}_{D^*}^T(k) \mathbf{W}_{D^*}(l)] = \min\{k, l\} D^*, \quad \forall k, l \in [0, B),$$

ist. Des Weiteren gilt für den Subtrahend in (E.98):

$$\sqrt{m} \sup_{k^* \leq k < \infty} \frac{\left| \frac{k-k^*+1}{m} (\boldsymbol{\theta} - \boldsymbol{\theta}^*)' \mathbb{E}(l''_0(\boldsymbol{\theta}^*)) \right|}{\left(1 + \frac{k}{m}\right) \mathbf{b}\left(\frac{k}{m}\right)} \xrightarrow{f.s.} \infty$$

für  $m \rightarrow \infty$ . Da die Grenzwerte in (E.96) und (E.99) stochastisch beschränkt sind und die variablen Teile der Funktion  $\mathbf{b}(\cdot)$  so gewählt werden, dass das Verfahren unter der Nullhypothese das Niveau einhält, divergieren die Detektorwerte für  $m \rightarrow \infty$ . Dies impliziert, dass jede Änderung im Vektor der Parameter nachgewiesen werden kann, wenn der historische Zeitraum lang genug ist. ■

## Anhang F

### Miscellanea

Dieses Kapitel umfasst verschiedene Aussagen, die in der Arbeit Anwendung finden.

**Theorem F.1** *Asymptotische Normalverteilung eines Extremumschätzers* (Amemiya (1985))

Sei  $\theta$  ein  $K$ -dimensionaler Parametervektor,  $\mathbf{y} = (y_1, \dots, y_T)'$  ein  $T$ -dimensionaler Zufallsvektor und  $Q_T(\mathbf{y}, \theta)$  die Zielfunktion. Außerdem sei  $\theta_0$  der den Zufallsvariablen  $y_i$ ,  $i = 1, \dots, T$ , zugrundeliegende wahre Parametervektor.

Seien die folgenden Bedingungen erfüllt:

- B1**  $\Theta$  ist eine offene Teilmenge des euklidischen Vektorraums der Dimension  $K$ .
- B2**  $Q_T(\mathbf{y}, \theta)$  ist eine messbare Funktion von  $\mathbf{y}$  für alle  $\theta \in \Theta$ . Außerdem existiert  $\frac{\partial Q_T(\theta)}{\partial \theta}$  in einer offenen Umgebung  $N_1(\theta_0)$  von  $\theta_0$ .
- B3** Es existiert eine offene Umgebung  $N_2(\theta_0)$  von  $\theta_0$  in der  $\frac{1}{T}Q_T(\theta)$  in Wahrscheinlichkeit und gleichmäßig in  $\theta$  gegen eine nichtstochastische Funktion  $Q(\theta)$  konvergiert. Außerdem besitzt  $Q(\theta)$  ein striktes lokales Maximum in  $\theta_0$ .
- B4**  $\frac{\partial^2 Q_T(\theta)}{\partial \theta \partial \theta'}$  existiert und ist in einer offenen und konvexen Umgebung von  $\theta_0$  stetig.
- B5**  $\frac{1}{T} \frac{\partial^2 Q_T(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta=\theta_T^*}$  konvergiert für jede Folge  $\{\theta_T^*, T \in \mathbb{N}\}$  mit  $\text{plim } \theta_T^* = \theta_0$  in Wahrscheinlichkeit gegen eine endliche nichtsinguläre Matrix

$$\mathbf{A}(\theta_0) := \lim_{T \rightarrow \infty} \mathbb{E} \left[ \frac{1}{T} \frac{\partial^2 Q_T(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta=\theta_0} \right].$$

- B6**  $\frac{1}{\sqrt{T}} \frac{\partial Q_T(\theta)}{\partial \theta} \Big|_{\theta=\theta_0} \rightarrow \mathcal{N}_K(\mathbf{0}_K, \mathbf{B}(\theta_0))$  mit

$$\mathbf{B}(\theta_0) := \lim_{T \rightarrow \infty} \mathbb{E} \left[ \frac{1}{T} \frac{\partial Q_T(\theta)}{\partial \theta} \Big|_{\theta=\theta} \frac{\partial Q_T(\theta)}{\partial \theta'} \Big|_{\theta=\theta} \right].$$

- B7** Sei  $\Theta_T$  die Menge der Lösungen der Gleichung

$$\frac{\partial Q_T(\theta)}{\partial \theta} = \mathbf{0}_K.$$

$\Theta_T$  entspricht somit der Menge der lokalen Maxima.

Sei  $\{\hat{\theta}_T, T \in \mathbb{N}\}$  eine Folge, die sich ergibt, indem jeweils ein Element aus  $\Theta_T$  gewählt wird, so dass  $\text{plim } \hat{\theta}_T = \theta_0$  gilt. Wenn die Bedingungen **B1-B7** erfüllt sind, gilt dann:

$$\sqrt{T}(\hat{\theta}_T - \theta) \xrightarrow{d} \mathcal{N}_K(\mathbf{0}_K, \mathbf{A}(\theta)^{-1} \mathbf{B}(\theta) \mathbf{A}(\theta)^{-1}).$$

Dieses Theorem wurde in Amemiya (1985) als Theorem 4.1.3 formuliert und bewiesen.

**Theorem F.2 Gleichmäßige Konvergenz der Zielfunktion eines Extremschätzers** (White (1994), S.351f)

Sei  $(\Omega, \mathcal{F}, P)$  ein Wahrscheinlichkeitsraum und sei  $T : \Omega \rightarrow \Omega$  eine injektive maßerhaltende Transformation. Sei außerdem  $\Theta$  eine kompakte Teilmenge von  $\mathbb{R}^p$ ,  $p \in \mathbb{N}$ , und sei  $q : \Omega \times \Theta \rightarrow \mathbb{R}$  eine Funktion, so dass  $q(\cdot, \theta)$  für alle  $\theta$  in  $\Theta$   $\mathcal{F}$ -messbar und  $q(\omega, \cdot)$  für alle  $\omega$  in  $\Omega$  stetig auf  $\Theta$  ist.

Es sei angenommen, dass eine  $\mathcal{F}$ -messbare Funktion  $D : \Omega \rightarrow \mathbb{R}^+$  existiert, so dass gilt:

$$|q(\omega, \theta)| \leq D(\omega)$$

für alle  $\theta$  in  $\Theta$  und  $\omega$  in  $\Omega$ , und  $E(D) \equiv \int_{\Omega} D dP < \infty$ .

Falls für alle  $\theta$  in  $\Theta$   $q_t(\theta) \equiv q(T^t \omega, \theta)$  ergodisch ist, gilt

(i)  $\bar{Q}(\theta) = E(q_t(\theta))$  ist stetig auf  $\Theta$ ; und

(ii)  $Q_n(\theta) - \bar{Q}(\theta) \xrightarrow{f.s.} 0$  gleichmäßig auf  $\Theta$  für  $n \rightarrow \infty$ , wobei  $Q_n(\theta) \equiv n^{-1} \sum_{t=1}^n q_t(\theta)$ .

Dieses Theorem wurde unter anderem als Theorem A.2.2 in White (1994) formuliert.

**Theorem F.3 Tweedies Theorem**

Sei  $E$  ein lokal kompakter separabler metrischer Raum mit  $\mathcal{F}$  einer Borel- $\sigma$ -Algebra auf  $E$ . Sei außerdem  $\{X_n, n \in \mathbb{N}_0\}$  eine Fellerkette mit Übergangswahrscheinlichkeit  $P(x, B) = P(X_t \in B | X_{t-1} = x)$  für  $x \in E$  und  $B \in \mathcal{F}$ .

(1) Wenn für eine kompakte Menge  $B \in \mathcal{B}$  mit Komplement  $B^C$  sowohl eine nicht-negative Funktion  $g(\cdot)$  als auch ein  $\varepsilon > 0$  existieren, für die gilt:

$$\int_{B^C} P(x, dy) g(y) \leq g(x) - \varepsilon, \text{ für } x \in B^C,$$

dann existiert ein  $\sigma$ -endliches Maß  $\mu$  für  $P$  mit  $0 < \mu(B) < \infty$ .

(2) Gilt außerdem

$$\int_B \mu(dx) \left[ \int_{A^C} P(x, dy) g(y) \right] < \infty,$$

dann ist  $\mu$  endlich und somit  $\pi = \frac{\mu}{\mu(E)}$  ein invariantes Wahrscheinlichkeitsmaß.

(3) Gilt außerdem

$$\int_{B^C} P(x, dy) g(y) \leq g(x) - f(x), \text{ für } x \in A^C,$$

dann besitzt das Maß  $\mu$  ein endliches  $f$ -Moment, das heißt, es gilt:

$$\int_E \mu(dy) f(y) < \infty.$$

Dieses Theorem wurde ursprünglich als Theorem 2 und 3 in Tweedie (1988) formuliert und bewiesen. Die hier angegebene Variante entspricht der von Fermanian und Malongo (2017) verwendeten und wird in der Literatur häufig als Tweedies Theorem bezeichnet.

**Definition F.1** *Vektortransformierte Brownsche Bewegung* (Davidson (1994), S.486 und 492)

Sei  $\{W(t), t \in [0, 1]\}$  eine Standard-Brownsche Bewegung und  $\mathcal{H}_S$  die Menge aller Homöomorphismen  $f: S \rightarrow S$ , für die gilt:  $f(0) = 0$  und  $f(1) = 1$ .

Für  $\eta \in \mathcal{H}_{[0,1]}$  wird  $\{W_\eta(t), t \in [0, 1]\}$  dann als vektortransformierte Brownsche Bewegung bezeichnet, wenn gilt:

$$W_\eta(t) \stackrel{d}{=} W(\eta(t)), \quad \forall t \in [0, 1].$$

Sei außerdem  $\boldsymbol{\eta} := (\eta^1, \dots, \eta^p)'$  ein Vektor von Homöomorphismen  $\eta_i \in \mathcal{H}_{[0,1]}$ , für  $i = 1, \dots, p$ , und  $\Omega$  eine  $(m \times m)$ -Kovarianzmatrix.

Dann wird der Prozess  $\{\mathbf{W}_{\Omega, \boldsymbol{\eta}}(t), t \in [0, 1]\}$  als multivariate vektortransformierte Brownsche Bewegung bezeichnet, wenn  $\{\mathbf{W}_{\Omega, \boldsymbol{\eta}}(t), t \in [0, 1]\}$  ein Gaußprozess ist mit Mittelwertfunktion  $\mathbb{E}[\mathbf{W}_{\Omega, \boldsymbol{\eta}}(t)] = \mathbf{0}, \forall t \in [0, 1]$ , und Kovarianzkern

$$\mathbb{E}[\mathbf{W}_{\Omega, \boldsymbol{\eta}}(t) \mathbf{W}_{\Omega, \boldsymbol{\eta}}(t)'] = DH(t)D',$$

wobei  $D$  eine spaltenreguläre  $(p \times m)$ -Matrix mit  $DD' = \Omega$  ist und

$$H(t) = \text{diag} \left\{ \eta^1(t), \dots, \eta^m(t) \right\}, \quad \forall t \in [0, 1].$$

**Theorem F.4** *Funktionaler Grenzwertsatz für near-epoch dependente Funktionen von mischenden Prozessen* (Davidson (1994), S.492)

Sei  $\{U_{ni}\}$  ein Schema beziehungsweise ein array von Zufallsvektoren, die jeweils den Erwartungsvektor  $\mathbf{0}_n$  besitzen. Für eine steigende rechtsstetige Funktion  $K_n(\cdot)$ , die nur ganzzahlige Werte annimmt, sei

$$X_n^K = \sum_{i=1}^{K_n(t)} U_{ni}$$

definiert. Seien außerdem die folgenden Bedingungen erfüllt:

- (a) Für jeden festen  $p$ -dimensionalen Vektor  $\boldsymbol{\lambda}$  mit  $\boldsymbol{\lambda}'\boldsymbol{\lambda} = 1$ , existieren eine Folge  $\{c_{ni}^\lambda\}$  und ein Homöomorphismus  $\eta^\lambda \in \mathcal{H}_{[0,1]}$ , so dass gilt:

$$\mathbb{E} \left( X_n^K(t)^2 \right) \rightarrow \eta(t), \quad \text{für } n \rightarrow \infty \quad \text{und } \forall t \in [0, 1].$$

- (b)  $\mathbb{E}(\boldsymbol{\lambda}'U_{ni}) = 0$ ;

- (c)  $\sup_{n,i} \left| \frac{\boldsymbol{\lambda}'U_{ni}}{c_{ni}^\lambda} \right|_r < \infty$  für ein  $r > 2$ ;

- (d) Der Prozess  $\{\boldsymbol{\lambda}'U_{ni}\}$  ist  $\mathcal{L}_2$ -near-epoch dependent der Ordnung  $-\gamma$  für  $\gamma \in [1/2, 1]$  auf einem stark mischenden Prozess  $\{V_{ni}\}$  der Ordnung  $-\frac{r}{r-2}$ ;

*Fortsetzung auf der nächsten Seite*

**Fortsetzung von Theorem F.4**

$$(e) \quad \sup_{t \in [0,1], \delta \in (0,1-t]} \left\{ \limsup_{n \rightarrow \infty} \frac{v_n^2(t, \delta)}{\delta} \right\} < \infty, \text{ für } v_n^2(t, \delta) = \sum_{i=K_n(t)+1}^{K_n(1+\delta)} c_{ni}^2;$$

$$(f) \quad \max_{1 \leq i \leq K_n(1)} c_{ni} = \mathcal{O} \left( \frac{1}{K_n(1)^{1-\gamma}} \right) \text{ für } \gamma \text{ aus (d);}$$

(g) Sei für die Funktion  $H(\cdot)$  aus Definition F.1 der Homöomorphismus  $\eta^j$  jeweils die Funktion  $\eta^\lambda$ , die sich für  $\lambda = e_j$  ergibt. Dabei ist  $e_j$  der  $j$ -te Einheitsvektor beziehungsweise die  $j$ -te Spalte von  $\mathbb{I}_p$ , für  $j = 1, \dots, p$ . Es sei angenommen, dass gilt:

$$\mathbb{E} \left[ X_n^K(t) X_n^K(t)' \right] = DH(t)D \quad \text{für } n \rightarrow \infty.$$

Dann gilt:

$$\left\{ X_n^K(t), t \in [0, 1] \right\} \xrightarrow{d} \left\{ \mathbf{W}_{\Omega, \eta}(t), t \in [0, 1] \right\}.$$

Dieses Theorem wurde als Theorem 29.18 in Davidson (1994) formuliert und bewiesen.

**Proposition F.1** Sei  $A = [a_{ij}]_{i,j=1,\dots,n}$  eine quadratische Matrix und

$$A^* := \text{diag} \left\{ a_{11}^{1/2}, \dots, a_{nn}^{1/2} \right\}.$$

Dann ist  $A$  genau dann positiv definit, wenn dies auch für

$$B = [A^*]^{-1} A [A^*]^{-1}$$

gilt.

Diese Aussage entspricht Proposition 1 in Engle und Sheppard (2001) und wurde dort auch bewiesen.

**Lemma F.1** Wenn  $\{\xi_k, 0 \leq k < \infty\}$  eine Folge von identisch verteilten Zufallsvariablen ist, die

$$\mathbb{E} \log^+ |\xi_0| < \infty$$

erfüllen, dann konvergiert  $\sum_{k=0}^{\infty} \xi_k z^k$  mit Wahrscheinlichkeit Eins für jedes  $|z| < 1$ .

Diese Aussage wurde als Lemma 2.2. in Berkes et al. (2003) formuliert und bewiesen.

**Lemma F.2 Borel-Cantelli-Lemma**

Sei  $(A_n)_{n \in \mathbb{N}}$  eine Folge von Ereignissen in einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$ .  
Dann gilt:

$$\sum_{n=1}^{\infty} P(A_n) < \infty \quad \Rightarrow \quad P\left(\limsup_{n \rightarrow \infty} A_n\right) = 0.$$

Dieses Lemma, das in dieser Arbeit als Borel-Cantelli-Lemma bezeichnet wird, obwohl es sich eigentlich nur um einen Teil der von Émile Borel und Francesco Paolo Cantelli formulierten Aussagen handelt, wurde unter anderem als Lemma 11.1 in Bauer (1991) festgehalten.

**Lemma F.3 Verallgemeinerte Hölder-Ungleichung** (Alt (2006), S.51)

Bezeichne  $\mathcal{L}^p(\mu)$  einen Lebesgue-Raum auf einem gegebenen Maßraum. Es sei  $m \in \mathbb{N}$  und  $f_i \in \mathcal{L}^{p_i}(\mu)$  für  $i = 1, \dots, m$  mit  $p_i \in [1, \infty]$ , sowie  $q \in [1, \infty]$  mit

$$\sum_{i=1}^m \frac{1}{p_i} = \frac{1}{q}.$$

Dann ist das Produkt  $f_1 \cdot \dots \cdot f_m \in \mathcal{L}^q(\mu)$  und es gilt

$$\left\| \prod_{i=1}^m f_i \right\|_q \leq \prod_{i=1}^m \|f_i\|_{p_i}.$$

Diese Aussage wurde unter anderem als Lemma 1.16 in Alt (2006) formuliert und wird in dieser Arbeit als verallgemeinerte Hölder-Ungleichung bezeichnet.



# Tabellenverzeichnis

3.1	Simulierte kritische Werte $c(\tilde{\alpha})$ (nichtparametrisch) . . . . .	18
3.2	Stoppzeiten und geschätzte Bruchpunkte bei der Anwendung des nichtparametrischen Verfahrens auf Aktienrenditen. . . . .	38
3.3	Geschätzte Volatilitäten der log-Renditen zwischen aufeinanderfolgenden Bruchpunkten. . . . .	38
5.1	Simulierte kritische Werte $c(\tilde{\alpha})$ (DCC-Modell). . . . .	55
5.2	Niveauvergleich für verschiedene Parameterkombinationen. . . . .	59
5.3	Güte gegenüber Parameteränderungen: Varianzreduktion ( $p = 3$ ). . . . .	60
5.4	Güte gegenüber Parameteränderungen: Varianzreduktion ( $p = 5$ ). . . . .	61
5.5	Güte gegenüber Parameteränderungen: Varianzanstieg ( $p = 3$ ). . . . .	62
5.6	Güte gegenüber Parameteränderungen: Varianzanstieg ( $p = 5$ ). . . . .	62
5.7	Stoppzeiten und geschätzte Bruchpunkte bei der Anwendung des parametrischen Verfahrens auf Aktienrenditen. . . . .	66
5.8	Geschätzte Parameter zwischen zwei aufeinanderfolgenden Bruchpunkten. . . . .	66
B.1	Niveau bei der Überwachung von normalverteilten Zufallsvektoren. . . . .	77
B.2	Niveau bei der Überwachung von $t$ -verteilten Zufallsvektoren. . . . .	77
B.3	Güte bei der Überwachung von normalverteilten Zufallsvektoren, wenn sich alle Varianzen ändern. . . . .	78
B.4	Güte bei der Überwachung von $t$ -verteilten Zufallsvektoren, wenn sich alle Varianzen ändern. . . . .	78
B.5	Güte bei der Überwachung von normalverteilten Zufallsvektoren, wenn sich nur eine der Varianzen ändert. . . . .	79
B.6	Güte bei der Überwachung von $t$ -verteilten Zufallsvektoren, wenn sich nur eine der Varianzen ändert. . . . .	79
B.7	Niveau bei der Überwachung von GARCH-Zeitreihen mit normalverteilten Innovationen. . . . .	80
B.8	Niveau bei der Überwachung von GARCH-Zeitreihen mit $t$ -verteilten Innovationen. . . . .	80
B.9	Güte bei der Überwachung von GARCH-Zeitreihen mit normalverteilten Innovationen, wenn sich alle Varianzen ändern. . . . .	81
B.10	Güte bei der Überwachung von GARCH-Zeitreihen mit $t$ -verteilten Innovationen, wenn sich alle Varianzen ändern. . . . .	81
B.11	Güte bei der Überwachung von GARCH-Zeitreihen mit normalverteilten Innovationen, wenn sich nur eine der Varianzen ändert. . . . .	82
B.12	Güte bei der Überwachung von GARCH-Zeitreihen mit $t$ -verteilten Innovationen, wenn sich nur eine der Varianzen ändert. . . . .	82
B.13	Güte bei der Überwachung von GARCH-Zeitreihen mit normalverteilten Innovationen durch univariate Verfahren, wenn sich alle Varianzen ändern. . . . .	83
B.14	Güte bei der Überwachung von GARCH-Zeitreihen mit $t$ -verteilten Innovationen durch univariate Verfahren, wenn sich alle Varianzen ändern. . . . .	83

---

B.15 Güte bei der Überwachung von GARCH-Zeitreihen mit normalverteilten Innovationen durch univariate Verfahren, wenn sich nur eine der Varianzen ändert. . . . .	84
B.16 Güte bei der Überwachung von GARCH-Zeitreihen mit $t$ -verteilten Innovationen durch univariate Verfahren, wenn sich nur eine der Varianzen ändert. . . . .	84
B.17 Güte bei der Überwachung von GARCH-Zeitreihen mit normalverteilten Innovationen, wenn sich die Innovationsvarianz ändert. . . . .	85
B.18 Güte bei der Überwachung von GARCH-Zeitreihen mit $t$ -verteilten Innovationen, wenn sich die Innovationsvarianz ändert. . . . .	85

# Abbildungsverzeichnis

3.1	Grenzfunktion $cw(\cdot)$ . . . . .	19
3.2	Niveauvergleich: u.i.v. Zufallsvektoren und GARCH-Zeitreihen. . . . .	24
3.3	Güte bei der Anwendung auf Zufallsvektoren, wenn sich alle Varianzen ändern. . . . .	25
3.4	Güte bei der Anwendung auf Zufallsvektoren, wenn sich nur eine der Varianzen ändert. . . . .	26
3.5	Niveauvergleich: Überwachung von GARCH-Zeitreihen mit multi- und univariaten Methoden. . . . .	30
3.6	Güte bei der Überwachung von skalaren BEKK-Zeitreihen unter der Verwendung des multivariaten Überwachungsverfahrens, wenn alle Varianzen von der Änderung betroffen sind. . . . .	31
3.7	Güte bei der Überwachung von GARCH-Zeitreihen unter der Verwendung des multivariaten Überwachungsverfahrens, wenn nur eine der Varianzen vom Strukturbruch betroffen ist. . . . .	31
3.8	Güte bei der Überwachung von GARCH-Zeitreihen mit normalverteilten Innovationen unter Verwendung von univariaten Überwachungsverfahren. . . . .	32
3.9	Eigenschaften der Verzögerungszeiten und des Lokationsschätzers bei der Überwachung von GARCH-Zeitreihen mit normalverteilten Innovationen, wenn alle Varianzen von der Änderung betroffen sind. . . . .	33
3.10	Eigenschaften der Verzögerungszeiten und des Lokationsschätzers bei der Überwachung von GARCH-Zeitreihen mit normalverteilten Innovationen, wenn nur eine der Varianzen von der Änderung betroffen ist. . . . .	33
3.11	Güte bei der Überwachung von GARCH-Zeitreihen mit normalverteilten Innovationen, wenn sich die Varianzen der Innovationen oder die bedingten Varianzen ändern. . . . .	34
3.12	Güte bei der Überwachung von GARCH-Zeitreihen mit normal- oder $t$ -verteilten Innovationen und Varianzerhöhung um den Wert $\Delta$ . . . . .	35
3.13	log-Renditen der Aktienkurswerte mit den ermittelten Bruchpunkten. . . . .	39
3.14	Geschätzte Volatilitäten der log-Renditen zwischen aufeinanderfolgenden Bruchpunkten. . . . .	40
5.1	Güte gegenüber Korrelationsänderungen im DCC-Modell. . . . .	63
5.2	log-Renditen der Aktienkurswerte mit den ermittelten Bruchpunkten. . . . .	65



# Symbolverzeichnis

## Allgemeines

$\ x\ _2$	Euklidische Norm eines realisierten Vektors $x$
$\ X\ _p$	$\mathcal{L}_p$ -Norm eines Zufallsvektors $X$
$ A $	Maximumsnorm einer Matrix $A =$ maximaler absoluter Eintrag
$[X]_{ij}$	Eintrag in der $i$ -ten Zeile und $j$ -ten Spalte der Matrix $X$
$[x_{ij}]_{\substack{i=1,\dots,m; \\ j=1,\dots,n}}$	$(m \times n)$ -Matrix, die durch ihre Einträge $x_{ij}$ definiert ist
$X^{(i,j)}$	Streichungsmatrix, die entsteht, wenn die $i$ -te Zeile und die $j$ -te Spalte aus der Matrix $X$ gelöscht werden
$f'(u)$	Ableitung einer Funktion $f(u)$ nach allen Argumenten $u$
$X'; X^T$	Transponierte eines Vektors oder einer Matrix $X$ ; in Fällen, in denen Transponierte und Ableitung verwechselt werden können, wird die Notation $X^T$ verwendet
$X^{-1}$	Inverse einer Matrix $X$
$[a]$	Untere Gaußklammer einer reellen Zahl $a$
$X \otimes Y$	Kroneckerprodukt der Matrizen $X$ und $Y$
$\otimes_k X$	$k$ -faches Kroneckerprodukt identischer Matrizen $X$
$X \stackrel{d}{=} Y$	Verteilungsgleichheit der Zufallsvariablen $X$ und $Y$
$X \xrightarrow{d} Y$	Konvergenz in Verteilung von $X$ gegen $Y$
$X \xrightarrow{p} Y$	Konvergenz in Wahrscheinlichkeit von $X$ gegen $Y$
$X \xrightarrow{f.s.} Y$	Fast sichere Konvergenz von $X$ gegen $Y$
$\stackrel{f.s.}{\leq}, \stackrel{f.s.}{\geq}, \stackrel{f.s.}{=}$	Fast sichere Größenrelationen
$\stackrel{f.s.}{<}, \stackrel{f.s.}{>}$	Fast sichere strikte Größenrelationen

## Numerische Ausdrücke

$\mathbf{0}_p$	Nullvektor der Länge $p$
$\mathbf{0}_{m \times n}$	Nullmatrix der Dimension $(m \times n)$

## Lateinische Symbole

$\mathbf{A}$	Matrix $\text{diag}(\alpha_1, \dots, \alpha_p)$ im DCC-Modell
$\mathbf{b}(\cdot); w(\cdot)$	Grenzfunktion in sequenziellen Verfahren
$b_{BT}$	Anzahl der Wiederholungen bei Bootstrapverfahren
$B$	Länge des Überwachungszeitraums als $B$ -faches der Länge des historischen Zeitraums
$\mathbf{B}$	Matrix $\text{diag}(\beta_1, \dots, \beta_p)$ im DCC-Modell
$c; c(\tilde{\alpha})$	Skalierungskonstante für die Grenzfunktion $w(\cdot)$ bzw. $\mathbf{b}(\cdot)$

$C_1; C_2; C_\lambda$	Abschätzungs-konstanten in DCC-Modellen
$\text{Cov}(X)$	Kovarianzmatrix eines Zufallsvektors $X$
$\text{Cov}(X Y)$	Auf $Y$ bedingte Kovarianzmatrix eines Zufallsvektors $X$
$\text{Cor}(X)$	Korrelationsmatrix eines Zufallsvektors $X$
$\text{Cor}(X Y)$	Auf $Y$ bedingte Korrelationsmatrix eines Zufallsvektors $X$
$d$	Gesamtzahl der Parameter im DCC-Modell
$\text{diag}\{a_1, \dots, a_p\}$	$(p \times p)$ -Diagonalmatrix mit Hauptdiagonale $a_1, \dots, a_p$
$D[d_1, d_2]$	Raum der càdlàg-Funktionen auf dem Intervall $[d_1, d_2]$ ausgestattet mit der Supremumsnorm
$D_t$	Diagonalmatrix der bedingten Standardabweichungen im DCC-Modell
$\mathbf{D}; \mathbf{D}_p$	Variations- bzw. Informationsmatrix zur Skalierung des Detektors
$\mathbf{D}_m(u)$	Empirisches Gegenstück zu $\mathbf{D}$ im DCC-Modell auf Basis von $m$ Gradientenwerten mit unendlicher Vergangenheit
$\widehat{\mathbf{D}}_m(u)$ mit	Empirisches Gegenstück zu $\mathbf{D}$ im DCC-Modell auf Basis von $m$ Gradientenwerten mit endlicher Vergangenheit
$\widehat{\mathbf{D}}_m(\hat{\theta}_m) = \widehat{\mathbf{D}}_m$	
$\widehat{\mathbf{D}}_m^{*(i)}(u)$	Schätzer für die Variationsmatrix aus der $i$ -ten Bootstrapstichprobe
$\mathbf{D}_p$	Duplikationsmatrix: $\mathbf{D}_p \text{vech}(A) = \text{vec}(A)$
$E(X)$	Erwartungswert einer Zufallsvariablen bzw. eines Zufallsvektors $X$
$E(X Y)$	Auf $Y$ bedingter Erwartungswert einer Zufallsvariablen bzw. eines Zufallsvektors $X$
$\mathbf{F}_{D_t}(u); \widehat{\mathbf{F}}_{D_t}(u)$	Diagonalmatrix der bedingten Standardabweichungen im DCC-Modell in Abhängigkeit von einem Element $u$ aus dem beschränkten Parameterraum $U$ mit un-/endlicher Vergangenheit
$\mathbf{F}_{Q_t}(u); \widehat{\mathbf{F}}_{Q_t}(u)$	Bedingte Hilfsmatrix im DCC-Modell in Abhängigkeit von $u$ mit un-/endlicher Vergangenheit
$\mathbf{F}_{Q_t^*}(u); \widehat{\mathbf{F}}_{Q_t^*}(u)$	Normierungsmatrix für $\mathbf{F}_{Q_t}(u)$ bzw. $\widehat{\mathbf{F}}_{Q_t}(u)$
$\mathbf{F}_{\widehat{Q}}(u)$	Unbedingte Korrelationsmatrix im DCC-Modell in Abhängigkeit von $u$
$\mathbf{F}_{R_t}(u); \widehat{\mathbf{F}}_{R_t}(u)$	Bedingte Korrelationsmatrix im DCC-Modell in Abhängigkeit von $u$ mit un-/endlicher Vergangenheit
$\mathbb{F}$	An den Prozess $\{y_t\}$ adaptierte Filtration
$\mathcal{F}_t$	Informationsmenge des Prozesses $\{y_t\}$ zum Zeitpunkt $t$
$\mathcal{F}_B$	Menge der beschränkten und auf dem Intervall $[0, B + 1]$ integrierbaren Funktionen
$\mathcal{G}_p(\cdot)$	$p$ -variater Gaußprozess gegen den jeweils die Folge der Detektorwerte der sequenziellen Verfahren strebt
$h_{1t}, \dots, h_{pt}$	Bedingte Varianzen im DCC-Modell
$h_t$	Vektor der bedingten Varianzen im DCC-Modell
$H$	Unbedingte Kovarianzmatrix im skalaren BEKK-Modell
$H_t$	Bedingte Kovarianzmatrix in multivariaten GARCH-Modellen
$H_0$	Nullhypothese
$H_1$	Alternative
$\mathcal{I}$	An den umformulierten Prozess $\{X_t\}$ adaptierte Filtration
$\mathcal{I}_t$	Informationsmenge des Prozesses $\{X_t\}$ zum Zeitpunkt $t$
$\mathbb{I}_p$	Einheitsmatrix der Dimension $(p \times p)$
$k^*$	Zeitpunkt des Strukturbruchs im Überwachungszeitraum
$\hat{k}$	Schätzer für die Lage des Strukturbruchs im Überwachungszeitraum
$\hat{k}_r$	Schätzer für die Lage eines retrospektiv entdeckten Strukturbruchs

$K_{mn}$	Kommutationsmatrix: $\text{vec}(A') = K_{mn}\text{vec}(A)$
$L_T(u); \hat{L}_T(u)$	Gaußsche Quasi-Loglikelihoodfunktion basierend auf $T$ Beobachtungen mit un-/endlicher Vergangenheit
$l_t(u); \hat{l}_t(u)$	Beitrag zur Gaußschen Quasi-Loglikelihoodfunktion mit un-/endlicher Vergangenheit
$l'_t(u); \hat{l}'_t(u)$	Gradient des Beitrags zur Gaußschen Quasi-Loglikelihoodfunktion mit un-/endlicher Vergangenheit
$\hat{l}^{*(i)'}(u)$	Gradient des Beitrags zur Gaußschen Quasi-Loglikelihoodfunktion aus der $i$ -ten Bootstrapstichprobe mit endlicher Vergangenheit
$L(\cdot)$	Grenzfunktion von $L_T(\cdot)$
$L_p$	Eliminationsmatrix: $L_p\text{vec}(A) = \text{vech}(A)$
$m$	Länge des historischen Zeitraums
$\mathcal{N}_p(\mu, \Sigma)$	$p$ -variate Normalverteilung mit Erwartungsvektor $\mu$ und Kovarianzmatrix $\Sigma$
$o_P; \mathcal{O}$	Landausymbole
$p$	Dimension der betrachteten Zufallsvektoren
$p^+$	Anzahl der Einträge auf und unter der Hauptdiagonalen einer quadratischen $(p \times p)$ -Matrix: $\frac{1}{2}p(p+1)$
$p^-$	Anzahl der Einträge unter der Hauptdiagonalen einer quadratischen $(p \times p)$ -Matrix: $\frac{1}{2}p(p-1)$
$P_{H_0}(A)$	Wahrscheinlichkeit, dass ein Ereignis $A$ unter der Nullhypothese eintritt
$P_{H_1}(A)$	Wahrscheinlichkeit, dass ein Ereignis $A$ unter der Alternative eintritt
$q_1, \dots, q_{p^-}$	Variable Einträge von $F_{\bar{Q}}(u)$ im DCC-Modell
$Q_t$	Bedingte Hilfsmatrix im DCC-Modell
$Q_t^*$	Normierungsmatrix für $Q_t$ im DCC-Modell
$\bar{Q}; \bar{Q}_p$	Unbedingte Korrelationsmatrix im DCC-Modell
$\bar{q}_{ij}$	Einträge von $\bar{Q}$ im DCC-Modell
$\bar{Q}^*$	Unbedingte Korrelationsmatrix im DCC-Modell nach einem Strukturbruch
$\bar{q}_{ij}^*$	Einträge von $\bar{Q}^*$ im DCC-Modell
$Q_k$	Detektor zur retrospektiven Detektion von Varianzänderungen
$r_i$	Teilvektor von $u$ , der einer allgemein parametrisierten Variante von $\phi_i$ entspricht
$R_t$	Bedingte Korrelationsmatrix im DCC-Modell
$s_1, \dots, s_p$	Allgemein parametrisierte Werte für $\alpha_1, \dots, \alpha_p$ im Vektor $u$
$S_p$	Menge aller Permutationen von $(1, \dots, p)$
$t_1, \dots, t_p$	Allgemein parametrisierte Werte für $\beta_1, \dots, \beta_p$ im Vektor $u$
$t_{p,\nu}(\mu, \Sigma)$	$p$ -variate $t$ -Verteilung mit Erwartungsvektor $\mu$ und Kovarianzmatrix $\Sigma$ sowie $\nu$ Freiheitsgraden
$T_t$	Matrix zur autoregressiven Darstellung der DCC-Modellgleichungen
$T^{*(i)}$	Maximum der Detektorwerte einer simulierten Bootstrapstichprobe
$U$	Beschränkter Parameterraum im DCC-Modell
$u$	Element aus dem beschränkten Parameterraum $U$
$\underline{u}$	Untere Schranke für einige Parameter im DCC-Modell
$\bar{u}$	Obere Schranke für einige Parameter im DCC-Modell

$V_k$	Detektorwerte bei sequenziellen Verfahren
$V_k^{*(i)}$	Detektorwerte basierend auf einer simulierten Bootstrapstichprobe
$\text{vec}(X)$	Vektor der spaltenweisen Einträge einer Matrix $X$
$\text{vech}(X)$	Vektor der Einträge auf und unter der Hauptdiagonalen einer quadratischen Matrix $X$
$\text{vecl}(X)$	Vektor der Einträge unter der Hauptdiagonalen einer quadratischen Matrix $X$
$\text{Var}(X)$	Varianz einer Zufallsvariablen $X$
$w(\cdot); \mathbf{b}(\cdot)$	Grenzfunktion in sequenziellen Verfahren
$w_{it}(u); \widehat{w}_{it}(u)$	Bedingte Varianz im DCC-Modell in Abhängigkeit von $u$ mit un-/endlicher Vergangenheit
$W(\cdot); W_i(\cdot)$	Univariate Brownsche Bewegung; ggfs. mit Index, um mehrere Brownsche Bewegungen zu unterscheiden
$\mathbf{W}_p(\cdot)$	$p$ -dimensionale Brownsche Bewegung
$x_1, \dots, x_p$	Allgemein parametrisierte Werte für $\omega_1, \dots, \omega_p$ im Vektor $u$
$y_t$	GARCH-Zeitreihenwerte (uni- oder multivariat)
$y_t^{(2)}$	Vektor der quadrierten Einträge von $y_t$
$Y^{*(i)}$	Simulierte Bootstrapstichprobe
$z_t$	„Standardisierte“ Vektoren im DCC-Modell
$\mathbf{z}(u); \widehat{\mathbf{z}}(u)$	„Standardisierte“ Vektoren in Abhängigkeit von $u$ mit un-/endlicher Vergangenheit
$Z_t$	Matrix $\text{diag} \{z_{1t}, \dots, z_{pt}\}$ im DCC-Modell

### Griechische Symbole

$\alpha, \alpha_1, \dots, \alpha_p$	Modellparameter in multivariaten GARCH-Modellen
$\alpha^*, \alpha_1^*, \dots, \alpha_p^*$	Modellparameter im DCC-Modell nach einem Strukturbruch
$\tilde{\alpha}$	Signifikanzniveau der sequenziellen Verfahren
$\beta, \beta_1, \dots, \beta_p$	Modellparameter in multivariaten GARCH-Modellen
$\beta^*, \beta_1^*, \dots, \beta_p^*$	Modellparameter im DCC-Modell nach einem Strukturbruch
$\delta_1$	Untere Schranke für die Eigenwerte von $\bar{Q}$ im DCC-Modell
$\delta_2$	Obere Schranke für die Eigenwerte von $Q_t$ im DCC-Modell
$\delta_*$	Obere Schranke für die Einträge von $F_{R_0}(u)^{-1}$ im DCC-Modell
$\varepsilon$	Positive Konstante
$\epsilon_t$	Innovation bzw. Innovationsvektor in GARCH-Modellen
$\zeta_t$	Vektor zur autoregressiven Darstellung der DCC-Modellgleichungen
$\eta_t$	Bedingte Innovationsvektoren im DCC-Modell
$\theta_t$	Parametervektor im DCC-Modell zum Zeitpunkt $t$
$\boldsymbol{\theta}$	Parametervektor im DCC-Modell (vor dem Strukturbruch)
$\boldsymbol{\theta}_1$	Vektor der Varianzparameter im DCC-Modell
$\boldsymbol{\theta}_2$	Vektor der Korrelationsparameter im DCC-Modell
$\boldsymbol{\theta}^*$	Parametervektor im DCC-Modell nach einem Strukturbruch
$\Theta$	Uneingeschränkter Parameterraum im DCC-Modell
$\lambda^*$	Relative Lage des Strukturbruchs im Überwachungszeitraum
$\lambda_{\min}(X)$	Kleinster Eigenwert einer Matrix $X$
$\lambda_{\max}(X)$	Größter Eigenwert einer Matrix $X$

$\mu$	Erwartungsvektor eines Zufallsvektors
$\nu$	Anzahl der Freiheitsgrade der $t$ -Verteilung
$\pi$	Konkrete Permutation von $(1, \dots, p)$
$\rho$	Obere Schranke für einige Parameter im DCC-Modell
$\sigma_t^2$	Bedingte Varianz in univariaten GARCH-Modellen
$\boldsymbol{\sigma}_t^2$	Vektor der Varianzen von $(X_{1t}, \dots, X_{pt})'$
$\Sigma; \Sigma_p$	Kovarianzmatrix eines Zufallsvektors
$\tau_m$	Stoppzeit eines sequenziellen Verfahrens
$\phi_i$	Parameter der $i$ -ten Varianzgleichung im DCC-Modell
$\phi_i^*$	Parameter der $i$ -ten Varianzgleichung im DCC-Modell nach einem Strukturbruch
$\psi$	Vektor der Einträge von $\bar{Q}$ im DCC-Modell
$\psi^*$	Vektor der Einträge von $\bar{Q}$ im DCC-Modell nach einem Strukturbruch
$\omega_1, \dots, \omega_p$	Parameter im DCC-Modell
$\omega_1^*, \dots, \omega_p^*$	Parameter im DCC-Modell nach einem Strukturbruch
$\boldsymbol{\omega}$	Vektor $(\omega_1, \dots, \omega_p)'$ im DCC-Modell



# Abkürzungsverzeichnis

<b>ARCH-Modell</b>	Modell mit autoregressiver bedingter ( <i>conditional</i> ) <b>H</b> eteroskedastie
<b>ARMA-Modell</b>	Modell mit autoregressivem und <i>moving average</i> -Einfluss
<b>BEKK-Modell</b>	GARCH-Modell, das auf die Arbeiten von <b>B</b> aba, <b>E</b> ngle, <b>K</b> raft und <b>K</b> roner zurückgeht
<b>CCC-Modell</b>	GARCH-Modell mit dynamischer bedingter Korrelation ( <i>constant conditional correlation</i> )
<b>DCC-Modell</b>	GARCH-Modell mit dynamischer bedingter Korrelation ( <i>dynamic conditional correlation</i> )
<b>GARCH-Modell</b>	Modell mit verallgemeinerter ( <i>generalized</i> ) autoregressiver bedingter ( <i>conditional</i> ) <b>H</b> eteroskedastie
<b>MWS</b>	Mittelwertsatz
<b>o.B.d.A.</b>	Ohne <b>B</b> eschränkung der <b>A</b> llgemeinheit
<b>QLL</b>	Quasi- <b>L</b> oglikelihood
<b>QMLS</b>	Quasi- <b>M</b> aximum- <b>L</b> ikelihood- <b>S</b> chätzer
<b>u.i.v.</b>	Unabhängig <b>i</b> dentisch <b>v</b> erteilt



# Literatur

- Alt, H.W.** (2006): *Lineare Funktionalanalysis: Eine anwendungsorientierte Einführung*. 5. Aufl. Springer-Verlag, Berlin Heidelberg.
- Amado, C. und T. Teräsvirta** (2014): Modelling Changes in the Unconditional Variance of Long Stock Return Series. *Journal of Empirical Finance* 25, S. 15–35.
- Amemiya, T.** (1985): *Advanced Econometrics*. Harvard University Press, Cambridge, MA, USA.
- Aue, A., L. Horváth, M. Hušková und P. Kokoszka** (2006): Change-point Monitoring in Linear Models. *Econometrics Journal* 9(3), S. 373–403.
- Aue, A., S. Hörmann, L. Horváth und M. Reimherr** (2009a): Break Detection in the Covariance Structure of Multivariate Time Series Models. *The Annals of Statistics* 37(6B), S. 4046–4087.
- Aue, A., L. Horváth und M. Reimherr** (2009b): Delay Times of Sequential Procedures for Multiple Time Series Regression Models. *Journal of Econometrics* 149(2), S. 174–190.
- Aue, A., S. Hörmann, L. Horváth, M. Hušková und J.G. Steinebach** (2012): Sequential Testing for the Stability of High Frequency Portfolio Betas. *Econometric Theory* 28(4), S. 804–837.
- Bartsch, R.** (2007): *Allgemeine Topologie I*. Oldenbourg Wissenschaftsverlag, München.
- Bauer, H.** (1991): *Wahrscheinlichkeitstheorie*. 4. Aufl. Walter de Gruyter, Berlin, New York.
- Bauer, P. und P. Hackl** (1978): The Use of MOSUMS for Quality Control. *Technometrics* 20(4), S. 431–436.
- Bauwens, L., S. Laurent und J.V.K. Rombouts** (2006): Multivariate GARCH Models: A Survey. *Journal of Applied Econometrics* 21(1), S. 79–109.
- Berkes, I., L. Horváth und P. Kokoszka** (2003): GARCH Processes: Structure and Estimation. *Bernoulli* 9(2), S. 201–227.
- Berkes, I., E. Gombay, L. Horváth und P. Kokoszka** (2004): Sequential Change-point Detection in GARCH(p,q) Models. *Econometric Theory* 20(6), S. 1140–1167.
- Bissantz, N., D. Ziggel und K. Bissantz** (2011): An Empirical Study of Correlation and Volatility Changes of Stock Indices and their Impact on Risk Figures. *Acta Universitatis Danubis. Oeconomica* 7(4), S. 127–141.
- Bollerslev, T.** (1986): Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 31(3), S. 307–327.
- (1990): Modelling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model. *The Review of Economics and Statistics* 72(3), S. 498–505.
- (2009): Glossary to ARCH (GARCH). *T. Bollerslev, J.R. Russell und M. Watson (Hgg.), Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle, Oxford University Press, New York*, S. 137–163.
- Bollerslev, T., R.F. Engle und J.M. Wooldridge** (1988): A Capital Asset Pricing Model with Time-varying Covariances. *Journal of Political Economy* 96(1), S. 116–131.
- Bosch, S.** (2006): *Lineare Algebra*. 3. Aufl. Springer, Berlin, Heidelberg, New York.
- Cameron, A.C. und P.K. Trivedi** (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press, New York.

- Capiello, L., R. Engle und K. Sheppard** (2006): Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns. *Journal of Financial Econometrics* 4(4), S. 537–572.
- Charles, A. und O. Darné** (2014): Large Shocks in the Volatility of the Dow Jones Industrial Average Index: 1928–2013. *Journal of Banking and Finance* 43, S. 188–199.
- Chu, C.-S. J., M. Stinchcombe und H. White** (1996): Monitoring Structural Change. *Econometrica* 64(5), S. 1045–1065.
- Davidson, J.** (1994): *Stochastic Limit Theory: An Introduction for Econometricians*. Oxford University Press, New York.
- Davidson, R. und J.G. MacKinnon** (2000): Bootstrap Tests: How Many Bootstraps? *Econometric Reviews* 19(1), S. 55–68.
- De Jong, R.M. und J. Davidson** (2000): Consistency of Kernel Estimators of Heteroscedastic and Autocorrelated Covariance Matrices. *Econometrica* 68(2), S. 407–423.
- De Santis, G. und B. Gerard** (1997): International Asset Pricing and Portfolio Diversification with Time-Varying Risk. *The Journal of Finance* 52(5), S. 1881–1912.
- Dette, H., W. Wu und Z. Zhou** (2015): Change Point Analysis of Second Order Characteristics in Non-Stationary Time Series. *SFB 823 Discussion Paper* 8.
- Ding, Z. und R.F. Engle** (2001): Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing. *Academia Economic Papers* 29(2), S. 157–184.
- Dolde, W. und D. Tirtiroglu** (2002): Housing Price Volatility Changes and Their Effects. *Real Estate Economics* 30(1), S. 41–66.
- Douc, R., E. Moulines und D.S. Stouffer** (2014): *Nonlinear Time Series: Theory, Methods, and Applications with R Examples*. CRC Press, Taylor & Francis Group, Boca Raton, FL, USA.
- Dufour, J.-M. und J.F. Kiviet** (1998): Exact Inference Methods for First-Order Autoregressive Distributed Lag Models. *Econometrica* 66(1), S. 79–104.
- Engle, R.** (1982): Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* 50(4), S. 987–1008.
- (2002): Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroscedasticity Models. *Journal of Business & Economic Statistics* 20(3), S. 339–350.
- Engle, R. und K. Sheppard** (2001): Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH. *NBER Working Paper No. 8554*.
- Engle, R.F. und T. Bollerslev** (1986): Modelling the Persistence of Conditional Variances. *Econometric Reviews* 5(1), S. 1–87.
- Engle, R.F. und K.F. Kroner** (1995): Multivariate Simultaneous Generalized ARCH. *Econometric Theory* 11(1), S. 122–150.
- Engle, R.F. und J. Mezrich** (1996): GARCH for Groups. *Risk Magazine* 9(8), S. 36–40.
- Fahrmeir, L., A. Hamerle und G. Tutz** (1996): *Multivariate Statistische Verfahren*. 2. Aufl. de Gruyter, Berlin.
- Fang, K.-T., S. Kotz und K.W. Ng** (1990): *Symmetric Multivariate and Related Distributions*. Chapman & Hall, New York, NY, USA.
- Fermanian, J.-D. und H. Malongo** (2017): On the Stationarity of Dynamic Conditional Correlation Models. *Econometric Theory* 33(3), S. 636–663.
- Florens, J.-P., V. Marimoutou und A. Péguin-Feissolle** (2007): *Econometric Modeling and Inference*. Cambridge University Press, New York, NY, USA.
- Francq, Ch. und J.-M. Zakoïan** (2010): *GARCH Models: Structure, Statistical Inference and Financial Applications*. John Wiley & Sons Ltd, Chichester, West Sussex, UK.
- Galeano, P. und D. Wied** (2014): Multiple Change Point Detection in the Correlation Structure of Financial Assets. *Computational Statistics and Data Analysis* 76, S. 262–282.

- Gallant, A.R. und H. White** (1988): *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*. Basil Blackwell, Oxford, UK.
- Golosnoy, V.** (2007): Sequential Monitoring of Minimum Variance Portfolio. *ASTA Advances in Statistical Analysis* 91(1), S. 39–55.
- Groen, J.J.J., G. Kapetanios und S. Price** (2013): Multivariate Methods for Monitoring Structural Change. *Journal of Applied Econometrics* 28(2), S. 250–274.
- Hafner, Ch.M.** (2003): Fourth Moment Structure of Multivariate GARCH Models. *Journal of Financial Econometrics* 1(1), S. 26–54.
- Hafner, Ch.M. und H. Herwatz** (2008): Analytical Quasi Maximum Likelihood Inference in Multivariate Volatility Models. *Metrika* 67(2), S. 219–239.
- Hansen, B.E.** (1991): GARCH(1,1) Processes are Near Epoch Dependent. *Economics Letters* 36(2), S. 181–186.
- Hayashi, F.** (2000): *Econometrics*. Princeton University Press, Princeton, NJ, USA.
- He, Ch. und T. Teräsvirta** (2004): An Extended Constant Conditional Correlation GARCH Model and its Fourth-Moment Structure. *Econometric Theory* 20(5), S. 904–926.
- Horn, R.A. und Ch.R. Johnson** (1985): *Matrix Analysis*. Cambridge University Press, Cambridge, UK.
- Horváth, L., M. Hušková, P. Kokoszka und J. Steinebach** (2004): Monitoring Changes in Linear Models. *Journal of Statistical Planning and Inference* 126(1), S. 225–251.
- Horváth, L., P. Kokoszka und A. Zhang** (2006): Monitoring Constancy of Variance in Conditionally Heteroskedastic Time Series. *Econometric Theory* 22(3), S. 373–402.
- Jeantheau, T.** (1998): Strong Consistency of Estimators for Multivariate ARCH Models. *Econometric Theory* 14(1), S. 70–86.
- Joe, H.** (2015): *Dependence Modeling with Copulas*. CRC Press, Taylor & Francis Group, Boca Raton, FL, USA.
- Kang, S.H., H.-G. Cho und S.-M. Yoon** (2009): Modeling Sudden Volatility Changes: Evidence from Japanese and Korean Stock Markets. *Physica A: Statistical Mechanics and its Applications* 388, S. 3543–3550.
- Karolyi, G.A. und R.M. Stulz** (1996): Why do Markets Move Together? An Investigation of U.S.-Japan Stock Return Comovements. *The Journal of Finance* 51(3), S. 951–986.
- Kijima, M.** (1997): *Markov Processes for Stochastic Modeling*. Chapman & Hall, London, UK.
- Klenke, A.** (2006): *Wahrscheinlichkeitstheorie*. Springer, Berlin und Heidelberg.
- Knabner, P. und W. Barth** (2013): *Lineare Algebra: Grundlagen und Anwendungen*. Springer-Verlag, Berlin und Heidelberg.
- Kotz, S. und S. Nadarajah** (2004): *Multivariate t Distributions and Their Applications*. Cambridge University Press, Cambridge, UK.
- Kulperger, R. und H. Yu** (2005): High Moment Partial Sum Processes of Residuals in GARCH Models and Their Applications. *The Annals of Statistics* 33(5), S. 2395–2422.
- Ling, S. und M. McAleer** (2003): Asymptotic Theory for a Vector ARMA-GARCH Model. *Econometric Theory* 19(2), S. 278–308.
- Longin, F. und B. Solnik** (1995): Is the Correlation in International Equity Returns Constant: 1960-1990? *Journal of International Money and Finance* 14(1), S. 3–26.
- Lütkepohl, H.** (1996): *Handbook of Matrices*. John Wiley & Sons, New York, NY, USA.
- Meyn, S. und R.L. Tweedie** (2009): *Markov Chains and Stochastic Stability*. 2. Aufl. Cambridge University Press, New York, NY, USA.
- Nakatani, T. und T. Teräsvirta** (2009): Testing for Volatility Interactions in the Constant Conditional Correlation GARCH Model. *Econometrics Journal* 12(1), S. 147–163.

- Nelson, D.B.** (1991): Conditional Heteroscedasticity in Asset Returns: A New Approach. *Econometrica* 59(2), S. 347–370.
- Nelson, D.B. und C.Q. Cao** (1992): Inequality Constraints in the Univariate GARCH model. *Journal of Business & Economic Statistics* 10(2), S. 229–235.
- Page, E.S.** (1954): Continuous Inspection Schemes. *Biometrika* 41(1/2), S. 100–115.
- Pape, K., D. Wied und P. Galeano** (2016): Monitoring Multivariate Variance Changes. *Journal of Empirical Finance* 39(A), S. 54–68.
- Pape, K., P. Galeano und D. Wied** (2017): Sequential Detection of Parameter Changes in Dynamic Conditional Correlation Models. *SFB 823 Discussion Paper* 7.
- Pedersen, R.S. und A. Rahbek** (2014): Multivariate Variance Targeting in the BEKK-GARCH Model. *Econometrics Journal* 17(1), S. 24–55.
- Robbins, H.** (1970): Statistical Methods Related to the Law of the Iterated Logarithm. *The Annals of Mathematical Statistics* 41(5), S. 1397–1409.
- Sandoval Jr., L. und I. De Paula Franca** (2012): Correlation of Financial Markets in Times of Crisis. *Physica A: Statistical Mechanics and its Applications* 391(1-2), S. 187–208.
- Scherrer, W. und E. Ribarits** (2007): On the Parametrization of Multivariate GARCH Models. *Econometric Theory* 23(3), S. 464–484.
- Schwert, G.W.** (1989): Why Does Stock Market Volatility Change Over Time? *The Journal of Finance* 44(5), S. 1115–1153.
- (2011): Stock Volatility During the Recent Financial Crisis. *European Financial Management* 17(5), S. 789–805.
- Seber, G.A.F.** (2008): *A Matrix Handbook for Statisticians*. John Wiley & Sons, Hoboken, NJ, USA.
- Segen, J. und A. Sanderson** (1980): Detecting Change in a Time-Series. *IEEE Transactions on Information Theory* 26(2), S. 249–254.
- Silvennoinen, A. und T. Teräsvirta** (2016): Testing Constancy of Unconditional Variance in Volatility Models by Misspecification and Specification Tests. *Studies in Non-linear Dynamics and Econometrics* 20(4), S. 347–364.
- Tse, Y.K. und A.K.C. Tsui** (2002): A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-Varying Correlations. *Journal of Business and Economic Statistics* 20(3), S. 351–362.
- Tweedie, R.L.** (1988): Invariant Measures for Markov Chains with no Irreducibility Assumptions. *Journal of Applied Probability* 25, A Celebration of Applied Probability, S. 275–285.
- Wald, A.** (1948): *Sequential Analysis*. John Wiley & Sons, New York, NY, USA.
- White, H.** (1994): *Estimation, Inference and Specification Analysis*. Cambridge University Press, New York.
- Wied, D.** (2015): A Nonparametric Test for a Constant Correlation Matrix. *erscheint in: Econometric Reviews*, DOI:10.1080/07474938.2014.998152.
- Wied, D. und P. Galeano** (2013): Monitoring Correlation Change in a Sequence of Random Variables. *Journal of Statistical Planning and Inference* 143(1), S. 186–196.
- Wied, D., M. Arnold, N. Bissantz und D. Ziggel** (2012a): A New Fluctuation Test for Constant Variances with Applications to Finance. *Metrika* 75(8), S. 1111–1127.
- Wied, D., W. Krämer und H. Dehling** (2012b): Testing for a Change in Correlation at an Unknown Point in Time Using an Extended Functional Delta Method. *Econometric Theory* 28(3), S. 570–589.
- Wilson, B., R. Aggarwal und C. Inclan** (1996): Detecting Volatility Changes Across the Oil Sector. *Journal of Futures Markets* 16(3), S. 313–330.
- Zakoïan, J.-M.** (1994): Threshold Heteroskedastic Models. *Journal of Economic Dynamics and Control* 18(5), S. 931–955.





# Eidesstattliche Versicherung

---

Name, Vorname

Matr.-Nr.

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel

## **Sequenzielle Erkennung von Strukturbrüchen in der Varianzstruktur von multivariaten Zeitreihen**

selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

---

Ort, Datum

Unterschrift

### **Belehrung:**

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00€ geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden (§63 Abs. 5 Hochschulgesetz - HG -).

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird gfls. elektronische Vergleichswerkzeuge (wie z.B. die Software „turnitin“) zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

---

Ort, Datum

Unterschrift