

# Nichtlineare GARCH-Modelle und Volatility Spillover

Modellbildung - Asymptotik - Anwendung

Uwe Wehrspohn

This version: October 10, 2005

Wehrspohn GmbH & Co. KG, Berwanger Strasse 4, 75031 Eppingen, Germany (e-mail: uwe.wehrspohn@wehrspohn.de).

**Abstract:** We extend the class of GARCH models to comprise asymmetric and nonlinear effects on volatility. In particular, we do not only explain future volatility of a time series on its own past, but allow for external influences and spillovers between capital markets. For this generalized class of models, the asymptotic behavior of the Quasi-Maximum-Likelihood estimator of model parameters is derived. The models are applied to time series of fx-rates. It is found that in particular the simple asymmetric models lead to improved performance.

**Key words:** GARCH, heteroskedasticity, volatility spillover

**JEL Classification:** C13, C22, C32

## 1 Einleitung

Es wird allgemein angenommen, daß Investoren bei Geldanlagen, die das Risiko in sich bergen, einen Teil des eingesetzten Kapitals zu verlieren, einen höheren erwarteten Ertrag fordern als bei risikofreien Geldanlagen. Seit Markowitz [1952] und in der Folge auch in der Optionspreistheorie von Black und Scholes [1973] wurde deshalb die Varianz der Erträge eines Wertpapiers als ein Maß für die Volatilität und das Risiko eingeführt. Hierbei wurden u.a. die Annahmen gemacht, daß die Renditen des Wertpapiers bei diskreter Zeit von Periode zu Periode i.i.d. normalverteilt sind, also insbesondere die Varianz konstant ist, bzw. bei kontinuierlicher Zeit einem Random Walk gehorchen.

Ein Vorteil dieses Vorgehens liegt neben der Vereinfachung der den Optionspreis beschreibenden Differentialgleichungen darin, daß aufgrund der Darstellbarkeit

der Normalverteilung als parametrischer Familie in  $\mu$  und  $\sigma^2$  mit den ersten beiden Momenten der Verteilung der Renditen bereits die Verteilung ganz bekannt ist. Konfidenzabschätzungen der Veränderungen des Preises des Wertpapiers in einem gewissen Zeitraum können so einfach vorgenommen und Aussagen über die abstrakte Kennzahl Varianz in Anlagestrategien umgesetzt werden.

Analysen von Daten haben jedoch immer wieder ergeben, daß die Annahmen des Modells von Black und Scholes bei vielen Anwendungen fragwürdig sind, so daß Anlageentscheidungen mit den erzielten Ergebnissen nicht ohne weiteres zuverlässig begründet werden konnten. So sind bei vielen Daten gerade die für Anlageentscheidungen so wichtigen extremen Abweichungen vom Erwartungswert der Renditen wesentlich häufiger als bei der Normalverteilung [vgl. auch Teil 5.4.2 unten]. Auch sind immer wieder Zweifel daran geäußert worden, daß die Verteilung der Renditen um den Erwartungswert symmetrisch ist [Saxinger 1997] oder daß die Renditen stochastisch unabhängig sind in den einzelnen Perioden [Lee und Hansen 1994]. Insbesondere jedoch hat sich gezeigt, daß bei den Zeitreihen oft Phasen geringer Schwankungen mit Phasen hoher Volatilität abwechseln, so daß die Annahme der Homoskedastizität der Verteilungen bestenfalls bei langfristigen Betrachtungen gerechtfertigt ist. Der Anlagehorizont der meisten Investoren ist dagegen eher auf kurze oder mittlere Fristigkeit hin ausgerichtet, so daß u.U. während der gesamten Laufzeit einer Anlage größere Abweichungen zwischen der aktuellen Varianz der Renditen und ihrem langfristigen Mittel auftreten können und hedging-Strategien somit ins Leere laufen.

Engle [1982] hat deshalb mit den ARCH-Prozessen (Auto-Regressive Conditional Heteroskedasticity) eine neue Klasse von Modellen vorgeschlagen, die es erlauben, bei diskreter Zeit langfristige Homoskedastizität und kurzfristige Heteroskedastizität simultan zu erfassen. Es handelt sich dabei um seriell unkorrelierte, bedingt normalverteilte Prozesse mit konstantem Erwartungswert, deren (bedingte) Varianz zum Zeitpunkt  $t$  in Abhängigkeit von den Realisationen des Prozesses zu den Zeitpunkten  $t - 1$  bis  $t - q$  für ein  $q \in \mathbf{N}$  beschrieben wird. Sei konkret  $\mathcal{F}_t$  die Menge der Informationen, die zum Zeitpunkt  $t$  über den Prozeß  $x_t$  zur Verfügung steht. Für den ARCH( $q$ )-Prozess  $x_t$  gelte dann:

$$x_t = \gamma_0 + \epsilon_t = \gamma_0 + z_t \sqrt{h_{0t}} \quad \text{mit } z_t \sim N(0,1) \quad \text{i.i.d.,}$$

wobei  $h_{0t}$  definiert sei als

$$h_{0t} = \omega_0 + \sum_{k=1}^q \alpha_{0k} \epsilon_{t-k}^2,$$

also als lineare Funktion der quadratischen Abweichungen vom Erwartungswert. Insgesamt ist also

$$x_t | \mathcal{F}_{t-1} \sim N(\gamma_0, h_{0t}).$$

Bollerslev [1986] hat diesen Ansatz zu den Generalized ARCH-Prozessen (GARCH) verallgemeinert. Hier wird die Varianz des Prozesses zum Zeitpunkt  $t$  als Funktion der letzten  $q$  Realisationen und der letzten  $p$  bedingten Varianzen dargestellt. GARCH-Prozesse können unter sehr allgemeinen Bedingungen

als ARCH( $\infty$ )-Prozesse interpretiert werden und erlauben somit eine flexiblere Lag-Struktur und einen längeren Einfluß der vergangenen Realisationen. Ein GARCH(p,q)-Prozeß ist definiert wie ein ARCH-Prozeß mit

$$\begin{aligned} h_{0t} &= \omega_0 + \sum_{k=1}^q \alpha_{0k} \epsilon_{t-k}^2 + \sum_{l=1}^p \beta_{0l} h_{0t-l} \\ &= \omega_0 \left( 1 - \sum_{l=1}^p \beta_{0l} \right) + \sum_{j=1}^{\infty} \delta_j \epsilon_{t-j}^2, \end{aligned}$$

wobei sich die  $\delta_j$  in Abhängigkeit von  $\omega_0$ ,  $\alpha_{0k}$  und  $\beta_{0l}$ ,  $k = 1, \dots, q$ ,  $l = 1, \dots, p$ , schreiben lassen.

In der Folge dieser bahnbrechenden Arbeiten von Engle und Bollerslev sind sehr viele Modelle entworfen worden, wie man ARCH-Effekte bei Zeitreihen beschreiben kann. Da die bedingte Varianz der jeweils nächsten Periode in den beiden Modellen von Engle und Bollerslev von den quadratischen Abweichungen der Realisationen vom Erwartungswert ( $\epsilon_t^2$ ) abhängt, haben positive wie negative Schocks gleicher Größe dieselbe Wirkung auf die Varianz der nächsten Periode. Für Investoren haben jedoch negative Schocks u.U. wesentlich gravierendere Konsequenzen als positive. Es liegt deshalb nahe zu vermuten, daß negative Schocks auch eine stärkere "Hebelwirkung" auf die Volatilität haben. Es sind verschiedene Vorschläge gemacht worden, diesen sog. leverage effect, der empirisch bereits 1976 von Black an manchen Zeitreihen nachgewiesen wurde [vgl. Darstellung bei Engle und Ng 1993], in Modifikationen von GARCH-Modellen zu erfassen [Nelson 1990, Glosten et al. 1993, Engle und Ng 1993, u.a.].

Da bei Finanzmarktzeitreihen wie bereits angedeutet viele Regularitätsbedingungen verletzt sind, gestaltete sich die Klärung der Asymptotik der Quasi-Maximum-Likelihood Schätzung (QML) von GARCH-Modellen als schwierig. Zu diesem Thema sind vor allem vier wichtige Arbeiten erschienen. Zunächst zeigte Weiss [1986] für das einfache ARCH-Modell, daß der QML-Schätzer konsistent und asymptotisch normalverteilt ist. Er konnte hierbei sogar die Annahme der bedingten Normalverteilung der Residuen  $\epsilon_t$  vermeiden und durch eine Momentenbedingung ersetzen. Lumsdaine [1996 erschienen] legte 1990 die erste Asymptotik für das GARCH(1,1)-Modell von Bollerslev vor. Sie konnte Prozesse mit unendlichem Erwartungswert zulassen, machte allerdings neben einigen schwächeren technischen Voraussetzungen (s.u. Teil 3) die starke Annahme, die standardisierten Residuen  $z_t$  seien unabhängig von der bedingten Varianz  $h_{0t}$ . Lee und Hansen [1994 erschienen] nahmen 1991 den Ansatz von Lumsdaine auf und konnten ihn so modifizieren, daß nur einige sehr schwache Momentbedingungen an die standardisierten Residuen erforderlich waren. Beide Arbeiten konzentrieren sich auf die Betrachtung des Standard GARCH-Modells. Bollerslev und Wooldridge [1992] versuchten deshalb, die Asymptotik des QML-Schätzers auch für flexiblere Modelle zu zeigen. Sie umgingen allerdings die Probleme des Beweises, indem sie wiederum, z.B. mit der gleichmäßigen stochastischen Konvergenz der Likelihood-Funktion, sehr starke Annahmen machten (siehe Kritik bei Lee und Hansen). Hieraus ergab sich die Schwierigkeit, daß verschiedene Autoren,

die in der Zwischenzeit eigene GARCH-Modelle vorgeschlagen haben, sich in der Frage der Asymptotik auf die Arbeit von Bollerslev und Wooldridge beriefen, ohne zu zeigen, daß ihre Modelle die dort geforderten Voraussetzungen erfüllen (vgl. etwa Glosten et al. [1993]).

Die GARCH-Familie ist insgesamt sehr heterogen, und es ist hier nicht möglich, die ganze Vielfalt auch nur ansatzweise zu beschreiben [vgl. dazu Bera und Higgins 1993, Bollerslev et al. 1992 und 1994, Hentschel 1995, Gouriéroux 1997]. Allen Modellen gemeinsam ist jedoch die Eigenschaft, daß die für die Darstellung der zukünftigen Varianz relevante Informationsmenge allein in der Vergangenheit des Prozesses besteht. Andere Einflüsse auf die Volatilität einer Geldanlage werden grundsätzlich nicht berücksichtigt, obwohl z.B. Korrelationen der Kursentwicklungen von Wertpapieren bereits im Capital Asset Pricing Modell (CAPM) enthalten sind. Ebenso ist bekannt, daß auch Börsenplätze nicht voneinander unabhängig funktionieren, sondern daß es zu volatility spillover Effekten kommen kann, wie Hamao et al. [1990] etwa zwischen verschiedenen internationalen Wertpapiermärkten nachgewiesen haben.

Anknüpfend an die Arbeit von Lee und Hansen [1994] werden wir im folgenden versuchen, diese Lücken zu schließen, indem wir eine um die Möglichkeit externer Einflüsse erweiterte, sehr allgemeine Form des GARCH-Modells formulieren, die es erlaubt, auch asymmetrische und nichtlineare Einflüsse der Schocks auf die zukünftige Varianz zu erfassen.

Die Arbeit ist wie folgt gegliedert: Teil 2 enthält die Darstellung des Modells, Teil 3 eine genaue Beschreibung der getroffenen - und teilweise weniger ökonomisch als mathematisch-technisch begründeten - Voraussetzungen. Die Hauptergebnisse über die asymptotische Verteilung der von uns benutzten Quasi-Maximum Likelihood Parameterschätzer sind in Teil 4 zusammengefaßt. Teil 5 schließlich enthält eine Anwendung der theoretischen Resultate auf die gegen den US-Dollar (US\$) gezeichneten Kassakurse von Kanadischem Dollar (CD) und Britischem Pfund (BP). Das Pfund wird hierbei als eine wichtige europäische Währung interpretiert, in deren Entwicklung sich möglicherweise für die kanadische Wirtschaft relevante europäische wirtschaftliche Vorgänge manifestieren, und in ihrer Wirkung auf die Volatilität des Kanadischen Dollars untersucht. Teil 6 faßt die Ergebnisse zusammen.

## 2 Das Modell

Wir bezeichnen unser Modell als Erweitertes GARCH(p,q,r)-Modell (Extended GARCH bzw. ExGARCH). Seine Beschreibung erfolgt in zwei Schritten. Zunächst definieren wir einen Erweiterten GARCH-Prozeß und nehmen an, daß die beobachteten Daten einem solchen Prozeß gehorchen. Um die unbekannt gebliebenen Größen des Prozesses zu schätzen, formulieren wir danach ein Modell für einen Erweiterten GARCH-Prozeß. Wir gehen also insbesondere davon aus, daß das Modell richtig spezifiziert ist.

Die wesentliche Modifikation des Erweiterten GARCH-Modells gegenüber dem oben beschriebenen Standard GARCH-Modell von Bollerslev besteht darin, daß einerseits die Parameter  $\alpha_{0k}$ ,  $k = 1, \dots, q$ , durch beliebige, als bekannt vorausgesetzte parametrische Funktionen  $g_k(\alpha_{0k}, \epsilon_t)$  ersetzt und zur Erfassung der externen Einflüsse auf die Varianz weitere, ebenfalls als bekannt angenommene parametrische Funktionen  $f_l(\delta_{0l}, r_t)$ ,  $l = 1, \dots, r$ , hinzugefügt werden. In unserem Kontext seien  $\delta_{0l}$  und  $\alpha_{0k}$  nicht nur reelle Parameter, sondern jeweils Vektoren von Parametern.  $r_t$  sei ein beobachtbarer Vektor von exogenen Prozessen.

Der Erweiterte GARCH(p,q,r)-Prozeß läßt sich dann für alle  $t \in \mathbf{Z}$  schreiben als

$$x_t = \gamma_0 + \epsilon_t = \gamma_0 + z_t \sqrt{h_{0t}},$$

wobei

$$h_{0t} := (1 - \beta_0)\omega_0 + \sum_{k=1}^q g_k(\alpha_{0k}, \epsilon_{t-k})\epsilon_{t-k}^2 + \sum_{l=1}^r f_l(\delta_{0l}, r_{t-l}) + \sum_{m=1}^p \beta_{0m} h_{0t-m}.$$

Durch Wahl von  $g_k(\alpha_{0k}, \epsilon_t) = \alpha_{0k} \in \mathbf{R}^1$ ,  $k = 1, \dots, q$ , und  $f_l \equiv 0$ ,  $l = 1, \dots, r$ , erhält man aus dem Erweiterten GARCH-Modell das Standard GARCH-Modell zurück.

Für  $q = 1$ ,  $g(\alpha_0, \epsilon_t) = \alpha_{01} + \alpha_{02} \text{sign}(\epsilon_t)$ , und  $f_l \equiv 0$ ,  $l = 1, \dots, r$ , erhält man ein von Glosten et al. [1993] formuliertes asymmetrisches GARCH-Modell, auf das wir uns als GJR-GARCH beziehen. Weitere Variationen dieses Themas sind denkbar.

Wie bereits angedeutet, wollen wir die technisch begründete Annahme der Unabhängigkeit und Normalverteilung der Residuen  $\epsilon_t$  vermeiden. Sei deswegen  $F_t$  die gemeinsame Verteilung von  $(z_t, \epsilon_t, r_t, z_{t-1}, \epsilon_{t-1}, r_{t-1}, \dots)$  und  $\mathcal{F}_t := \sigma(z_t, \epsilon_t, r_t, z_{t-1}, \epsilon_{t-1}, r_{t-1}, \dots)$  die von der Vergangenheit des Prozesses  $(z_t, \epsilon_t, r_t)_{t \in \mathbf{Z}}$  erzeugte  $\sigma$ -Algebra. "Fast sicher" stehe grundsätzlich für " $F_t$ -fast sicher." Es gelte:

$$E(z_t | \mathcal{F}_{t-1}) = 0 \quad \text{und} \quad E(z_t^2 | \mathcal{F}_{t-1}) = 1 \quad \text{fast sicher}$$

oder äquivalent

$$E(\epsilon_t | \mathcal{F}_{t-1}) = 0 \quad \text{und} \quad E(\epsilon_t^2 | \mathcal{F}_{t-1}) = h_{0t} \quad \text{fast sicher.}$$

Wenn wir davon ausgehen, daß die Beobachtungen  $x_t$  die Struktur eines Erweiterten GARCH-Prozesses haben und daß die Funktionen  $f_l$  und  $g_k$  bekannt sind, reduziert sich das Problem, die Varianz des Prozesses  $x_t$  zu schätzen, darauf, den unbekannt Parametervektor  $\theta_0 := (\alpha_{01}, \dots, \alpha_{0q}, \beta_0, \gamma_0, \delta_{01}, \dots, \delta_{0r}, \omega_0)'$  zu schätzen. Die Varianz der jeweils nächsten Periode läßt sich dann mit den o.g. Formeln aus den Beobachtungen berechnen.

Um die formalen Komplikationen so gering wie möglich zu halten, nehmen wir im folgenden an, daß die Beobachtungen einem Erweiterten GARCH(1,1,1)-Prozeß

gehorchen.  $\theta_0$  hat dann die Form  $\theta_0 := (\alpha_0, \beta_0, \gamma_0, \delta_0, \omega_0)'$ . Wir schreiben deshalb das Modell für den Prozeß  $x_t$  in Abhängigkeit von dem Parametervektor  $\theta := (\alpha, \beta, \gamma, \delta, \omega)'$  als

$$x_t = \gamma + e_t(\theta),$$

wobei  $e_t(\theta)$  als Differenz von  $x_t$  und  $\gamma$  definiert ist. Da wir den Prozeß  $x_t$  nicht für alle  $t \in \mathbf{Z}$ , sondern o.E. nur für  $t = 1, \dots, n$ ,  $n \in \mathbf{N}$ , beobachten können und da die bedingte Varianz  $h_{0t}$  des Prozesses eine rekursive Form hat, müssen wir eine Anfangsbedingung annehmen, um  $h_{0t}$  modellieren zu können. Die von dieser Anfangsbedingung abhängigen empirischen Prozesse kennzeichnen wir mit \*. Sei also analog zur Form von  $h_{0t}$

$$h_t^*(\theta) := (1 - \beta)\omega + g(\alpha, e_{t-1}(\theta))e_{t-1}^2(\theta) + f(\delta, r_{t-1}) + \beta h_{t-1}^*(\theta)$$

das Modell für  $h_{0t}$  und  $h_1^*(\theta) := \omega$  die Anfangsbedingung. Dann läßt sich  $h_t^*(\theta)$  darstellen als

$$h_t^*(\theta) = \omega + \sum_{k=0}^{t-2} \beta^k [g(\alpha, e_{t-k-1}(\theta))e_{t-k-1}^2(\theta) + f(\delta, r_{t-k-1})].$$

Um den Parametervektor  $\theta_0$  zu schätzen, benutzen wir das Quasi-Maximum-Likelihood Verfahren (QML). Wegen

$$z_t = \frac{\epsilon_t}{\sqrt{h_{0t}}}, \quad E(z_t | \mathcal{F}_{t-1}) = 0 \quad \text{und} \quad E(z_t^2 | \mathcal{F}_{t-1}) = 1$$

definieren wir mit  $(X_n, R_n) := (x_n, r_n, \dots, x_1, r_1)$  die empirische Quasi-log-Likelihood Funktion  $L_n^*(\theta)$  als:

$$\begin{aligned} L_n^*(\theta)(X_n, R_n) &:= \frac{2}{n} \ln \left( \prod_{t=1}^n \frac{1}{\sqrt{2\pi h_t^*(\theta)}} \exp\left(-\frac{e_t^2(\theta)}{2h_t^*(\theta)}\right) \right) - 2 \ln\left(\frac{1}{\sqrt{2\pi}}\right) \\ &= \frac{2}{n} \sum_{t=1}^n \left( \ln\left(\frac{1}{\sqrt{2\pi}}\right) + \frac{1}{2} \ln\left(\frac{1}{h_t^*(\theta)}\right) - \frac{e_t^2(\theta)}{2h_t^*(\theta)} \right) - 2 \ln\left(\frac{1}{\sqrt{2\pi}}\right) \\ &= \frac{1}{n} \sum_{t=1}^n l_t^*(\theta), \end{aligned}$$

wobei

$$l_t^*(\theta) := - \left( \ln h_t^*(\theta) + \frac{e_t^2(\theta)}{h_t^*(\theta)} \right).$$

Für  $z_t$  iid  $\sim N(0, 1)$  wäre dies bis auf den Faktor  $\frac{1}{2}$  die richtige log-Likelihood Funktion. Wir bezeichnen  $L_n^*(\theta)$  in unserem Kontext als "Quasi"-log-Likelihood Funktion, da wir auf die Annahme der iid Normalverteilung der  $z_t$  verzichten.

### 3 Technische Voraussetzungen

Die folgenden Annahmen ergeben sich im wesentlichen nicht aus ökonomischen Überlegungen, sondern vielmehr aus technischen Notwendigkeiten im Zusam-

menhang mit der eingeschlagenen Beweisstrategie für die in Teil 4 unten zusammengefaßten Theoreme 1- 3. Es sei jedoch angemerkt, daß diese Voraussetzungen vom ökonomischen Standpunkt aus gesehen sehr schwach sind und den Kreis der enthaltenen GARCH-Prozesse gegenüber bisherigen Ansätzen stark erweitern.

Für unseren Beweis der Konsistenz des QML-Schätzers ist es entscheidend, daß der wahre Parametervektor  $\theta_0$  in einem kompakten Parameterraum  $\Theta$  liegt.

- Seien  $l_1, l_2 \in \mathbf{N}$ ,  $\Theta_\alpha \subset \mathbf{R}^{l_1}$  kompakt und  $\Theta_\delta \subset \mathbf{R}^{l_2}$  kompakt. Dann sei  $\Theta \subset \mathbf{R}^{l_1+l_2+3}$  definiert als kompakter Parameterraum

$$\Theta := \{ \theta = (\alpha, \beta, \gamma, \delta, \omega)' : \alpha \in \Theta_\alpha; \beta \in [\beta_l, \beta_u] \subset (0, 1) \subset \mathbf{R}^1; \\ \gamma \in [\gamma_l, \gamma_u] \subset \mathbf{R}^1; \delta \in \Theta_\delta; \omega \in [\omega_l, \omega_u] \subset \mathbf{R}^+ \setminus \{0\} \}.$$

Um die Likelihood-Funktion und ihre Ableitungen abschätzen zu können, machen wir folgende Annahmen an die Funktionen  $g$  und  $f$  :

- Sei  $g(\alpha, y), y \in \mathbf{R}$ , eine parametrische Funktion mit den Eigenschaften:

- $g(\alpha, y)$  ist dreimal partiell differenzierbar bezüglich der Koeffizienten von  $\alpha$  und bezüglich  $y$ .

Wird der Erwartungswert der Beobachtungen als identisch 0 angenommen, hat der Parametervektor  $\theta$  wegen  $x_t = \epsilon_t = e_t(\theta)$  keinen Einfluß auf die zweite Variable von  $g(\alpha, \epsilon_t)$ . Es reicht deshalb in diesem Fall, alternativ die dreimalige partielle Differenzierbarkeit von  $g$  nach den Koeffizienten von  $\alpha$  und die Meßbarkeit in  $y$  von  $g$  und der ersten beiden partiellen Ableitungen vorauszusetzen (vgl. Lemma 12 (1) im Appendix).

- Es existieren positive Konstanten  $\tau_1, \tau_2$ , so daß für alle  $\alpha \in \Theta_\alpha$  und alle  $y \in \mathbf{R}$  gilt:

$$0 < \tau_1 \leq g(\alpha, y) \leq \tau_2.$$

- Es existiert ein  $\tau_3 > 0$ , so daß alle partiellen Ableitungen erster, zweiter und dritter Ordnung von  $g(\alpha, y)$  für alle  $\alpha \in \Theta_\alpha$  und alle  $y \in \mathbf{R}$  betragsmäßig kleiner sind als  $\tau_3$ .

- Sei  $0 \leq f(\delta, y), y \in \mathbf{R}^m$ , eine parametrische Funktion mit den Eigenschaften:

- Für festes  $y$  ist  $f(\delta, y)$  dreimal partiell differenzierbar in den Koeffizienten von  $\delta$ .  $f$  selbst und die ersten beiden partiellen Ableitungen nach den Koeffizienten von  $\delta$  sind meßbar in  $y$ .

- Es existiert ein  $S \in \mathbf{R}$ , so daß  $\frac{f(\delta_1, y)}{f(\delta_2, y)} \leq S$  für alle  $\delta_1, \delta_2 \in \Theta_\delta$  und  $y \in \mathbf{R}^m$ .

Um die rekursiv definierte bedingte Varianz  $h_{0t}$  in einer Rekursionsformel zusammenfassend darstellen zu können, machen wir folgende Voraussetzung:

- Es existiert eine Verteilung  $P_{0-\infty}$  und eine Zufallsvariable  $h_{0-\infty} := \lim_{t \rightarrow -\infty} h_{0t}$  mit  $\mathcal{L}(h_{0-\infty}) = P_{0-\infty}$  und

$$P_{0-\infty}\{h_{0-\infty} \in (0, \infty)\} = 1.$$

Das heißt, der Startwert des Prozesses ( $h_{0t}$ ) ist fast sicher strikt positiv und endlich.

Beim Beweis der Konsistenzaussagen in den folgenden Theoremen spielt die Anwendbarkeit eines schwachen Gesetzes der großen Zahl eine entscheidende Rolle. Bei den meisten Aufsätzen zu ARCH- und GARCH-Modellen wird deshalb vorausgesetzt, daß es sich bei dem Prozeß ( $z_t$ ) um einen iid-Prozeß handelt [vgl. z.B. Engle 1982, Bollerslev 1986, Weiss 1986, Liu und Brorsen 1995, Lumsdaine 1996 und andere]. Es ist jedoch nicht klar, ob diese Annahme bei ökonomischen Zeitreihen erfüllt ist. Lee und Hansen [1994] haben schließlich gezeigt, daß die Unabhängigkeit des generierenden Prozesses nicht wesentlich ist für das asymptotische Verhalten des QML-Schätzers, und sie durch die schwächere Annahme der Ergodizität ersetzt. Wir übertragen diesen allgemeineren Ansatz in unseren Kontext:

- $(z_t, \epsilon_t, r_t)$  ist strikt stationär und ergodisch.

Je nach Wahl der konkreten Funktionen  $f$  und  $g$  kann diese Annahme u.U. weiter abgeschwächt werden, etwa wenn sich der Prozess der  $(\epsilon_t)$  vollständig als Funktion des Prozesses  $(z_t, r_t)$  darstellen läßt.

Den Einfluß der exogenen Prozesse auf die Volatilität des endogenen Prozesses präzisiert folgende Annahme:

- Es existiert ein  $K < \infty$ , so daß für  $1 \leq q \leq 6$  die  $q$ -ten Momente von  $f(\delta, r_t)$  und der ersten drei partiellen Ableitungen von  $f$  nach den Koeffizienten von  $\delta$  kleinergleich  $K^q$  sind.

Eine konkrete Funktion  $f$  vor Augen wäre es u.U. möglich, diese Momentbedingungen abzuschwächen. Es erscheint jedoch auch ökonomisch als durchaus sinnvoll anzunehmen, daß die exogenen Prozesse nur einen in gewissem Sinne beschränkten Einfluß auf die Volatilität haben.

Die folgende Annahme ist im wesentlichen verantwortlich dafür, welche Restriktionen an den Parameterraum  $\Theta$  sich ergeben werden, um die Likelihood-Funktion und ihre Ableitungen in der Norm abzuschätzen zu können (siehe unten

und Lemma 7). Wir beweisen im Anhang eine Abwandlung von Theorem 1, die eine alternative Annahme macht.

- Für ein  $\nu > 0$  existiert ein  $S_\nu < \infty$ , so daß

$$E(z_t^{2+\nu} | \mathcal{F}_{t-1}) \leq S_\nu < \infty \quad \text{fast sicher.}$$

Lumsdaine [1996] etwa nimmt an dieser Stelle an, der Prozeß  $(z_t)$  sei iid mit einer symmetrischen, unimodalen Dichte, endlichem 32. Moment und Unabhängigkeit vom Prozeß der  $(h_{0t})$ . Diese starken Voraussetzungen ermöglichen es ihr, zusätzliche Parameterrestriktionen zu vermeiden.

Um zu gewährleisten, daß  $h_{0t}$  für alle  $t \in \mathbf{Z}$  fast sicher endlich ist, machen wir die Voraussetzung

- $\sup_{t \in \mathbf{Z}} E(\ln(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2) | \mathcal{F}_{t-1}) < 0$  fast sicher.

Wie aus dem Beweis von Lemma 1 im Anhang hervorgeht, gewährleistet diese Annahme die Gültigkeit einer hinreichenden und notwendigen Bedingung für die fast sichere Endlichkeit von  $h_{0t}$ . Ebenso ermöglicht sie in Lemma 5, die Existenz eines positiven Momentes von  $h_{0t}$  zu zeigen.

Im Verlauf der Argumentation im Hinblick auf Konsistenz und asymptotische Normalität des QML-Schätzers wird die Anwendbarkeit des Ergodentheorems eine ausschlaggebende Rolle spielen. Aufgrund der Anfangsbedingung  $h_1^*(\theta) = \omega$  ist der Prozeß  $h_t^*(\theta)$  - und damit  $L_n^*(\theta)$  - aber auch unter Annahme der strikten Stationarität von  $(z_t, r_t, \epsilon_t)$  nicht strikt stationär. Wir definieren deshalb die beiden hypothetischen Prozesse

$$h_t(\theta) := \omega + \sum_{k=0}^{\infty} \beta^k [g(\alpha, e_{t-k-1}(\theta))e_{t-k-1}^2(\theta) + f(\delta, r_{t-k-1})]$$

und

$$h_t^\epsilon(\theta) := \omega + \sum_{k=0}^{\infty} \beta^k [g(\alpha, \epsilon_{t-k-1})\epsilon_{t-k-1}^2 + f(\delta, r_{t-k-1})]$$

mit der hypothetischen Quasi-log-Likelihood Funktion

$$L_n(\theta) := \frac{1}{n} \sum_{t=1}^n l_t(\theta)$$

mit

$$l_t(\theta) := - \left( \ln h_t(\theta) + \frac{e_t^2(\theta)}{h_t(\theta)} \right),$$

die - wie sich in Lemma 3 zeigen wird - unter den gemachten Voraussetzungen strikt stationär und ergodisch sind.

Um die Eindeutigkeit der Parameterschätzung und die Existenz der Kovarianzmatrix des Parameterschätzers sicherzustellen, nehmen wir an,

- daß für alle  $\theta \neq \theta_0 \in \Theta$  gilt:

$$h_t(\theta) \neq h_{0t} \quad \text{fast sicher}$$

(vgl. Lemma 13). Da, wie bereits angedeutet, der Prozeß  $h_t(\theta)$  strikt stationär ist, gilt die Annahme für alle  $t \in \mathbf{Z}$ .

- daß für alle  $t \in \mathbf{Z}$  mit

$$x_t := \left(1, \frac{\partial g}{\partial \alpha_1}(\alpha_0, \epsilon_t)\epsilon_t^2, \dots, \frac{\partial g}{\partial \alpha_{l_1}}(\alpha_0, \epsilon_t)\epsilon_t^2, \frac{\partial g}{\partial y}(\alpha_0, \epsilon_t)\epsilon_t^2, \right. \\ \left. g(\alpha_0, \epsilon_t)\epsilon_t, \frac{\partial f}{\partial \delta_1}(\delta_0, r_t), \dots, \frac{\partial f}{\partial \delta_{l_2}}(\delta_0, r_t), f(\delta_0, r_t)\right)'$$

gilt

$$\det Cov(x_t, x_t) > 0.$$

In den Beweisen der folgenden Theoreme wird ein wichtiges Argument darin bestehen, daß die Likelihood-Funktion  $L_n(\theta)$  und ihre partiellen Ableitungen in der  $q$ -Norm endlich sind ( $\|f\|_q := (E|f|^q)^{1/q}$ ) für  $1 \leq q \leq 6$ . Wie die Beweise zeigen werden, ist eine hinreichende Bedingung hierfür die folgende Beschränkung des Parameterraums, die wir in Analogie zu Lee und Hansen [1994] formulieren.

Sei

$$\mathcal{R}_l := 1 - \frac{\tau_1}{(2T_1 + \tau_1)2^{\frac{2+\nu}{\nu}} S_\nu^{2/\nu}} < 1 \quad \text{und} \quad \mathcal{R}_u := 1 - \frac{\tau_1}{(2 + \tau_1)2^{\frac{2+\nu}{\nu}} S_\nu^{2/\nu}} < 1$$

mit  $T_1 := \frac{\omega_u}{\omega_0} + \frac{T_2}{\tau_1} + S$  (vgl. Lemma 8(3)). Wähle dazu möglichst kleine Konstanten  $\eta_l, \eta_u > 0$  mit der Eigenschaft:

$$\eta_l < \beta_0(1 - \mathcal{R}_l^{\frac{1}{6}}) \quad \text{und} \quad \eta_u < \beta_0(1 - \mathcal{R}_u^{\frac{1}{6}}),$$

und definiere für  $1 \leq q \leq 6$  Konstanten

$$\beta_{ql} := \beta_0 \mathcal{R}_l^{\frac{1}{q}} + \eta_l < \beta_0,$$

$$\beta_{qu} := \frac{\beta_0 - \eta_u}{\mathcal{R}_u^{\frac{1}{q}}} > \beta_0.$$

$\beta_{ql}$  und  $\beta_{qu}$  sind insbesondere so gewählt worden, daß

$$\frac{\beta_0 \mathcal{R}_l^{\frac{1}{q}}}{\beta_{ql}} < 1 \quad \text{und} \quad \frac{\beta_{qu} \mathcal{R}_u^{\frac{1}{q}}}{\beta_0} < 1,$$

um zu gewährleisten, daß geometrische Reihen mit diesen Gewichten konvergieren (vgl. die Beweise von Lemma 12 (2b) und Lemma 10 (1)).

Definiere schließlich die Parameterräume

$$\Theta_l^q := \{\theta \in \Theta : \beta_{ql} \leq \beta_0\},$$

$$\Theta_u^q := \{\theta \in \Theta : \beta_0 \leq \beta_{qu}\}$$

sowie

$$\Theta_q := \Theta_l^q \cup \Theta_u^q \subset \Theta.$$

Beachte, daß die  $\Theta_q$  mit wachsendem  $q$  eine absteigende Folge von Mengen bilden, da  $\beta_{ql}$  streng monoton wachsend und  $\beta_{qu}$  streng monoton fallend in  $q$  ist. Die Parameterrestriktion kann insgesamt umso schwächer gehalten werden, je kleiner  $\eta_l, \eta_u > 0$  gewählt werden.

Wir definieren noch folgende Schreibweise:

$\nabla^i l_t(\theta)$ ,  $i = 1, 2, 3$ , bezeichne die  $i$ -te partielle Ableitung der Funktion  $l_t(\theta)$  nach den Koeffizienten von  $\theta$ .

## 4 Konsistenz und Asymptotische Normalität des Quasi-Maximum-Likelihood-Schätzers

Wir können nun die wesentlichen Resultate anführen:

**Theorem 1** *Unter den gemachten Voraussetzungen gilt für jedes  $\hat{\theta}_n \in C_n := \arg \max_{\theta \in \Theta_2} L_n^*(\theta)$ :*

$$\hat{\theta}_n \rightarrow_p \theta_0$$

**Theorem 2** *Unter den gemachten Voraussetzungen und den zusätzlichen Annahmen, daß  $E(z_t^4 | \mathcal{F}_{t-1}) \leq \mathcal{K} < \infty$  und daß  $\theta_0$  im Inneren von  $\Theta_6$  liegt, gilt:*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow_{\mathcal{L}} N(0, V_0),$$

wobei  $V_0 := B_0^{-1} A_0 B_0^{-1}$ ,  $A_0 = E(\nabla l_t(\theta_0) \nabla l_t(\theta_0)')$  und  $B_0 := B(\theta_0) = -E \nabla^2 l_t(\theta_0)$ .

Bemerkung: Mit der zusätzlichen Annahme, daß die standardisierten Residuen  $z_t$  bedingt normalverteilt sind,

$$z_t | \mathcal{F}_{t-1} \sim N(0, 1),$$

gilt unter den Voraussetzungen von Theorem 2:

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow_{\mathcal{L}} N(0, B_0^{-1}) = N(0, A_0^{-1})$$

(vgl. Lemma 16 im Appendix).

Die unbekannte Kovarianzmatrix  $V_0$  kann mit Hilfe von

$$B_n^*(\theta) := -\frac{1}{n} \sum_{t=1}^n \nabla^2 l_t^*(\theta),$$

$$A_n^*(\theta) := \frac{1}{n} \sum_{t=1}^n \nabla l_t^*(\theta) \nabla l_t^*(\theta)'$$

geschätzt werden.

**Theorem 3** *Unter den Voraussetzungen von Theorem 2 ist die empirische Kovarianzmatrix*

$$V_n^* := B_n^*(\hat{\theta}_n)^{-1} A_n^*(\hat{\theta}_n) B_n^*(\hat{\theta}_n)^{-1}$$

*ein konsistenter Schätzer von  $V_0$ .*

Die Beweise befinden sich im Appendix.

## 5 Umsetzung der Ergebnisse: Wechselkurse des Kanadischen Dollars

Um zu zeigen, daß die Erweiterung der GARCH-Familie um externe Einflüsse auf die Volatilität auch praktische Relevanz haben kann und ggf. Prädiktionen verbessert, wenden wir verschiedene GARCH-Modelle mit und ohne Erweiterung auf die täglichen Veränderungen des Wechselkurses des Kanadischen Dollars (CD) unter Einfluß des Britischen Pfundes (BP) an.

Als Gegenüber zur GARCH-Familie betrachten wir ein einfaches stochastisch-homoskedastisches Modell.

Alle Modelle werden gefittet und Out of Sample-Prognosen durchgeführt.

### 5.1 Verschiedene GARCH-Modelle

Wir betrachten folgende acht Modelle:

- Das GARCH-Modell von Bollerslev [1986] ist bereits angeführt worden. Als GARCH(1,1)-Modell hat es die Form:

$$x_t = \gamma_0 + \epsilon_t$$

$$h_{0t} = \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 h_{0t-1}.$$

Die gemachten Parameterrestriktionen erfordern, daß  $\omega_0, \alpha_0, \beta_0 > 0$ .

Wir erweitern dieses Modell auf zwei Arten. Wir beschreiben zunächst die täglichen Veränderungen des BP-Kurses ( $y_t$ ) durch ein GARCH(1,1)-Modell wie oben:

$$\begin{aligned} y_t &= \tilde{\gamma}_0 + \eta_t \\ s_{0t} &= \tilde{\omega}_0 + \tilde{\alpha}_0 \eta_{t-1}^2 + \tilde{\beta}_0 s_{0t-1}. \end{aligned}$$

- Das erste Modell für die täglichen Schwankungen des CD betrachtet die Entwicklung des BP bis zum Vortag als Einflußfaktor auf die aktuelle Volatilität des CD. Die gesamte Vergangenheit des BP wird zusammengefaßt in der Volatilität  $s_{0t}$  des Pfundes von heute:

$$\begin{aligned} x_t &= \gamma_0 + \epsilon_t \\ h_{0t} &= \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 h_{0t-1} + \delta_0 s_{0t}. \end{aligned}$$

Dieses Modell hat den Vorteil, von Anlegern ohne weiteres benutzt werden zu können, da alle notwendigen Informationen jeweils vor Beginn des Börsentages zur Verfügung stehen. Wir bezeichnen es als ExGARCH(VT) für Extended GARCH mit Einflüssen vom Vortag.

- Da sich Anpassungsprozesse an den internationalen Finanzplätzen sehr schnell vollziehen, ist nicht ohne weiteres klar, ob die Entwicklung des Pfundes bis zum Vortag als Information über Vorgänge in der deutschen bzw. europäischen Wirtschaft nicht jeweils schon veraltet ist. Aus theoretischen Gründen modellieren wir deshalb:

$$\begin{aligned} x_t &= \gamma_0 + \epsilon_t \\ h_{0t} &= \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 h_{0t-1} + \delta_0 s_{0t+1}, \end{aligned}$$

mit der aktuellen Entwicklung des BP als Einfluß auf den CD zusammengefaßt in der Größe  $s_{0t+1}$ . Wir bezeichnen dieses Modell als ExGARCH(ST) für Extended GARCH mit Einflüssen vom selben Tag. In beiden Modellen muß  $\delta_0 > 0$  sein.

Den drei Modellen ist gemeinsam, daß sie symmetrisch sind in der Richtung der Schocks. Ein  $\epsilon_t > 0$  hat dieselbe Wirkung auf die Varianz der nächsten Periode wie ein  $\epsilon_t < 0$  desselben Betrages. Der leverage effect kann also von diesen Modellen nicht erfaßt werden. Wir betrachten deshalb zusätzlich

- das von Glosten, Jagannathan und Runkle [1993] vorgeschlagene asymmetrische Modell

$$\begin{aligned} x_t &= \gamma_0 + \epsilon_t \\ h_{0t} &= \omega_0 + (\alpha_{01} + \alpha_{02} \text{sign}(\epsilon_{t-1})) \epsilon_{t-1}^2 + \beta_0 h_{0t-1}, \end{aligned}$$

das wir als GJR-GARCH(1,1)-Modell bezeichnen. Für  $\epsilon_t > 0$  nimmt die Funktion  $g(\alpha_{01}, \alpha_{02}, \epsilon_t) := \alpha_{01} + \alpha_{02} \text{sign}(\epsilon_t)$  den Wert  $\alpha_{01} + \alpha_{02}$ , für  $\epsilon_t < 0$

den Wert  $\alpha_{01} - \alpha_{02}$  an. Für  $\alpha_{02} \neq 0$  ist also die Darstellung eines leverage effects möglich. Die Parameterrestriktion erfordert  $|\alpha_{01}| > |\alpha_{02}|$  und  $\alpha_{01} > 0$ .

Wie oben verallgemeinern wir dieses Modell, indem wir ein GJR-GARCH(1,1)-Modell auf die BP-Zeitreihe anwenden:

$$\begin{aligned} y_t &= \tilde{\gamma}_0 + \eta_t \\ s_{0t} &= \tilde{\omega}_0 + (\tilde{\alpha}_{01} + \tilde{\alpha}_{02} \text{sign}(\eta_{t-1}))\eta_{t-1}^2 + \tilde{\beta}_0 s_{0t-1}, \end{aligned}$$

und die Volatilität  $s_{0t}$  des BP als relevante Information für die Entwicklung des CD geltend machen. Wir erhalten

- das Ex-GJR-GARCH(VT)-Modell [Extended GARCH mit Einflüssen vom Vortag]

$$\begin{aligned} x_t &= \gamma_0 + \epsilon_t \\ h_{0t} &= \omega_0 + (\alpha_{01} + \alpha_{02} \text{sign}(\epsilon_{t-1}))\epsilon_{t-1}^2 + \beta_0 h_{0t-1} + \delta_0 s_{0t}, \end{aligned}$$

- und das Ex-GJR-GARCH(ST)-Modell, in dem die Entwicklung des BP bis zum selben Tag zu Buche schlägt:

$$\begin{aligned} x_t &= \gamma_0 + \epsilon_t \\ h_{0t} &= \omega_0 + (\alpha_{01} + \alpha_{02} \text{sign}(\epsilon_{t-1}))\epsilon_{t-1}^2 + \beta_0 h_{0t-1} + \delta_0 s_{0t+1}. \end{aligned}$$

Um die Positivität der Varianz  $h_{0t}$  zu gewährleisten, muß wieder  $\delta_0 > 0$  erfüllt sein.

Da  $g(\alpha_{01}, \alpha_{02}, x) := \alpha_{01} + \alpha_{02} \text{sign}(x)$  nicht differenzierbar in  $x$  ist, erfüllen die GJR-GARCH-Modelle die im ersten Teil der Arbeit gemachten Voraussetzungen an sich nur für  $\gamma \equiv 0$ . Diese Annahme wäre ökonomisch durchaus gerechtfertigt, da der Erwartungswertparameter  $\gamma$  in allen Fällen als praktisch 0 und auch nicht immer signifikant geschätzt wurde. Um die Vergleichbarkeit mit den anderen Modellen aber zu gewährleisten und eben weil  $\gamma$  als so klein geschätzt wurde, haben wir den Parameter in den GJR-Modellen und aus denselben Gründen im noch folgenden PGARCH-Modell unten belassen.

Bei diesen Modellen hat die parametrische Funktion  $g$ , die bestimmt, wie stark die quadrierten Residuen  $\epsilon_t$  die Varianz der nächsten Periode beeinflussen, jeweils eine sehr einfache Form. Bei den drei GARCH-Modellen handelt es sich um die konstante Funktion  $g(\alpha_0, \epsilon_t) = \alpha_0$ , bei den GJR-GARCH-Modellen nimmt  $g$  nur zwei Werte an, wie oben bereits dargestellt. Mit der erwähnten Einschränkung bieten die an  $g$  gemachten Voraussetzungen jedoch wesentlich größere Spielräume zur Gestaltung dieser Funktion. Exemplarisch modellieren wir deshalb in Anlehnung an das partiell nichtparametrische Modell (PNP) von Engle und Ng [1993]  $g$  nicht als stückweise konstante, sondern als stückweise lineare Funktion, deren Knickpunkte wir vorgeben.

- Sei zunächst  $\tilde{g}$  definiert als

$$\begin{aligned}\tilde{g}(\alpha_0, \epsilon_t) &:= \alpha_{01} + \alpha_{02}(\epsilon_t + 0, 004)1_{\{\epsilon_t \leq -0,004\}} \\ &\quad + \alpha_{03}(\epsilon_t + 0, 002)1_{\{\epsilon_t \leq -0,002\}} \\ &\quad + \alpha_{04}\epsilon_t 1_{\{\epsilon_t \leq 0\}} + \alpha_{05}\epsilon_t 1_{\{\epsilon_t \geq 0\}} \\ &\quad + \alpha_{06}(\epsilon_t - 0, 002)1_{\{\epsilon_t \geq 0,002\}} \\ &\quad + \alpha_{07}(\epsilon_t - 0, 004)1_{\{\epsilon_t \geq 0,004\}}\end{aligned}$$

und ausgehend davon  $g$  als

$$g(\alpha_0, \epsilon_t) := \begin{cases} \tilde{g}(\alpha_0, -0, 02) & \text{für } \epsilon_t < -0, 02 \\ \tilde{g}(\alpha_0, \epsilon_t) & \text{für } -0, 02 \leq \epsilon_t \leq 0, 02 \\ \tilde{g}(\alpha_0, 0, 02) & \text{für } 0, 02 < \epsilon_t. \end{cases}$$

Die Knickpunkte sind hierbei suggestiv gewählt worden als näherungsweise 5, 20, 50, 80 und 95%-Quantile der täglichen Schwankungen des CD. Es liegen keine Beobachtungen außerhalb von  $[-0, 02, +0, 02]$ . Man beachte, daß  $g$  nach oben beschränkt ist.

Wir erhalten damit das vollständige Modell als

$$\begin{aligned}x_t &= \gamma_0 + \epsilon_t \\ h_{0t} &= \omega_0 + g(\alpha_0, \epsilon_{t-1}) \epsilon_{t-1}^2 + \beta_0 h_{0t-1}.\end{aligned}$$

Da  $g$  linear ist, ist  $g(\alpha_0, \epsilon_t) \epsilon_t^2$  ein Polynom 3. Grades in  $\epsilon_t$ . Wir bezeichnen deshalb dieses Modell als Power GARCH-Modell (PGARCH).

Schließlich betrachten wir folgendes bedingt wie unbedingt homoskedastische Modell, das wir als

- Statisches Modell bezeichnen:

$$x_t = \gamma_0 + \epsilon_t$$

mit

$$E(\epsilon_t) = 0 \quad \text{und} \quad E(\epsilon_t^2) = \sigma^2.$$

## 5.2 Beschreibung der Daten

Die Ergebnisse der Auswertung basieren auf den im Zeitraum vom 01.06.1973 bis zum 13.11.1986 börsentäglich erhobenen Kassakursen der US-Dollarnotierungen von Britischem Pfund und Kanadischem Dollar. Die genannte Zeitspanne umfaßt insgesamt 3511 Datensätze. Sie wurde ausgewählt, da sie weitgehend frei ist von Strukturbrüchen und BP und CD sich sehr unterschiedlich gegenüber dem US-Dollar entwickeln (vgl. Abb.1). Wir lassen uns von der Hypothese leiten, daß in den Schwankungen des Pfundes genuin europäische Vorgänge zum

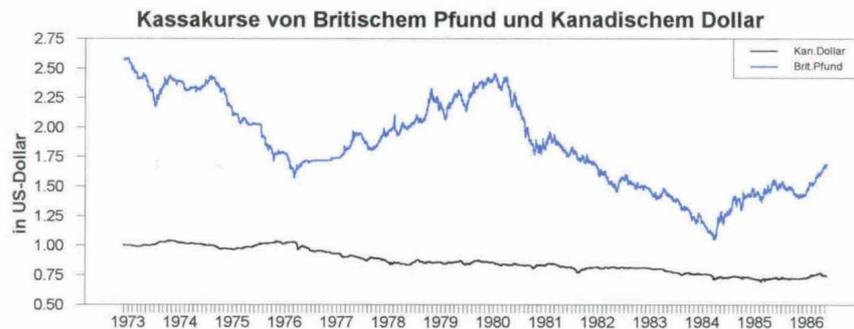


Abbildung 1: Kassakurse von Britischem Pfund und Kanadischem Dollar

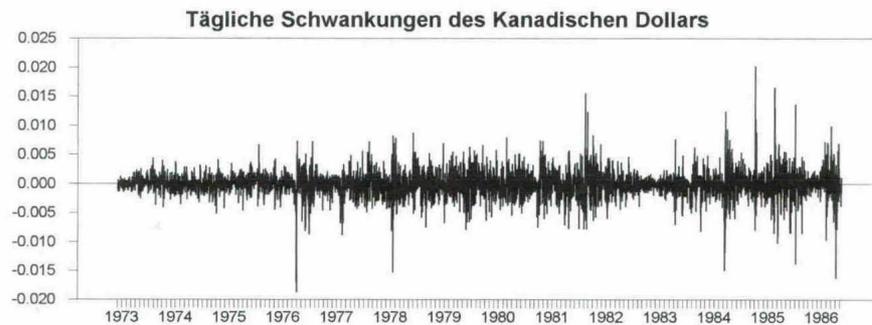


Abbildung 2: Tägliche Schwankungen des Kanadischen Dollars

Ausdruck kommen, die aufgrund internationaler Verbindungen - Europa ist Kanadas zweitwichtigster Wirtschaftspartner - in der Entwicklung des Kanadischen Dollars Spuren hinterlassen und einen Teil zur Erklärung der Varianz des CD beitragen können.

Da die durch GARCH-Modelle dargestellten Zeitreihen stationär sein müssen, betrachten wir nicht die Wechselkurse selbst, sondern ihre täglichen relativen Änderungen. Der vorherrschenden Praxis in der Literatur folgend, approximieren wir die relativen Änderungen durch die Differenz der logarithmierten Kassakurse und erhalten

$$x_t := \ln X_{t+1} - \ln X_t \approx \frac{X_{t+1} - X_t}{X_t}$$

(vgl. Tschernig 1994, Kaiser 1996 u.v.a.). Abbildung 2 zeigt die transformierte Zeitreihe für den Kanadischen Dollar. Das Phänomen des "volatility clustering", d.h. des geballten Auftretens von Perioden großer bzw. kleiner Kursschwankungen, kann gut beobachtet werden.

### 5.3 Ergebnisse der Parameterschätzungen

Sämtliche Schätzungen wurden mit dem Programm *Regression Analysis for Time Series*, Version 4.30, durchgeführt. Die Likelihood-Funktionen wurden mit dem Simplex-Algorithmus maximiert. Der Simplex-Algorithmus setzt nicht die Differenzierbarkeit der Likelihood-Funktion voraus. Über den geschätzten Parameterwerten sind jeweils die approximativen Standardabweichungen des Schätzwertes in Klammern angegeben.

#### GARCH-Modell:

$$\begin{aligned} x_t &= -7,66 \cdot 10^{-5} + \hat{\epsilon}_t \\ \hat{h}_{0t} &= 2,13 \cdot 10^{-6} + 0,348 \hat{\epsilon}_{t-1}^2 + 0,635 \hat{h}_{0t-1} \end{aligned}$$

#### ExGARCH(VT):

Für das BP erhält man:

$$\begin{aligned} y_t &= -6,96 \cdot 10^{-5} + \hat{\eta}_t \\ \hat{s}_{0t} &= 1,59 \cdot 10^{-5} + 0,301 \hat{\eta}_{t-1}^2 + 0,692 \hat{s}_{0t-1} \end{aligned}$$

Für den CD ergibt sich damit

$$\begin{aligned} x_t &= -7,72 \cdot 10^{-5} + \hat{\epsilon}_t \\ \hat{h}_{0t} &= 2,09 \cdot 10^{-6} + 0,347 \hat{\epsilon}_{t-1}^2 + 0,601 \hat{h}_{0t-1} + 0,0253 \hat{s}_{0t} \end{aligned}$$

#### ExGARCH(ST):

Mit dem bereits genannten Resultat für das Pfund ergibt sich für den CD weiter:

$$\begin{aligned} x_t &= -8,62 \cdot 10^{-5} + \hat{\epsilon}_t \\ \hat{h}_{0t} &= 1,25 \cdot 10^{-6} + 0,320 \hat{\epsilon}_{t-1}^2 + 0,584 \hat{h}_{0t-1} + 0,0284 \hat{s}_{0t+1} \end{aligned}$$

#### GJR-GARCH:

$$\begin{aligned} x_t &= -1,01 \cdot 10^{-4} + \hat{\epsilon}_t \\ \hat{h}_{0t} &= 1,83 \cdot 10^{-7} + (0,182 - 0,0951 \text{sign}(\hat{\epsilon}_{t-1})) \hat{\epsilon}_{t-1}^2 + 0,699 \hat{h}_{0t-1} \end{aligned}$$

**Ex-GJR-GARCH(VT):**

Für das BP:

$$\begin{aligned}
y_t &= -1,81 \cdot 10^{-4} + \hat{\eta}_t \\
\hat{s}_{0t} &= 6,71 \cdot 10^{-7} + (0,1449 - 0,1342 \operatorname{sign}(\hat{\eta}_{t-1})) \hat{\eta}_{t-1}^2 + 0,82 \hat{s}_{0t-1}
\end{aligned}$$

Für den CD erhält man:

$$\begin{aligned}
x_t &= -9,2 \cdot 10^{-5} + \hat{\epsilon}_t \\
\hat{h}_{0t} &= 1,29 \cdot 10^{-7} + (0,1127 - 0,0921 \operatorname{sign}(\hat{\epsilon}_{t-1})) \hat{\epsilon}_{t-1}^2 + 0,662 \hat{h}_{0t-1} \\
&\quad + 0,00617 \hat{s}_{0t}
\end{aligned}$$

**Ex-GJR-GARCH(ST):**

Für den CD erhält man wiederum mit dem o.g. Ergebnis für das BP:

$$\begin{aligned}
x_t &= -1,01 \cdot 10^{-4} + \hat{\epsilon}_t \\
\hat{h}_{0t} &= 9,93 \cdot 10^{-8} + (0,1296 - 0,0963 \operatorname{sign}(\hat{\epsilon}_{t-1})) \hat{\epsilon}_{t-1}^2 + 0,674 \hat{h}_{0t-1} \\
&\quad + 0,0124 \hat{s}_{0t+1}
\end{aligned}$$

**PGARCH:**

$$\begin{aligned}
\tilde{g}(\hat{\alpha}_0, \hat{\epsilon}_t) &= 0,03498 - 7,151 (\hat{\epsilon}_t + 0,004) 1_{\{\hat{\epsilon}_t \leq -0,004\}} \\
&\quad - 6,682 (\hat{\epsilon}_t + 0,002) 1_{\{\hat{\epsilon}_t \leq -0,002\}} \\
&\quad - 5,572 \hat{\epsilon}_t 1_{\{\hat{\epsilon}_t \leq 0\}} - 1,141 \hat{\epsilon}_t 1_{\{\hat{\epsilon}_t \geq 0\}} \\
&\quad - 0,013 (\hat{\epsilon}_t - 0,002) 1_{\{\hat{\epsilon}_t \geq 0,002\}} \\
&\quad + 4,62 (\hat{\epsilon}_t - 0,004) 1_{\{\hat{\epsilon}_t \geq 0,004\}} \\
g(\hat{\alpha}_0, \hat{\epsilon}_t) &:= \begin{cases} \tilde{g}(\hat{\alpha}_0, -0,02) & \text{für } \hat{\epsilon}_t < -0,02 \\ \tilde{g}(\hat{\alpha}_0, \hat{\epsilon}_t) & \text{für } -0,02 \leq \hat{\epsilon}_t \leq 0,02 \\ \tilde{g}(\hat{\alpha}_0, 0,02) & \text{für } 0,02 < \hat{\epsilon}_t. \end{cases}
\end{aligned}$$

$$\begin{aligned}
x_t &= -7,6 \cdot 10^{-5} + \hat{\epsilon}_t \\
\hat{h}_{0t} &= 4,57 \cdot 10^{-5} + g(\hat{\alpha}_0, \hat{\epsilon}_{t-1}) \hat{\epsilon}_{t-1}^2 + 0,462 \hat{h}_{0t-1}.
\end{aligned}$$

### Statisches Modell:

$$x_t = -8,6 \cdot 10^{-5} + \hat{\epsilon}_t$$

mit

$$\text{Var}(\hat{\epsilon}_t) = \hat{\sigma}^2 = 5,7 \cdot 10^{-6}.$$

Die Parameter des Statischen Modells sind durch Stichprobenmittel und Stichprobenvarianz geschätzt worden.

Ein Vergleich zeigt unmittelbar, daß die geforderten Parameterrestriktionen in allen Modellen sowohl für die CD- als auch für die BP-Zeitreihe erfüllt sind.

In allen GJR-GARCH-Modellen wurde der den leverage effect abbildende Parameter  $\alpha_{02}$  signifikant negativ geschätzt. Abbildung 3 zeigt die News Impact Curve (NIC) des GJR-GARCH-Modells [vgl. Engle und Ng 1993], d.h. den Graphen der Funktion  $(\hat{\alpha}_{01} + \hat{\alpha}_{02} \text{sign}(\epsilon)) \epsilon^2$ . Engle und Ng haben den Namen "News Impact Curve" gewählt, da sie die Wirkung von Schocks, also neuen Informationen, auf die Volatilität der nächsten Periode darstellt. Die geschätzten Parameterwerte ergeben eine deutliche Asymmetrie der NIC um den Nullpunkt. Negative Abweichungen der Daten von ihrem Erwartungswert beeinflussen die zukünftige Varianz deutlich stärker, es scheint also ein leverage effect tatsächlich vorzuliegen. Die Konfidenzbänder ergeben sich aus gemeinsamen Konfidenzintervallen für die Parametersummen  $\alpha_{01} + \alpha_{02}$  und  $\alpha_{01} - \alpha_{02}$  durch Multiplikation mit  $\epsilon^2$ . Das dargestellte Intervall  $[-0,006; +0,006]$  umfaßt ca. 97% der Beobachtungen.

Das PGARCH-Modell kommt hinsichtlich des leverage effectes zu einem vergleichbaren Ergebnis wie die GJR-Modelle. Abbildung 4 und 5 zeigen die gefittete stückweise lineare Funktion  $g$  dieses Modells bzw. die sich daraus ergebende News Impact Curve. Hier haben offensichtlich negative Schocks einen größeren Einfluß auf die Volatilität als positive. Gleichzeitig wird die Asymmetrie der Wirkung noch subtiler dargestellt als in den GJR-Modellen, da nicht nur die bloße Richtung der Schocks über den Wert der Funktion  $g$  entscheidet, sondern auch ihr Betrag. Die News Impact Curve ist im PGARCH-Modell etwas symmetrischer als in den GJR-Modellen und somit der leverage effect etwas schwächer ausgeprägt. Bemerkenswerterweise nimmt  $g$  sein Minimum nicht im Nullpunkt, sondern rechts davon in  $\epsilon = 0,004$  an. Die Konfidenzbänder sind nach der Methode von Scheffé bestimmt worden. Die empirische Kovarianzmatrix hatte vollen Rang.

Der den Einfluß des Pfundes auf die Volatilität des CD ausdrückende Parameter  $\delta_0$  wird in allen erweiterten Modellen signifikant geschätzt. Er ist jedoch grundsätzlich von viel kleinerer Größenordnung als die Parameter  $\alpha_0$  und  $\beta_0$ , durch die die vergangene Entwicklung des CD in seine Varianz einfließt. Dies scheint für diese Anwendung die im theoretischen Teil der Arbeit gemachte Annahme zu bestätigen, daß die externen Einflüsse auf die Volatilität relativ klein sind in Verhältnis zur eigenen Dynamik der Zeitreihe.

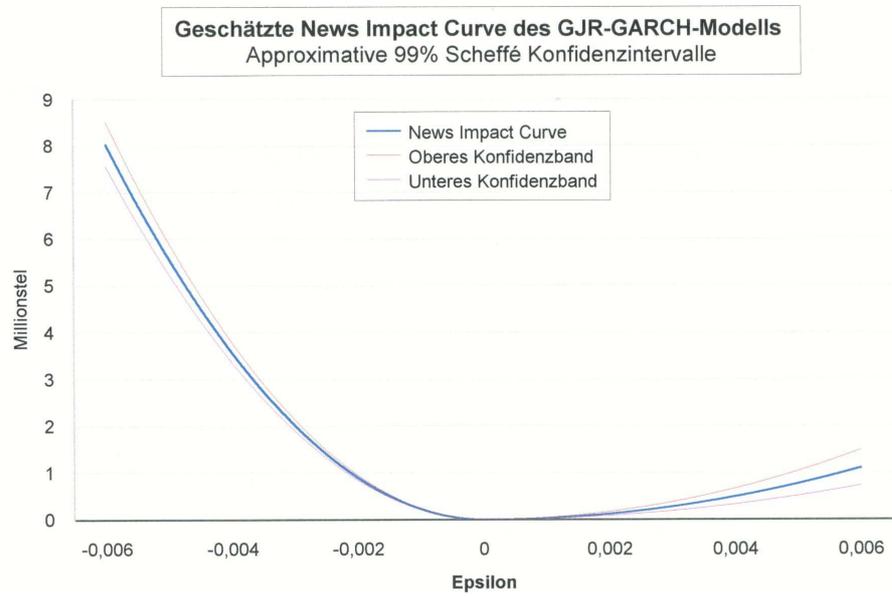


Abbildung 3: News Impact Curve des GJR-GARCH-Modells

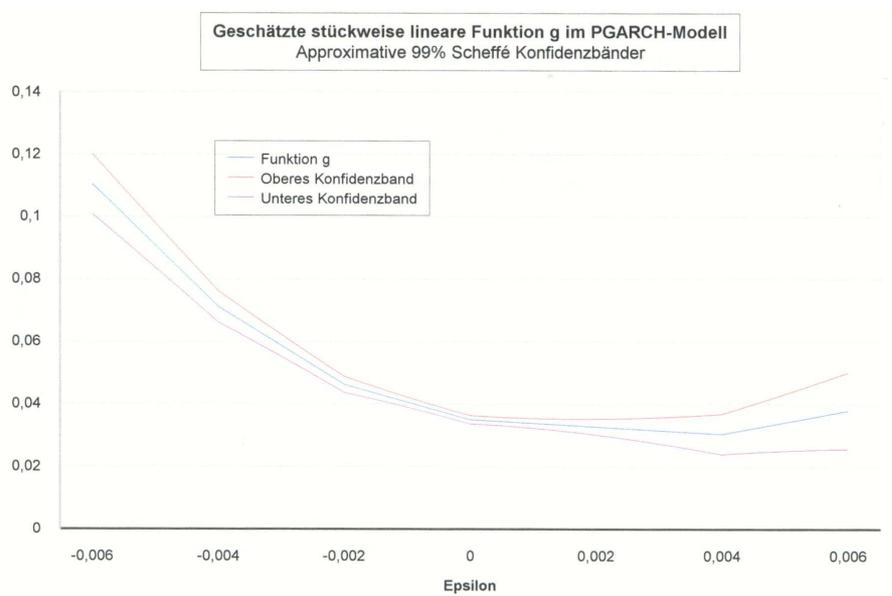


Abbildung 4: Stückweise lineare Funktion g im PGARCH-Modell

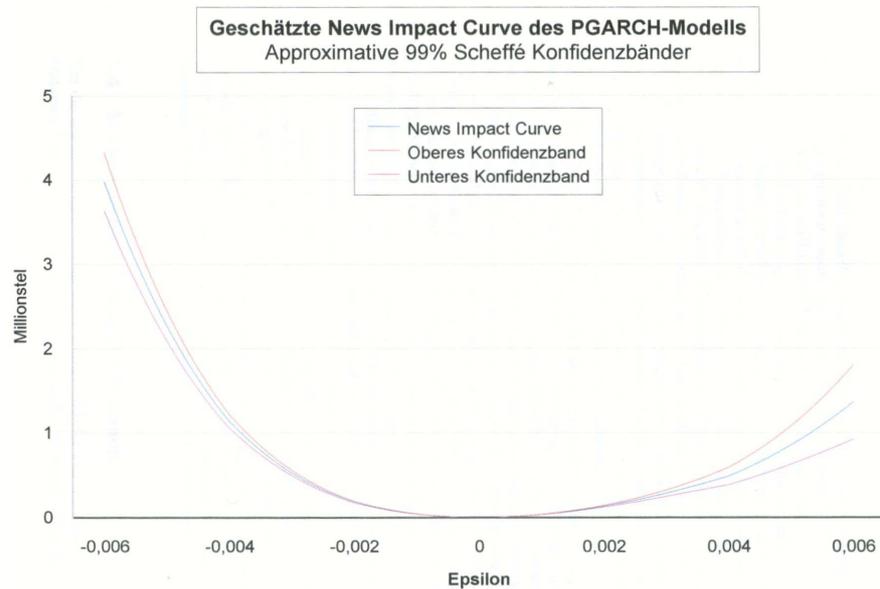


Abbildung 5: News Impact Curve des PGARCH-Modells

## 5.4 Modellvergleich durch Prädiktionen

Um die zukünftige Entwicklung einer Anlage und der von ihr abgeleiteten Optionen besser abschätzen zu können, sind möglichst präzise Vorhersagen der Varianzen ihrer Preisschwankungen wünschenswert. Wir machen deshalb mit den oben angeführten Modellen Out of Sample Prognosen, um das den Daten in dieser Hinsicht am besten angepaßte Modell zu lokalisieren.

Bei den “Out of Sample“ Prognosen sind die Parameter der einzelnen Modelle mit den ersten 1500 Beobachtungen der Stichprobe geschätzt und dann Ein-Schritt- bzw. Fünf-Schritt-Prädiktionen der Volatilität durchgeführt worden, das heißt, von den Zeitpunkten  $T = 1500, T + 1, T + 2, \dots$  aus sind jeweils die bedingten Varianzen des nächsten Tages bzw. von  $T = 1500, T + 5, T + 10, \dots$  aus die bedingten Varianzen jeweils für die folgenden 5 Tage vorhergesagt worden.

### 5.4.1 Prädiktions- und Bewertungsmodelle

Da aufgrund der den GARCH-Modellen immanenten Rekursionsstruktur jeweils die Beobachtung der Vorperiode zur Bestimmung der Varianz eingeht, müssen Ein-Schritt- und Mehr-Schritt-Prädiktionen mit unterschiedlichen Formeln berechnet werden. Wir tragen bei der Bestimmung der Formeln der Tatsache Rechnung, daß  $\epsilon_t^2$  sich schreiben läßt als  $\epsilon_t^2 = z_t^2 h_{0t}$  und daß  $Ez_t^2 = 1$ .

In allen Modellen ist  $\hat{\epsilon}_t = x_t - \hat{\gamma}_0$ .

**GARCH-Modell:**

Ein-Schritt-Prädiktion [ESP]:

$$\hat{h}_{0t+1} = \hat{\omega}_0 + \hat{\alpha}_0 \hat{\epsilon}_t^2 + \hat{\beta}_0 \hat{h}_{0t}$$

Mehr-Schritt-Prädiktion ( $k > 1$ ) [MSP]:

$$\hat{h}_{0t+k} = \hat{\omega}_0 + (\hat{\alpha}_0 + \hat{\beta}_0) \hat{h}_{0t+k-1}.$$

**ExGARCH(VT):**

ESP:

$$\hat{h}_{0t+1} = \hat{\omega}_0 + \hat{\alpha}_0 \hat{\epsilon}_t^2 + \hat{\beta}_0 \hat{h}_{0t} + \hat{\delta}_0 \hat{s}_{0t+1}$$

MSP:

$$\hat{h}_{0t+k} = \hat{\omega}_0 + (\hat{\alpha}_0 + \hat{\beta}_0) \hat{h}_{0t+k-1} + \hat{\delta}_0 \hat{s}_{0t+k},$$

wobei  $\hat{s}_{0t+k-1}$  mit dem o.g. Ansatz berechnet wurde.**ExGARCH(ST):**

ESP:

$$\hat{h}_{0t+1} = \hat{\omega}_0 + \hat{\alpha}_0 \hat{\epsilon}_t^2 + \hat{\beta}_0 \hat{h}_{0t} + \hat{\delta}_0 \hat{s}_{0t+2}$$

MSP:

$$\hat{h}_{0t+k} = \hat{\omega}_0 + (\hat{\alpha}_0 + \hat{\beta}_0) \hat{h}_{0t+k-1} + \hat{\delta}_0 \hat{s}_{0t+k+1}.$$

**GJR-GARCH:**

Bei den GJR-Modellen ergibt sich das Problem, daß bei Prädiktionen über mehr als einen Schritt das Vorzeichen des Schocks nicht bekannt ist. Wir machen deshalb hier die zusätzliche Annahme, daß  $\epsilon_t$  symmetrisch um den Nullpunkt verteilt ist.

ESP:

$$\hat{h}_{0t+1} = \hat{\omega}_0 + (\hat{\alpha}_{01} + \hat{\alpha}_{02} \text{sign}(\hat{\epsilon}_t)) \hat{\epsilon}_t^2 + \hat{\beta}_0 \hat{h}_{0t}$$

MSP:

$$\hat{h}_{0t+k} = \hat{\omega}_0 + (\hat{\alpha}_{01} + \hat{\beta}_0) \hat{h}_{0t+k-1}.$$

**Ex-GJR-GARCH(VT):**

ESP:

$$\hat{h}_{0t+1} = \hat{\omega}_0 + (\hat{\alpha}_{01} + \hat{\alpha}_{02} \text{sign}(\hat{\epsilon}_t)) \hat{\epsilon}_t^2 + \hat{\beta}_0 \hat{h}_{0t} + \hat{\delta}_0 \hat{s}_{0t+1}$$

MSP:

$$\hat{h}_{0t+k} = \hat{\omega}_0 + (\hat{\alpha}_{01} + \hat{\beta}_0) \hat{h}_{0t+k-1} + \hat{\delta}_0 \hat{s}_{0t+k}.$$

**Ex-GJR-GARCH(ST):**

ESP:

$$\hat{h}_{0t+1} = \hat{\omega}_0 + (\hat{\alpha}_{01} + \hat{\alpha}_{02} \text{sign}(\hat{\epsilon}_t)) \hat{\epsilon}_t^2 + \hat{\beta}_0 \hat{h}_{0t} + \hat{\delta}_0 \hat{s}_{0t+2}$$

MSP:

$$\hat{h}_{0t+k} = \hat{\omega}_0 + (\hat{\alpha}_{01} + \hat{\beta}_0) \hat{h}_{0t+k-1} + \hat{\delta}_0 \hat{s}_{0t+k+1}.$$

**PGARCH:**

Um mit dem PGARCH-Modell Mehr-Schritt-Prädiktionen durchführen zu können, wären noch weitere Annahmen an die Verteilung der  $\epsilon_t$  nötig, da neben dem Vorzeichen der Schocks auch noch ihr konkreter Wert zu Buche schlägt. Wir beschränken uns deshalb bei diesem Modell auf die Ein-Schritt-Prädiktion.

ESP:

$$\begin{aligned} \tilde{g}(\hat{\alpha}_0, \hat{\epsilon}_t) &= \hat{\alpha}_{01} + \hat{\alpha}_{02}(\hat{\epsilon}_t + 0,004)1_{\{\hat{\epsilon}_t \leq -0,004\}} \\ &\quad + \hat{\alpha}_{03}(\hat{\epsilon}_t + 0,002)1_{\{\hat{\epsilon}_t \leq -0,002\}} \\ &\quad + \hat{\alpha}_{04}\hat{\epsilon}_t 1_{\{\hat{\epsilon}_t \leq 0\}} + \hat{\alpha}_{05}\hat{\epsilon}_t 1_{\{\hat{\epsilon}_t \geq 0\}} \\ &\quad + \hat{\alpha}_{06}(\hat{\epsilon}_t - 0,002)1_{\{\hat{\epsilon}_t \geq 0,002\}} \\ &\quad + \hat{\alpha}_{07}(\hat{\epsilon}_t - 0,004)1_{\{\hat{\epsilon}_t \geq 0,004\}} \\ g(\hat{\alpha}_0, \hat{\epsilon}_t) &:= \begin{cases} \tilde{g}(\hat{\alpha}_0, -0,02) & \text{für } \hat{\epsilon}_t < -0,02 \\ \tilde{g}(\hat{\alpha}_0, \hat{\epsilon}_t) & \text{für } -0,02 \leq \hat{\epsilon}_t \leq 0,02 \\ \tilde{g}(\hat{\alpha}_0, 0,02) & \text{für } 0,02 < \hat{\epsilon}_t. \end{cases} \\ \hat{h}_{0t+1} &= \hat{\omega}_0 + g(\hat{\alpha}_0, \hat{\epsilon}_t) \hat{\epsilon}_t^2 + \hat{\beta}_0 \hat{h}_{0t} \end{aligned}$$

**Statisches Modell:**

$$\hat{\sigma}_{t+1}^2 = \hat{\sigma}_{t+k}^2 = \hat{\sigma}^2,$$

wobei  $\hat{\sigma}^2$  wieder durch die jeweilige Stichprobenvarianz geschätzt wird.

Als ‘wahre Varianz’ der Zeitreihe in jeder Periode, die es vorherzusagen gilt, haben wir die beobachtete Varianz  $\tilde{h}_{0t} := (x_t - \bar{x})^2$  gewählt (Abb.6), wenn  $\bar{x}$  das Mittel der gesamten Stichprobe bezeichnet. Diese Größe hat den Vorteil, daß sie unabhängig ist von dem Modell, mit dessen Prädiktionen sie verglichen wird.

In der Literatur werden vielfältige Möglichkeiten vorgeschlagen, die Präzision der jeweiligen Prognosen zu messen [vgl. Pagan und Schwert 1990, Hentschel 1995, Kaiser 1996, u.a.]. Wir messen sie anhand von zwei Kennzahlen:

- Der Wurzel der mittleren quadratischen Abweichung der prognostizierten Varianz von der beobachteten Varianz (RMSE = Root Mean Squared Error):

$$\text{RMSE}(\hat{h}_0) := \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{h}_{0t} - \tilde{h}_{0t})^2},$$

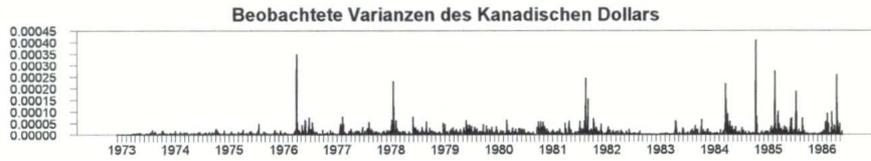


Abbildung 6: Beobachtete Varianzen des Kanadischen Dollars

wenn insgesamt  $N$  Werte vorhergesagt wurden. Aufgrund der Quadrierung werden große Abweichungen bei diesem Mittelwert stärker gewichtet als kleine. Der RMSE ist deshalb insbesondere ein Kriterium, das angibt, wie gut Perioden großer Volatilität von den Prognosemodellen erfaßt werden.

- Dem Median der quadratischen Abweichungen (MedSE = Median Squared Error):

$$\text{MedSE}(\hat{h}_0) := \text{Median}(\hat{h}_{0t} - \tilde{h}_{0t})^2.$$

Da der Median lediglich zählt, wieviele Realisationen links und rechts von einem Wert liegen, kommen bei diesem Kriterium kleine Abweichungen von der beobachteten Varianz stärker zur Geltung als bei dem RMSE. Der MedSE indiziert also stärker, wie gut Perioden kleiner Volatilität von den Prognosemodellen dargestellt werden.

#### 5.4.2 Ergebnisse der Prädiktionen

Das Vorgehen bei den Out of Sample Prädiktionen ist bereits beschrieben worden. Insgesamt sind vom Zeitpunkt  $T = 1500$  an 2010 Tage bzw. 402 Wochen zu je 5 Tagen vorhergesagt worden. Die einzelnen Modelle haben folgende Ergebnisse geliefert:

Tabelle 1: Ergebnisse der 1-Schritt-Out of Sample-Prädiktion

Modell	RMSE [ $\cdot 10^{-5}$ ]	Rang	MedSE [ $\cdot 10^{-11}$ ]	Rang
GARCH	1,796	5	0,557	6
GJR-GARCH	1,806	6	0,389	2
Stat.Modell	1,955	8	1,303	8
ExGARCH(ST)	1,762	3	0,629	7
ExGARCH(VT)	1,777	4	0,525	5
Ex-GJR-GARCH(ST)	1,736	1	0,404	3
Ex-GJR-GARCH(VT)	1,744	2	0,496	4
PGARCH	1,942	7	0,368	1

Tabelle 2: Ergebnisse der 5-Schritt-Out of Sample-Prädiktion

Modell	RMSE [ $\cdot 10^{-5}$ ]	Rang	MedSE [ $\cdot 10^{-11}$ ]	Rang
GARCH	1,911	2	1,296	6
GJR-GARCH	1,923	6	1,153	3
Stat.Modell	1,955	7	1,302	7
ExGARCH(ST)	1,912	4	1,266	5
ExGARCH(VT)	1,919	5	1,190	4
Ex-GJR-GARCH(ST)	1,899	1	1,048	2
Ex-GJR-GARCH(VT)	1,904	2	0,949	1

Zunächst fällt auf, daß das Statische Modell bei beiden Prädiktionen und in beiden Kriterien schlechter abschneidet als die Modelle der GARCH-Familie. Dieses Resultat ist zu erwarten gewesen, da das Statische Modell den volatility clustern nicht folgen kann, sondern zwischen Phasen großer und Phasen kleiner Schwankungen ausmitteln muß. Es weicht deshalb in seiner Prädiktion sowohl bei großen Schocks als auch bei kleinen Schocks mehr von der beobachteten Varianz ab als die Modelle der GARCH-Familie, da es große Schocks zu stark unter- und kleine Schocks zu stark überschätzt. Da das statische Modell homoskedastisch ist und da wir nur einmal mit den ersten 1500 Beobachtungen die Parameter geschätzt haben, sind die Prädiktionsergebnisse für alle Zeithorizonte identisch. Die GARCH-Modelle sind demgegenüber wegen ihrer rekursiven Struktur, die die Kenntnis der Realisation des Vortages erfordert, auf die 1-Schritt-Prognose spezialisiert. Es ist deshalb nicht erstaunlich, daß der Abstand zwischen GARCH-Modellen und Statischem Modell bei der 5-Schritt-Prädiktion wesentlich kleiner ausfällt als bei der 1-Schritt-Prognose.

Innerhalb der Gruppe der heteroskedastischen Modelle schneiden die asymmetrischen GARCH-Modelle beim Median der quadratischen Abweichungen (MedSE) deutlich besser ab als die symmetrischen GARCH-Modelle. Bei Betrachtung der mittleren quadratischen Abweichungen (RMSE) liegen symmetrische und asymmetrische Modelle dichter zusammen. Die Gruppe der symmetrischen GARCH-Modelle scheint vor allem in Phasen größerer Schwankungen die beobachteten Varianzen gut vorherzusagen, während sie in ruhigeren Zeiten die Varianz überschätzt.

Von den betrachteten Modellen passen sich die GJR-GARCH-Modelle den qualitativen Veränderungen der Varianzen - vermutlich wegen ihrer Fähigkeit, den leverage effect darzustellen - am flexibelsten an. Überparametrisierung scheint hier noch kein gravierendes Problem zu sein.

Das Verhalten des PGARCH-Modells ist bei beiden Kriterien sehr unterschiedlich. Daß es beim RMSE so schlecht abschneidet im Vergleich mit den anderen GARCH-Modellen, deutet darauf hin, daß es gerade Phasen großer Volatilität

schlecht modelliert. So wird etwa die in Abbildung 5 oben dargestellte News Impact Curve des PGARCH-Modells gerade im negativen Bereich deutlich von der NIC des GJR-GARCH-Modells (Abb. 3) majorisiert. Auch der "Erinnerungsparameter"  $\beta$  ist vergleichsweise niedrig geschätzt worden, so daß größere Varianzen nicht sehr persistent sind. Aus demselben Grund schneidet das PGARCH-Modell beim Median der quadratischen Abweichungen besser ab, da auch die beobachteten Varianzen in den meisten Fällen sehr klein sind, da der Kanadische Dollar mit dem US-Dollar stark korreliert ist. Insgesamt ist das PGARCH-Modell von den heteroskedastischen Modellen das unflexibelste.

Die Erweiterung der Modelle um Einflüsse des Pfundes auf die Schwankungen des CD macht sich insbesondere in Zusammenhang mit der Möglichkeit, Asymmetrie zu erfassen, bemerkbar. Sowohl in der 1-Schritt- wie in der 5-Schritt-Prädiktion liegen die erweiterten GJR-Modelle bei beiden Kriterien vorn, während das einfache GJR-Modell bei den großen Abweichungen der Prognose von der beobachteten Varianz betonenden RMSE in beiden Zeithorizonten von den symmetrischen Modellen dominiert wird.

Das symmetrische GARCH-Modell profitiert ebenso von der Erweiterung, aber längst nicht so ausgeprägt. Alle symmetrischen Modelle sind durchweg auf den hinteren Rängen plazierte. Eine qualitative Verbesserung wie beim GJR-Modell wird hier nicht erzielt.

Die Frage, ob erweiterten Modellen mit Einflüssen vom selben Tag oder mit Einflüssen vom Vortag der Vorzug zu geben ist, läßt sich mit dem hier erzielten Befund nicht eindeutig beantworten. Tendenziell erzielen die Modelle mit Einflüssen vom selben Tag etwas bessere Ergebnisse bei der Prognose der zukünftigen Varianz. Dies bestätigt die Vermutung, daß sich die Reaktionen an den Kapitalmärkten so schnell vollziehen, daß Informationen vom Vortag u.U. schon veraltet sind.

Insgesamt sind die beiden Ex-GJR-GARCH-Modelle aus dem Kreis der von uns betrachteten Modelle den Daten am besten angepaßt. Sie erfassen sowohl die Asymmetrie in der Wirkung von positiven bzw. negativen Schocks auf die Varianz als auch die Einflüsse der europäischen Währung auf die Volatilität des Kanadischen Dollars.

Abbildung 7 zeigt die beobachteten Varianzen des Kanadischen Dollars zusammen mit dem jeweiligen Fit von symmetrischen und asymmetrischen GARCH-Modellen. Der Vergleich der täglichen Schwankungen des CD mit dem Fit des GJR-GARCH-Modells zeigt deutlich, daß hier negative Schocks eine größere Wirkung auf die geschätzte Varianz haben als positive (siehe insbesondere den Schock Ende 1984).

## 6 Zusammenfassung

Ausgehend vom Standard GARCH-Modell von Bollerslev haben wir eine weitere Verallgemeinerung dieses Modells vorgeschlagen. Einerseits läßt sie vielfältigere

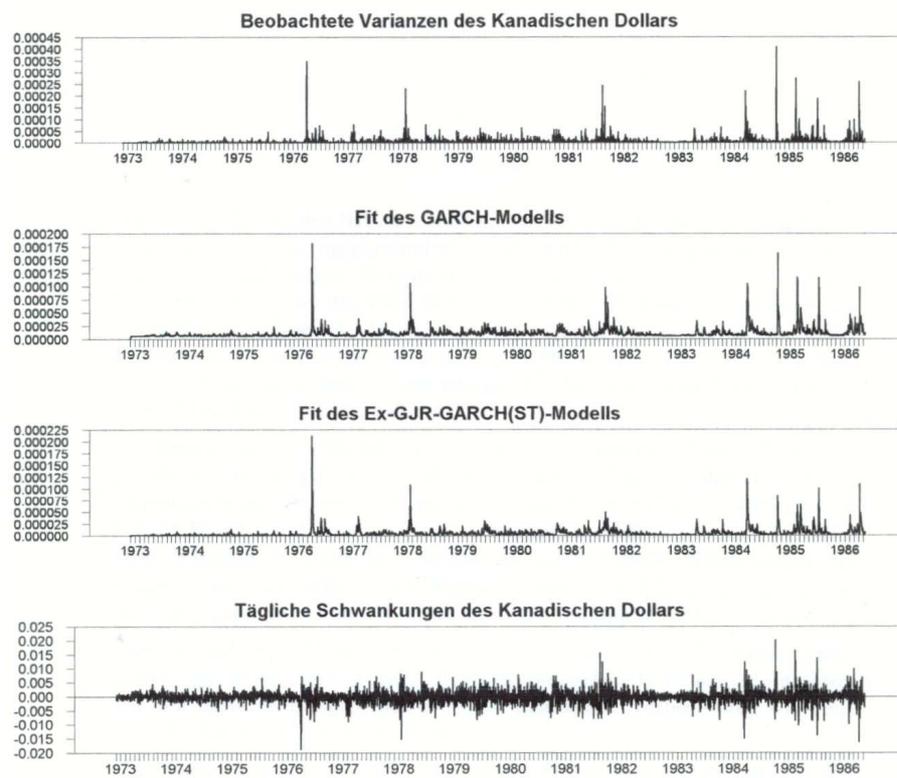


Abbildung 7: Beobachtete Varianzen vs. Schätzungen

Wirkungen der Schocks auf die bedingte Varianz der nächsten Periode zu - insbesondere ermöglicht sie asymmetrische und nichtlineare Einflüsse; zum anderen erlaubt unser Ansatz, die Volatilität einer Zeitreihe nicht nur auf Basis ihrer eigenen Vergangenheit zu erklären, sondern auch externe Einflüsse auf die Volatilität geltend zu machen. Damit wird eine wesentlich breitere Klasse von GARCH-Prozessen erfaßt. Schließlich haben wir für dieses Modell unter gewissen Regularitätsbedingungen das asymptotische Verhalten des Quasi-Maximum-Likelihood Schätzers gezeigt. Wie die obige Auswertung der Kursschwankungen des Kanadischen Dollars und des Britischen Pfundes gezeigt hat, kann die Berücksichtigung von äußeren Einflüssen auf die Volatilität des Preises von Geldanlagen durchaus sinnvoll sein. In unserem Fall vertrat das Britische Pfund den europäischen und der Kanadische Dollar den amerikanischen Wirtschaftsraum, die beide für sich ihre Dynamik haben, aber über internationale Beziehungen auch Verknüpfungen untereinander aufweisen. Die erweiterten GARCH-Modelle erlauben es, beide Informationskreise miteinander zu verbinden und so die Entwicklung der einzelnen Kurse präziser zu prognostizieren.

Bei den von uns betrachteten Zeitreihen brachte mit Heteroskedastizität, Asymmetrie und externen Effekten das Erfassen weiterer Eigenschaften der Daten auch jeweils eine Verbesserung der Vorhersagegenauigkeit. Die genaue Formulierung der einzelnen Modelle bleibt jedoch von Bedeutung. So kann das PGARCH-Modell zwar Asymmetrie nuancierter beschreiben als die GJR-Modelle, erkauft diesen Vorteil aber mit einem Verlust an Flexibilität, der den Gewinn an Genauigkeit überkompensiert. Den besten Kompromiß zwischen Komplexität und Einfachheit bieten für die Analyse der Wechselkurse des Kanadischen Dollars die beiden erweiterten GJR-GARCH-Modelle.

## 7 Appendix

Bemerkung: Der folgende Gesamtbeweis ist eine Verallgemeinerung der Arbeit von Lee und Hansen [1994]. Insbesondere die Gliederung des Beweises von Lee und Hansen ist zum Vorbild genommen worden. Es werden jedoch eine Reihe von Lücken geschlossen und die Aussagen präzisiert und auf unseren Kontext erweitert.

Bemerkung: Lemma 1 liefert zwei Rekursionsformeln für die Darstellung von  $h_{0t}$ , die wir bei fast allen folgenden Beweisen als Ausgangspunkt nehmen werden. Lemma 1 (1) ist notwendige Bedingung für die Endlichkeit der Likelihood-Funktion im Erwartungswert.

**Lemma 1** *Unter den gemachten Voraussetzungen gilt:*

- (1) Für alle  $t \in \mathbf{Z}$  ist  $h_{0t}$  fast sicher endlich.
- (2)  $h_{0t}$  läßt sich fast sicher darstellen als:

(a)

$$h_{0t} = (1 - \beta_0)\omega_0 \left(1 + \sum_{k=1}^{\infty} \prod_{i=1}^k (g(\alpha_0, \epsilon_{t-i})z_{t-i}^2 + \beta_0)\right) + \sum_{k=1}^{\infty} f(\delta_0, r_{t-k}) \prod_{i=2}^k (g(\alpha_0, \epsilon_{t-i})z_{t-i}^2 + \beta_0)$$

(b)

$$h_{0t} = \omega_0 + \sum_{k=0}^{\infty} \beta_0^k [g(\alpha_0, \epsilon_{t-k-1})\epsilon_{t-k-1}^2 + f(\delta_0, r_{t-k-1})]$$

**Beweis von Lemma 1:**

Für alle  $n \in \mathbf{N}$  läßt sich  $h_{0t}$  schreiben als

$$\begin{aligned} h_{0t} &:= (1 - \beta_0)\omega_0 + g(\alpha_0, \epsilon_{t-1})\epsilon_{t-1}^2 + f(\delta_0, r_{t-1}) + \beta_0 h_{0t-1} \\ &= (1 - \beta_0)\omega_0 + f(\delta_0, r_{t-1}) + (g(\alpha_0, \epsilon_{t-1})z_{t-1}^2 + \beta_0)h_{0t-1} \\ &= (1 - \beta_0)\omega_0 \left(1 + \sum_{k=1}^{n-1} \prod_{i=1}^k (g(\alpha_0, \epsilon_{t-i})z_{t-i}^2 + \beta_0)\right) \\ &\quad + \sum_{k=1}^n f(\delta_0, r_{t-k}) \prod_{i=2}^k (g(\alpha_0, \epsilon_{t-i})z_{t-i}^2 + \beta_0) \\ &\quad + h_{0t-n} \prod_{k=1}^n (g(\alpha_0, \epsilon_{t-k})z_{t-k}^2 + \beta_0) \end{aligned}$$

Nach Voraussetzung ist  $\zeta := E(\ln(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)) \leq E(\sup_t E(\ln(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2) | \mathcal{F}_{t-1})) < 0$  und  $(z_t, \epsilon_t, r_t)_{t \in \mathbf{Z}}$  ist strikt stationär und ergodisch.

Damit folgt aus dem Ergodentheorem, daß ein  $M < \infty$  existiert, so daß für alle  $n > M$  fast sicher gilt:

(a)

$$\left| \frac{1}{n} \sum_{k=1}^n \ln(g(\alpha_0, \epsilon_{t-k}) z_{t-k}^2 + \beta_0) - \zeta \right| < \left| \frac{\zeta}{2} \right|$$

und

$$\sum_{k=1}^n \ln(g(\alpha_0, \epsilon_{t-k}) z_{t-k}^2 + \beta_0) < \frac{\zeta}{2} n.$$

(b) Wegen  $\lim_{n \rightarrow \infty} \frac{1}{n} f(\delta_0, r_{t-n}) = 0$  fast sicher und  $\ln f \leq f$  ist fast sicher

$$\begin{aligned} & \ln(f(\delta_0, r_{t-n})) + \sum_{k=2}^n \ln(g(\alpha_0, \epsilon_{t-k}) z_{t-k}^2 + \beta_0) \\ & \leq f(\delta_0, r_{t-n}) + \sum_{k=2}^n \ln(g(\alpha_0, \epsilon_{t-k}) z_{t-k}^2 + \beta_0) < \frac{\zeta}{2} (n-1). \end{aligned}$$

Aus der Stetigkeit der Exponentialfunktion folgt dann

$$\prod_{k=1}^n (g(\alpha_0, \epsilon_{t-k}) z_{t-k}^2 + \beta_0) < \exp\left(\frac{\zeta}{2} n\right) \rightarrow 0 \quad \text{für } n \rightarrow \infty \quad \text{fast sicher,}$$

da  $\zeta < 0$ , sowie

$$f(\delta_0, r_{t-n}) \prod_{k=2}^n (g(\alpha_0, \epsilon_{t-k}) z_{t-k}^2 + \beta_0) < \exp\left(\frac{\zeta}{2} (n-1)\right) \rightarrow 0 \quad \text{für } n \rightarrow \infty \quad \text{fast sicher}$$

und somit

$$\sum_{k=1}^{\infty} \prod_{i=1}^k (g(\alpha_0, \epsilon_{t-i}) z_{t-i}^2 + \beta_0) < \infty \quad \text{fast sicher,}$$

sowie

$$\sum_{k=1}^{\infty} f(\delta_0, r_{t-k}) \prod_{i=2}^k (g(\alpha_0, \epsilon_{t-i}) z_{t-i}^2 + \beta_0) < \infty \quad \text{fast sicher,}$$

da  $f(\delta_0, r_t)$  und die ersten  $M$  Terme der Summe mit Wahrscheinlichkeit eins endlich sind und die übrigen von der geometrischen Reihe mit Gewichten  $\exp(\frac{\zeta}{2})$  majorisiert werden. Ebenso zeigt man, daß

$$h_{0t-n} \prod_{k=1}^n (g(\alpha_0, \epsilon_{t-k}) z_{t-k}^2 + \beta_0) \rightarrow 0 \quad \text{für } n \rightarrow \infty \quad \text{fast sicher,}$$

da  $h_{0-\infty} = \lim_{n \rightarrow \infty} h_{0t}$  fast sicher endlich ist. Das heißt, es ergibt sich gleichmäßig in  $t$ :

$$h_{0t} := (1 - \beta_0) \omega_0 \left( 1 + \sum_{k=1}^{\infty} \prod_{i=1}^k (g(\alpha_0, \epsilon_{t-i}) z_{t-i}^2 + \beta_0) \right)$$

$$+ \sum_{k=1}^{\infty} f(\delta_0, r_{r-k}) \prod_{i=2}^k (g(\alpha_0, \epsilon_{t-i}) z_{t-i}^2 + \beta_0) < \infty \quad \text{fast sicher.}$$

und damit die Teile (1) und (2a) des Lemmas.

Nach Definition läßt sich  $h_{0t}$  für alle  $n \in \mathbf{N}$  schreiben als

$$h_{0t} = (1 - \beta_0^{n+1})\omega_0 + \sum_{k=0}^n \beta_0^k (g(\alpha_0, \epsilon_{t-k-1}) \epsilon_{t-k-1}^2 + f(\delta_0, r_{t-k-1})) + \beta_0^{n+1} h_{0t-n-1}.$$

Da nach Teil (1) dieses Lemmas  $h_{0t}$  für alle  $t \in \mathbf{Z}$  mit Wahrscheinlichkeit eins endlich ist, ergibt sich, daß

$$\lim_{n \rightarrow \infty} \sum_{k=0}^n \beta_0^k (g(\alpha_0, \epsilon_{t-k-1}) \epsilon_{t-k-1}^2 + f(\delta_0, r_{t-k-1})) \leq h_{0t} < \infty$$

fast sicher. Insbesondere geht also

$$\sum_{k=n+1}^{\infty} \beta_0^k (g(\alpha_0, \epsilon_{t-k-1}) \epsilon_{t-k-1}^2 + f(\delta_0, r_{t-k-1})) \rightarrow 0 \quad (n \rightarrow \infty) \quad \text{fast sicher.}$$

Die fast sichere Konvergenz von  $\beta_0^n h_{0t-n}$  gegen 0 ist bereits begründet worden. Damit folgt Teil (2b) des Lemmas.

(Vgl. auch Nelson [1991], der in seinem Theorem 1 ein Pendant zu Lemma 1 (1) und (2a) für das Standard GARCH-Modell bei Unabhängigkeit der  $z_t$  beweist.)

□

Bemerkung: Lemma 2 liefert eine Abschätzung von  $h_t^\epsilon(\theta)$  gegen  $h_t(\theta)$ , die die beiden Prozesse in Beziehung setzt, die nachfolgenden Beweise vereinfacht und es insbesondere ermöglicht, sich bei den Beweisen der eigentlich interessanten Sätze immer wieder auf die Aussagen von Lemma 8 bis 10 zurückzuziehen.

**Lemma 2** *Fast sicher ist:*

$$B^{-1} h_t^\epsilon(\theta) \leq h_t(\theta) \leq B h_t^\epsilon(\theta)$$

wobei

$$B := 2 + \frac{\tau_2}{\tau_1} (1 + 2\tilde{m}) + \frac{\tau_2 \tilde{m}}{\omega_l (1 - \beta_u)} (2 + \tilde{m}) \quad \text{mit } \tilde{m} := \gamma_u - \gamma_l.$$

**Beweis von Lemma 2:**

Sei  $m := \gamma_0 - \gamma$ . Es ist  $|m| \leq \tilde{m}$  und  $\epsilon_t = e_t(\theta) + \gamma - \gamma_0 = e_t(\theta) - m$ .

Dann gilt fast sicher:

$$\begin{aligned}
& \frac{h_t^\epsilon(\theta)}{h_t(\theta)} \\
= & \frac{\omega + \sum_{k=0}^{\infty} \beta^k [g(\alpha, \epsilon_{t-k-1}) \epsilon_{t-k-1}^2 + f(\delta, r_{t-k-1})]}{h_t(\theta)} \\
\leq & \frac{\omega + \tau_2 \sum_{k=0}^{\infty} \beta^k (e_{t-k-1}(\theta) - m)^2}{h_t(\theta)} \\
& + \frac{\sum_{k=0}^{\infty} \beta^k f(\delta, r_{t-k-1})}{\underbrace{\omega + \sum_{k=0}^{\infty} \beta^k [g(\alpha, e_{t-k-1}(\theta)) \epsilon_{t-k-1}^2(\theta) + f(\delta, r_{t-k-1})]}_{\leq 1}} \\
\leq & \frac{\omega + \tau_2 \sum_{k=0}^{\infty} \beta^k e_{t-k-1}^2(\theta)}{\omega + \tau_1 \sum_{k=0}^{\infty} \beta^k e_{t-k-1}^2(\theta)} + \frac{2\tau_2 |m| \sum_{k=0}^{\infty} \beta^k |e_{t-k-1}(\theta)|}{h_t(\theta)} + \frac{\tau_2 m^2}{(1-\beta)h_t(\theta)} + 1 \\
\leq & 1 + \frac{\tau_2}{\tau_1} + \frac{2\tau_2 |m| \sum_{k=0}^{\infty} \beta^k e_{t-k-1}^2(\theta) + 2\tau_2 |m| \sum_{k=0}^{\infty} \beta^k}{\omega + \tau_1 \sum_{k=0}^{\infty} \beta^k e_{t-k-1}^2(\theta)} + \frac{\tau_2 m^2}{\omega(1-\beta)} + 1, \\
& \text{da } |e_{t-k-1}(\theta)| \leq e_{t-k-1}^2(\theta) + 1 \\
\leq & 2 + \frac{\tau_2}{\tau_1} + 2|m| \frac{\tau_2}{\tau_1} + \frac{2\tau_2 |m|}{\omega(1-\beta)} + \frac{\tau_2 m^2}{\omega(1-\beta)} \\
\leq & 2 + \frac{\tau_2}{\tau_1} (1 + 2\tilde{m}) + \frac{\tau_2 \tilde{m}}{\omega_l(1-\beta_u)} (2 + \tilde{m}) = B.
\end{aligned}$$

Entsprechend gilt umgekehrt:

$$\begin{aligned}
& \frac{h_t(\theta)}{h_t^\epsilon(\theta)} \\
= & \frac{\omega + \tau_2 \sum_{k=0}^{\infty} \beta^k (\epsilon_{t-k-1} + m)^2}{h_t^\epsilon(\theta)} + \frac{\sum_{k=0}^{\infty} \beta^k f(\delta, r_{t-k-1})}{\omega + \sum_{k=0}^{\infty} \beta^k [g(\alpha, \epsilon_{t-k-1}) \epsilon_{t-k-1}^2 + f(\delta, r_{t-k-1})]} \\
\leq & \frac{\omega + \tau_2 \sum_{k=0}^{\infty} \beta^k \epsilon_{t-k-1}^2}{\omega + \tau_1 \sum_{k=0}^{\infty} \beta^k \epsilon_{t-k-1}^2} + \frac{2\tau_2 |m| \sum_{k=0}^{\infty} \beta^k |\epsilon_{t-k-1}|}{h_t^\epsilon(\theta)} + \frac{\tau_2 m^2}{h_t^\epsilon(\theta)(1-\beta)} + 1 \\
\leq & 2 + \frac{\tau_2}{\tau_1} + \frac{2\tau_2 |m| \sum_{k=0}^{\infty} \beta^k \epsilon_{t-k-1}^2 + 2\tau_2 |m| \sum_{k=0}^{\infty} \beta^k}{\omega + \tau_1 \sum_{k=0}^{\infty} \beta^k \epsilon_{t-k-1}^2} + \frac{\tau_2 m^2}{\omega(1-\beta)}, \\
& \text{da } |\epsilon_{t-k-1}| \leq \epsilon_{t-k-1}^2 + 1 \\
\leq & 2 + \frac{\tau_2}{\tau_1} + 2|m| \frac{\tau_2}{\tau_1} + 2 \frac{\tau_2 |m|}{\omega(1-\beta)} + \frac{\tau_2 m^2}{\omega(1-\beta)} \\
\leq & 2 + \frac{\tau_2}{\tau_1} (1 + 2\tilde{m}) + \frac{\tau_2 \tilde{m}}{\omega_l(1-\beta_u)} (2 + \tilde{m}).
\end{aligned}$$

□

Bemerkung: Lemma 3 zeigt, daß sich die in den Voraussetzungen geforderte strikte Stationarität und Ergodizität von  $(z_t, \epsilon_t, r_t)$  bei der Konstruktion weiterer Prozesse daraus überträgt.

**Lemma 3** *Unter den gemachten Voraussetzungen gilt:*

*Die Prozesse  $(h_{0t}), (e_t(\theta)), (h_t(\theta)), (l_t(\theta)), (\nabla l_t(\theta))$  und  $(\nabla^2 l_t(\theta))$  sind strikt stationär und ergodisch für alle  $\theta \in \Theta$ .*

**Beweis von Lemma 3:**

Nach Lemma 1 (2) läßt sich  $h_{0t}$  auffassen als eine Funktion

$$h_{0t} : \mathbf{R}^\infty \rightarrow \mathbf{R}^+$$

$$(\epsilon_{t-1}, r_{t-1}, \epsilon_{t-2}, r_{t-2}, \dots) \mapsto h_{0t}(\epsilon_{t-1}, r_{t-1}, \epsilon_{t-2}, r_{t-2}, \dots).$$

Da der Prozeß  $(\epsilon_t, r_t)$  nach Voraussetzung strikt stationär und ergodisch ist, folgt nach einem Standardresultat die strikte Stationarität und Ergodizität von  $h_{0t}$ , wenn  $h_{0t}$  meßbar ist.

Es bleibt also lediglich zu zeigen, daß  $h_{0t} = \omega_0 + \sum_{k=0}^{\infty} \beta_0^k [g(\alpha_0, \epsilon_{t-k-1}) \epsilon_{t-k-1}^2 + f(\delta_0, r_{t-k-1})]$  meßbar ist.

Sei hierzu  $c_n = \omega_0 + \sum_{k=0}^n \beta_0^k [g(\alpha_0, \epsilon_{t-k-1}) \epsilon_{t-k-1}^2 + f(\delta_0, r_{t-k-1})]$ . Dann ist  $h_{0t} = c_\infty = \sup_{n \geq 0} c_n$ .

$r_n$  ist als endliche Summe der meßbaren Funktionen  $g$  und  $f$  meßbar, und  $h_{0t}$  ist als Supremum einer Folge meßbarer Funktionen meßbar.

Ebenso gilt:

$e_t(\theta) = \epsilon_t + \gamma$  meßbar.

Für  $h_t(\theta)$  gilt das über  $h_{0t}$  Gesagte entsprechend.

$l_t(\theta) = \ln h_t(\theta) + \frac{\epsilon_t(\theta)^2}{h_t(\theta)}$  ist meßbar, da  $\ln(\cdot)$  meßbar und  $h_t(\theta)$  von 0 weg beschränkt ist. Auch  $\nabla l_t(\theta)$  und  $\nabla^2 l_t(\theta)$  sind meßbare Bilder des strikt stationären und ergodischen Prozesses  $(\epsilon_t, r_{t-1})$  wie aus den Formeln im Beweis von Lemma 12 und Lemma 15 hervorgeht.

□

**Lemma 4** *Für  $x, y \geq 0$  und  $0 < p < 1$ :*

$$(x + y)^p \leq x^p + y^p.$$

**Beweis von Lemma 4:**

Sei o.E.  $y \geq x$ . Wir entwickeln  $f(y) = (x + y)^p$  um 0:

$$(x + y)^p = x^p + px^{p-1}y + \underbrace{p(p-1) \frac{(x+\xi)^{p-2}}{2} y^2}_{\leq 0} \quad \text{für ein } \xi \in [0, x-y]$$

$$\begin{aligned}
&\leq x^p + px^{p-1}y \\
&\leq x^p + y^p
\end{aligned}$$

da  $y \geq x$  und  $p < 1$ .

□

Bemerkung: Um zum Beweis der punktweisen Konvergenz der Likelihood-Funktion  $L_n(\theta)$  gegen die nichtstochastische Funktion  $L(\theta) := E(l_t(\theta))$  das Ergodentheorem einsetzen zu können, muß gezeigt werden, daß  $l_t(\theta)$  endlichen Erwartungswert hat. Eine Schwierigkeit dabei ist, daß nicht für alle  $\theta \in \Theta$  und möglichen Funktionen  $g$   $Eh_t(\theta) < \infty$ . Jedoch läßt sich zeigen, daß ein niedrigeres Moment von  $h_t(\theta)$  existiert.

**Lemma 5** *Unter den gemachten Voraussetzungen gilt:*

*Es gibt ein  $0 < p < \frac{1}{2}$  und Zahlen  $C < \infty$  und  $T_p < \infty$ , so daß für alle  $\theta \in \Theta$  gilt:*

(1)

$$E(h_{0t}^p) \leq C < \infty$$

(2)

$$E(h_t(\theta)^p) \leq E(\sup_{\theta \in \Theta} h_t(\theta)^p) \leq T_p < \infty$$

**Beweis von Lemma 5:**

(1) Offensichtlich ist  $E(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2) \leq \beta_0 + \tau_2 < \infty$ . Aus dem Satz von der majorisierten Konvergenz folgt dann für  $0 < r < 1$  wegen

$$(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)^r \leq \max[1, (\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)],$$

daß

$$\begin{aligned}
\lim_{r \rightarrow p} E[(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)^r] &= E[\lim_{r \rightarrow p} (\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)^r] \\
&= E[(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)^p]
\end{aligned}$$

für alle  $0 < p < 1$ , und damit, daß  $\|(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)\|_p$  stetig ist in  $p$ . Weiter gilt:

$$\begin{aligned}
&\ln \|(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)\|_r \\
&= \ln [E(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)^r]^{\frac{1}{r}} \\
&= \frac{1}{r} \ln [E(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)^r] \\
&\leq \frac{1}{r} (E[(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)^r] - 1), \quad \text{da } \ln(x) \leq x - 1 \\
&= E\left(\frac{(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)^r - 1}{r}\right) \rightarrow E(\ln(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)) \quad \text{für } r \rightarrow 0 \\
&\leq E(\sup_t E(\ln(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2)) | \mathcal{F}_{t-1}) =: \ln(\kappa) < 0
\end{aligned}$$

Denn es ist

$$\begin{aligned}\lim_{r \rightarrow 0} \frac{x^r - 1}{r} &= \lim_{r \rightarrow 0} \frac{e^{r \ln(x)} - 1}{r} \\ &= \frac{d}{dr} (e^{r \ln(x)})|_{r=0} = \ln(x)\end{aligned}$$

und  $\frac{x^r - 1}{r}$  ist monoton fallend in  $r$  für  $r$  klein genug (vgl. Hardy et al. [1934] Theorem 36). Damit erhalten wir:

$$\lim_{r \rightarrow 0} [E((\beta_0 + g(\alpha_0, \epsilon_t) z_t^2)^r)]^{\frac{1}{r}} \leq \exp(\ln(\kappa)) = \kappa < 1.$$

Wegen der Stetigkeit des Ausdrucks in  $r$  gibt es also ein  $0 < \tilde{p} < 1$  und  $0 < \mu < 1$ , so daß für alle  $0 < p < \tilde{p}$  gilt:

$$E(\beta_0 + g(\alpha_0, \epsilon_t) z_t^2)^p \leq \mu < 1.$$

Dieselbe Argumentation gilt für  $E((\beta_0 + g(\alpha_0, \epsilon_t) z_t^2) | \mathcal{F}_{t-1})$ .

Wähle nun  $p$  so, daß  $0 < p < \frac{\tilde{p}}{2}$ . Dann ergibt sich mit Lemma 4, der zu Anfang gemachten Voraussetzung, daß  $E(f(\delta, r_t)) \leq K$ , und Adams Regel:

$$\begin{aligned}E(h_{0t}^p) &= E\left[\left((1 - \beta_0)\omega_0\left(1 + \sum_{k=1}^{\infty} \prod_{i=1}^k (g(\alpha_0, \epsilon_{t-i}) z_{t-i}^2 + \beta_0)\right)\right.\right. \\ &\quad \left.\left.+ \sum_{k=1}^{\infty} f(\delta_0, r_{t-k}) \prod_{i=2}^k (g(\alpha_0, \epsilon_{t-i}) z_{t-i}^2 + \beta_0)\right)^p\right] \\ &\leq \omega_0^p + \omega_0^p \sum_{k=1}^{\infty} E\left[\prod_{i=1}^k (g(\alpha_0, \epsilon_{t-i}) z_{t-i}^2 + \beta_0)^p\right] \\ &\quad + \sum_{k=1}^{\infty} E[f(\delta_0, r_{t-k})^p \prod_{i=2}^k (g(\alpha_0, \epsilon_{t-i}) z_{t-i}^2 + \beta_0)^p] \\ &\leq \omega_0^p + \omega_0^p \sum_{k=1}^{\infty} \mu^k + 1 + K \\ &\quad + \sum_{k=2}^{\infty} \mu^{k-2} E[f(\delta_0, r_{t-k})^p (g(\alpha_0, \epsilon_{t-k}) z_{t-k}^2 + \beta_0)^p] \quad \text{Adams Regel} \\ &\leq \omega_0^p \sum_{k=0}^{\infty} \mu^k + 1 + K + \sum_{k=0}^{\infty} \mu^k \tilde{C} \\ &= 1 + K + (\omega_0^p + \tilde{C}) \frac{1}{1 - \mu} =: C < \infty,\end{aligned}$$

da  $\mu < 1$ , mit

$$E[f(\delta_0, r_{t-k})^p (g(\alpha_0, \epsilon_{t-k}) z_{t-k}^2 + \beta_0)^p]$$

$$\begin{aligned}
&\leq \|f(\delta_0, r_{t-k})^p\|_2 \| (g(\alpha_0, \epsilon_{t-k}) z_{t-k}^2 + \beta_0)^p \|_2 \\
&\leq [E(1 + f(\delta_0, r_{t-k}))]^{\frac{1}{2}} \mu^{\frac{1}{2}} \\
&\leq (1 + K) \mu^{\frac{1}{2}} =: \tilde{C} < \infty.
\end{aligned}$$

(2) Mit Teil (1) dieses Lemmas gilt:

$$\begin{aligned}
E(h_t(\theta)^p) &\leq E(\sup_{\theta \in \Theta} h_t(\theta)^p) \leq B^p E(\sup_{\theta \in \Theta} h_t^\epsilon(\theta)^p) \\
&\leq B^p (\omega_u^p + \tau_2^p \sum_{k=0}^{\infty} \beta^{kp} (E\epsilon_{t-k-1}^{2p} + Ef(\tilde{\delta}, r_{t-k-1}))) \quad \text{mit Lemma 4} \\
&\leq B^p (\omega_u^p + \frac{\tau_2^p}{(1 - \beta_u^p)} \left( \frac{Eh_{0t}^p}{\tau_1^p} + 1 + K \right)) =: T_p < \infty
\end{aligned}$$

□

Bemerkung: Für die Konsistenz des QML-Schätzers muß im wesentlichen gezeigt werden, daß die empirische Likelihood-Funktion  $L_n^*(\theta)$  gleichmäßig stochastisch gegen die nichtstochastische Funktion  $L(\theta)$  konvergiert. Wir benutzen dazu die Zerlegung

$$\sup_{\theta \in \Theta} |L_n^*(\theta) - L(\theta)| \leq \sup_{\theta \in \Theta} |L_n^*(\theta) - L_n(\theta)| + \sup_{\theta \in \Theta} |L_n(\theta) - L(\theta)| \rightarrow_p 0$$

und zeigen in Lemma 6 bzw. Lemma 11(5) die uniforme Konsistenz für beide Terme separat.

**Lemma 6** *Unter den gemachten Voraussetzungen gilt:*

$$\sup_{\theta \in \Theta} |L_n^*(\theta) - L_n(\theta)| \rightarrow_p 0.$$

**Beweis von Lemma 6:**

Es ist zu zeigen, daß es ein  $M(\epsilon, \eta)$  gibt, so daß für alle  $n \geq M$  gilt:

$$\begin{aligned}
&P\left\{ \sup_{\theta \in \Theta} |L_n^*(\theta) - L_n(\theta)| \geq \epsilon \right\} \leq \eta \\
&P\left\{ \sup_{\theta \in \Theta} |L_n^*(\theta) - L_n(\theta)| \geq \epsilon \right\} \\
&= P\left\{ \sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n \ln\left(\frac{h_t(\theta)}{h_t^*(\theta)}\right) - \frac{1}{n} \sum_{t=1}^n \left( \frac{e_t^2(\theta)}{h_t^*(\theta)} - \frac{e_t^2(\theta)}{h_t(\theta)} \right) \right| \geq \epsilon \right\} \\
&\leq P\left\{ \sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n \ln\left(\frac{h_t(\theta)}{h_t^*(\theta)}\right) \right| \geq \frac{\epsilon}{2} \right\} \\
&\quad + P\left\{ \sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n \left( \frac{e_t^2(\theta)}{h_t^*(\theta)} - \frac{e_t^2(\theta)}{h_t(\theta)} \right) \right| \geq \frac{\epsilon}{2} \right\}
\end{aligned}$$

Für den ersten Term erhält man:

$$\begin{aligned}
& P \left\{ \sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n \ln \left( \frac{h_t(\theta)}{h_t^*(\theta)} \right) \right| \geq \frac{\epsilon}{2} \right\} \\
&= P \left\{ \sup_{\theta \in \Theta} \left( \left| \frac{1}{n} \sum_{t=1}^n \ln \left( \frac{h_t(\theta)}{h_t^*(\theta)} \right) \right| \right)^p \geq \left( \frac{\epsilon}{2} \right)^p \right\} \\
&\leq \left( \frac{2}{\epsilon n} \right)^p E \left( \sup_{\theta \in \Theta} \left( \sum_{t=1}^n \ln \left( \frac{\omega + \sum_{k=0}^{\infty} \beta^k [g(\alpha, e_{t-k-1}(\theta)) e_{t-k-1}^2(\theta) + f(\delta, r_{t-k-1})]}{\omega + \sum_{k=0}^{t-2} \beta^k [g(\alpha, e_{t-k-1}(\theta)) e_{t-k-1}^2(\theta) + f(\delta, r_{t-k-1})]} \right) \right)^p \right) \\
&\leq \left( \frac{2}{\epsilon n} \right)^p E \left( \sup_{\theta \in \Theta} \left( \sum_{t=1}^n \ln \left( 1 + \frac{\beta^{t-1} \sum_{k=0}^{\infty} \beta^k [g(\alpha, e_{-k}(\theta)) e_{-k}(\theta)^2 + f(\delta, r_{-k})]}{\omega} \right) \right)^p \right) \\
&\leq \left( \frac{2}{\epsilon n} \right)^p E \left( \sup_{\theta \in \Theta} \left( \sum_{t=1}^n \ln \left( 1 + \frac{\beta^{t-1} (h_1(\theta) - \omega)}{\omega} \right) \right)^p \right) \\
&\leq \left( \frac{2}{\epsilon n} \right)^p E \left( \sup_{\theta \in \Theta} \left( \sum_{t=1}^n \ln \left( 1 + \frac{\beta^{t-1} h_1(\theta)}{\omega} \right) \right)^p \right) \\
&\leq \left( \frac{2}{\epsilon n} \right)^p E \left( \sup_{\theta \in \Theta} \left( \sum_{t=1}^n \frac{\beta^{t-1} h_1(\theta)}{\omega} \right)^p \right) \\
&\leq \left( \frac{2}{\epsilon n \omega} \right)^p E \left( \sup_{\theta \in \Theta} h_1(\theta)^p \right) \frac{1}{(1 - \beta_u)^p} \\
&\leq \left( \frac{2}{\epsilon n \omega_l (1 - \beta_u)} \right)^p T_p \rightarrow 0 \quad (n \rightarrow \infty)
\end{aligned}$$

Für den zweiten Term ergibt sich:

$$\begin{aligned}
& P \left\{ \sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n \left( \frac{e_t^2(\theta)}{h_t^*(\theta)} - \frac{e_t^2(\theta)}{h_t(\theta)} \right) \right| \geq \frac{\epsilon}{2} \right\} \\
&= P \left\{ \sup_{\theta \in \Theta} \left| \left( \frac{1}{n} \sum_{t=1}^n \left( \frac{e_t^2(\theta)}{h_t^*(\theta)} - \frac{e_t^2(\theta)}{h_t(\theta)} \right) \right)^{\frac{p}{2}} \right| \geq \left( \frac{\epsilon}{2} \right)^{\frac{p}{2}} \right\} \\
&\leq \left( \frac{2}{\epsilon n} \right)^{\frac{p}{2}} E \left( \sup_{\theta \in \Theta} \left( \sum_{t=1}^n \frac{(h_t(\theta) - h_t^*(\theta))}{h_t(\theta) h_t^*(\theta)} e_t^2(\theta) \right)^{\frac{p}{2}} \right) \\
&\leq \left( \frac{2}{\epsilon n} \right)^{\frac{p}{2}} E \left( \sup_{\theta \in \Theta} \left( \sum_{t=1}^n \frac{(h_1(\theta) - \omega)}{\tau_1 \omega^2} \beta^{t-1} h_{t+1}(\theta) \right)^{\frac{p}{2}} \right) \\
&\quad \text{wegen } h_t(\theta) - h_t^*(\theta) = \beta^{t-1} (h_1(\theta) - \omega) \text{ und } e_t^2(\theta) \leq \frac{h_{t+1}(\theta)}{\tau_1}
\end{aligned}$$

$$\begin{aligned}
&\leq \left( \frac{2}{\epsilon n \tau_1 \omega_l^2} \right)^{\frac{p}{2}} E \left( \sup_{\theta \in \Theta} \left( h_1(\theta) \sum_{t=1}^{\infty} \beta^{t-1} h_{t+1}(\theta) \right)^{\frac{p}{2}} \right) \\
&\leq \left( \frac{2}{\epsilon n \tau_1 \omega_l^2} \right)^{\frac{p}{2}} \left( E(\sup_{\theta \in \Theta} h_1^p(\theta)) \right)^{\frac{1}{2}} \left( \sum_{t=1}^{\infty} \beta_u^{(t-1)p} E(\sup_{\theta \in \Theta} h_{t+1}^p(\theta)) \right)^{\frac{1}{2}}, \\
&\quad \text{mit der Cauchy-Schwarz-Ungleichung und Lemma 4} \\
&\leq \left( \frac{2}{\epsilon n \tau_1 \omega_l^2} \right)^{\frac{p}{2}} T_p^{\frac{1}{2}} \frac{T_p^{\frac{1}{2}}}{(1 - \beta_u^p)^{\frac{1}{2}}} \rightarrow 0 \quad (n \rightarrow \infty)
\end{aligned}$$

□

Bemerkung: Beim Beweis von Lemma 8ff. wird immer wieder gebraucht werden, Quotienten von Erweiterten GARCH-Prozessen durch geometrische Reihen zu majorisieren. Lemma 7 liefert in diesem Zusammenhang die Möglichkeit, die geometrischen Gewichte dieser Reihen von 1 weg zu beschränken und die bereits angegebenen Parameterrestriktionen zu formulieren, unter denen die folgenden Lemmata gültig sind.

**Lemma 7** (Lee und Hansen, 1994)

Unter den gemachten Voraussetzungen gilt für alle  $\psi > 0$  und alle  $q \geq 1$ :

$$\begin{aligned}
\beta^q E \left( \left( \frac{1}{\beta + \psi z_t^2} \right)^q \middle| \mathcal{F}_{t-1} \right) &\leq E \left( \frac{1}{1 + \psi z_t^2} \middle| \mathcal{F}_{t-1} \right) \\
&\leq 1 - \frac{\psi}{2^{\frac{(2+\nu)}{\nu}} S_\nu^{2/\nu} (2 + \psi)} \\
&=: \mathcal{R}(\psi) < 1 \quad \text{fast sicher.}
\end{aligned}$$

Bemerkung: Lemma 8-10 liefern eine Reihe technischer Resultate, mit denen schließlich der Prozeß  $l_t(\theta)$  und seine ersten, zweiten und dritten Ableitungen im Erwartungswert allgemein abgeschätzt werden können.

**Lemma 8** Unter den gemachten Voraussetzungen gilt:

(1) Für alle  $1 \leq q \leq 6$  gilt:

$$E \left( \left( \frac{h_{0t-k}}{h_{0t}} \right)^q \middle| \mathcal{F}_{t-k-1} \right) \leq \left( \frac{\mathcal{R}_u}{\beta_0^q} \right)^k \quad \text{fast sicher,}$$

wobei  $\mathcal{R}_u = \mathcal{R}(\tau_1) < 1$ .

(2) Für alle  $1 \leq q \leq 6$  gilt:

$$E \left( \left( \frac{h_{0t-k}}{h_{0t}} \right)^q \right) \leq \left( \frac{\mathcal{R}_u}{\beta_0^q} \right)^k$$

(3) Wenn  $\beta \leq \beta_0$ ,

$$\sup_{\theta \in \Theta} \frac{h_t^\epsilon(\theta)}{h_{0t}} \leq T_1 := \frac{\omega_u}{\omega_0} + \frac{\tau_2}{\tau_1} + S < \infty \quad \text{fast sicher.}$$

(4) Wenn  $\beta \geq \beta_0$ ,

$$\sup_{\theta \in \Theta} \frac{h_{0t}}{h_t^\epsilon(\theta)} \leq T_2 := \frac{\omega_0}{\omega_l} + \frac{\tau_2}{\tau_1} + S < \infty \quad \text{fast sicher.}$$

**Beweis von Lemma 8:**

(1) Fast sicher gilt:

$$\begin{aligned} \frac{h_{0t-1}}{h_{0t}} &= \frac{h_{0t-1}}{(1-\beta_0)\omega_0 + \beta_0 h_{0t-1} + g(\alpha_0, \epsilon_{t-1})\epsilon_{t-1}^2 + f(\delta_0, r_{t-1})} \\ &= \frac{h_{0t-1}}{(1-\beta_0)\omega_0 + h_{0t-1}[\beta_0 + g(\alpha_0, \epsilon_{t-1})z_{t-1}^2] + f(\delta_0, r_{t-1})} \\ &\leq \frac{1}{\beta_0 + g(\epsilon_{t-1})z_{t-1}^2} \\ &\leq \frac{1}{\beta_0 + \tau_1 z_{t-1}^2} \end{aligned}$$

Weiter gilt:

$$\frac{h_{0t-k}}{h_{0t}} = \frac{h_{0t-1}}{h_{0t}} \frac{h_{0t-2}}{h_{0t-1}} \cdots \frac{h_{0t-k}}{h_{0t-k+1}} \leq \prod_{i=1}^k \left( \frac{1}{\beta_0 + \tau_1 z_{t-i}^2} \right)$$

Und somit fast sicher:

$$\begin{aligned} E \left( \left( \frac{h_{0t-k}}{h_{0t}} \right)^q \middle| \mathcal{F}_{t-k-1} \right) &\leq E \left( \prod_{i=1}^k E \left( \left( \frac{1}{\beta_0 + \tau_1 z_{t-i}^2} \right)^q \middle| \mathcal{F}_{t-i-1} \right) \middle| \mathcal{F}_{t-k-1} \right) \\ &\leq \prod_{i=1}^k \frac{\mathcal{R}(\tau_1)}{\beta_0^q} = \left( \frac{\mathcal{R}_u}{\beta_0^q} \right)^k \end{aligned}$$

(2)

$$E \left( \left( \frac{h_{0t-k}}{h_{0t}} \right)^q \right) = E \left( E \left( \left( \frac{h_{0t-k}}{h_{0t}} \right)^q \middle| \mathcal{F}_{t-k-1} \right) \right) \leq \left( \frac{\mathcal{R}_u}{\beta_0^q} \right)^k$$

(3) Für  $\beta \leq \beta_0$  ist fast sicher:

$$\begin{aligned} \sup_{\theta \in \Theta} \frac{h_t^\epsilon(\theta)}{h_{0t}} &\leq \frac{\omega_u + \tau_2 \sum_{k=0}^{\infty} \beta_0^k \epsilon_{t-k-1}^2 + \sum_{k=0}^{\infty} \beta_0^k f(\delta, r_{t-k-1})}{\omega_0 + \tau_1 \sum_{k=0}^{\infty} \beta_0^k \epsilon_{t-k-1}^2 + \sum_{k=0}^{\infty} \beta_0^k f(\delta_0, r_{t-k-1})} \\ &\leq \frac{\omega_u + \tau_2 \sum_{k=0}^{\infty} \beta_0^k \epsilon_{t-k-1}^2 + \sum_{k=0}^{\infty} \beta_0^k \left( \frac{f(\delta, r_{t-k-1})}{f(\delta_0, r_{t-k-1})} f(\delta_0, r_{t-k-1}) \right)}{\omega_0 + \tau_1 \sum_{k=0}^{\infty} \beta_0^k \epsilon_{t-k-1}^2 + \sum_{k=0}^{\infty} \beta_0^k f(\delta_0, r_{t-k-1})} \end{aligned}$$

$$\begin{aligned}
&\leq \frac{\omega_u + \tau_2 \sum_{k=0}^{\infty} \beta_0^k \epsilon_{t-k-1}^2 + S \sum_{k=0}^{\infty} \beta_0^k f(\delta_0, r_{t-k-1})}{\omega_0 + \tau_1 \sum_{k=0}^{\infty} \beta_0^k \epsilon_{t-k-1}^2 + \sum_{k=0}^{\infty} \beta_0^k f(\delta_0, r_{t-k-1})} \\
&\leq \frac{\omega_u}{\omega_0} + \frac{\tau_2}{\tau_1} + S =: T_1
\end{aligned}$$

(4) Für  $\beta \geq \beta_0$  ist fast sicher:

$$\begin{aligned}
\sup_{\theta \in \Theta} \frac{h_{0t}}{h_t^\epsilon(\theta)} &\leq \frac{\omega_0 + \tau_2 \sum_{k=0}^{\infty} \beta_0^k \epsilon_{t-k-1}^2 + S \sum_{k=0}^{\infty} \beta_0^k f(\delta, r_{t-k-1})}{\omega + \tau_1 \sum_{k=0}^{\infty} \beta_0^k \epsilon_{t-k-1}^2 + \sum_{k=0}^{\infty} \beta_0^k f(\delta, r_{t-k-1})} \\
&\leq \frac{\omega_0}{\omega_l} + \frac{\tau_2}{\tau_1} + S =: T_2
\end{aligned}$$

□

**Lemma 9** *Unter den gemachten Voraussetzungen gilt für  $1 \leq q \leq 6$ :*

(1) Für  $\theta \in \Theta_l^q$  ist:

$$\left\| \sup_{\theta \in \Theta_l^q} \frac{h_{t-k}^\epsilon(\theta)}{h_t^\epsilon(\theta)} \right\|_q \leq \left( \frac{\mathcal{R}_l^{\frac{1}{q}}}{\beta_{ql}} \right)^k$$

wobei  $\mathcal{R}_l := \mathcal{R}(\frac{\tau_1}{T_1})$

(2) Für  $\theta \in \Theta_q$  ist:

$$\left\| \sup_{\theta \in \Theta_q} \frac{h_{0t}}{h_t^\epsilon(\theta)} \right\|_q \leq T_3 := \frac{\omega_0}{\omega_l} + \frac{\tau_2}{\tau_1 \eta_l} + \frac{K}{\omega_0(1 - \beta_0)} + S < \infty$$

**Beweis von Lemma 9:**

(1) Da  $\theta \in \Theta_l^q$ , ist  $\beta \leq \beta_0$ . Somit gilt mit Lemma 8(3) fast sicher:

$$\begin{aligned}
&\sup_{\theta \in \Theta_l^q} \frac{h_t^\epsilon(\theta)}{h_{t+1}^\epsilon(\theta)} \\
&= \sup_{\theta \in \Theta_l^q} \frac{h_t^\epsilon(\theta)}{(1 - \beta)\omega + \beta h_t^\epsilon(\theta) + g(\alpha, \epsilon_t)\epsilon_t^2 + f(\delta, r_t)} \\
&\leq \sup_{\theta \in \Theta_l^q} \frac{h_t^\epsilon(\theta)}{\beta h_t^\epsilon(\theta) + \tau_1 \frac{h_{0t}}{h_t^\epsilon(\theta)} h_t^\epsilon(\theta) z_t^2} \\
&\leq \sup_{\theta \in \Theta_l^q} \frac{1}{\beta + \tau_1 \frac{h_{0t}}{h_t^\epsilon(\theta)} z_t^2} \\
&\leq \frac{1}{\beta_{ql} + \frac{\tau_1}{T_1} z_t^2}
\end{aligned}$$

Und damit fast sicher:

$$E \left( \sup_{\theta \in \Theta_l^q} \left( \frac{h_t^\epsilon(\theta)}{h_{t+1}^\epsilon(\theta)} \right)^q \middle| \mathcal{F}_{t-1} \right) \leq E \left( \left( \frac{1}{\beta_{ql} + \frac{\tau_1}{T_1} z_t^2} \right)^q \middle| \mathcal{F}_{t-1} \right)$$

$$\begin{aligned}
&\leq \frac{\mathcal{R}(\frac{\tau_1}{T_1})}{\beta_{ql}^q} \quad \text{mit Lemma 7(2)} \\
&= \frac{\mathcal{R}_l}{\beta_{ql}^q}
\end{aligned}$$

So daß schließlich gilt:

$$\begin{aligned}
E \left( \sup_{\theta \in \Theta_t^q} \left( \frac{h_{t-k}^\epsilon(\theta)}{h_t^\epsilon(\theta)} \right)^q \right) &= E \left( \sup_{\theta \in \Theta_t^q} \prod_{i=1}^k \left( \frac{h_{t-i}^\epsilon(\theta)}{h_{t-i+1}^\epsilon(\theta)} \right)^q \right) \\
&\leq E \left( \prod_{i=1}^k E \left( \sup_{\theta \in \Theta_t^q} \left( \frac{h_{t-i}^\epsilon(\theta)}{h_{t-i+1}^\epsilon(\theta)} \right)^q \mid \mathcal{F}_{t-i-1} \right) \right) \\
&\leq \left( \frac{\mathcal{R}_l}{\beta_{ql}^q} \right)^k
\end{aligned}$$

(2) Für  $\beta \leq \beta_0$  gilt:

$$\begin{aligned}
&\left\| \sup_{\theta \in \Theta_t^q} \frac{h_{0t}}{h_t^\epsilon(\theta)} \right\|_q \\
&\leq \left\| \sup_{\theta \in \Theta_t^q} \left( \frac{\omega_0}{\omega} + \frac{\tau_2 \sum_{k=0}^{\infty} \beta_0^k \epsilon_{t-k-1}^2}{h_t^\epsilon(\theta)} + \frac{\sum_{k=0}^{\infty} \beta_0^k f(\delta_0, r_{t-k-1})}{h_t^\epsilon(\theta)} \right) \right\|_q \\
&\leq \frac{\omega_0}{\omega_l} + \frac{\tau_2}{\tau_1} \sum_{k=0}^{\infty} \beta_0^k \left\| \sup_{\theta \in \Theta_t^q} \frac{h_{t-k}^\epsilon(\theta)}{h_t^\epsilon(\theta)} \right\|_q + \frac{K}{\omega_0(1-\beta_0)} \\
&\leq \frac{\omega_0}{\omega_l} + \frac{\tau_2}{\tau_1} \sum_{k=0}^{\infty} \left( \frac{\beta_0 \mathcal{R}_l^{\frac{1}{q}}}{\beta_{ql}} \right)^k + \frac{K}{\omega_0(1-\beta_0)} \quad \text{mit Lemma 8(1)} \\
&= \frac{\omega_0}{\omega_l} + \frac{\tau_2}{\tau_1} \frac{\beta_{ql}}{\beta_{ql} - \beta_0 \mathcal{R}_l^{\frac{1}{q}}} + \frac{K}{\omega_0(1-\beta_0)} \\
&= \frac{\omega_0}{\omega_l} + \frac{\tau_2}{\tau_1} \frac{\beta_{ql}}{\eta_l} + \frac{K}{\omega_0(1-\beta_0)} \\
&\leq \frac{\omega_0}{\omega_l} + \frac{\tau_2}{\tau_1 \eta_l} + \frac{K}{\omega_0(1-\beta_0)} + S =: T_3
\end{aligned}$$

Für  $\beta \geq \beta_0$  gilt:

$$\sup_{\theta \in \Theta_u^q} \frac{h_{0t}}{h_t^\epsilon(\theta)} \leq \frac{\omega_0}{\omega_l} + \frac{\tau_2}{\tau_1} + S = T_2 \leq T_3$$

mit Lemma 8(4) und  $\eta_l < 1$

□

**Lemma 10** *Unter den gemachten Voraussetzungen gilt für  $1 \leq q \leq 6$  und  $\theta \in \Theta_u^q$ :*

$$(1) \quad \left\| \sup_{\theta \in \Theta_u^q} \frac{h_t^\epsilon(\theta)}{h_{0t}} \right\|_q \leq T_4 := \frac{\omega_u}{\omega_0} + \frac{\tau_2}{\tau_1 \eta_u} + \frac{K}{\omega_0(1-\beta_u)} < \infty$$

$$(2) \quad \left\| \sup_{\theta \in \Theta_u^q} \frac{h_{t-k}^\epsilon(\theta)}{h_t^\epsilon(\theta)} \right\|_q \leq T_4 T_2 \left( \frac{\mathcal{R}_u^{\frac{1}{q}}}{\beta_0} \right)^k$$

**Beweis von Lemma 10:**

Da  $\theta \in \Theta_u^q$ , ist insbesondere  $\beta \geq \beta_0$ . Damit erhalten wir:

$$(1) \quad \begin{aligned} & \left\| \sup_{\theta \in \Theta_u^q} \frac{h_t^\epsilon(\theta)}{h_{0t}} \right\|_q \\ & \leq \left\| \sup_{\theta \in \Theta_u^q} \left( \frac{\omega + \tau_2 \sum_{k=0}^{\infty} \beta^k \epsilon_{t-k-1}^2 + \sum_{k=0}^{\infty} \beta^k f(\delta, r_{t-k-1})}{h_{0t}} \right) \right\|_q \\ & \leq \frac{\omega_u}{\omega_0} + \frac{\tau_2}{\tau_1} \sum_{k=0}^{\infty} \beta_{qu}^k \left\| \sup_{\theta \in \Theta_u^q} \frac{h_{0t-k}}{h_{0t}} \right\|_q + \frac{K}{\omega_0(1-\beta_u)} \\ & \leq \frac{\omega_u}{\omega_0} + \frac{\tau_2}{\tau_1} \sum_{k=0}^{\infty} \underbrace{\left( \frac{\beta_{qu} \mathcal{R}_u^{\frac{1}{q}}}{\beta_0} \right)^k}_{< 1 \text{ nach Def. von } \beta_{qu}} + \frac{K}{\omega_0(1-\beta_u)} \quad \text{mit Lemma 8(2)} \\ & \leq \frac{\omega_u}{\omega_0} + \frac{\tau_2}{\tau_1} \frac{\beta_0}{\beta_0 - \beta_{qu} \mathcal{R}_u^{\frac{1}{q}}} + \frac{K}{\omega_0(1-\beta_u)} \\ & \leq \frac{\omega_u}{\omega_0} + \frac{\tau_2}{\tau_1 \eta_u} + \frac{K}{\omega_0(1-\beta_u)} =: T_4, \quad \text{da } \eta_u = \beta_0 - \beta_{qu} \mathcal{R}_u^{\frac{1}{q}} \end{aligned}$$

$$(2) \quad \begin{aligned} & \left\| \sup_{\theta \in \Theta_u^q} \frac{h_{t-k}^\epsilon(\theta)}{h_t^\epsilon(\theta)} \right\|_q \\ & = \left\| \sup_{\theta \in \Theta_u^q} \left( \frac{h_{0t}}{h_t^\epsilon(\theta)} \frac{h_{t-k}^\epsilon(\theta)}{h_{0t-k}} \frac{h_{0t-k}}{h_{0t}} \right) \right\|_q \\ & \leq T_2 \left( E \left( \sup_{\theta \in \Theta_u^q} \left( \frac{h_{t-k}^\epsilon(\theta)}{h_{0t-k}} \right)^q \left( \frac{h_{0t-k}}{h_{0t}} \right)^q \right) \right)^{\frac{1}{q}} \quad \text{mit Lemma 8(4)} \\ & \leq T_2 \left( E \left( \sup_{\theta \in \Theta_u^q} \left( \frac{h_{t-k}^\epsilon(\theta)}{h_{0t-k}} \right)^q E \left( \sup_{\theta \in \Theta_u^q} \left( \frac{h_{0t-k}}{h_{0t}} \right)^q \mid \mathcal{F}_{t-k-1} \right) \right) \right)^{\frac{1}{q}} \\ & \leq T_2 \left( \frac{\mathcal{R}_u^{\frac{1}{q}}}{\beta_0} \right)^k \left( E \left( \sup_{\theta \in \Theta_u^q} \left( \frac{h_{t-k}^\epsilon(\theta)}{h_{0t-k}} \right)^q \right) \right)^{\frac{1}{q}} \quad \text{mit Lemma 8(1)} \end{aligned}$$

$$\leq T_2 \left( \frac{\mathcal{R}_u^{\frac{1}{q}}}{\beta_0} \right)^k T_4 \quad \text{mit Lemma 10(1)}$$

□

Bemerkung: Das folgende Lemma stellt die punktweise stochastische Konvergenz von  $L_n(\theta)$  gegen  $L(\theta)$  fest und bereitet damit Lemma 12(5) vor.

**Lemma 11** *Unter den gemachten Voraussetzungen gilt für alle  $\theta \in \Theta_1$ :*

(1)

$$E \left( \sup_{\theta \in \Theta_1} \frac{e_t^2(\theta)}{h_t(\theta)} \right) \leq T_5 := B \left( \frac{2\tilde{m} + \tilde{m}^2}{\omega_l} + (1 + 2\tilde{m})T_3 \right) < \infty$$

(2)

$$\sup_{\theta \in \Theta_1} E \left( \frac{e_t^2(\theta)}{h_t(\theta)} \right) \leq T_5$$

(3)

$$L_n(\theta) \rightarrow_p L(\theta) := El_t(\theta) \quad \text{punktweise}$$

**Beweis von Lemma 11:**

(1)

$$\begin{aligned} & E \left( \sup_{\theta \in \Theta_1} \frac{e_t^2(\theta)}{h_t(\theta)} \right) \\ & \leq BE \left( \sup_{\theta \in \Theta_1} \frac{\epsilon_t^2 + 2\tilde{m}|\epsilon_t| + \tilde{m}^2}{h_t^\epsilon(\theta)} \right) \\ & \leq BE \left( \sup_{\theta \in \Theta_1} \frac{\epsilon_t^2 + 2\tilde{m}\epsilon_t^2 + 2\tilde{m} + \tilde{m}^2}{h_t^\epsilon(\theta)} \right), \quad \text{da } |\epsilon_t| \leq \epsilon_t^2 + 1 \\ & \leq B \left( \frac{2\tilde{m} + \tilde{m}^2}{\omega_l} + (1 + 2\tilde{m})E \left( \sup_{\theta \in \Theta_1} \frac{h_{0t}}{h_t^\epsilon(\theta)} \right) \right), \quad \text{da } \epsilon_t^2 = h_{0t}z_t^2 \\ & \leq B \left( \frac{2\tilde{m} + \tilde{m}^2}{\omega_l} + (1 + 2\tilde{m})T_3 \right) \quad \text{mit Lemma 9(2)}. \end{aligned}$$

(2) folgt unmittelbar aus (1), da  $\sup_{\theta \in \Theta_1} E \left( \frac{e_t^2(\theta)}{h_t(\theta)} \right) \leq E \left( \sup_{\theta \in \Theta_1} \frac{e_t^2(\theta)}{h_t(\theta)} \right)$ .

(3) Die Behauptung folgt aus dem Ergodentheorem, da

$$\begin{aligned} & \sup_{\theta \in \Theta_1} E|l_t(\theta)| \\ & \leq \sup_{\theta \in \Theta_1} E|\ln(h_t(\theta))| + \sup_{\theta \in \Theta_1} E \left( \frac{e_t^2(\theta)}{h_t(\theta)} \right) \\ & \leq \frac{1}{p} \sup_{\theta \in \Theta_1} E|\ln(h_t(\theta)^p)| + T_5 \end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{p} \sup_{\theta \in \Theta_1} (E[(-1) \ln(h_t(\theta)^p) 1_{\{h_t(\theta) \leq 1\}}] + E[\ln(h_t(\theta)^p) 1_{\{h_t(\theta) \geq 1\}}]) + T_5 \\
&\leq \frac{1}{p} |\ln(\omega_l^p)| + \frac{1}{p} \sup_{\theta \in \Theta_1} E[(h_t(\theta)^p - 1) 1_{\{h_t(\theta) \geq 1\}}] + T_5 \\
&\leq \frac{1}{p} |\ln(\omega_l^p)| + \frac{1}{p} \sup_{\theta \in \Theta_1} E h_t(\theta)^p + T_5 \\
&\leq \frac{1}{p} |\ln(\omega_l^p)| + \frac{1}{p} T_p + T_5 \\
&= T_6 + T_5 < \infty
\end{aligned}$$

mit  $T_6 := \frac{1}{p} |\ln(\omega_l^p)| + \frac{1}{p} T_p$ .

□

**Lemma 12** *Unter den gemachten Voraussetzungen gilt mit*

$$h_{\xi t}(\theta) := \frac{\partial h_t(\theta)}{\partial \xi} \frac{1}{h_t(\theta)}$$

für  $\xi \in \{\alpha, \beta, \gamma, \delta, \omega\}$ :

(1) *Es gibt eine universelle Konstante  $M < \infty$ , so daß*

$$\sup_{\theta \in \Theta} \left| \frac{\partial h_t(\theta)}{\partial \xi} \frac{1}{h_t(\theta)} \right| \leq M \quad \text{fast sicher}$$

für  $\xi \in \{\alpha_i, \gamma, \omega\}, i = 1, \dots, l_1$ .

(2) *Für  $1 \leq q \leq 6$  und  $i = 1, \dots, l_2$  gilt:*

$$\left\| \sup_{\theta \in \Theta_q} |h_{\beta t}(\theta)| \right\|_q \leq \frac{B^2 \tau_2 \beta_0}{\tau_1} \max[T_4 T_2 \frac{1}{\eta_u^2}, \frac{1}{\eta_l^2}] + \frac{K}{\omega_l (1 - \beta_u)^2} =: H_1 < \infty$$

$$\left\| \sup_{\theta \in \Theta_q} |h_{\delta_i t}(\theta)| \right\|_q \leq \frac{K}{\omega_l (1 - \beta_u)} =: T_\delta$$

(3)

$$E(\sup_{\theta \in \Theta_2} |\nabla l_t(\theta)|) < \infty$$

(4) *Die Funktion  $L(\theta)$  ist gleichmäßig stetig in  $\theta \in \Theta_2$ .*

(5)

$$\sup_{\theta \in \Theta_2} |L_n(\theta) - L(\theta)| \rightarrow_p 0$$

**Beweis von Lemma 12:**

(1)  $h_t(\theta) = \omega + \sum_{k=0}^{\infty} \beta^k [g(\alpha, e_{t-k-1}(\theta)) e_{t-k-1}^2(\theta) + f(\delta, r_{t-k-1})]$  fast sicher.  
Daraus folgt für  $\omega$ :

$$\sup_{\theta \in \Theta} \left| \frac{\partial h_t(\theta)}{\partial \omega} \frac{1}{h_t(\theta)} \right| = \sup_{\theta \in \Theta} \left| \frac{1}{h_t(\theta)} \right| \leq \frac{1}{\omega_l} < \infty$$

für  $\alpha_i$ :

$$\begin{aligned} \sup_{\theta \in \Theta} \left| \frac{\partial h_t(\theta)}{\partial \alpha_i} \frac{1}{h_t(\theta)} \right| &= \sup_{\theta \in \Theta} \left| \frac{\sum_{k=0}^{\infty} \beta^k \frac{\partial g(\alpha, e_{t-k-1}(\theta))}{\partial \alpha_i} e_{t-k-1}^2(\theta)}{h_t(\theta)} \right| \\ &\leq \sup_{\theta \in \Theta} \left| \frac{\tau_3 \sum_{k=0}^{\infty} \beta^k e_{t-k-1}^2(\theta)}{\tau_1 \sum_{k=0}^{\infty} \beta^k e_{t-k-1}^2(\theta)} \right| = \frac{\tau_3}{\tau_1} < \infty \end{aligned}$$

für  $\gamma$ :

$$\begin{aligned} &\sup_{\theta \in \Theta} \left| \frac{\partial h_t(\theta)}{\partial \gamma} \frac{1}{h_t(\theta)} \right| \\ &= \sup_{\theta \in \Theta} \left| \frac{-2 \sum_{k=0}^{\infty} \beta^k g(\alpha, e_{t-k-1}(\theta)) e_{t-k-1}(\theta)}{h_t(\theta)} \right. \\ &\quad \left. - \frac{\sum_{k=0}^{\infty} \beta^k \frac{\partial g}{\partial e_{t-k-1}(\theta)}(\alpha, e_{t-k-1}(\theta)) e_{t-k-1}^2(\theta)}{h_t(\theta)} \right| \\ &\leq \frac{2\tau_2}{\omega_l(1-\beta_u)} + \sup_{\theta \in \Theta} \left| \frac{2(\tau_2 + \tau_3) \sum_{k=0}^{\infty} \beta^k e_{t-k-1}^2(\theta)}{\tau_1 \sum_{k=0}^{\infty} \beta^k e_{t-k-1}^2(\theta)} \right| \\ &= \frac{2\tau_2}{\omega_l(1-\beta_u)} + \frac{2(\tau_2 + \tau_3)}{\tau_1} \end{aligned}$$

$M < \infty$  kann als Maximum dieser oberen Schranken gewählt werden.

(2) Allgemein gilt:

$$h_{\beta t}(\theta) = \frac{\partial h_t(\theta)}{\partial \beta} \frac{1}{h_t(\theta)} = \left( \sum_{k=1}^{\infty} k \beta^{k-1} [g(\alpha, e_{t-k-1}(\theta)) e_{t-k-1}^2(\theta) + f(\delta, r_{t-k-1})] \right) \frac{1}{h_t(\theta)}$$

$$\begin{aligned} \Rightarrow \left\| \sup_{\theta \in \Theta_q} |h_{\beta t}(\theta)| \right\|_q &\leq B^2 \frac{\tau_2}{\tau_1} \left\| \sup_{\theta \in \Theta_q} \left( \sum_{k=1}^{\infty} k \beta^{k-1} \left( \frac{h_{t-k}^\epsilon(\theta)}{h_t^\epsilon(\theta)} \right) \right) \right\|_q + \frac{K}{\omega_l(1-\beta_u)^2} \\ &\leq B^2 \frac{\tau_2}{\tau_1} \sum_{k=1}^{\infty} k \beta^{k-1} \left\| \sup_{\theta \in \Theta_q} \left( \frac{h_{t-k}^\epsilon(\theta)}{h_t^\epsilon(\theta)} \right) \right\|_q + \frac{K}{\omega_l(1-\beta_u)^2} \end{aligned}$$

(a) Für  $\beta \geq \beta_0$  folgt mit Lemma 10 (2):

$$\begin{aligned} &\left\| \sup_{\theta \in \Theta_q} |h_{\beta t}(\theta)| \right\|_q + \frac{K}{\omega_l(1-\beta_u)^2} \\ &\leq B^2 \frac{\tau_2}{\tau_1} T_4 T_2 \sum_{k=1}^{\infty} k \beta^{k-1} \left( \frac{\mathcal{R}_u^{\frac{1}{q}}}{\beta_0} \right)^k + \frac{K}{\omega_l(1-\beta_u)^2} \\ &\leq B^2 \frac{\tau_2}{\tau_1} T_4 T_2 \left( \frac{\mathcal{R}_u^{\frac{1}{q}}}{\beta_0} \right) \sum_{k=1}^{\infty} k \left( \frac{\beta_u \mathcal{R}_u^{\frac{1}{q}}}{\beta_0} \right)^{k-1} + \frac{K}{\omega_l(1-\beta_u)^2} \end{aligned}$$

$$\begin{aligned}
&\leq B^2 \frac{\tau_2}{\tau_1} T_4 T_2 \frac{\mathcal{R}_u^{\frac{1}{q}}}{\beta_0} \frac{1}{\left(1 - \frac{\beta_u \mathcal{R}_u^{\frac{1}{q}}}{\beta_0}\right)^2} + \frac{K}{\omega_l (1 - \beta_u)^2} \\
&\leq B^2 \frac{\tau_2}{\tau_1} T_4 T_2 \frac{\beta_0}{\eta_u^2} + \frac{K}{\omega_l (1 - \beta_u)^2}
\end{aligned}$$

(b) Für  $\beta \leq \beta_0$  folgt mit Lemma 9 (1):

$$\begin{aligned}
&\| \sup_{\theta \in \Theta_q} |h_{\beta t}(\theta)| \|_q \\
&\leq B^2 \frac{\tau_2}{\tau_1} \sum_{k=1}^{\infty} k \beta_0^{k-1} \left( \frac{\mathcal{R}_l^{\frac{1}{q}}}{\beta_{ql}} \right)^k + \frac{K}{\omega_l (1 - \beta_u)^2} \\
&\leq B^2 \frac{\tau_2}{\tau_1} \frac{\mathcal{R}_l^{\frac{1}{q}}}{\beta_{ql}} \frac{1}{\left(1 - \frac{\beta_0 \mathcal{R}_l^{\frac{1}{q}}}{\beta_{ql}}\right)^2} + \frac{K}{\omega_l (1 - \beta_u)^2} \\
&\leq B^2 \frac{\tau_2}{\tau_1} \frac{\beta_0}{\eta_l^2} + \frac{K}{\omega_l (1 - \beta_u)^2}
\end{aligned}$$

Für die Ableitung nach  $\delta_i$  erhalten wir:

$$\begin{aligned}
&\| \sup_{\theta \in \Theta_q} |h_{\delta_i t}(\theta)| \|_q \\
&\leq \| \sup_{\theta \in \Theta_q} \left| \frac{\sum_{k=0}^{\infty} \beta^k \frac{\partial}{\partial \delta_i} f(\delta, r_{t-k-1})}{h_t(\theta)} \right| \|_q \\
&= \frac{K}{\omega_l (1 - \beta_u)} < \infty
\end{aligned}$$

(3) Für  $\xi \in \{\alpha, \omega\}$  gilt für die partiellen Ableitungen von  $l_t(\theta)$ :

$$\begin{aligned}
E \left( \sup_{\theta \in \Theta_2} \left| \frac{\partial l_t(\theta)}{\partial \xi} \right| \right) &= E \left( \sup_{\theta \in \Theta_2} \left| \left( \frac{e_t^2(\theta)}{h_t(\theta)} - 1 \right) \frac{\partial h_t(\theta)}{\partial \xi} \frac{1}{h_t(\theta)} \right| \right) \\
&\leq E \left( \sup_{\theta \in \Theta_2} \left| \left( \frac{e_t^2(\theta)}{h_t(\theta)} - 1 \right) M \right| \right) \\
&\leq M(T_5 + 1) < \infty \quad \text{mit Lemma 11(1)}
\end{aligned}$$

Für die Ableitung nach  $\gamma$  ergibt sich wegen

$$E \left( \sup_{\theta \in \Theta} \left| \frac{e_t(\theta)}{h_t(\theta)} \right| \right) \leq \frac{1}{\omega_l} + E \left( \sup_{\theta \in \Theta} \left| \frac{e_t^2(\theta)}{h_t(\theta)} \right| \right) \leq \frac{1}{\omega_l} + T_5$$

die Abschätzung:

$$E \left( \sup_{\theta \in \Theta_2} \left| \frac{\partial l_t(\theta)}{\partial \gamma} \right| \right) = E \left( \sup_{\theta \in \Theta_2} \left| \left( \frac{e_t^2(\theta)}{h_t(\theta)} - 1 \right) \frac{\partial h_t(\theta)}{\partial \gamma} \frac{1}{h_t(\theta)} - 2 \frac{e_t(\theta)}{h_t(\theta)} \right| \right)$$

$$\begin{aligned}
&\leq E \left( \sup_{\theta \in \Theta_2} \left| \left( \frac{e_t^2(\theta)}{h_t(\theta)} - 1 \right) M + 2 \left( \frac{1}{\omega_l} + T_5 \right) \right| \right) \\
&\leq M(H_1 + 1) + 2 \left( \frac{1}{\omega_l} + T_5 \right) < \infty
\end{aligned}$$

Für die Ableitung nach  $\beta$  erhalten wir:

$$\begin{aligned}
&E \left( \sup_{\theta \in \Theta_2} \left| \frac{\partial l_t(\theta)}{\partial \beta} \right| \right) \\
&= E \left( \sup_{\theta \in \Theta_2} \left| \left( \frac{e_t^2(\theta)}{h_t(\theta)} - 1 \right) \frac{\partial h_t(\theta)}{\partial \beta} \frac{1}{h_t(\theta)} \right| \right) \\
&= E \left( \sup_{\theta \in \Theta_2} \left| \frac{\epsilon_t^2 + 2m\epsilon_t + m^2}{h_t(\theta)} h_{\beta t}(\theta) + h_{\beta t}(\theta) \right| \right) \\
&\leq E \left( \sup_{\theta \in \Theta_2} \left| \frac{\epsilon_t^2(1 + 2\tilde{m}) + (2\tilde{m} + \tilde{m}^2)}{h_t(\theta)} h_{\beta t}(\theta) \right| \right) + H_1 \\
&\leq (1 + 2\tilde{m}) E \left( \sup_{\theta \in \Theta_2} \left| \frac{h_{0t}}{h_t(\theta)} h_{\beta t}(\theta) \right| \right) + H_1 \left( 1 + \frac{2\tilde{m} + \tilde{m}^2}{\omega_l} \right) \\
&\leq B(1 + 2\tilde{m}) \left\| \sup_{\theta \in \Theta_2} \frac{h_{0t}}{h_t^\epsilon(\theta)} \right\|_2 \left\| \sup_{\theta \in \Theta_2} h_{\beta t}(\theta) \right\|_2 + H_1 \left( 1 + \frac{2\tilde{m} + \tilde{m}^2}{\omega_l} \right) \\
&\leq B(1 + 2\tilde{m}) T_3 H_1 + H_1 \left( 1 + \frac{2\tilde{m} + \tilde{m}^2}{\omega_l} \right) \\
&\quad \text{mit Lemma 9 (2) und Lemma 12 (2)}
\end{aligned}$$

Schließlich erhalten wir für die Ableitungen nach  $\delta_i, i = 1, \dots, l_2$ :

$$\begin{aligned}
&E \left( \sup_{\theta \in \Theta_2} \left| \frac{\partial l_t(\theta)}{\partial \delta_i} \right| \right) \\
&\leq B(1 + 2\tilde{m}) \left\| \sup_{\theta \in \Theta_2} \frac{h_{0t}}{h_t^\epsilon(\theta)} \right\|_2 \left\| \sup_{\theta \in \Theta_2} h_{\delta_i t}(\theta) \right\|_2 + T_\delta \left( 1 + \frac{2\tilde{m} + \tilde{m}^2}{\omega_l} \right) \\
&\leq B(1 + 2\tilde{m}) T_3 T_\delta + T_\delta \left( 1 + \frac{2\tilde{m} + \tilde{m}^2}{\omega_l} \right)
\end{aligned}$$

Insgesamt ergibt sich also  $E(\sup_{\theta \in \Theta_2} |\nabla l_t(\theta)|) < \infty$ .

(4) Allgemein gilt, daß

$$|l_t(\theta) - l_t(\theta')| \leq \sup_{\theta \in \Theta_2} |\nabla l_t(\theta)| |\theta - \theta'| \quad \text{fast sicher.}$$

Mit Punkt (3) dieses Lemmas impliziert dies für  $\theta \in \Theta_2$ , daß die Funktion  $L(\theta) = El_t(\theta)$  gleichmäßig stetig ist in  $\theta$ , denn

$$|El_t(\theta) - El_t(\theta')| \leq E \sup_{\theta \in \Theta_2} |\nabla l_t(\theta)| |\theta - \theta'|.$$

(5) Wie in Lemma 6 ist zu zeigen, daß es ein  $M(\epsilon, \eta)$  gibt, so daß für alle  $n \geq M$  gilt:

$$P\{\sup_{\theta \in \Theta_2} |L_n(\theta) - L(\theta)| > \epsilon\} < \eta$$

Wähle hierzu zunächst ein  $\delta > 0$  mit

$$\delta < \frac{\eta\epsilon}{2E(\sup_{\theta \in \Theta_2} |\nabla l_t(\theta)|)}.$$

Sei weiter

$$B_\delta(\theta') := \{\theta \in \Theta_2 : |\theta - \theta'| < \delta\}$$

die offene  $\delta$ -Kugel um  $\theta'$  in  $\Theta_2$ . Da  $\Theta_2$  kompakt ist gibt es eine endliche Menge  $\{\theta_1, \dots, \theta_J\} \subset \Theta_2$  derart, daß  $\Theta_2 \subset \bigcup_{j=1}^J B_\delta(\theta_j)$ . Damit erhalten wir:

$$\begin{aligned} & P\{\sup_{\theta \in \Theta_2} \left| \frac{1}{n} \sum_{t=1}^n l_t(\theta) - El_t(\theta) \right| > \epsilon\} \\ & \leq \frac{1}{\epsilon n} E \left( \sup_{\theta \in \Theta_2} \left| \sum_{t=1}^n (l_t(\theta) - El_t(\theta)) \right| \right) \\ & = \frac{1}{\epsilon n} E \left[ \max_{j=1, \dots, J} \sup_{\theta \in B_\delta(\theta_j)} \left| \sum_{t=1}^n (l_t(\theta) - El_t(\theta)) - \sum_{t=1}^n (l_t(\theta_j) - El_t(\theta_j)) \right. \right. \\ & \quad \left. \left. + \sum_{t=1}^n (l_t(\theta_j) - El_t(\theta_j)) \right| \right] \\ & \leq \frac{1}{\epsilon n} E \left( \sup_{\theta \in \Theta_2} \sup_{\theta' \in B_\delta(\theta)} \left| \sum_{t=1}^n (l_t(\theta) - El_t(\theta)) - \sum_{t=1}^n (l_t(\theta') - El_t(\theta')) \right| \right) \\ & \quad + \frac{1}{\epsilon n} E \left( \max_{j=1, \dots, J} \left| \sum_{t=1}^n (l_t(\theta_j) - El_t(\theta_j)) \right| \right) \\ & \leq \frac{1}{\epsilon n} E \left( \sup_{\theta \in \Theta_2} \sup_{\theta' \in B_\delta(\theta)} \left| \sum_{t=1}^n (l_t(\theta) - l_t(\theta')) \right| \right) \\ & \quad + \frac{1}{\epsilon} E \left( \sup_{\theta \in \Theta_2} \sup_{\theta' \in B_\delta(\theta)} |El_t(\theta) - El_t(\theta')| \right) \\ & \quad + \frac{1}{\epsilon n} E \left( \max_{j=1, \dots, J} \left| \sum_{t=1}^n (l_t(\theta_j) - El_t(\theta_j)) \right| \right) \\ & \leq \frac{1}{\epsilon n} \sum_{t=1}^n E \sup_{\theta \in \Theta_2} |\nabla l_t(\theta)| \delta + \frac{1}{\epsilon} E \sup_{\theta \in \Theta_2} |\nabla l_t(\theta)| \delta \\ & \quad + \frac{1}{\epsilon} E \left( \max_{j=1, \dots, J} \left| \left( \frac{1}{n} \sum_{t=1}^n l_t(\theta_j) \right) - El_t(\theta_j) \right| \right) \\ & = \frac{2}{\epsilon} E \sup_{\theta \in \Theta_2} |\nabla l_t(\theta)| \delta + \frac{1}{\epsilon} E \left( \max_{j=1, \dots, J} \left| \left( \frac{1}{n} \sum_{t=1}^n l_t(\theta_j) \right) - El_t(\theta_j) \right| \right) \end{aligned}$$

$$< \eta + \frac{1}{\epsilon} E \left( \max_{j=1, \dots, J} \left| \left( \frac{1}{n} \sum_{t=1}^n l_t(\theta_j) \right) - E l_t(\theta_j) \right| \right).$$

Da der zweite Term nach dem Ergodentheorem gegen 0 geht, existiert also eine Zahl  $M$  wie gefordert (vgl. auch Andrews 1992).

□

Bemerkung: Lemma 13 ist entscheidend für die Konsistenz des QML-Schätzers und wird nach einer Idee von Lumsdaine (1996, Theorem 1) bewiesen.

**Lemma 13** *Unter den gemachten Voraussetzungen gilt:  
 $L(\theta)$  hat ein eindeutiges Maximum in  $\theta_0$*

**Beweis von Lemma 13:**

Da  $\sup_{\theta \in \Theta} E(|l_t(\theta)|) < \infty$ , ist  $-\infty < L(\theta_0), L(\theta) < \infty$ .

Die Aussage des Lemmas impliziert also

$$\sup_{\theta \in \Theta} [L(\theta) - L(\theta_0)] = 0.$$

Also

$$\begin{aligned} & \sup_{\theta \in \Theta} [L(\theta) - L(\theta_0)] \\ &= -E \left( \ln \frac{h_t(\theta)}{h_{0t}} \right) - E \left( \frac{e_t^2(\theta)}{h_t(\theta)} - \frac{\epsilon_t^2}{h_{0t}} \right) \\ &= -E \left( \ln \frac{h_t(\theta)}{h_{0t}} \right) - E \left( \frac{\epsilon_t^2 + 2m\epsilon_t + m^2}{h_t(\theta)} - \frac{\epsilon_t^2}{h_{0t}} \right) \\ &= -E \left( \ln \frac{h_t(\theta)}{h_{0t}} \right) - E \left( \frac{m^2}{h_t(\theta)} \right) - E \left( \frac{\epsilon_t^2(h_{0t} - h_t(\theta))}{h_{0t}h_t(\theta)} \right) \\ &= -E \left( \ln \frac{h_t(\theta)}{h_{0t}} \right) - E \left( \frac{m^2}{h_t(\theta)} \right) - E \left( h_{0t} E(z_t^2 | \mathcal{F}_{t-1}) \frac{(h_{0t} - h_t(\theta))}{h_{0t}h_t(\theta)} \right) \\ &= -E \left( \ln \frac{h_t(\theta)}{h_{0t}} \right) - E \left( \frac{(\gamma_0 - \gamma)^2}{h_t(\theta)} \right) - E \left( \frac{h_{0t}}{h_t(\theta)} \right) + 1 \\ &= E \left( \underbrace{\ln \frac{h_{0t}}{h_t(\theta)} - \frac{h_{0t}}{h_t(\theta)}}_{\leq 0} \right) + 1 - E \left( \underbrace{\frac{(\gamma_0 - \gamma)^2}{h_t(\theta)}}_{\leq 0} \right) \end{aligned}$$

Der erste Term ist  $\leq 0$ , da  $f(x) = \ln x - x \leq 0$ . Dies wird maximiert für  $x = 1$ , da  $f'(x) = \frac{1}{x} - 1$  und  $f''(x) = -\frac{1}{x^2} < 0$ . Da nach Voraussetzung fast sicher  $h_t(\theta) \neq h_{0t}$  für  $\theta \neq \theta_0 \in \Theta$ , nimmt  $\frac{h_{0t}}{h_t(\theta)}$  nur für  $\theta = \theta_0$  fast sicher den Wert 1 an.

Der zweite Term wird offenbar für  $\gamma = \gamma_0$  fast sicher maximiert.

□

Bemerkung: Wir können nun Theorem 1 beweisen.

**Beweis von Theorem 1:**

Nach Lemma 13 nimmt  $L(\theta)$  an der Stelle des wahren Parametervektors  $\theta_0$  sein eindeutiges Maximum an. Wähle dann eine beliebige offene Umgebung  $U$  von  $\theta_0$ . Wir zeigen, daß  $\lim_{n \rightarrow \infty} P\{\hat{\theta}_n \in U\} = 1$ . Definiere dazu  $\eta := \min_{\theta \in U^c \cap \Theta_2} (L(\theta_0) - L(\theta))$  mit  $\min_{\emptyset} := \infty$ . Da die Menge  $U^c \cap \Theta_2$  kompakt und da  $L(\theta)$  nach Lemma 12 (4) stetig ist, wird das Minimum angenommen.

Bezeichne  $A_n$  das Ereignis:  $\sup_{\theta \in \Theta_2} |L(\theta) - L_n^*(\theta)| < \frac{\eta}{2}$ . Da  $L_n^*(\theta)$  nach Voraussetzung meßbar ist, ist auch  $A_n$  meßbar.

Beachte zunächst, daß sich die Ergebnisse von Lemma 6 und 12 (5) zusammenfassen lassen zu der Aussage, daß die empirische Likelihoodfunktion  $L_n^*(\theta)$  gleichmäßig stochastisch gegen  $L(\theta)$  konvergiert, denn es gilt:

$$\sup_{\theta \in \Theta_2} |L_n^*(\theta) - L(\theta)| \leq \sup_{\theta \in \Theta_2} |L_n^*(\theta) - L_n(\theta)| + \sup_{\theta \in \Theta_2} |L_n(\theta) - L(\theta)| \rightarrow_p 0.$$

Daraus folgt unmittelbar:

$$1 = \lim_{n \rightarrow \infty} P\{A_n\}.$$

Das Eintreten von  $A_n$  impliziert:

$$L(\hat{\theta}_n) > L_n^*(\hat{\theta}_n) - \frac{\eta}{2} \geq L_n^*(\theta_0) - \frac{\eta}{2},$$

nach Definition von  $\hat{\theta}_n$ . Ebenso erhalten wir

$$0 > L(\theta_0) - L_n^*(\theta_0) - \frac{\eta}{2}.$$

Gegeben  $A_n$  erhalten wir also insgesamt:

$$\eta > L(\theta_0) - L(\hat{\theta}_n),$$

also  $\hat{\theta}_n \in U$  nach Definition von  $\eta$ .

Damit gilt also

$$1 = \lim_{n \rightarrow \infty} P\{A_n\} \leq \lim_{n \rightarrow \infty} P\{\hat{\theta}_n \in U\}$$

und somit die im Theorem gemachte Konsistenzaussage, da  $U$  beliebig war. (Vgl. z.B. Amemiya 1985, Theorem 4.1.1.).

□

Bemerkung: Theorem 4 beweist die Konsistenzaussage unter einer schwächeren Parameterrestriktion, aber unter der stärkeren Annahme an den wahren Prozeß, daß  $Eh_{0t} < \infty$ .

**Theorem 4** Sei  $\sup_{t \in \mathbf{Z}} E[(\beta_0 + g(\alpha_0, \epsilon_t)z_t^2) | \mathcal{F}_{t-1}] =: \tilde{\kappa} < 1$  und  $f \equiv 0$ . Dann gilt unter Voraussetzung 1 und 2 für jedes  $\tilde{\theta}_n \in \tilde{A} := \arg \max_{\theta \in \Theta} L_n^*(\theta)$ :

$$\tilde{\theta}_n \rightarrow_p \theta_0$$

**Beweis von Theorem 4:**

Die Parameterrestriktion von Theorem 1 war nur für die punktweise und gleichmäßige stochastische Konvergenz von  $L_n(\theta)$  gegen  $L(\theta)$  von Bedeutung. Im Rahmen von Theorem 4 folgt die punktweise stochastische Konvergenz wieder aus dem Ergodentheorem, denn mit der im Beweis von Lemma 11(3) hergeleiteten Abschätzung  $\sup_{\theta \in \Theta} E |\ln(h_t(\theta))| \leq \frac{1}{p} |\ln(\omega_l^p)| + \frac{1}{p} T_p = T_6$  ist:

$$\begin{aligned} \sup_{\theta \in \Theta} E |l_t(\theta)| &\leq \sup_{\theta \in \Theta} E |\ln(h_t(\theta))| + \sup_{\theta \in \Theta} E \left( \frac{e_t^2(\theta)}{h_t(\theta)} \right) \\ &\leq T_6 + \frac{1}{\omega_l} \sup_{\theta \in \Theta} E e_t^2(\theta) \\ &\leq T_6 + \frac{1}{\omega_l} \sup_{\theta \in \Theta} E (\epsilon_t + m)^2 \\ &\leq T_6 + \frac{1}{\omega_l} (\tilde{m}^2 + E \epsilon_t^2) \\ &\leq T_6 + \frac{1}{\omega_l} \left( \tilde{m}^2 + \frac{\omega_0(1-\beta_0)}{1-\tilde{\kappa}} \right) < \infty, \end{aligned}$$

wobei verwendet wurde, daß für alle  $t \in \mathbf{Z}$  gilt:

$$\begin{aligned} E \epsilon_t^2 &= E(z_t^2 h_{0t}) = E(h_{0t} E(z_t^2 | \mathcal{F}_{t-1})) \\ &= E(h_{0t}) = E[(1-\beta_0)\omega_0 + (g(\alpha_0, \epsilon_{t-1})z_{t-i}^2 + \beta_0)h_{0t-1}] \\ &\leq (1-\beta_0)\omega_0 + \tilde{\kappa} E h_{0t-1} \\ &= \frac{\omega_0(1-\beta_0)}{1-\tilde{\kappa}} < \infty, \end{aligned}$$

Damit erhält man auch:

$$\begin{aligned} E \sup_{\theta \in \Theta} h_t(\theta) &= E \sup_{\theta \in \Theta} \left( \omega + \sum_{k=0}^{\infty} \beta^k g(\alpha, \epsilon_{t-k-1}) \epsilon_{t-k-1}^2 \right) \\ &\leq \sup_{\theta \in \Theta} \left( \omega + \tau_2 \sum_{k=0}^{\infty} \beta^k E (\epsilon_{t-k-1} + m)^2 \right) \\ &= \sup_{\theta \in \Theta} \left( \omega + \tau_2 \sum_{k=0}^{\infty} \beta^k (m^2 + E \epsilon_{t-k-1}^2) \right) \\ &\leq \omega_u + \tau_2 \frac{\tilde{m}^2 + \frac{\omega_0(1-\beta_0)}{1-\tilde{\kappa}}}{1-\beta_u} =: \bar{h} \end{aligned}$$

Es bleibt zu zeigen, daß  $E(\sup_{\theta \in \Theta} |\nabla l_t(\theta)|) < \infty$ :

Für  $\xi \in \{\alpha, \gamma, \delta, \omega\}$  folgt mit Lemma 12  $E \left( \sup_{\theta \in \Theta} \left| \frac{\partial l_t(\theta)}{\partial \xi} \right| \right) < \infty$ , denn

$$E\left(\sup_{\theta \in \Theta} \left| \frac{e_t^2(\theta)}{h_t(\theta)} \right| \right) \leq \frac{1}{\omega_l \tau_1} E(\sup_{\theta \in \Theta} h_{t+1}(\theta)) \leq \frac{\bar{h}}{\omega_l \tau_1} < \infty.$$

$$E\left(\sup_{\theta \in \Theta} \left| \frac{\partial l_t(\theta)}{\partial \beta} \right| \right) < \infty \text{ folgt so:}$$

$$E\left(\sup_{\theta \in \Theta} \left| \frac{\partial l_t(\theta)}{\partial \beta} \right| \right) \leq E(\sup_{\theta \in \Theta} |h_{\beta t}(\theta)|) + E\left(\sup_{\theta \in \Theta} \left| \frac{e_t^2(\theta)}{h_t(\theta)} h_{\beta t}(\theta) \right| \right)$$

Für den ersten Term erhält man:

$$\begin{aligned} E(\sup_{\theta \in \Theta} |h_{\beta t}(\theta)|) &= E\left(\sup_{\theta \in \Theta} \left( \sum_{k=0}^{\infty} k \beta^{k-1} g(e_{t-k-1}(\theta)) e_{t-k-1}^2(\theta) \right) \frac{1}{h_t(\theta)} \right) \\ &\leq \frac{\tau_2}{\tau_1 \omega_l} E\left(\sup_{\theta \in \Theta} \left( \sum_{k=0}^{\infty} k \beta^{k-1} h_{t-k}(\theta) \right)\right) \\ &\leq \frac{\tau_2}{\tau_1 \omega_l} \sum_{k=0}^{\infty} k \beta^{k-1} E(\sup_{\theta \in \Theta} h_{t-k}(\theta)) \\ &\leq \frac{\tau_2 \bar{h}}{\tau_1 \omega_l (1 - \beta_u)^2} < \infty \end{aligned}$$

Für den 2. Term hat man:

$$\begin{aligned} E\left(\sup_{\theta \in \Theta} \left| \frac{e_t^2(\theta)}{h_t(\theta)} h_{\beta t}(\theta) \right| \right) &= E\left(\sup_{\theta \in \Theta} \left| \frac{\epsilon_t^2 + m^2}{h_t(\theta)} h_{\beta t}(\theta) \right| \right) \\ &\leq \underbrace{\frac{\tilde{m}^2}{\omega_l} E(\sup_{\theta \in \Theta} |h_{\beta t}(\theta)|)}_{< \infty} + BE\left(\sup_{\theta \in \Theta} \left| \frac{h_{0t}}{h_t^\epsilon(\theta)} h_{\beta t}(\theta) \right| \right) \end{aligned}$$

Es bleibt also zu zeigen:  $E\left(\sup_{\theta \in \Theta} \left| \frac{h_{0t}}{h_t^\epsilon(\theta)} h_{\beta t}(\theta) \right| \right) < \infty$ .

Für  $\beta \geq \beta_0$  gilt mit Lemma 8(4):

$$E\left(\sup_{\theta \in \Theta} \left| \frac{h_{0t}}{h_t^\epsilon(\theta)} h_{\beta t}(\theta) \right| \right) < T_2 E(\sup_{\theta \in \Theta} |h_{\beta t}(\theta)|) < \infty$$

Für  $\beta \leq \beta_0$  können wir  $\frac{h_{0t}}{h_t^\epsilon(\theta)}$  abschätzen durch

$$\begin{aligned} &\frac{h_{0t}}{h_t^\epsilon(\theta)} \\ &\leq \frac{\omega_0(1 - \beta_0^k) + \tau_2 \sum_{i=0}^{k-1} \beta_0^i \epsilon_{t-i-1}^2 + \beta_0^k h_{0t-k}}{h_t^\epsilon(\theta)} \\ &\leq \frac{\omega_0(1 - \beta_0^k)}{\omega} + \frac{\tau_2}{\tau_1} \sum_{i=0}^{k-1} \left( \frac{\beta_0}{\beta} \right)^i + \beta_0^k \frac{h_{0t-k}}{\omega}, \quad \text{da } h_t^\epsilon(\theta) \geq \beta^i \epsilon_{t-i-1}^2 \\ &\leq \frac{\omega_0}{\omega} + \frac{\tau_2}{\tau_1} k \left( \frac{\beta_0}{\beta} \right)^k + \beta_0^k \frac{h_{0t-k}}{\omega}, \quad \text{da } \frac{\beta_0}{\beta} \geq 1 \end{aligned}$$

Damit ergibt sich:

$$\begin{aligned}
& E \left( \sup_{\theta \in \Theta} \left| \frac{h_{0t}}{h_t^\xi(\theta)} h_{\beta t}(\theta) \right| \right) \\
&= E \left( \sup_{\theta \in \Theta} \left| \frac{h_{0t}}{h_t^\xi(\theta)} \tau_2 \sum_{k=1}^{\infty} k \beta^{k-1} \frac{e_{t-k-1}^2(\theta)}{h_t(\theta)} \right| \right) \\
&\leq E \left( \sup_{\theta \in \Theta} \left| \frac{h_{0t}}{h_t^\xi(\theta)} \frac{\tau_2}{\tau_1} \sum_{k=1}^{\infty} k \beta^{k-1} \frac{h_{t-k}(\theta)}{h_t(\theta)} \right| \right) \\
&\leq E \sup_{\theta \in \Theta} \left( \frac{\tau_2}{\tau_1} \sum_{k=1}^{\infty} \left( k \beta^{k-1} \frac{h_{t-k}(\theta)}{h_t(\theta)} \left( \frac{\omega_0}{\omega} + \frac{\tau_2}{\tau_1} k \left( \frac{\beta_0}{\beta} \right)^k + \beta_0^k \frac{h_{0t-k}}{\omega} \right) \right) \right) \\
&\leq \frac{\omega_0 \tau_2}{\omega_l^2 \tau_1} \sum_{k=1}^{\infty} k \beta_0^{k-1} E(\sup_{\theta \in \Theta} h_{t-k}(\theta)) \\
&\quad + \frac{\tau_2^2}{\omega_l \tau_1^2 \beta_l} \sum_{k=1}^{\infty} k^2 \beta_0^k E(\sup_{\theta \in \Theta} h_{t-k}(\theta)) + \frac{\tau_2}{\omega_l \tau_1 \beta_l} \sum_{k=1}^{\infty} k \beta_0^k E(h_{0t-k}), \quad \text{da } \beta^k \frac{h_{t-k}(\theta)}{h_t(\theta)} \leq 1 \\
&\leq \frac{\tau_2 \omega_0 \bar{h}}{\omega_l^2 \tau_1 (1 - \beta_0)^2} \\
&\quad + \frac{\tau_2^2 \bar{h}}{\omega_l \tau_1^2 \beta_l} \left( \beta_0^2 \sum_{k=1}^{\infty} (k-1) k \beta_0^{k-2} + \beta_0 \sum_{k=1}^{\infty} k \beta_0^{k-1} \right) \\
&\quad + \frac{\tau_2 \bar{h} \beta_0}{\omega_l \tau_1 (1 - \beta_0)^2 \beta_l} \\
&= \frac{\tau_2 \omega_0 \bar{h}}{\omega_l^2 \tau_1 (1 - \beta_0)^2} + \frac{2 \tau_2^2 \bar{h} \beta_0^2}{\omega_l \tau_1^2 (1 - \beta_0)^3 \beta_l} \\
&\quad + \frac{\tau_2^2 \bar{h} \beta_0}{\omega_l \tau_1^2 (1 - \beta_0)^2 \beta_l} + \frac{\tau_2 \bar{h} \beta_0}{\omega_l \tau_1 (1 - \beta_0)^2 \beta_l} \\
&= \frac{\tau_2 \bar{h}}{\tau_1 \omega_l (1 - \beta_0)^2} \left( \frac{\omega_0}{\omega_l} + \frac{2 \tau_2 \beta_0^2}{\tau_1 (1 - \beta_0) \beta_l} + \frac{\tau_2 \beta_0}{\tau_1 \beta_l} + \frac{\beta_0}{\beta_l} \right) < \infty
\end{aligned}$$

Damit folgt die Konsistenz durch dieselbe Argumentation wie in Theorem 1.

□

Bemerkung: Lemma 14 und 15 liefern wieder vorbereitende Zwischenergebnisse für Theorem 2.

**Lemma 14** *Unter den gemachten Voraussetzungen und der zusätzlichen Annahme, daß  $E(z_t^4 | \mathcal{F}_{t-1}) \leq \mathcal{K} < \infty$ , gilt:*

(1)

$$E|\nabla l_t(\theta) \nabla l_t(\theta)'| < \infty \quad \text{für alle } \theta \in \Theta_4$$

(2)

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \nabla l_t^*(\theta_0) \rightarrow_{\mathcal{L}} N(0, A_0),$$

wobei  $A_0 = E(\nabla l_t(\theta_0) \nabla l_t(\theta_0)')$ .

**Beweis von Lemma 14:**

Der Beweis läuft parallel zu dem Beweis von Lee und Hansen (1994) Lemma 9. Die Abschätzungen der Ableitungen nach  $\delta_i$ ,  $i = 1, \dots, l_2$ , erhält man mit ähnlichen Mitteln wie die Abschätzungen der Ableitungen nach  $\beta$ .

□

Bemerkung: In Anlehnung an die Bezeichnungen von Lemma 12 führen wir folgende Abkürzungen ein:

Für  $\xi_1, \xi_2, \xi_3 \in \{\alpha_i, \beta, \gamma, \delta_j, \omega\}$   $i = 1, \dots, l_1$ ,  $j = 1, \dots, l_2$  sei

$$h_{\xi_1 \xi_2 t}(\theta) := \frac{\partial^2 h_t(\theta)}{\partial \xi_1 \partial \xi_2} \frac{1}{h_t(\theta)} \quad \text{und} \quad h_{\xi_1 \xi_2 \xi_3 t}(\theta) := \frac{\partial^3 h_t(\theta)}{\partial \xi_1 \partial \xi_2 \partial \xi_3} \frac{1}{h_t(\theta)}.$$

Des weiteren sei

$$\begin{aligned} B_n^*(\theta) &:= -\frac{1}{n} \sum_{t=1}^n \nabla^2 l_t^*(\theta) \\ B_n(\theta) &:= -\frac{1}{n} \sum_{t=1}^n \nabla^2 l_t(\theta) \\ B(\theta) &:= -E \nabla^2 l_t(\theta) \end{aligned}$$

**Lemma 15** *Unter den gemachten Voraussetzungen gilt:*

- (1) Für alle  $\theta \in \Theta_4$  ist  $E|\nabla^2 l_t(\theta)| < \infty$ .
- (2)  $E(\sup_{\theta \in \Theta_6} |\nabla^3 l_t(\theta)|) < \infty$ .
- (3)  $|B_n^*(\theta) - B(\theta)| \rightarrow_p 0$  gleichmäßig in  $\theta \in \Theta_6$ , und  $B(\theta)$  ist stetig in  $\Theta_6$ .

**Beweis von Lemma 15:**

(1) und (2):

Beachte, daß für  $\xi_1, \xi_2 \in \{\alpha, \beta, \gamma, \delta, \omega\}$  gilt:

$$\begin{aligned} &\frac{\partial l_t(\theta)}{\partial \xi_1} \\ = &\left( \frac{e_t^2(\theta)}{h_t(\theta)^2} - \frac{1}{h_t(\theta)} \right) \frac{\partial h_t(\theta)}{\partial \xi_1} - \frac{2e_t(\theta)}{h_t(\theta)} \frac{\partial e_t(\theta)}{\partial \xi_1}, \end{aligned}$$

wie bereits in Lemma 12(3) verwendet.

$$\begin{aligned}
& \frac{\partial^2 l_t(\theta)}{\partial \xi_1 \partial \xi_2} \\
&= \left( \frac{e_t^2(\theta)}{h_t(\theta)^2} - \frac{1}{h_t(\theta)} \right) \frac{\partial^2 h_t(\theta)}{\partial \xi_1 \partial \xi_2} + \frac{\partial h_t(\theta)}{\partial \xi_1} \left( (-2) \frac{e_t^2(\theta)}{h_t^3(\theta)} + \frac{1}{h_t^2(\theta)} \right) \frac{\partial h_t(\theta)}{\partial \xi_2} \\
&\quad - 2 \frac{\partial^2 e_t(\theta)}{\partial \xi_1 \partial \xi_2} \frac{e_t(\theta)}{h_t(\theta)} - 2 \frac{\partial e_t(\theta)}{\partial \xi_1} \frac{\partial e_t(\theta)}{\partial \xi_2} \frac{1}{h_t(\theta)} \\
&\quad + 2 \frac{\partial e_t(\theta)}{\partial \xi_1} \frac{\partial h_t(\theta)}{\partial \xi_2} \frac{e_t(\theta)}{h_t^2(\theta)} + 2 \frac{\partial e_t(\theta)}{\partial \xi_2} \frac{\partial h_t(\theta)}{\partial \xi_1} \frac{e_t(\theta)}{h_t^2(\theta)} \\
&= \left( \frac{e_t^2(\theta)}{h_t(\theta)} - 1 \right) h_{\xi_1 \xi_2 t}(\theta) + \left( (-2) \frac{e_t^2(\theta)}{h_t(\theta)} + 1 \right) h_{\xi_1 t}(\theta) h_{\xi_2 t}(\theta) \\
&\quad + \left( \frac{\partial e_t(\theta)}{\partial \xi_1} \frac{\partial h_t(\theta)}{\partial \xi_2} + \frac{\partial e_t(\theta)}{\partial \xi_2} \frac{\partial h_t(\theta)}{\partial \xi_1} \right) \frac{2e_t(\theta)}{h_t^2(\theta)} - 2 \frac{\partial e_t(\theta)}{\partial \xi_1} \frac{\partial e_t(\theta)}{\partial \xi_2} \frac{1}{h_t(\theta)}.
\end{aligned}$$

Die dritten partiellen Ableitungen geben wir nur für  $\xi_1, \xi_2, \xi_3 \in \{\alpha, \beta, \delta, \omega\}$  an:

$$\begin{aligned}
& \frac{\partial^3 l_t(\theta)}{\partial \xi_1 \partial \xi_2 \partial \xi_3} \\
&= \left( \frac{e_t^2(\theta)}{h_t(\theta)^2} - \frac{1}{h_t(\theta)} \right) \frac{\partial^3 h_t(\theta)}{\partial \xi_1 \partial \xi_2 \partial \xi_3} + \left( (-2) \frac{e_t^2(\theta)}{h_t(\theta)^3} + \frac{1}{h_t(\theta)^2} \right) \\
&\quad \left( \frac{\partial^2 h_t(\theta)}{\partial \xi_1 \partial \xi_2} \frac{\partial h_t(\theta)}{\partial \xi_3} + \frac{\partial^2 h_t(\theta)}{\partial \xi_1 \partial \xi_3} \frac{\partial h_t(\theta)}{\partial \xi_2} + \frac{\partial h_t(\theta)}{\partial \xi_1} \frac{\partial^2 h_t(\theta)}{\partial \xi_2 \partial \xi_3} \right) \\
&\quad + \frac{\partial h_t(\theta)}{\partial \xi_1} \frac{\partial h_t(\theta)}{\partial \xi_2} \left( 6 \frac{e_t^2(\theta)}{h_t(\theta)^4} + \frac{-2}{h_t(\theta)^3} \right) \frac{\partial h_t(\theta)}{\partial \xi_3} \\
&= \left( \frac{e_t^2(\theta)}{h_t(\theta)} - 1 \right) h_{\xi_1 \xi_2 \xi_3 t}(\theta) + \left( (-2) \frac{e_t^2(\theta)}{h_t(\theta)} + \right) \\
&\quad (h_{\xi_1 \xi_2 t}(\theta) h_{\xi_3 t}(\theta) + h_{\xi_1 \xi_3 t}(\theta) h_{\xi_2 t}(\theta) + h_{\xi_2 \xi_3 t}(\theta) h_{\xi_1 t}(\theta)) \\
&\quad + \left( 6 \frac{e_t^2(\theta)}{h_t(\theta)} - 2 \right) h_{\xi_1 t}(\theta) h_{\xi_2 t}(\theta) h_{\xi_3 t}(\theta),
\end{aligned}$$

mit ähnlichen Ergebnissen für die partiellen Ableitungen nach  $\gamma$ .

Alle Ableitungen lassen sich auf vergleichbarem Wege wie in Lemma 12 abschätzen. Dies sei für die Ableitungen nach  $\beta$  exemplarisch vorgeführt, da für  $\xi_1, \xi_2, \xi_3 \in \{\alpha, \gamma, \delta, \omega\}$  die Terme  $h_{\xi_1 t}(\theta)$ ,  $h_{\xi_1 \xi_2 t}(\theta)$  und  $h_{\xi_1 \xi_2 \xi_3 t}(\theta)$  gleichmäßig beschränkt sind und sich für  $\xi_i = \beta$ ,  $i = 1, 2, 3$ , die stärksten Einschränkungen ergeben.

Wir erhalten für  $h_{\beta \beta t}(\theta)$  und  $h_{\beta \beta \beta t}(\theta)$ :

$$\begin{aligned}
h_{\beta \beta t}(\theta) &= \frac{\partial^2 h_t(\theta)}{\partial \beta^2} \frac{1}{h_t(\theta)} \\
&= \frac{\partial}{\partial \beta} \left( \sum_{k=1}^{\infty} k \beta^{k-1} g(e_{t-k-1}(\theta)) e_{t-k-1}^2(\theta) \right) \frac{1}{h_t(\theta)}
\end{aligned}$$

$$= \left( \sum_{k=2}^{\infty} k(k-1)\beta^{k-2} g(e_{t-k-1}(\theta)) e_{t-k-1}^2(\theta) \right) \frac{1}{h_t(\theta)} \quad (1)$$

$$h_{\beta\beta\beta t}(\theta) = \left( \sum_{k=3}^{\infty} k(k-1)(k-2)\beta^{k-3} g(e_{t-k-1}(\theta)) e_{t-k-1}^2(\theta) \right) \frac{1}{h_t(\theta)} \quad (2)$$

Damit ergibt sich:

$$\| \sup_{\theta \in \Theta_q} |h_{\beta\beta t}(\theta)| \|_q \leq B^2 \frac{\tau_2}{\tau_1} \sum_{k=1}^{\infty} k(k-1)\beta^{k-2} \| \sup_{\theta \in \Theta_q} \frac{h_{t-k}^{\epsilon}(\theta)}{h_t^{\epsilon}(\theta)} \|_q$$

(a) Für  $\beta \geq \beta_0$ :

$$\begin{aligned} &\leq B^2 \frac{\tau_2}{\tau_1} \sum_{k=2}^{\infty} k(k-1)\beta^{k-2} \left( \frac{\mathcal{R}_u^{\frac{1}{q}}}{\beta_0} \right)^k T_4 T_2 \quad \text{mit Lemma 10(2)} \\ &\leq B^2 \frac{\tau_2}{\tau_1} \left( \frac{\mathcal{R}_u^{\frac{1}{q}}}{\beta_0} \right)^2 T_4 T_2 \sum_{k=2}^{\infty} k(k-1) \left( \frac{\beta_u \mathcal{R}_u^{\frac{1}{q}}}{\beta_0} \right)^k \\ &\leq B^2 \frac{\tau_2}{\tau_1} \left( \frac{\mathcal{R}_u^{\frac{1}{q}}}{\beta_0} \right)^2 T_4 T_2 \frac{2}{\left( 1 - \frac{\beta_u \mathcal{R}_u^{\frac{1}{q}}}{\beta_0} \right)^3} \\ &\leq 2B^2 \frac{\tau_2}{\tau_1} T_4 T_2 \beta_0 \frac{1}{\eta_u^3}, \quad \text{da } \mathcal{R}_u \leq 1 \text{ und } \beta_0 - \beta_u \mathcal{R}_u^{\frac{1}{q}} = \eta_u \end{aligned}$$

(b) Für  $\beta \leq \beta_0$ :

$$\begin{aligned} &\leq B^2 \frac{\tau_2}{\tau_1} \left( \frac{\mathcal{R}_l^{\frac{1}{q}}}{\beta_{ql}} \right)^2 \frac{2}{\left( 1 - \frac{\beta_0 \mathcal{R}_l^{\frac{1}{q}}}{\beta_{ql}} \right)^3} \quad \text{mit Lemma 9(1)} \\ &\leq 2B^2 \frac{\tau_2}{\tau_1} \frac{1}{\eta_l^4} \quad \text{da } \mathcal{R}_l \leq 1, \beta_{ql} - \beta_0 \mathcal{R}_l^{\frac{1}{q}} = \eta_l \text{ und } \beta_{ql} < \beta_0 \end{aligned}$$

Insgesamt ergibt sich also:

$$\| \sup_{\theta \in \Theta_q} |h_{\beta\beta t}(\theta)| \|_q \leq 2B^2 \frac{\tau_2}{\tau_1} \beta_0 \max \left( T_4 T_2 \frac{1}{\eta_u^3}, \frac{1}{\eta_l^4} \right) =: H_2$$

Für  $h_{\beta\beta\beta t}(\theta)$  erhalten wir analog:

$$\| \sup_{\theta \in \Theta_q} |h_{\beta\beta\beta t}(\theta)| \|_q \leq 6B^2 \frac{\tau_2}{\tau_1} \beta_0 \max \left( T_4 T_2 \frac{1}{\eta_u^4}, \frac{1}{\eta_l^4} \right) =: H_3$$

Wir können nun die zweiten und dritten Ableitungen von  $l_t(\theta)$  nach  $\beta$  abschätzen.

Aus Gleichung (2.1) folgt für die zweite Ableitung:

$$\begin{aligned}
& \sup_{\theta \in \Theta_4} E \left| \frac{\partial^2 l_t(\theta)}{\partial \beta^2} \right| \\
& \leq E \sup_{\theta \in \Theta_4} \left| \frac{\partial^2 l_t(\theta)}{\partial \beta^2} \right| \\
& = E \sup_{\theta \in \Theta_4} \left| \frac{e_t^2(\theta)}{h_t(\theta)} (h_{\beta\beta t}(\theta) - 2h_{\beta t}(\theta)^2) + h_{\beta t}(\theta)^2 - h_{\beta\beta t}(\theta) \right| \\
& \leq E \sup_{\theta \in \Theta_4} \left( \frac{e_t^2(\theta)}{h_t(\theta)} (h_{\beta\beta t}(\theta) + 2h_{\beta t}(\theta)^2) \right) \\
& \quad + E \sup_{\theta \in \Theta_4} h_{\beta t}(\theta)^2 + E \sup_{\theta \in \Theta_4} h_{\beta\beta t}(\theta)^2 \\
& \leq E \sup_{\theta \in \Theta_4} \left( \left( \frac{\epsilon_t^2 + 2\epsilon_t m + m^2}{h_t(\theta)} \right) (h_{\beta\beta t}(\theta) + 2h_{\beta t}(\theta)^2) \right) \\
& \quad + H_1^2 + H_2 \\
& \quad \text{mit Lemma 12(2) und da } E \sup_{\theta \in \Theta_4} h_{\beta t}(\theta)^2 = \left\| \sup_{\theta \in \Theta_4} h_{\beta t}(\theta) \right\|_2^2 \\
& = E \sup_{\theta \in \Theta_4} \left( \left( \frac{(1 + 2\tilde{m})h_{0t} E(z_t^2 | \mathcal{F}_{t-1}) + (2\tilde{m} + \tilde{m}^2)}{h_t(\theta)} \right) (h_{\beta\beta t}(\theta) + 2h_{\beta t}(\theta)^2) \right) \\
& \quad + H_1^2 + H_2 \\
& \leq \left( (1 + 2\tilde{m}) \left\| \sup_{\theta \in \Theta_4} \frac{h_{0t}}{h_t(\theta)} \right\|_2 + \frac{2\tilde{m} + \tilde{m}^2}{\omega_l} \right) \\
& \quad \left( \left\| \sup_{\theta \in \Theta_4} h_{\beta\beta t}(\theta) \right\|_2 + 2 \left\| \sup_{\theta \in \Theta_4} h_{\beta t}(\theta) \right\|_4 \right) + H_1^2 + H_2 \\
& \leq ((1 + 2\tilde{m})BT_3 + \frac{2\tilde{m} + \tilde{m}^2}{\omega_l})(H_2 + 2H_1^2) + H_1^2 + H_2 < \infty, \quad \text{mit Lemma 9(2)}
\end{aligned}$$

Entsprechend folgt aus Gleichung (2.2) für die dritte Ableitung:

$$\begin{aligned}
& E \sup_{\theta \in \Theta_6} \left| \frac{\partial^3 l_t(\theta)}{\partial \beta^3} \right| \\
& = E \sup_{\theta \in \Theta_6} \left| \frac{e_t^2(\theta)}{h_t(\theta)} (h_{\beta\beta\beta t}(\theta) - 6h_{\beta\beta t}(\theta)h_{\beta t}(\theta) + 6h_{\beta t}(\theta)^3) \right. \\
& \quad \left. - h_{\beta\beta\beta t}(\theta) + 3h_{\beta\beta t}(\theta)h_{\beta t}(\theta) - 2h_{\beta t}(\theta)^3 \right| \\
& \leq \left( (1 + 2\tilde{m}) \left\| \sup_{\theta \in \Theta_6} \frac{h_{0t}}{h_t(\theta)} \right\|_2 + \frac{2\tilde{m} + \tilde{m}^2}{\omega_l} \right) \\
& \quad \left( \left\| \sup_{\theta \in \Theta_6} h_{\beta\beta\beta t}(\theta) \right\|_2 + 6 \left\| \sup_{\theta \in \Theta_6} h_{\beta\beta t}(\theta) \right\|_4 \left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta) \right\|_4 \right. \\
& \quad \left. + 6 \left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta)^3 \right\|_6^3 + \left\| \sup_{\theta \in \Theta_6} h_{\beta\beta\beta t}(\theta) \right\|_1 \right. \\
& \quad \left. + 3 \left\| \sup_{\theta \in \Theta_6} h_{\beta\beta t}(\theta) \right\|_2 \left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta) \right\|_2 + 2 \left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta) \right\|_3^3 \right)
\end{aligned}$$

$$\begin{aligned} &\leq ((1 + 2\tilde{m})BT_3 + \frac{2\tilde{m} + \tilde{m}^2}{\omega_l})(H_3 + 6H_2H_1 + 6H_1^3) \\ &\quad + H_3 + 3H_2H_1 + 2H_1^3 < \infty \end{aligned}$$

Damit sind für die wichtigsten Spezialfälle die Behauptungen (1) und (2) des Lemmas bewiesen.

(3) Wie im Beweis von Theorem 1 benutzen wir

$$\sup_{\theta \in \Theta_6} |B_n^*(\theta) - B(\theta)| \leq \sup_{\theta \in \Theta_6} |B_n^*(\theta) - B_n(\theta)| + \sup_{\theta \in \Theta_6} |B_n(\theta) - B(\theta)|$$

Mit Lemma 3(2) und Lemma 15(1) folgt aus dem Ergodentheorem, daß

$$B_n(\theta) = -\frac{1}{n} \sum_{t=1}^n \nabla^2 l_t(\theta) \xrightarrow{p} -E \nabla^2 l_t(\theta) = B(\theta) \quad \text{punktweise}$$

für alle  $\theta \in \Theta_6$ , da  $\Theta_4 \subset \Theta_6$ .

Wie im Beweis von Lemma 12(5) zeigt man mit Lemma 15(2) die gleichmäßige stochastische Konvergenz von  $B_n(\theta)$  gegen  $B(\theta)$  und die gleichmäßige Stetigkeit von  $B(\theta)$ .

Die gleichmäßige stochastische Konvergenz von  $B_n^*(\theta)$  gegen  $B_n(\theta)$  folgt aus ähnlichen Argumenten wie Lemma 6.

□

### Beweis von Theorem 2:

Wir machen eine Taylor-Entwicklung von  $\nabla L_n^*(\theta)$  um  $\theta_0$ :

$$\nabla L_n^*(\theta)|_{\theta=\hat{\theta}_n} = \nabla L_n^*(\theta)|_{\theta=\theta_0} + \nabla^2 L_n^*(\theta)|_{\theta=\theta^*}(\hat{\theta}_n - \theta_0)$$

mit  $\theta^* := \theta_0 + \lambda(\hat{\theta}_n - \theta_0)$  für ein  $\lambda \in [0, 1]$ .

Nach Definition von  $\hat{\theta}_n$  ist  $\nabla L_n^*(\theta)|_{\theta=\hat{\theta}_n} = 0$ . Wir erhalten also

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = -[\nabla^2 L_n^*(\theta)|_{\theta=\theta^*}]^{-1} \sqrt{n} \nabla L_n^*(\theta)|_{\theta=\theta_0}.$$

Die Konsistenz von  $\hat{\theta}_n$  zusammen mit der Stetigkeit von  $\nabla^2 L_n^*(\theta)$  und Lemma 15 (3) liefert, daß  $-\nabla^2 L_n^*(\theta)|_{\theta=\theta^*}$  gleichmäßig stochastisch gegen  $B(\theta_0)$  konvergiert.

Lemma 14 (2) wiederum liefert, daß  $\sqrt{n} \nabla L_n^*(\theta)|_{\theta=\theta_0}$  für  $n \rightarrow \infty$  normalverteilt ist mit Erwartung  $\mathbf{0}$  und Kovarianzmatrix  $A_0$ . Insgesamt erhalten wir also, daß die Verteilung von  $\sqrt{n}(\hat{\theta}_n - \theta_0)$  für  $n \rightarrow \infty$  schwach gegen eine Normalverteilung mit Erwartung  $\mathbf{0}$  und Kovarianzmatrix  $V_0 := B_0^{-1} A_0 B_0^{-1}$  konvergiert, wie im Theorem angegeben.

Die Existenz der Kovarianzmatrix  $V_0$  muß noch begründet werden. Es ist hierzu noch zu zeigen, daß die Matrix  $B_0$  invertierbar ist. Wir zeigen, daß  $B_0$  sich als Summe einer positiv definiten und einer positiv semidefiniten Matrix schreiben läßt.

Sei  $\lambda \in \mathbf{R}^{1+1_2+3}$  ein Koeffizientenvektor. Dann gilt mit den im Beweis von Lemma 15 hergeleiteten Formeln:

$$\begin{aligned}
& B_0 \text{ positiv definit} \\
\Leftrightarrow & \lambda' B_0 \lambda > 0 \quad \text{für alle } \lambda \neq 0 \\
\Leftrightarrow & -\lambda' E \left( \frac{\partial^2 l_t(\theta_0)}{\partial \xi_1 \partial \xi_2} \right)_{\xi_1, \xi_2} \lambda = 0 \quad \text{nur für } \lambda = 0 \\
\Leftrightarrow & \lambda' E \left( \frac{1}{h_t^2(\theta_0)} \frac{\partial h_t(\theta_0)}{\partial \xi_1} \frac{\partial h_t(\theta_0)}{\partial \xi_2} + 21_{\xi_1 = \xi_2 = \gamma} \frac{1}{h_t(\theta)} \right)_{\xi_1, \xi_2} \lambda = 0 \\
& \text{nur für } \lambda = 0 \\
\Leftrightarrow & \lambda' E \left( \frac{1}{h_t^2(\theta_0)} \frac{\partial h_t(\theta_0)}{\partial \xi_1} \frac{h_t(\theta_0)}{\partial \xi_2} \right)_{\xi_1, \xi_2} \lambda + 2\lambda' E \left( 1_{\xi_1 = \xi_2 = \gamma} \frac{1}{h_t(\theta)} \right)_{\xi_1, \xi_2} \lambda = 0 \\
& \text{nur für } \lambda = 0 \\
\Leftrightarrow & \lambda' B_1 \lambda + 2\lambda' B_2 \lambda = 0 \quad \text{nur für } \lambda = 0
\end{aligned}$$

Da  $\frac{1}{h_t(\theta)}$  fast sicher positiv ist und da alle Hauptunterdeterminanten von  $B_2$  nicht negativ sind, ist  $B_2$  positiv semidefinit.

Für  $B_1$  gilt:

$$\begin{aligned}
& \lambda' E \left( \frac{1}{h_t^2(\theta_0)} \nabla h_t(\theta)|_{\theta=\theta_0} \nabla h_t(\theta)'|_{\theta=\theta_0} \right) \lambda = 0 \quad \text{nur für } \lambda = 0 \\
\Leftrightarrow & E \left( \frac{1}{h_t^2(\theta_0)} \lambda' \nabla h_t(\theta)|_{\theta=\theta_0} \nabla h_t(\theta)'|_{\theta=\theta_0} \lambda \right) = 0 \quad \text{nur für } \lambda = 0 \\
\Leftrightarrow & \frac{1}{h_t^2(\theta_0)} \lambda' \nabla h_t(\theta)|_{\theta=\theta_0} \nabla h_t(\theta)'|_{\theta=\theta_0} \lambda = 0 \quad \text{fast sicher, nur für } \lambda = 0 \\
\Leftrightarrow & \lambda' \nabla h_t(\theta)|_{\theta=\theta_0} = 0 \quad \text{fast sicher, nur für } \lambda = 0, \\
& \text{da } h_t(\theta_0) \text{ fast sicher positiv ist.} \\
\Leftrightarrow & \lambda' \nabla h_{t-1}(\theta)|_{\theta=\theta_0} = 0 \quad \text{fast sicher, nur für } \lambda = 0, \\
& \text{da } \nabla h_t(\theta_0) \text{ strikt stationär ist.}
\end{aligned}$$

Angenommen,  $\lambda' \nabla h_{t+1}(\theta)|_{\theta=\theta_0} = 0$  fast sicher für ein  $\lambda \neq 0$ . Dann können wir schreiben

$$\begin{aligned}
& \lambda' \nabla h_{t+1}(\theta)|_{\theta=\theta_0} \\
= & \lambda_\omega \left( 1 + \beta_0 \frac{\partial h_t(\theta_0)}{\partial \omega} \right) + \sum_{i=1}^{l_1} \lambda_{\alpha_i} \left( \frac{\partial g(\alpha_0, \epsilon_t)}{\partial \alpha_i} \epsilon_t^2 + \beta_0 \frac{\partial h_t(\theta_0)}{\partial \alpha_i} \right) \\
& + \lambda_\gamma \left( \left( \frac{\partial g(\alpha_0, \epsilon_t)}{\partial y} \epsilon_t^2 - 2g(\alpha_0, \epsilon_t) \epsilon_t + \beta_0 \frac{\partial h_t(\theta_0)}{\partial \gamma} \right) \right) \\
& + \sum_{i=1}^{l_2} \lambda_{\delta_i} \left( \frac{\partial f(\delta_0, r_t)}{\partial \delta_i} + \beta_0 \frac{\partial h_t(\theta_0)}{\partial \delta_i} \right)
\end{aligned}$$

$$\begin{aligned}
&= \beta_0 \lambda' \nabla h_t(\theta)|_{\theta=\theta_0} + \lambda' \left( 1, \frac{\partial g}{\partial \alpha_1}(\alpha_0, \epsilon_t) \epsilon_t^2, \dots, \frac{\partial g}{\partial \alpha_{l_1}}(\alpha_0, \epsilon_t) \epsilon_t^2, \frac{\partial g}{\partial y}(\alpha_0, \epsilon_t) \epsilon_t^2, \right. \\
&\quad \left. g(\alpha_0, \epsilon_t) \epsilon_t, \frac{\partial f}{\partial \delta_1}(\delta_0, r_t), \dots, \frac{\partial f}{\partial \delta_{l_2}}(\delta_0, r_t), f(\delta_0, r_t) \right)' \\
&= \lambda' \left( 1, \frac{\partial g}{\partial \alpha_1}(\alpha_0, \epsilon_t) \epsilon_t^2, \dots, \frac{\partial g}{\partial \alpha_{l_1}}(\alpha_0, \epsilon_t) \epsilon_t^2, \frac{\partial g}{\partial y}(\alpha_0, \epsilon_t) \epsilon_t^2, \right. \\
&\quad \left. g(\alpha_0, \epsilon_t) \epsilon_t, \frac{\partial f}{\partial \delta_1}(\delta_0, r_t), \dots, \frac{\partial f}{\partial \delta_{l_2}}(\delta_0, r_t), f(\delta_0, r_t) \right)' =: \lambda' x_t
\end{aligned}$$

Da die Kovarianzmatrix von  $x_t$  nach Voraussetzung für alle  $t \in \mathbf{Z}$  vollen Rang hat, ist  $\lambda' x_t$  nur für  $\lambda = 0$  fast sicher  $= 0$ . Damit ist  $B_1$  positiv definit und somit auch  $B_0$  (vgl. Weiss [1986] Lemmas 3.1-3.3).

□

**Lemma 16** *Unter den Voraussetzungen von Theorem 2 und der zusätzlichen Annahme, daß*

$$z_t | \mathcal{F}_{t-1} \sim N(0, 1)$$

*gilt:*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow_{\mathcal{L}} N(0, B_0^{-1}) = N(0, A_0^{-1}).$$

**Beweis von Lemma 16:**

Wegen  $z_t | \mathcal{F}_{t-1} \sim N(0, 1)$  ist

$$E(e_t^3 | \mathcal{F}_{t-1}) = E(\epsilon_t^3 | \mathcal{F}_{t-1}) = 0$$

und

$$E(e_t^4 | \mathcal{F}_{t-1}) = E(\epsilon_t^4 | \mathcal{F}_{t-1}) = 3h_t^2(\theta_0).$$

Mit den Formeln aus dem Beweis von Lemma 15,  $\xi_1, \xi_2 \in \{\alpha, \beta, \gamma, \delta, \omega\}$  und  $\frac{\partial e_t(\theta)}{\partial \xi_1} = \text{const}$  gilt dann unmittelbar:

$$A_0 = B_0 = E \left( \frac{1}{2h_t^2(\theta_0)} \frac{\partial h_t}{\partial \xi_1}(\theta_0) \frac{\partial h_t}{\partial \xi_2}(\theta_0) + \frac{1}{h_t(\theta_0)} \frac{\partial e_t}{\partial \xi_1}(\theta_0) \frac{\partial e_t}{\partial \xi_2}(\theta_0) \right)_{\xi_1, \xi_2}$$

□

**Beweis von Theorem 3:**

Wenn (a)  $A_n^*(\hat{\theta}_n) \rightarrow_p A_0$  und (b)  $B_n^*(\hat{\theta}_n) \rightarrow_p B_0$ , folgt nach dem Theorem von Slutsky, daß

$$V_n^* \rightarrow_{\mathcal{L}} V_0.$$

Da mit  $A_0$  und  $B_0$  auch  $V_0$  eine konstante Matrix ist, folgt dann die Behauptung von Theorem 3, da in diesem Fall die Konvergenz in Verteilung die stochastische Konvergenz impliziert.

Aussage (b) ist bereits im Beweis von Theorem 2 gezeigt worden.

Aussage (a) folgt ebenso, wenn als Pendant zu Lemma 15 (3) gezeigt werden kann, daß  $A_n^*(\theta)$  gleichmäßig stochastisch gegen  $A(\theta)$  konvergiert. Betrachte dazu wieder folgende Abschätzung:

$$\begin{aligned} & \sup_{\theta \in \Theta_6} \frac{1}{n} \left| \sum_{t=1}^n (\nabla l_t^*(\theta) \nabla l_t^*(\theta)' - E(\nabla l_t(\theta) \nabla l_t(\theta)')) \right| \\ & \leq \sup_{\theta \in \Theta_6} \frac{1}{n} \left| \sum_{t=1}^n \nabla l_t^*(\theta) \nabla l_t^*(\theta)' - \sum_{t=1}^n \nabla l_t(\theta) \nabla l_t(\theta)' \right| \\ & + \sup_{\theta \in \Theta_6} \frac{1}{n} \left| \sum_{t=1}^n (\nabla l_t(\theta) \nabla l_t(\theta)' - E(\nabla l_t(\theta) \nabla l_t(\theta)')) \right| \end{aligned}$$

Die stochastische Konvergenz des ersten Terms gegen 0 folgt aus ähnlichen Argumenten wie Lemma 6.

Für den zweiten Term folgt die punktweise stochastische Konvergenz aus dem Ergodentheorem, da  $A_n^*(\theta)$  strikt stationär und ergodisch ist und wegen Lemma 15 (1). Die gleichmäßige stochastische Konvergenz folgt wie Lemma 12 (5), wenn

$$E(\sup_{\theta \in \Theta_6} |\nabla(\nabla l_t(\theta) \nabla l_t(\theta)')|) < \infty$$

gezeigt werden kann:

Für die Abschätzung der partiellen Ableitungen gilt das im Beweis von Lemma 15 Gesagte entsprechend. Es ist deshalb sinnvoll, sich auf die Betrachtung der Ableitungen nach  $\beta$  zu beschränken.

Zunächst ist

$$\begin{aligned} & \frac{1}{2} \frac{\partial}{\partial \beta} \left( \frac{\partial l_t(\theta)}{\partial \beta} \right)^2 \\ & = \frac{\partial l_t(\theta)}{\partial \beta} \frac{\partial^2 l_t(\theta)}{\partial \beta^2} \\ & = \frac{e_t^4(\theta)}{h_t^2(\theta)} (h_{\beta t}(\theta) h_{\beta \beta t}(\theta) - 2h_{\beta t}(\theta)^3) \\ & \quad + \frac{e_t^2(\theta)}{h_t(\theta)} (3h_{\beta t}(\theta)^3 - 2h_{\beta t}(\theta) h_{\beta \beta t}(\theta)) - h_{\beta t}(\theta)^3 + h_{\beta t}(\theta) h_{\beta \beta t}(\theta) \end{aligned}$$

wie aus den im Beweis von Lemma 12 und Lemma 15 hergeleiteten Formeln hervorgeht.

Beachte, daß

$$\begin{aligned} e_t^4(\theta) & \leq \epsilon_t^4 + 3\tilde{m}\epsilon_t^3 + 6\tilde{m}^2\epsilon_t^2 + 3\tilde{m}^3\epsilon_t + \tilde{m}^4 \\ & \leq (3\tilde{m} + 1)\epsilon_t^4 + (6\tilde{m}^2 + 3\tilde{m}^3)\epsilon_t^2 + 3\tilde{m} + 3\tilde{m}^3 + \tilde{m}^4 \end{aligned}$$

Also erhalten wir:

$$\frac{1}{2} E \sup_{\theta \in \Theta_6} \left| \frac{\partial}{\partial \beta} \left( \frac{\partial l_t(\theta)}{\partial \beta} \right)^2 \right|$$

$$\begin{aligned}
&\leq E \sup_{\theta \in \Theta_6} \left( \frac{\mathcal{K}(3\tilde{m} + 1)h_{0t}^2 + (6\tilde{m}^2 + 3\tilde{m}^3)h_{0t} + 3\tilde{m} + 3\tilde{m}^3 + \tilde{m}^4}{h_t^2(\theta)} (h_{\beta t}(\theta)h_{\beta\beta t}(\theta) + 2h_{\beta t}(\theta)^3) \right) \\
&\quad + E \sup_{\theta \in \Theta_6} \left( (1 + 2\tilde{m}) \frac{h_{0t} + 2\tilde{m} + \tilde{m}^2}{h_t^2(\theta)} (2h_{\beta t}(\theta)h_{\beta\beta t}(\theta) + 3h_{\beta t}(\theta)^3) \right) \\
&\quad + E \sup_{\theta \in \Theta_6} (h_{\beta t}(\theta)h_{\beta\beta t}(\theta) + h_{\beta t}(\theta)^3) \\
&\leq \left( \mathcal{K}(3\tilde{m} + 1) \left\| \sup_{\theta \in \Theta_6} \frac{h_{0t}}{h_t(\theta)} \right\|_4^2 + \frac{6\tilde{m}^2 + 3\tilde{m}^3}{\omega_l} \left\| \sup_{\theta \in \Theta_6} \frac{h_{0t}}{h_t(\theta)} \right\|_2 \right) \\
&\quad (\left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta) \right\|_4 \left\| \sup_{\theta \in \Theta_6} h_{\beta\beta t}(\theta) \right\|_4 + 2 \left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta) \right\|_6^3) \\
&\quad + \frac{3\tilde{m} + 3\tilde{m}^3 + \tilde{m}^4}{\omega_l^2} (\left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta) \right\|_2 \left\| \sup_{\theta \in \Theta_6} h_{\beta\beta t}(\theta) \right\|_2 + 2 \left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta) \right\|_3^3) \\
&\quad + (1 + 2\tilde{m}) \left\| \sup_{\theta \in \Theta_6} \frac{h_{0t}}{h_t(\theta)} \right\|_2 (2 \left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta) \right\|_4 \left\| \sup_{\theta \in \Theta_6} h_{\beta\beta t}(\theta) \right\|_4 + 3 \left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta) \right\|_6^3) \\
&\quad + \frac{2\tilde{m} + \tilde{m}^2}{\omega_l} (2 \left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta) \right\|_2 \left\| \sup_{\theta \in \Theta_6} h_{\beta\beta t}(\theta) \right\|_2 + 3 \left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta) \right\|_3^3) \\
&\quad + \left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta) \right\|_2 \left\| \sup_{\theta \in \Theta_6} h_{\beta\beta t}(\theta) \right\|_2 + \left\| \sup_{\theta \in \Theta_6} h_{\beta t}(\theta) \right\|_3^3 \\
&\leq \left( \mathcal{K}(3\tilde{m} + 1)(BT_3)^2 + \frac{6\tilde{m}^2 + 3\tilde{m}^3}{\omega_l} BT_3 + \frac{3\tilde{m} + 3\tilde{m}^3 + \tilde{m}^4}{\omega_l^2} \right) (H_1 H_2 + 2H_1^3) \\
&\quad + \left( (1 + 2\tilde{m})BT_3 + \frac{2\tilde{m} + \tilde{m}^2}{\omega_l} \right) (2H_1 H_2 + 3H_1^3) \\
&\quad + (H_1 H_2 + H_1^3) < \infty \quad \text{mit Lemma 9(2)}
\end{aligned}$$

□

## 8 Literaturverzeichnis

- Amemiya, T.**, 1985, *Advanced Econometrics*, Oxford, Basil Blackwell
- Andrews, D.W.K.**, 1992, Uniform Convergence, *Econometric Theory* 8, 241-257
- Bera, A.K., Higgins, M.L.**, 1993, ARCH Models: Properties, Estimation and Testing, *Journal of Economic Surveys* 7, 305-366
- Biermann, B.**, 1997, Modernes Risikomanagement in Banken, in: Eller, R., *Handbuch des Risikomanagements*, 3-25
- Billingsley, P.**, 1968, *Convergence of Probability Measures*, New York, Wiley
- Billingsley, P.**, 1979, *Probability and Measure*, New York, Wiley
- Black, F., Scholes, M.**, 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81, 637-654
- Bollerslev, T.**, 1986, Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* 31, 307-327
- Bollerslev, T., Chou, R.Y., Kroner, K.F.**, 1992, ARCH modeling in finance, *Journal of Econometrics* 52, 5-59
- Bollerslev, T., Engle, R.F., Nelson, D.B.**, 1994, ARCH Models, *Handbook of Econometrics, Volume IV*, 2959-3038
- Bollerslev, T., Wooldridge, J.M.**, 1992, Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances, *Econometric Reviews* 11(2), 143-172
- Breiman, L.**, 1992, *Probability, Classics in Applied Mathematics* 7, Addison Wesley, Reading, Massachusetts
- Deb, P.**, 1996, Finite sample properties of maximum likelihood and quasi-maximum likelihood estimators of EGARCH models, *Econometric reviews* 15, 51-68
- Dümbgen, L.**, 1997, *Empirische Prozesse, Skriptum zur Vorlesung im Sommersemester 1997*

- Engle, R.F.**, 1982, Autoregressive Conditional Heteroskedasticity with Estimation of the Variance of United Kingdom Inflation, *Econometrica* 50, 987-1007
- Engle, R.F., Lilien, D.M., Robins, R.P.**, 1987, Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model, *Econometrica* 55, 391-407
- Engle, R.F., Ng, V.K.**, 1993, Measuring and Testing the Impact of News on Volatility, *The Journal of Finance* 48, 1749-1777
- Glosten, L.R., Jagannathan, R., Runkle**, 1993, On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *The Journal of Finance*, 1779-1801
- Gouriéroux, C.**, 1997, ARCH Models and Financial Applications, Springer, New York
- Gouriéroux, C., Montfort, A.**, 1992, Qualitative threshold ARCH models, *Journal of Econometrics* 52, 159-199
- Hamao, Y., Masulis, R.W., Ng, V.**, 1990, Correlations in price changes and volatility across international stock markets, *Review of Financial Studies* 3, 281-308
- Hardy, G.H., Littlewood, J.E., Pólya, G.**, 1934, Inequalities, Cambridge at the University Press
- Hentschel, L.**, 1995, All in the family. Nesting symmetric and asymmetric GARCH models, *Journal of Financial Economics* 39, 71-104
- Kaiser, T.**, 1996, One-Factor-GARCH Models for German Stocks - Estimation and Forecasting, Mimeo, Universität Tübingen
- Lee, S.W., Hansen, B.E.**, 1994, Asymptotic Theory for the GARCH(1,1) Quasi-Maximum Likelihood Estimator, *Econometric Theory* 10, 29-52
- Lin, W.L.**, 1992, Alternative Estimators for Factor GARCH Models - A Monte Carlo Comparison, *Journal of Applied Econometrics* 7, 259-279
- Liu, S.M., Brorsen, B.W.**, 1995, Maximum Likelihood Estimation of a GARCH-Stable Model, *Journal of Applied Econometrics* 10, 273-285
- Lütkepohl, H.**, 1996, Statistische Modellierung von Volatilitäten, Mimeo,

Universität Berlin

**Lumsdaine, R.L.**, 1996, Consistency and asymptotic Normality of the Quasi-Maximum Likelihood Estimator in IGARCH(1,1) and Covariance Stationary GARCH(1,1) Models, *Econometrica* 64, 575-596

**Maercker, G.**, 1996, Statistical Inference in Conditional Heteroskedastic Autoregressive Models, Dissertation, Universität Braunschweig

**Markowitz, H.M.**, 1952, Portfolio Selection, *Journal of Finance Vol. VII, No. 1*

**Nelson, D.B.**, 1991, Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica* 59, 347-370

**Nelson, D.B.**, 1990, Stationarity and Persistence in the GARCH(1,1) Model, *Econometric Theory* 6, 318-334

**Pagan, A.R., Schwert, G.W.**, 1990, Alternative Models for Conditional Stock Volatility, *Journal of Econometrics* 50, 270-290

**Saxinger, R.**, 1997, Traditionelle und neuere Risikomaße im Asset-Management, in: Eller, R., *Handbuch des Risikomanagements*, 3-25

**Searle, S.R.**, 1982, Matrix Algebra Useful for Statistics, New York, Wiley

**Schmitt, Christian**, 1996, Option Pricing Using EGARCH Models, Discussion Paper No. 96-20, Zentrum für Europäische Wirtschaftsforschung

**Stout, W.F.**, 1974, Almost Sure Convergence, New York, Academic Press

**Tschernig, R.**, 1994, Wechselkurse, Unsicherheit und Long Memory, Heidelberg, Physica-Verlag

**Weiss, A.A.**, 1986, Asymptotic Theory for ARCH Models: Estimation and Testing, *Econometric Theory* 2, 107-131