

# **Otto-Friedrich-Universität Bamberg**

**Sozial- und Wirtschaftswissenschaftliche Fakultät**



**Freie wissenschaftliche Arbeit  
zur Erlangung des akademischen Grades**

**Doktor der Sozial- und Wirtschaftswissenschaften**

**-Dr. rer. pol.-**

## **Bayessche Betrachtung von erweiterten GARCH-Modellen**

**eingereicht am Lehrstuhl für Statistik und Ökonometrie, Prof. Dr. S. Rässler**

Erstgutachterin: Prof. Dr. S. Rässler

Zweitgutachter: Prof. Dr. I. Klein

vorgelegt von: Dipl. Kfm. D. Wan Hussin

---

# Inhaltsverzeichnis

|  |           |
|--|-----------|
| Abbildungsverzeichnis . . . . .  | V         |
| Tabellenverzeichnis . . . . .  | XI        |
| Abkürzungsverzeichnis . . . . .  | XIII      |
| <b>1 Einleitung und Motivation</b>   | <b>1</b>  |
| 1.1 Motivation . . . . .   | 3         |
| 1.2 Optionen . . . . .   | 4         |
| 1.2.1 Einflussfaktoren auf den Optionspreis . . . . .                                | 6         |
| 1.2.2 Volatilität . . . . .  | 7         |
| 1.2.3 VDAX - ein Index für Volatilität . . . . .                                     | 9         |
| 1.3 Stylized Facts . . . . .   | 10        |
| 1.4 Kombinationsstrategien auf Volatilität . . . . .                                 | 15        |
| <b>2 GARCH-Modelle und deren Erweiterungen</b>                                       | <b>20</b> |
| 2.1 ARCH-Modell . . . . .  | 21        |
| 2.2 GARCH-Modell . . . . .   | 22        |
| 2.2.1 Schätzung des Modells mit der (Quasi-)Maximum-Likelihood-<br>Methode . . . . . | 24        |
| 2.2.2 Test auf GARCH-Effekte . . . . .   | 25        |
| 2.2.3 Prognose mit GARCH-Modellen . . . . .  | 25        |
| 2.3 Alternative Residualverteilungen . . . . .                                       | 26        |
| 2.3.1 t-Verteilung . . . . .   | 27        |
| 2.3.2 GED-Verteilung . . . . .   | 28        |

---

|          |   |           |
|----------|---|-----------|
| 2.4      | Erweiterungen des GARCH-Modells . . . . .                       | 30        |
| 2.4.1    | Erweiterungen der Mittelwertgleichung . . . . .                 | 30        |
| 2.4.1.1  | ARMA-GARCH . . . . .  | 31        |
| 2.4.1.2  | GARCH-in-Mean . . . . .   | 31        |
| 2.4.2    | H-GARCH-Modell nach Hentschel . . . . .                         | 31        |
| 2.4.3    | EGARCH-Modell nach Nelson . . . . .                             | 33        |
| 2.4.4    | GJR-GARCH-Modell nach Glosten, Jagannathan und Runkle . . . . . | 34        |
| 2.4.5    | AP-ARCH-Modell nach Ding, Granger und Engle . . . . .           | 35        |
| 2.4.6    | TS-GARCH-Modell nach Taylor und Schwert . . . . .               | 35        |
| 2.4.7    | TGARCH-Modell nach Zakoian . . . . .                            | 36        |
| 2.5      | Multivariate GARCH-Modelle . . . . .                            | 36        |
| 2.5.1    | Constant-Conditional-Correlation-GARCH-Modell . . . . .         | 38        |
| 2.5.2    | Dynamic-Conditional-Correlation-GARCH-Modell . . . . .          | 40        |
| <b>3</b> | <b>Bayes-Ansatz und Markov Chain Monte Carlo-Methoden</b>       | <b>42</b> |
| 3.1      | Bayes-Ansatz . . . . .  | 42        |
| 3.1.1    | Einführung . . . . .  | 43        |
| 3.1.2    | Die a priori Verteilung . . . . .                               | 44        |
| 3.1.3    | Die a posteriori Prädiktivverteilung . . . . .                  | 45        |
| 3.2      | Markov Chain Monte Carlo-Verfahren . . . . .                    | 45        |
| 3.2.1    | Theorie der Markovketten . . . . .                              | 46        |
| 3.2.1.1  | Definition und Eigenschaften einer Markovkette . . . . .        | 46        |
| 3.2.1.2  | Vorgehensweise . . . . .  | 51        |
| 3.2.2    | Der Metropolis-Hastings-Algorithmus . . . . .                   | 52        |
| 3.2.3    | Spezialfälle . . . . .  | 54        |
| 3.2.4    | Mischung und Konvergenz . . . . .                               | 57        |
| 3.2.5    | Konvergenzanalyse . . . . .                                     | 57        |
| 3.3      | Modellselektion . . . . .                                       | 58        |
| 3.3.1    | Bayes-Faktor . . . . .  | 59        |
| 3.3.2    | Harmonoic Mean Estimator . . . . .                              | 60        |
| 3.3.3    | Reciprocal Importance Estimator . . . . .                       | 60        |
| 3.3.4    | Chib' Candidate's Estimator . . . . .                           | 60        |

|          |   |            |
|----------|---|------------|
| 3.3.5    | Reversible-Jump MCMC . . . . .  | 61         |
| 3.4      | Theorie des Model Averaging . . . . .   | 61         |
| <b>4</b> | <b>Ergebnisse der Simulationsstudien</b>  | <b>63</b>  |
| 4.1      | Ablauf der Schätzung . . . . .  | 64         |
| 4.2      | Schätzer . . . . .  | 66         |
| 4.3      | ML- und MCMC-Schätzung eines GARCH(1,1)-Modells . . . . .   | 68         |
| 4.3.1    | Wichtige Aspekte bei der Simulation . . . . .   | 73         |
| 4.3.2    | Asymptotische Normalverteilung der ML-Schätzer . . . . .  | 84         |
| 4.3.3    | Vergleich der MCMC-Schätzung und der ML-Schätzer . . . . .  | 96         |
| 4.4      | ML- und MCMC-Simulationen diverser GARCH-Varianten . . . . .  | 110        |
| 4.4.1    | Simulationen zu EGARCH-Modellen . . . . .   | 110        |
| 4.4.2    | Simulationen zu GJR-GARCH-Modellen . . . . .  | 112        |
| 4.4.3    | Simulation zu einem AP-ARCH-Modell . . . . .  | 122        |
| 4.4.4    | Simulationen zu TS-GARCH-Modellen . . . . .   | 123        |
| 4.4.5    | Simulation zu einem T-GARCH-Modell . . . . .  | 131        |
| 4.4.6    | Zusammenfassung . . . . .   | 131        |
| 4.5      | Prognosen mittels der a posteriori Prädiktivverteilung . . . . .                                    | 132        |
| 4.6      | Simulationen mit multivariaten Zeitreihen . . . . .   | 133        |
| 4.7      | Anwendung der MCMC-Methodik auf reale univariate Datensätze . . . . .                               | 139        |
| 4.8      | Anwendungsbeispiel zur Modellselektion . . . . .  | 148        |
| <b>5</b> | <b>Zusammenfassung und Ausblick</b>   | <b>151</b> |
|          | <b>Literaturverzeichnis</b>   | <b>153</b> |
| <b>A</b> | <b>Anhang</b>   | <b>162</b> |
| A.1      | Grafiken zu den Ergebnissen der robusten Schätzung der Varianzen bei<br>der ML-Simulation . . . . . | 162        |
| A.2      | Grafiken zu den Simulationsstudien zur ML-Schätzung eines<br>GARCH(1,1)-Modells . . . . .           | 164        |
| A.2.1    | Variante 2 . . . . .  | 164        |
| A.2.2    | Variante 4 . . . . .  | 166        |

---

|   |     |
|---|-----|
| A.2.3 Variante 5 . . . . .  | 169 |
| A.3 Ergebnisse der Simulationsstudie zum TS-GARCH(1,1)-Modell . . . . . | 171 |

---

# Abbildungsverzeichnis

|      |  |    |
|------|--|----|
| 1.1  | Gewinn- und Verlustdiagramm eines Long (a) bzw. Short (b) Call . . . . .   | 4  |
| 1.2  | Gewinn- und Verlustdiagramm eines Long (a) bzw. Short (b) Put . . . . .  | 5  |
| 1.3  | Darstellung der Tagesschlusskurse des VDAX vom 16.11.2005 bis<br>18.09.2009 (n=972 Handelstage) . . . . .                  | 9  |
| 1.4  | Tagesschlusskurse des DAX vom 26.11.1990 bis 17.09.2009 (n=4750<br>Handelstage) . . . . .                                  | 10 |
| 1.5  | Kerndichteschätzung für die Rendite des DAX und eine an die Daten an-<br>gepasste Normalverteilung . . . . .               | 12 |
| 1.6  | Tagesschlusskurse der Aktie des Volkswagen AG im Zeitraum vom<br>01.01.03 bis 19.09.2009 . . . . .                         | 13 |
| 1.7  | Darstellung der stetigen Renditen des DAX vom 26.11.1990 bis<br>17.09.2009 (n=4750 Handelstage) . . . . .                  | 14 |
| 1.8  | Darstellung der Absolutwerte der stetigen Renditen des DAX vom<br>26.11.1990 bis 17.09.2009 (n=4750 Handelstage) . . . . . | 15 |
| 1.9  | Gewinn- und Verlustdiagramm eines Long (a) bzw. Short (b) Straddle . .   | 16 |
| 1.10 | Gewinn- und Verlustdiagramm eines Long (a) bzw. Short (b) Strangle . .   | 17 |
| 1.11 | Gewinn- und Verlustdiagramm eines Long (a) bzw. Short (b) Strip . . . .  | 18 |
| 1.12 | Gewinn- und Verlustdiagramm eines Long (a) bzw. Short (b) Strap . . . .  | 18 |
| 1.13 | Gewinn- und Verlustdiagramm eines Butterfly-Spread . . . . .   | 19 |
| 2.1  | Angepasste t-Verteilung und Dichteschätzung . . . . .  | 28 |
| 2.2  | Angepasste GED-Verteilung und Dichteschätzung . . . . .  | 29 |
| 2.3  | H-GARCH und genestete Modelle . . . . .  | 32 |
| 2.4  | Tagesschlusskurse vom 26.11.1990 bis 17.09.2009 für DAX, Dow Jones<br>und Nikkei . . . . .                                 | 37 |

|      |   |    |
|------|---|----|
| 2.5  | Renditen vom 26.11.1990 bis 17.09.2009 für DAX, Dow Jones und Nikkei  | 38 |
| 4.1  | Markovketten des ersten Laufs für $y5000garch\_norm$ . . . . .  | 70 |
| 4.2  | Histogramme des ersten Laufs für $y5000garch\_norm$ . . . . .   | 71 |
| 4.3  | Markovketten des zweiten Laufs für $y5000garch\_norm$ . . . . .   | 72 |
| 4.4  | Histogramme des zweiten Laufs für $y5000garch\_norm$ . . . . .  | 73 |
| 4.5  | Markovketten des ersten Laufs bei ungünstigen Startwerten . . . . .   | 74 |
| 4.6  | Markovketten des zweiten Laufs bei ungünstigen Startwerten im ersten<br>Lauf . . . . .  | 74 |
| 4.7  | Folge ungünstiger Startwerte und Vorschlagsdichten . . . . .  | 76 |
| 4.8  | Vergleich der a posteriori Verteilung und der Log-Likelihood . . . . .  | 83 |
| 4.9  | Verteilung des ML-Schätzers des Parameters $\alpha_0$ : 0.2 bei Variante 1, bei<br>unterschiedlichen Stichprobenumfängen . . . . .  | 91 |
| 4.10 | Verteilung des ML-Schätzers des Parameters $\alpha_1$ : 0.2 bei Variante 1, bei<br>unterschiedlichen Stichprobenumfängen . . . . .  | 92 |
| 4.11 | Verteilung des ML-Schätzers des Parameters $\beta_1$ : 0.6 bei Variante 1, bei<br>unterschiedlichen Stichprobenumfängen . . . . .   | 92 |
| 4.12 | Grafische Darstellung der robusten Schätzer für die Varianz der 1000 Re-<br>plikationen für die Stichprobenumfänge $n=100, 250$ und $500$ bei Variante<br>2 . . . . .   | 94 |
| 4.13 | Verteilung des ML-Schätzers des Parameters $\alpha_0$ : 0.05 bei Variante 3, bei<br>unterschiedlichen Stichprobenumfängen . . . . .   | 95 |
| 4.14 | Verteilung des ML-Schätzers des Parameters $\alpha_1$ : 0.05 bei Variante 3, bei<br>unterschiedlichen Stichprobenumfängen . . . . .   | 95 |
| 4.15 | Verteilung des ML-Schätzers des Parameters $\beta_1$ : 0.94 bei Variante 3, bei<br>unterschiedlichen Stichprobenumfängen . . . . .  | 96 |
| 4.16 | Randverteilungen der ML-Schätzer bei unterschiedlichen Stichprobe-<br>numfängen im Vergleich zu der a posteriori Verteilung der MCMC-<br>Schätzung eines einzelnen Datensatzes ( $n=100, 250, 500$ und $1000$ ) . . . | 97 |
| 4.17 | Randverteilungen der ML-Schätzer bei unterschiedlichen Stichprobe-<br>numfängen im Vergleich zu der a posteriori Verteilung der MCMC-<br>Schätzung eines einzelnen Datensatzes ( $n=2500$ und $5000$ ) . . . . .      | 98 |

---

|      |   |     |
|------|---|-----|
| 4.18 | Vergleich der Histogramme der Verteilung der ML-Schätzer und der a posteriori Verteilung auf Basis eines Datensatzes ( $n=100, 250, 500$ und $1000$ ) . . . . . | 100 |
| 4.19 | Vergleich der Histogramme der Verteilung der ML-Schätzer und der a posteriori Verteilung auf Basis eines Datensatzes ( $n=2500$ und $5000$ ) . . .              | 101 |
| 4.20 | Vergleich der Randverteilungen der ML-Schätzung und einer MCMC-Schätzung an einem Datensatz ( $n=100, 250, 500$ und $1000$ ) . . . . .                          | 102 |
| 4.21 | Vergleich der Randverteilungen der ML-Schätzung und einer MCMC-Schätzung an einem Datensatz ( $n=2500$ und $5000$ ) . . . . .                                   | 103 |
| 4.22 | Markovkette und Histogramm für die Anzahl der Freiheitsgrade einer $t$ -Verteilung . . . . .  | 106 |
| 4.23 | Verteilung des ML-Schätzers für die Anzahl der Freiheitsgrade einer $t$ -Verteilung bei unterschiedlichen Stichprobenumfängen . . . . .                         | 107 |
| 4.24 | Markovkette und Histogramm für den Parameter $\nu$ einer $GED$ -Verteilung  | 108 |
| 4.25 | Verteilung des ML-Schätzers für den Parameter $\nu$ einer $GED$ -Verteilung bei unterschiedlichen Stichprobenumfängen . . . . .                                 | 109 |
| 4.26 | Verteilung des ML-Schätzers des Parameters $\alpha_0: 0.1$ bei unterschiedlichen Stichprobenumfängen . . . . .  | 113 |
| 4.27 | Verteilung des ML-Schätzers des Parameters $\alpha_1: 0.1$ bei unterschiedlichen Stichprobenumfängen . . . . .  | 113 |
| 4.28 | Verteilung des ML-Schätzers des Parameters $c: -0.1$ bei unterschiedlichen Stichprobenumfängen . . . . .  | 114 |
| 4.29 | Verteilung des ML-Schätzers des Parameters $\beta_1: 0.90$ bei unterschiedlichen Stichprobenumfängen . . . . .  | 114 |
| 4.30 | Verteilung der ML-Schätzer der Parameter bei unterschiedlichen Stichprobenumfängen . . . . .  | 116 |
| 4.31 | Verteilung des ML-Schätzers des Parameters $\alpha_0: 0.03$ bei unterschiedlichen Stichprobenumfängen . . . . .   | 119 |
| 4.32 | Verteilung des ML-Schätzers des Parameters $\alpha_1: 0.03$ bei unterschiedlichen Stichprobenumfängen . . . . .   | 119 |



|      |   |     |
|------|---|-----|
| 4.33 | Verteilung des ML-Schätzers des Parameters $c$ : 0.1 bei unterschiedlichen Stichprobenumfängen . . . . .  | 120 |
| 4.34 | Verteilung der ML-Schätzers des Parameters $\beta_1$ : 0.9 bei unterschiedlichen Stichprobenumfängen . . . . .  | 120 |
| 4.35 | Vergleich der Histogramme der Verteilung der ML-Parameterschätzer und der a posteriori Verteilung auf Basis eines Datensatzes (Spalte 1: $n=250$ , Spalte 2: $n=500$ und Spalte 3: $n=1000$ ) . . . . . | 121 |
| 4.36 | Verteilung des ML-Schätzers des Parameters bei Variante 1, $\alpha_0$ : 0.1 bei unterschiedlichen Stichprobenumfängen . . . . .   | 128 |
| 4.37 | Verteilung des ML-Schätzers des Parameters bei Variante 1, $\alpha_1$ : 0.1 bei unterschiedlichen Stichprobenumfängen . . . . .   | 128 |
| 4.38 | Verteilung des ML-Schätzers des Parameters bei Variante 1, $\beta_1$ : 0.8 bei unterschiedlichen Stichprobenumfängen . . . . .  | 129 |
| 4.39 | Vergleich der Histogramme der Verteilung der ML-Schätzer und der a posteriori Verteilung auf Basis eines Datensatzes ( $n=100, 250, 500$ und $1000$ ) . . . . .   | 130 |
| 4.40 | Prädiktivverteilung . . . . .   | 132 |
| 4.41 | Prognosewerte für ein GARCH(1,1)-Modell . . . . .   | 133 |
| 4.42 | CCC-GARCH(1,1): Parameter Modell 1 . . . . .  | 134 |
| 4.43 | CCC-GARCH(1,1): Parameter Modell 2 . . . . .  | 135 |
| 4.44 | CCC-GARCH(1,1)-Modell $\rho$ . . . . .  | 136 |
| 4.45 | DCC-GARCH(1,1)-Modell $\alpha^*$ . . . . .  | 138 |
| 4.46 | DCC-GARCH(1,1)-Modell $\beta^*$ . . . . .   | 138 |
| A.1  | Ergebnisse der robusten Schätzung der Varianzen bei der ML-Simulationsstudie (GARCH) ( $n=100, 250$ ) . . . . .   | 162 |
| A.2  | Verteilung der ML-Parameterschätzer bei unterschiedlichen Stichprobenumfängen im Vergleich zu der a posteriori Verteilung der MCMC-Schätzung ( $n=500, 1000, 2500$ und $5000$ ) . . . . .               | 163 |
| A.3  | Verteilung des ML-Schätzers des Parameters $\alpha_0$ : 0.1 bei Variante 2, bei unterschiedlichen Stichprobenumfängen . . . . .   | 164 |

---

|      |  |     |
|------|--|-----|
| A.4  | Verteilung des ML-Schätzers des Parameters $\alpha_1$ : 0.1 bei Variante 2, bei unterschiedlichen Stichprobenumfängen . . . . .  | 165 |
| A.5  | Verteilung des ML-Schätzers des Parameters $\beta_1$ : 0.8 bei Variante 2, bei unterschiedlichen Stichprobenumfängen . . . . .   | 165 |
| A.6  | Verteilung des ML-Schätzers des Parameters $\alpha_0$ : 0.4 bei Variante 4, bei unterschiedlichen Stichprobenumfängen . . . . .  | 166 |
| A.7  | Verteilung des ML-Schätzers des Parameters $\alpha_1$ : 0.49 bei Variante 4, bei unterschiedlichen Stichprobenumfängen . . . . . | 167 |
| A.8  | Verteilung des ML-Schätzers des Parameters $\beta_1$ : 0.49 bei Variante 4, bei unterschiedlichen Stichprobenumfängen . . . . .  | 167 |
| A.9  | Verteilung des ML-Schätzers des Parameters $\alpha_0$ : 0.1 bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . .  | 169 |
| A.10 | Verteilung des ML-Schätzers des Parameters $\alpha_1$ : 0.3 bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . .  | 170 |
| A.11 | Verteilung des ML-Schätzers des Parameters $\beta_1$ : 0.3 bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . .   | 170 |
| A.12 | Verteilung des ML-Schätzers des Parameters $\alpha_0$ : 0.2 bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . .  | 171 |
| A.13 | Verteilung des ML-Schätzers des Parameters $\alpha_1$ : 0.2 bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . .  | 172 |
| A.14 | Verteilung des ML-Schätzers des Parameters $\beta_1$ : 0.6 bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . .   | 172 |
| A.15 | Verteilung des ML-Schätzers des Parameters $\alpha_0$ : 0.05 bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . . | 173 |
| A.16 | Verteilung des ML-Schätzers des Parameters $\alpha_1$ : 0.05 bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . . | 173 |
| A.17 | Verteilung des ML-Schätzers des Parameters $\beta_1$ : 0.94 bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . .  | 174 |
| A.18 | Verteilung des ML-Schätzers des Parameters $\alpha_0$ : 0.4 bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . .  | 174 |

---

|  |     |
|--|-----|
| A.19 Verteilung des ML-Schätzers des Parameters $\alpha_1: 0.49$ bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . . | 175 |
| A.20 Verteilung des ML-Schätzers des Parameters $\beta_1: 0.49$ bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . .  | 175 |
| A.21 Verteilung des ML-Schätzers des Parameters $\alpha_0: 0.1$ bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . .  | 176 |
| A.22 Verteilung des ML-Schätzers des Parameters $\alpha_1: 0.3$ bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . .  | 176 |
| A.23 Verteilung des ML-Schätzers des Parameters $\beta_1: 0.3$ bei Variante 1, bei unterschiedlichen Stichprobenumfängen . . . . .   | 177 |

# Tabellenverzeichnis

|      |  |     |
|------|--|-----|
| 2.1  | Genestete Modelle in der Übersicht . . . . .   | 33  |
| 4.1  | GARCH-Modell mit y5000garch_norm . . . . .   | 70  |
| 4.2  | GARCH-Modell mit y5000garch_norm . . . . .   | 71  |
| 4.3  | Ergebnisse bei ungünstigen Startwerten . . . . .   | 75  |
| 4.4  | Einfluss der Schätzung von $\sigma_0^2$ . . . . .  | 79  |
| 4.5  | Auswahl unterschiedlicher Werte für die Analyse der a posteriori Verteilung: (1) alle Werte, (2) jeder 10. Wert und (3) akzeptierte Kandidaten . . . . . | 82  |
| 4.6  | Vergleich der Schätzer für die Varianzen (ML versus MCMC) . . . . .  | 85  |
| 4.7  | Vergleich der GARCH-Parameter für die Renditen des DAX für vier Zeiträume mit jeweils $n = 1000$ . . . . .   | 86  |
| 4.8  | Mittelwerte und Varianzen der ML-Schätzer bei unterschiedlichen Stichprobenumfängen (mäßige Persistenz) . . . . .  | 88  |
| 4.9  | Mittelwerte und Varianzen der ML-Schätzer bei unterschiedlichen Stichprobenumfängen (hohe Persistenz) . . . . .  | 89  |
| 4.10 | Stichprobenumfang: 100 . . . . .   | 104 |
| 4.11 | Stichprobenumfang: 250 . . . . .   | 104 |
| 4.12 | Stichprobenumfang: 500 . . . . .   | 105 |
| 4.13 | Stichprobenumfang: 1000 . . . . .  | 105 |
| 4.14 | GARCH <sub>t</sub> -Modell mit y5000garch_t . . . . .  | 106 |
| 4.15 | GARCH <sub>GED</sub> -Modell mit y5000garch_ged . . . . .  | 108 |
| 4.16 | GARCH <sub>t</sub> -Modell mit normalverteilten Residuen (unter der Annahme einer $t$ -Verteilung geschätzt) . . . . .                                   | 110 |
| 4.17 | GARCH <sub>GED</sub> -Modell mit normalverteilten Residuen (unter der Annahme einer $GED$ -Verteilung geschätzt) . . . . .                               | 110 |

---

|      |  |     |
|------|--|-----|
| 4.18 | ML-Schätzergebnisse für ein EGARCH-Modell mit simulierten Daten . . .  | 111 |
| 4.19 | EGARCH-Modell mit normalverteilten Residuen . . . . .  | 115 |
| 4.20 | ML-Schätzergebnisse (GJR-GARCH) für y5000gjrgarch . . . . .  | 117 |
| 4.21 | GJR-GARCH-Modell mit normalverteilten Residuen . . . . .   | 118 |
| 4.22 | Ergebnisse für 1000 Replikationen eines GJR-GARCH(1,1)-Modells mit<br>$\alpha_0 = 0.03$ , $\alpha_1 = 0.03$ , $c = 0.1$ und $\beta_1 = 0.9$ bei unterschiedlichen<br>Stichprobenumfängen $n$ . . . . . | 122 |
| 4.23 | ML-Schätzergebnisse (AP-ARCH) für y5000aparch . . . . .  | 123 |
| 4.24 | ML-Schätzergebnisse (TS-GARCH) für y1000tsgarch . . . . .  | 124 |
| 4.25 | Mittelwerte und Varianzen der ML-Schätzer bei unterschiedlichen Stich-<br>probenumfängen (mäßige Persistenz) . . . . .   | 126 |
| 4.26 | Mittelwerte und Varianzen der ML-Schätzer bei unterschiedlichen Stich-<br>probenumfängen (hohe Persistenz) . . . . .   | 127 |
| 4.27 | ML-Schätzergebnisse für die simulierten Daten eines T-GARCH(1,1)-<br>Modells . . . . .   | 131 |
| 4.28 | Ergebnisse der Schätzung der Parameter eines CCC-GARCH(1,1)-Modells  | 136 |
| 4.29 | Ergebnisse der Schätzung der GARCH-Parameter eines DCC-GARCH-<br>Modells . . . . .   | 137 |
| 4.30 | Ergebnisse der Schätzung der DCC-Parameter eines DCC-GARCH-Modell  | 138 |
| 4.31 | Schätzergebnisse verschiedener GARCH(1,1) und TS-GARCH(1,1)-<br>Modelle für um den Mittelwert bereinigte, stetige Renditen des DAX . . .   | 141 |
| 4.32 | Schätzergebnisse verschiedener T-GARCH(1,1)- und GJR-GARCH(1,1)-<br>Modelle für um den Mittelwert bereinigte, stetige Renditen des DAX . . .   | 142 |
| 4.33 | Schätzergebnisse verschiedener EGARCH(1,1)-Modelle für um den Mit-<br>telwert bereinigte, stetige Renditen des DAX . . . . .   | 143 |
| 4.34 | Schätzergebnisse verschiedener AP-ARCH(1,1)-Modelle für um den Mit-<br>telwert bereinigte, stetige Renditen des DAX . . . . .  | 144 |
| 4.35 | Schätzergebnisse verschiedener GARCH(1,1) und TS-GARCH(1,1)-<br>Modelle für um den Mittelwert bereinigte, stetige Renditen des Dow<br>Jones . . . . .  | 145 |

---

|      |  |     |
|------|--|-----|
| 4.36 | Schätzergebnisse verschiedener T-GARCH(1,1)- und GJR-GARCH(1,1)-<br>Modelle für um den Mittelwert bereinigte, stetige Renditen des Dow Jones                           | 146 |
| 4.37 | Schätzergebnisse verschiedener EGARCH(1,1)-Modelle für um den Mit-<br>telwert bereinigte, stetige Renditen des Dow Jones . . . . .                                     | 147 |
| 4.38 | Schätzergebnisse verschiedener AP-ARCH(1,1)-Modelle für um den Mit-<br>telwert bereinigte, stetige Renditen des DAX . . . . .  | 148 |
| 4.39 | Verschiedene Kriterien zur Modellselektion . . . . .   | 149 |
| 4.40 | Jeffrey's Skala zur Interpretation von Bayes-Faktoren . . . . .  | 150 |
| 4.41 | Bayes-Faktoren . . . . .   | 150 |
| A.1  | Ergebnisse für 1000 Replikationen eines GARCH(1,1)-Modells mit $\alpha_0 =$<br>0.1, $\alpha_1 = 0.3$ und $\beta_1 = 0.3$ bei unterschiedlichen Stichprobenumfängen $n$ | 168 |

# Abkürzungsverzeichnis

## I. Generelle Bezeichnungen

|                     |                       |
|---------------------|-----------------------|
| $\sim$              | verteilt              |
| $\overset{a}{\sim}$ | asymptotisch verteilt |
| $\approx$           | approximativ gleich   |
| $:=$                | definiert durch       |

## II. Mengen und Räume

|                |                              |
|----------------|------------------------------|
| $\mathbb{N}$   | natürliche Zahlen            |
| $\mathbb{N}_0$ | natürliche Zahlen inkl. Null |
| $\mathbb{R}$   | reelle Zahlen                |
| $\mathbb{R}^+$ | positive reelle Zahlen       |

## III. Funktionen und Zufallsvariablen

|                         |  |
|-------------------------|--|
| $Cov$                   | Kovarianz  |
| $E$                     | Erwartungswert   |
| $F(x)$                  | Verteilungsfunktion  |
| $f(x)$                  | Wahrscheinlichkeits- bzw. Dichtefunktion                               |
| $g^{(n)}(x)$            | n-te Ableitung von $g(x)$  |
| $\ln(x)$                | natürlicher Logarithmus  |
| $\log N(\mu, \sigma^2)$ | Lognormalverteilung mit Erwartungswert $\mu$<br>und Varianz $\sigma^2$ |

---

|                       |   |
|-----------------------|---|
| $\Phi(\mu, \sigma^2)$ | Verteilungsfunktion der Normalverteilung<br>mit Erwartungswert $\mu$ und Varianz $\sigma^2$ |
| $N(\mu, \sigma^2)$    | Normalverteilung mit Erwartungswert $\mu$<br>und Varianz $\sigma^2$                         |
| $Pr[A]$               | Wahrscheinlichkeit des Ereignisses A  |
| $Pr[A B]$             | Bedingte Wahrscheinlichkeit von A, gegeben B  |
| $t(n)$                | $t$ -Verteilung mit $n$ Freiheitsgraden   |
| $Var$                 | Varianz   |

#### IV. Abkürzungen

|         |  |
|---------|--|
| AIC     | Akaike Informationskriterium   |
| AG      | Aktiengesellschaft   |
| AP-ARCH | Asymmetrisches Power ARCH  |
| ARCH    | Autoregressive Conditional Heteroscedastic                               |
| ARMA    | Autoregressive Moving Average  |
| BF      | Bayes-Faktor   |
| BIC     | Bayesian Information Criterion<br>(Bayesianisches Informationskriterium) |
| bzw.    | beziehungsweise  |
| bspw.   | beispielsweise   |
| CCC     | Constant-Conditional-Correlation   |
| engl.   | englisch   |
| et al.  | (lat.) et alii (und andere)  |
| etc.    | etcetera   |
| iid     | identical independently distributed<br>(unabhängig identisch verteilt)   |
| DAX     | Deutscher Aktienindex  |
| DCC     | Dynamic-Conditional-Correlation  |



---

|          |   |
|----------|---|
| EGARCH   | Exponential Generalized Autoregressive<br>Conditional Heteroscedastic |
| FIGARCH  | Fraktionelles Integriertes GARCH                                      |
| GARCH    | Generalized Autoregressive Conditional<br>Heteroscedastic             |
| GED      | Generalized Error Distribution  |
| ggT      | größter gemeinsamer Teiler  |
| GMM      | Generalized Method of Moments   |
| H-GARCH  | GARCH-Modell nach Hentschel   |
| lat.     | lateinisch  |
| LM       | Lagrange-Multiplikator  |
| MAP      | Maximum-A-Posteriori  |
| MC       | Monte Carlo   |
| MCMC     | Markov Chain Monte Carlo  |
| MDE      | Minimum Distance Estimator  |
| ML       | Maximum Likelihood  |
| MSE      | Mean Square Error (mittlerer quadratischer Fehler)                    |
| OLS      | Ordinary Least Square (Kleinst-Quadrat Schätzung)                     |
| QML      | Quasi Maximum Likelihood  |
| S.       | Seite oder Seiten   |
| sog.     | so genannt/er   |
| SoWi     | Sozial- und Wirtschaftswissenschaftliche Fakultät                     |
| T-GARCH  | Threshold-GARCH   |
| TS-GARCH | GARCH-Modell nach Taylor und Schwert                                  |
| u.a.     | unter anderem   |
| usw.     | und so weiter   |
| u.U.     | unter Umständen   |
| VDAX     | Volatilitätsindex für den Deutschen Aktienindex                       |
| vgl.     | vergleiche  |
| vs.      | versus, gegen   |
| z.B.     | zum Beispiel  |

# Kapitel 1

## Einleitung und Motivation

Die Entwicklung von Finanzmarktdaten wie Renditen bzw. Änderungsraten von Wechselkursen oder Indizes im Zeitverlauf durch ein geeignetes Modell zu beschreiben und zu prognostizieren, stellt in vielen Bereichen eine herausfordernde Aufgabe dar. Mit Hilfe der Zeitreihenanalyse wird versucht Muster in der zeitlichen Entwicklung der Daten zu erkennen und für die Vorhersage zukünftiger Werte zu nutzen. In der Regel lässt sich bei Renditen im Zeitverlauf keine Autokorrelation feststellen, was dazu führt, dass sich zukünftige Renditen nicht aus vergangenen Renditen vorhersagen lassen. Betrachtet man allerdings die Absolutbeträge oder Quadrate der Renditen, so findet man durchaus Autokorrelation. Daher werden häufig die quadrierten Renditen für die Analyse der Zeitreihen verwendet.

Zukünftige Schwankungen, denen solche Zeitreihen unterliegen, beispielsweise die Entwicklung des Preises für Rohöl, stellen aus ökonomischer Sicht ein Risiko dar. Nicht zuletzt durch gesetzliche Vorschriften wie Basel II und Solvency II für die Banken- und Versicherungsbranche, sind Unternehmen daran interessiert Risiken zu kontrollieren. Diverse Finanzmarktinstrumente wurden entwickelt, um sich gegen die unerwünschten Schwankungen einer Investition zu schützen. Andererseits bieten spezielle Anlageformen auch Möglichkeiten die Schwankungen für Spekulationen zu nutzen.

Eine gängige Methode die zeitabhängige Standardabweichung darzustellen, bestand lange Zeit darin eine feste Anzahl von Beobachtungen zu wählen und für diese die Standardabweichung zu bestimmen. Diese feste Anzahl von Beobachtungen, auch "Fenster" genannt, wird dann über die gesamte Zeitreihe geschoben und jeweils die Standardabweichung bestimmt. Diese Methode wird als *rolling window* bezeichnet. Alle betrachteten Beobach-

tungen haben dabei das gleiche Gewicht. Die Annahme, dass neuere Beobachtungen ein höheres Gewicht haben sollten als ältere, veranlasste Engle 1982 zur Entwicklung der autoregressiven, bedingt heteroskedastischen (ARCH-) Modelle. Sie haben sich als äußerst hilfreich erwiesen bei der Modellierung der Volatilität von Zeitreihen- und Finanzmarktdaten. Zahlreiche Erweiterungen dieser Modelle werden heutzutage eingesetzt, um die verschiedenen Eigenschaften der Volatilität abzubilden. Die in Finanzmarktzeitreihen auftretende Clusterbildung der zeitabhängigen Volatilität, bei der Perioden relativer Ruhe von Phasen mit großen Kurschwankungen gefolgt werden, lässt sich durch eine Funktion aller momentan verfügbaren Informationen modellieren. Es existiert bereits eine umfangreiche Literatur über diese Modellklasse und die Forschung entwickelt die Möglichkeiten und Anwendungsgebiete stetig weiter.

Seit einiger Zeit werden GARCH-Modelle auch unter bayesianischen Gesichtspunkten betrachtet. Die fortschreitende Rechenleistung moderner Computer ermöglicht es die komplexen Berechnungen, die bei der Bayes-Analyse notwendig werden, in angemessener Zeit durchzuführen. Dies war in Vergangenheit häufig ein großes Problem der Bayes-Schätzung. Ziel dieser Arbeit ist es die Anwendung der Bayes-Statistik auf GARCH-Modelle und deren Erweiterungen zu demonstrieren und dabei wichtige Aspekte hervorzuheben. Die Ergebnisse werden mit der Maximum-Likelihoodschätzung verglichen, dabei steht die a posteriori Verteilung im Mittelpunkt.

Die Arbeit ist folgendermaßen aufgebaut. Zunächst wird in diesem Kapitel die Verwendung von ARCH- und GARCH Modellen motiviert. Dazu werden zunächst die Einflussfaktoren auf den Preis einer Finanzoption erläutert und die besondere Bedeutung der Volatilität als einzige nicht direkt verfügbare Information dargestellt. Außerdem werden Besonderheiten von Finanzmarktzeitreihen gezeigt, u.a. die bedingte Heteroskedastie, die zu den so genannten Volatilitätsclustern führt, deren Modellierung und Prognose durch GARCH-Modelle möglich ist.

Kapitel 2 gibt einen Überblick über die in dieser Arbeit verwendeten Varianten der ARCH- und GARCH-Modelle. Dabei werden anschließend an das Standard-GARCH-Modell diverse Erweiterungen gezeigt und auf den multivariaten Fall ausgedehnt. Von besonderer Bedeutung für diese Arbeit ist die, alternativ zur frequentistischen Schätzung, Bayes-Schätzung der Modelle mittels Markov Chain Monte Carlo (MCMC)-Methoden.

Der Bayes-Ansatz wird gemeinsam mit den MCMC-Methoden in Kapitel 3 behandelt. Besondere Bedeutung hat dabei die Modellselektion, d.h. die Wahl eines geeigneten Modells, da diese maßgeblich für die Qualität der Modellierung und damit auch der Prognose der Volatilität ist. Für den Fall, dass die Modellselektion kein einzelnes Modell favorisiert, besteht die Möglichkeit mit Hilfe des "Model Averaging", also der Mittelung der Ergebnisse verschiedener Modelle, die Prognose zu verbessern.

Den Kern der Arbeit bildet Kapitel 4. Hier wird die grundsätzliche Vorgehensweise bei der bayesianischen Schätzung von GARCH-Modellen an simulierten Datensätzen erläutert und durch die Betrachtung verschiedener Besonderheiten vertieft.

Ebenso erfolgt ein Vergleich der Schätzung der Modelle mittels Maximum-Likelihood-Methode und Bayes, dabei stehen die Eigenschaften der Schätzer in kleinen Stichproben und die asymptotischen Eigenschaften im Vordergrund. Weitere Berechnungen demonstrieren die Anwendung der Bayes-Schätzung auf diverse Varianten des GARCH-Modells unter Berücksichtigung verschiedener Verteilungsannahmen sowie zwei multivariate Erweiterungen der GARCH-Modelle. Die Anwendung auf reale Datensätze in Verbindung mit Berechnungen zur Modell-Selektion beschließen Kapitel 4.

Die Arbeit endet mit einer Zusammenfassung der Ergebnisse und einem Ausblick.

## 1.1 Motivation

Dieses Kapitel gibt eine kurze Einführung in das Grundkonzept einer Option im Sinne eines derivativen Finanzinstruments und motiviert diese Arbeit. Dazu beschreibt Abschnitt 1.2 zunächst das Wesen einer Finanzoption und Abschnitt 1.2.1 dann die wesentlichen Einflussfaktoren auf den Preis oder Wert einer Option. Unterschiedliche Begriffe der Volatilität als einer dieser Einflussfaktoren werden in Abschnitt 1.2.2 speziell erläutert. Dabei soll die Notwendigkeit der Modellierung und Prognose von Volatilität hervorgehoben werden, was dann mit den in Kapitel 2 vorgestellten Modellen mit bedingter Heteroskedastie geschieht. Einen weiteren Einblick in die Volatilität der deutschen Aktienlandschaft ermöglicht der VDAX, welcher in Abschnitt 1.2.3 separat vorgestellt wird. Abschnitt 1.3 zeigt einige Besonderheiten auf, die Zeitreihen aus Kapitalmärkten von anderen Zeitreihen unterscheiden, dies ist von besonderer Bedeutung, da für viele weitere Überlegungen Annahmen bezüglich der Zeitreihen getroffen werden, die sich empirisch nicht belegen

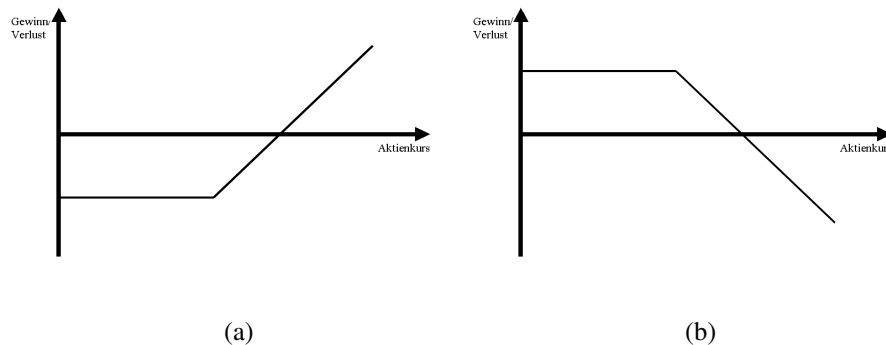


Abbildung 1.1: Gewinn- und Verlustdiagramm eines Long (a) bzw. Short (b) Call

lassen.

Die praktische Relevanz der Modellierung der Schwankungsbreite von Investitionsobjekten wird besonders deutlich bei Betrachtung von Kombinationsstrategien, weshalb in Abschnitt 1.4 die wichtigsten Strategien aufgeführt werden.

Die Optionspreistheorie kann und soll hier nicht umfassend behandelt werden, deshalb sei hier an Copeland und Weston (1992) [27] und Hull (2003) [57] verwiesen, die noch weitere Aspekte von Optionen und Futures erläutern.

## 1.2 Optionen

Unter Optionen im Sinne von derivativen Finanzmarktinstrumenten versteht man das Recht, an oder bis zu einem bestimmten Verfallstermin ein bestimmtes Basisinstrument (Underlying) zu einem vorher bestimmten Preis (Basispreis) zu kaufen oder zu verkaufen (Call-Option oder Put-Option), ohne sich dabei zu verpflichten. Die jeweilige Option wird dabei durch die Optionsart (Call oder Put), das Basisinstrument (z.B. Aktie, Rohstoff, Wahrung Index, etc.), den Basispreis, den Verfallstermin sowie die Borse, an der sie gehandelt wird, charakterisiert.

Einleitend werden die verschiedenen Rechte und Pflichten betrachtet. Die so genannte Long-Position entspricht dabei dem Kauf einer Option. Der Verkauf einer Option wird auch Short-Position genannt, daher wird der Leerverkauf, d.h. der Verkauf eines Gutes, welches nicht im eigenen Besitz ist, auch "short gehen" oder "shorten" genannt. Die Long-Position berechtigt zum Ausuben der Option, wahrend der Stillhalter (Short-

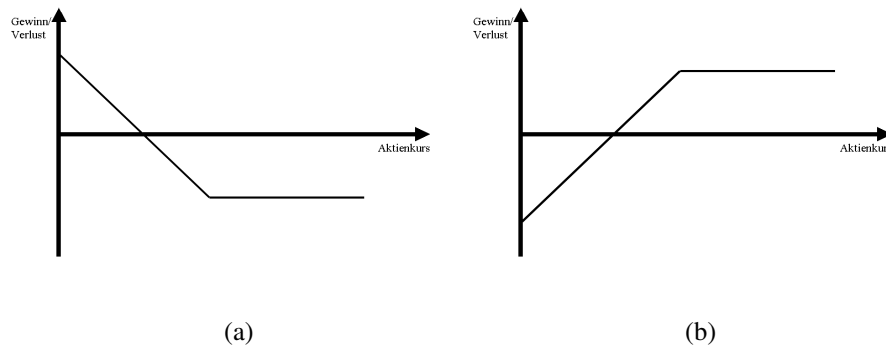


Abbildung 1.2: Gewinn- und Verlustdiagramm eines Long (a) bzw. Short (b) Put

Position) bei Ausübung der Option verpflichtet ist das entsprechende Underlying zu liefern.

Die Abbildung 1.1 zeigt die Gewinn- und Verlustdiagramme eines Long Call bzw. eines Short Call. Die entsprechenden Diagramme für eine Put-Option zeigt die Abbildung 1.2. Es wird dabei deutlich, dass der maximale Verlust einer Long-Positionen begrenzt ist, wohingegen der potentielle Gewinn unbegrenzt bleibt. Entsprechend Gegensätzliches gilt für die Short-Positionen. Dies wird klar aus der Tatsache, dass der Käufer der Option den Preis der Option bezahlt, um sich das entsprechende Recht zu sichern. Dies stellt also den maximalen Verlust dar, da im Falle einer ungünstigen Entwicklung des Underlyings das Recht nicht ausgeübt wird.

In der Folge der Krise auf den internationalen Finanzmärkten 2008 hat sich der Begriff "Casino-Kapitalismus" etabliert, der die Investition in derivative Finanzmarktinstrumente beschreibt, welche ausschließlich zu spekulativen Zwecken durchgeführt wird. Dies geschieht unter anderem mittels der Hebelwirkung oder Leverage-Effekt (nicht zu verwechseln mit dem Leverage-Effekt aus Abschnitt 2.4.3). Die Hebelwirkung resultiert aus der Tatsache, dass die Option relativ zum Underlying einen niedrigeren Preis hat. So haben Kursschwankungen einen prozentual größeren Einfluss auf die Option als auf das Underlying. Daneben gibt es verschiedene andere Gründe für ein Engagement in Optionen und Futures, bspw. zur Absicherung des eigenen Bestands oder Bedarfs an Rohstoffen. Diese Motive tragen eher den Charakter einer Versicherung als den der Spekulation.

### 1.2.1 Einflussfaktoren auf den Optionspreis

Nachdem das Wesen einer Option sowie die damit verbundenen Gewinn- und Verlustdiagramme dargestellt wurden, erläutert dieser Abschnitt die Faktoren, die maßgeblich für die Bestimmung des Preises der Option sind.

#### **Preis des Underlyings**

Es ist unmittelbar einsichtig, dass der Preis des Underlyings der wichtigste Faktor für den Preis der Option selbst ist. Dabei gilt, je höher (niedriger) der Kurs des Underlyings, desto teurer der Preis der entsprechenden Call-Option (Put-Option).

#### **Basispreis**

Eine weitere Möglichkeit den Preis einer Option zu verändern, ist die Festlegung des Basispreises. Dabei handelt es sich um den Preis, zu dem die Option bei Verfall ausgeübt werden kann. Man unterscheidet zwischen amerikanischen und europäischen Optionen. Eine europäische Option kann nur am Verfallstag ausgeübt werden, eine amerikanische Option kann während der gesamten Laufzeit ausgeübt werden. Für eine Kaufoption gilt, je höher der Basispreis, also der Preis zu dem das Underlying erworben werden kann, umso geringer die Wahrscheinlichkeit, dass der Preis der Aktie den Basispreis überschreitet und so Gewinn bringt. Also erhöht ein niedriger (hoher) Basispreis den Wert einer Call (Put)-Option.

Man unterscheidet in den Gewinn- und Verlustdiagrammen verschiedene Phasen. Solange der Preis des Underlyings den Basispreis nicht erreicht, wird der Käufer der Option sein Recht nicht ausüben, die Option ist in dieser Situation "aus dem Geld". Sobald das Underlying den Basispreis über- bzw. unterschreitet (Call- bzw. Put-Option) befindet sich die Option "im Geld".

Der innere Wert einer Option ist die Differenz zwischen dem aktuellen Kurs des Underlyings und dem Basispreis.

#### **Laufzeit**

Die Laufzeit ist ebenfalls wichtig für die Bestimmung des Wertes einer Option. Dabei ist die Restlaufzeit, also die Zeit, die bis zum Verfall der Option verbleibt, gemeint. Eine lange Restlaufzeit bedeutet für den Stillhalter der Option, dass er das Underlying länger vorhalten muss und zusätzlich, dass die Option mehr Zeit hat, um "ins Geld" zu kommen. Dies stellt für den Stillhalter ein erhöhtes Risiko dar, welches er sich vergüten lässt.

Der Zeitwert errechnet sich aus der Differenz zwischen dem Kurswert und dem inneren Wert. Da sich der innere Wert zum Ende der Laufzeit dem Kurswert nähert, geht der Zeit-

wert im Zeitverlauf gegen null.

### Zinsniveau

Dazu kommt noch das allgemeine Zinsniveau als preisbestimmende Größe. Der Stillhalter der Option muss das Underlying bereithalten, um im Falle der Ausübung der Option liefern zu können. Dabei muss er auf die Anlage des Gegenwerts verzichten. Dieser Zinssausfall ist umso höher, je höher das allgemeine Zinsniveau ist.

### Volatilität

Alle bisher genannten Einflussfaktoren können direkt am Markt beobachtet werden. Das bedeutet der Käufer einer Option kann sich aus der Fülle des Angebots diejenige Option auswählen, die ihm bezüglich der Einflussfaktoren am besten geeignet erscheint, bzw. ihm im Fall des Zinssatzes zumindest Sicherheit gibt, da er die Höhe kennt.

Der letzte, und für diese Arbeit wichtigste Einflussfaktor, ist die Volatilität. Im Gegensatz zu den oben genannten Faktoren ist dieser Parameter nicht direkt zu beobachten. Dies erklärt die Bemühungen diese Größe möglichst gut zu quantifizieren und zu prognostizieren. Der nächste Abschnitt erklärt den Begriff und die Bedeutung für diese Arbeit.

## 1.2.2 Volatilität

Der Kurs oder Preis eines Finanzmarktprodukts zum Zeitpunkt  $t$  sei  $K_t$ , dann ergibt sich die diskrete Rendite zum Zeitpunkt  $t$  durch

$$r_t = \frac{K_t - K_{t-1}}{K_{t-1}}. \quad (1.1)$$

Nimmt man an, dass die Wertzuwächse in  $n$  äquidistanten Subperioden wieder angelegt werden, so ergibt sich die die stetige Rendite durch

$$r_t = \ln \left( \frac{K_t}{K_{t-1}} \right) = \ln(K_t) - \ln(K_{t-1}). \quad (1.2)$$

Sehr häufig wird für die Modellierung der Verteilung der Preise die Log-Normalverteilung verwendet, da dies dazu führt, dass die logarithmierten Differenzen  $\ln(K_t) - \ln(K_{t-1})$  als normalverteilt angenommen werden können. Dies geht bereits auf die Arbeit von Bachelier [2] aus dem Jahre 1900 zurück. Als Volatilität einer Zeitreihe  $r$  bezeichnet man die Varianz

$$\sigma_t^2 = \text{Var}(r_t | I_{t-1}).$$

Es gibt verschiedene Möglichkeiten die Volatilität zu messen. Eine Möglichkeit besteht darin die Volatilität eines bestimmten Betrachtungszeitraums (bspw. 1 Monat), mittels



Volatilitäten mit einer höheren Frequenz zu schätzen. Die Volatilität der Zeitreihe  $r_t$  zum Zeitpunkt  $t + 1$  wird geschrieben als:

$$\sigma_t^2 = \frac{1}{m} \sum_{j=1}^m r_{t_j}^2 - \left( \frac{1}{m} \sum_{j=1}^m r_{t_j} \right)^2,$$

wobei  $r_{t_j}$  die Zeitreihe gemessen mit der höheren Frequenz zu den Zeitpunkten  $j = 1, \dots, m$  ist.

In der Optionspreistheorie unterscheidet man zwei Arten von Volatilität. Unter der historischen Volatilität versteht man die über einen gewissen Zeitraum gemessene Standardabweichung der Kurse eines bestimmten Wertpapiers. Wird eine Option bereits gehandelt, so interessiert meist die im Preis enthaltene Volatilität (implizite Volatilität), also die Volatilität, die vom Markt für die Preisbestimmung unterstellt wird. Als Maß für die implizite Volatilität dient z.B. der VDAX, der im folgenden Abschnitt vorgestellt wird. Sie ist in hohem Maße zukunftsorientiert, da sie den zukünftigen Erwartungen der Marktteilnehmer entspricht.

Je höher die Volatilität, desto höher der Preis der Option, und zwar unabhängig davon, ob es sich um einen Call- oder Put-Option handelt. Die Begründung ist ganz einfach: Je höher die Volatilität, desto größer die Wahrscheinlichkeit, dass die Option am Ende der Laufzeit im Geld ist und damit einen inneren Wert hat. Geht die Volatilität zurück, wird diese Entwicklung unwahrscheinlicher. Man muss bei diesen Überlegungen immer berücksichtigen, dass aus Sicht des Stillhalters der Gewinn begrenzt ist, der Verlust aber nicht. Für den Stillhalter spielt daher auch nur die Wahrscheinlichkeit eine Rolle, die für ihn ungünstig ist. Dies stellt für den Käufer einer Option eine Chance dar, bedeutet für den Stillhalter allerdings ein höheres Risiko. Daher hat die Option einen höheren Wert.

Die Volatilität spiegelt nichts anderes als das Risiko wider. Erwarten die professionellen Marktteilnehmer in einem Basiswert in der Zukunft größere Unsicherheiten, dann erhöht sich die Volatilität und umgekehrt. Die Volatilität ändert sich während der Laufzeit einer Option. Optionen, welche sich auf das gleiche Underlying beziehen und gleiche Merkmale aufweisen (Laufzeit, Basispreis), lassen sich über die implizite Volatilität vergleichen. Die in Kapitel 2 erläuterten GARCH-Modelle bieten eine in der Praxis verbreitete Methode zur Modellierung der Volatilität.

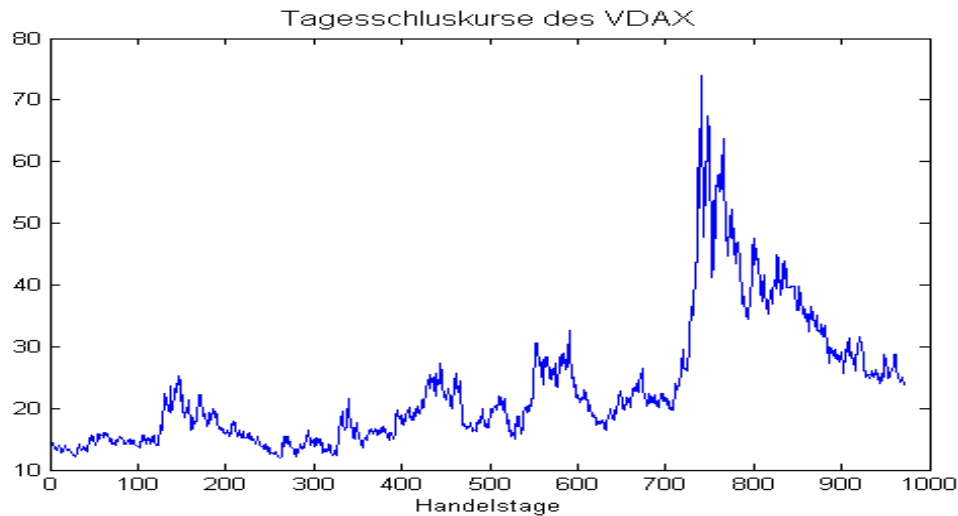


Abbildung 1.3: Darstellung der Tagesschlusskurse des VDAX vom 16.11.2005 bis 18.09.2009 (n=972 Handelstage)

### 1.2.3 VDAX - ein Index für Volatilität

Die Deutsche Börse AG berechnet seit 1997 minütlich einen Index für die Volatilität des Deutschen Aktienindex (DAX), den sog. VDAX und seit 2005 daneben auch den VDAX-NEW. Letzterer misst die erwartete (implizite) Volatilität für den DAX innerhalb der darauf folgenden 30 Tage und soll den VDAX mittelfristig ersetzen. Die Vorteile des VDAX-NEW gegenüber dem VDAX, der lediglich die Optionen „am Geld“ verwendet, liegen vor allem in der zusätzlichen Berücksichtigung der Optionen, die „aus dem Geld“ gehandelt werden. Dadurch ist er noch repräsentativer als der VDAX. Des Weiteren wird er nicht aus fiktiven sondern tatsächlich an der Börse EUREX gehandelten Optionen berechnet - und das über eine fixe Laufzeit von 30 Tagen und nicht über 45 Tage wie sein Vorgänger. Beide Volatilitätsindizes ermöglichen eine Aussage über die von den Marktteilnehmern erwartete Volatilität innerhalb eines festgelegten Zeitraums. Da die Wertsteigerung eines Marktes mit der Schwankungsbreite negativ korreliert ist, eignet sich der VDAX-NEW zur Diversifikation eines Portfolios<sup>1</sup>.

Abbildung 1.3 zeigt die Entwicklung des VDAX. Man kann sehen, dass der VDAX seit November 2005 recht ruhig verlaufen ist und sich meist in dem Bereich zwischen 10 und 25 bewegt hat. Die letzten ca. 250 (also ca. 1 Handelsjahr) sind gekennzeichnet von der Krise auf den Finanzmärkten, mit einem enormen Anstieg der Volatilität, welche sich

<sup>1</sup>vgl. dazu <http://deutsche-boerse.com/>.

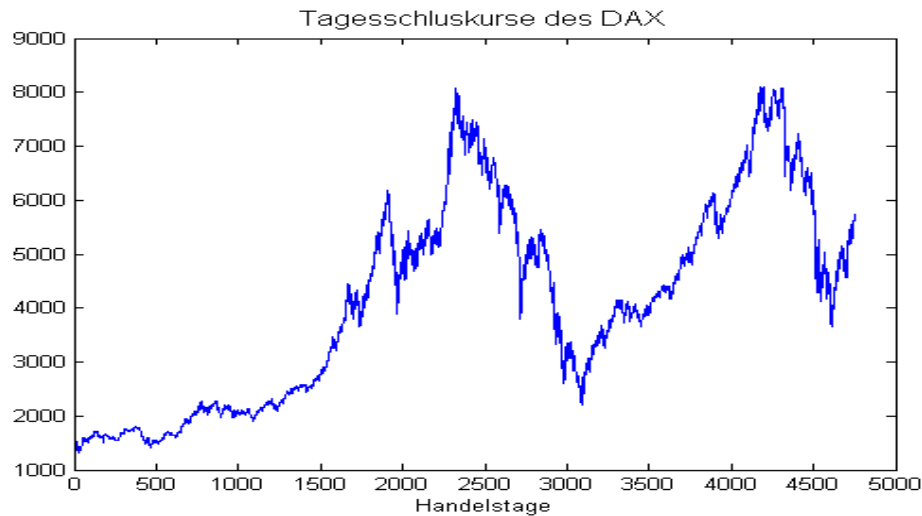


Abbildung 1.4: Tagesschlusskurse des DAX vom 26.11.1990 bis 17.09.2009 (n=4750 Handelstage)

über einen längeren Zeitraum nur langsam abbaut.

Die aus dem VDAX gewonnenen Informationen können hilfreich sein bei der Analyse der Volatilität der Werte im DAX oder anderer Indizes. Detaillierte Angaben über den VDAX und dessen Berechnung gibt die Deutsche Börse AG in ihrem "Leitfaden zu den Volatilitätsindizes der Deutschen Börse" [31].

### 1.3 Stylized Facts

Zeitreihen aus dem Finanzmarktbereich weisen einige empirisch belegte Eigenschaften auf, die häufig unter dem Begriff "Stylized Facts" zusammengefasst werden. Abbildung 1.4 zeigt die Entwicklung des DAX seit dem 26. November 1990. Es liegen die Werte von 4750 Tagesschlusskursen vor. Dabei entsprechen etwa 250 Handelstage einem Handelsjahr. Eine detaillierte Übersicht über die Stylized Facts von Renditen bietet Cont (2001) [26].

Sehr deutlich ist der Anstieg zu erkennen, der durch das gestiegene Interesse von Privatanlegern u.a. auf Grund der "Volksaktie" Deutsche Telekom Mitte der neunziger Jahre aufkam. Auch der rasante Anstieg durch Aktien aus dem Internetbereich und die New Economy inklusive des Platzens der sog. "Dot-Com-Blase" zur Jahrtausendwende ist klar erkennbar. Ebenso deutlich sieht man die Folgen des Anschlags auf das World Trade Center am 11.09.2001 und den anschließenden anhaltenden Anstieg bis Herbst 2008. Danach

dominiert die Finanzkrise den DAX. Man kann also festhalten, dass unterschiedlich lange Phasen zu beobachten sind, in denen die Zeitreihe einen mehr oder weniger ausgeprägten lokalen Trend beinhaltet.

Bevor weitere Stylized Facts angesprochen werden, sollen zunächst einige Grundbegriffe im Zusammenhang mit Zeitreihen erklärt werden.

Ein stochastischer Prozess ist definiert als eine Folge von Zufallsvariablen  $\{X(t, \omega_i), t \in T\}$  (kurz  $\{X_t\}$ ) auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P)$ .  $\Omega$  ist dabei die Menge aller möglichen Realisationen des stochastischen Prozesses,  $\mathcal{F}$  eine Sigma-Algebra und  $P$  ein Wahrscheinlichkeitsmaß.  $T$  steht bei Zeitreihen für die Zeitpunkte, wobei diese sowohl in diskreter als auch in stetiger Form vorliegen können. Der stochastische Prozess kann im Allgemeinen beliebig viele Realisationen haben,  $\omega_i$  bezeichnet eine Realisation oder Trajektorie dieses stochastischen Prozesses. Damit stellt die Zeitreihe, die durch den Kursverlauf einer Aktie oder eines Index gebildet wird eine Realisation eines stochastischen Prozesses dar, dessen exakte Spezifikation im Allgemeinen unbekannt ist. Der Prozess dem diese Zeitreihe als eine Realisation entspringt, wird häufig als datengenerierender Prozess bezeichnet. Aus diesem Beispiel ist leicht zu ersehen, dass Zeitreihen aus dem Finanzmarktbereich, beispielsweise Tagesschlusskurse von Aktien, in aller Regel zu Prozessen mit diskreten Zeitpunkten führen.

Eine Eigenschaft, die einen stochastischen Prozess charakterisieren kann, ist die Stationarität. Man unterscheidet dabei zwei Arten der Stationarität. Von *strenger Stationarität* spricht man, wenn die gemeinsamen Verteilungsfunktionen des stochastischen Prozesses invariant gegenüber einer zeitlichen Verschiebung sind. Die *schwache Stationarität* bezieht sich auf die Momente der Verteilung. Als *mittelwertstationär* wird der Prozess bezeichnet, falls

$$E(X_t) = \mu_t = \mu \quad \text{für alle } t.$$

Das heißt der Mittelwert bleibt über die Zeit konstant. Der Prozess heißt *varianzstationär*, falls

$$Var(X_t) = E[(X_t - \mu_t)^2] = \sigma^2 \quad \text{für alle } t.$$

Die Varianz ist also über die Zeit konstant und endlich. Kovarianzstationarität bedeutet, dass die Kovarianz zwischen zwei Zufallsvariablen  $X_t$  und  $X_s$  nur vom zeitlichen Abstand  $|t - s|$  abhängt, mithin also gilt:

$$Cov(X_t, X_s) = E[(X_t - \mu_t)(X_s - \mu_s)] = \gamma(|t - s|).$$

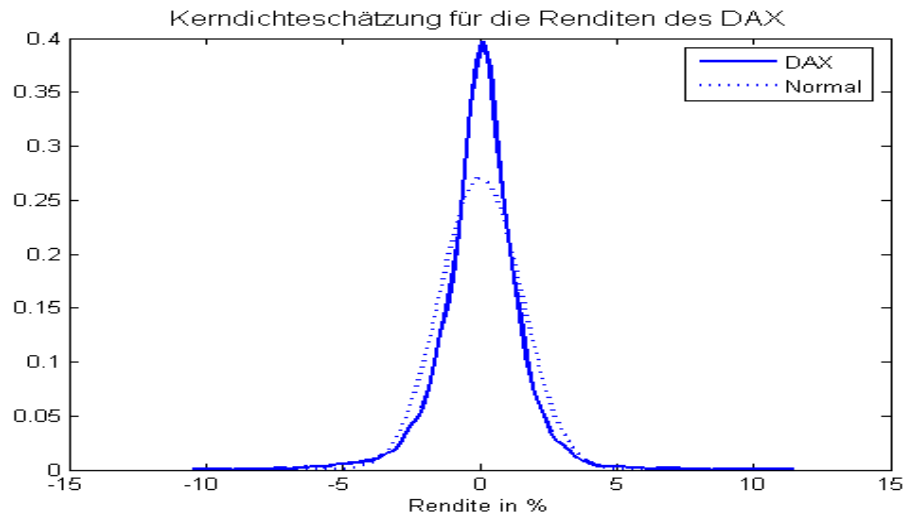


Abbildung 1.5: Kerndichteschätzung für die Rendite des DAX und eine an die Daten angepasste Normalverteilung

Es ist leicht zu sehen, dass für  $t = s$  aus der Kovarianzstationarität die Varianzstationarität folgt. Ein stochastischer Prozess, der sowohl mittelwertstationär als auch kovarianzstationär ist, wird als *schwach stationär* bezeichnet, häufig wird in diesem Fall auch nur der Begriff stationär verwendet.

Ein schwach stationärer Prozess heißt mittelwertergodisch, falls

$$\lim_{T \rightarrow \infty} E \left[ \left( \frac{1}{T} \sum_{t=1}^T X_t - \mu \right)^2 \right] = 0,$$

und varianzergodisch, falls

$$\lim_{T \rightarrow \infty} E \left[ \left( \frac{1}{T} \sum_{t=1}^T (X_t - \mu)^2 - \sigma^2 \right)^2 \right] = 0.$$

Allerdings müssen zusätzlich zur schwachen Stationarität weitere Voraussetzungen erfüllt sein, vgl. hierzu beispielsweise Schlittgen und Streitberg (1997) [82].

Die bisherigen Aussagen über die Zeitreihe des DAX sind allein mit Hilfe der Abbildung der Kurse, also hier dem Stand des Index, möglich. Um weitere Erkenntnisse zu erhalten, betrachten wir nun die Renditen des DAX. Dazu können sowohl diskrete als auch stetige Renditen verwendet werden. Abbildung 1.5 zeigt die Kerndichteschätzung der Renditen des Index und eine angepasste Normalverteilung. Man kann erkennen, dass die Renditen

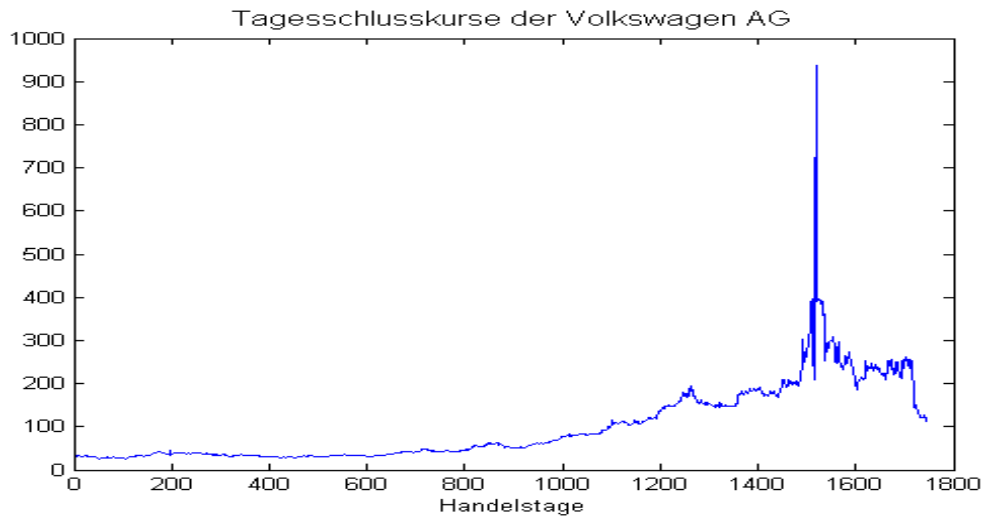


Abbildung 1.6: Tagesschlusskurse der Aktie des Volkswagen AG im Zeitraum vom 01.01.03 bis 19.09.2009

in den Rändern mehr Wahrscheinlichkeitsmasse haben als die Normalverteilung. Ausserdem ist die Verteilung der Renditen spitzgipfliger als die Normalverteilung, d.h. sie ist leptokurtisch, weist also eine Kurtosis auf, die größer ist als drei. Bereits 1963 untersuchte Mandelbrot [72] Zeitreihen bei Rohstoffen hinsichtlich ihrer empirischen Phänomene. Die meisten Stylized Facts gehen auf diese Arbeit zurück.

Für die Renditen bedeuten die schweren Ränder (Heavy Tails), dass häufiger als durch die Annahme der Normalverteilung unterstellt, extreme Werte, d.h. große Kurssprünge vorkommen. Dies ist insbesondere nach Einsetzen der Finanzkrise im Herbst 2008 leicht nachzuvollziehen, da in dieser Phase sehr große Kurssprünge sehr häufig waren. Insbesondere die Aktie der Volkswagen AG zeigte Kurssprünge, die bei Gültigkeit der Normalverteilung nicht hätten vorkommen dürfen, da sie deutlich außerhalb des  $3\sigma$  oder  $4\sigma$  Bereichs lagen, vgl. Abbildung 1.6. Diese großen Kursveränderungen sind allerdings nicht die Regel. Über einen langen Zeitraum betrachtet, kommen häufig Handelstage mit eher geringer Veränderung des Kurses vor, dies führt zu der leptokurtischen Form der Verteilung, die auch High Peakedness genannt wird. Die sog. aggregierte Normalität steht im Zusammenhang mit der Normalverteilungsannahme und besagt, dass sich die Verteilung mit steigendem Stichprobenumfang der Normalverteilung nähert.

Eine weitere empirische Tatsache bei Aktienkursen oder Indizes ist die Asymmetrie des Einflusses von positiven und negativen Kursänderungen. So werden starke Abwärtsbewegungen häufiger beobachtet als entsprechend starke Aufwärtsbewegungen. Allerdings gilt

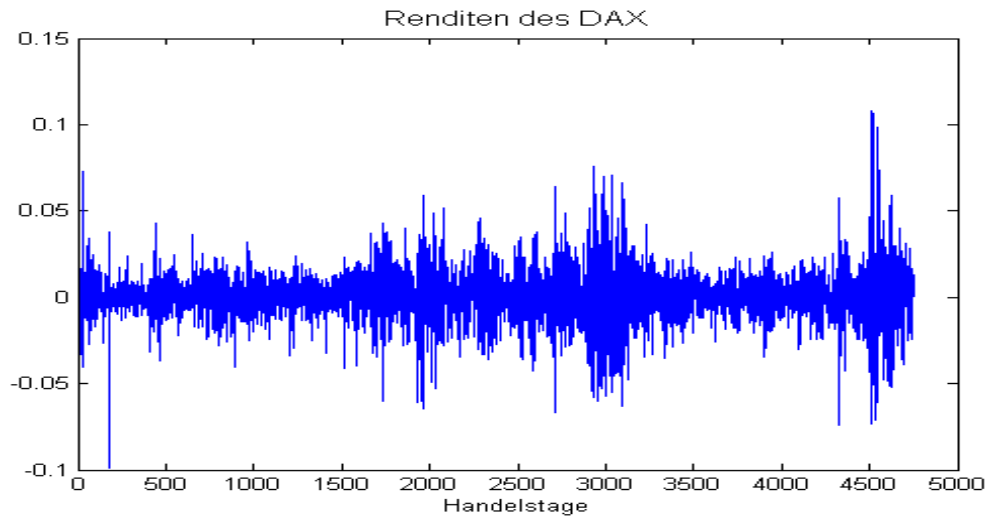


Abbildung 1.7: Darstellung der stetigen Renditen des DAX vom 26.11.1990 bis 17.09.2009 (n=4750 Handelstage)

dies nicht im gleichen Maße für Wechselkurse.

Die unterschiedliche Wirkung positiver und negativer Nachrichten auf die Renditen wird als Leverage-Effekt bezeichnet. Er besagt, dass die Volatilität negativ mit den Renditen korreliert ist, so führen negative Informationen zu einem stärkeren Anstieg der Volatilität als positive Nachrichten.

Zur Überprüfung der Effizienz von Finanzmärkten kann unter anderem untersucht werden, ob die Renditen einem Random-Walk folgen. Es gibt verschiedene Definitionen für Kapitalmarkteffizienz. Nach der Definition der "schwachen Kapitalmarkteffizienz" darf die Kenntnis aller vorhandenen Informationen keine Möglichkeit bieten durch Prognose der zukünftigen Renditen einen risikolosen Gewinn zu erzielen. Somit ist der beste Schätzer für den nächsten Kurs der jetzige Kurs. Bei der Betrachtung der Renditen fällt aber auf, dass sich Phasen hoher Renditeschwankungen mit Phasen geringer Volatilität abwechseln, d.h. auf niedrige (hohe) Schwankungen folgen wieder niedrige (hohe) Schwankungen. Man kann in Abbildung 1.7 deutlich die Volatilitätscluster sehen, d.h. Phasen in denen der DAX stark schwankt wechseln sich mit Phasen ab, in denen der DAX nur leichte Schwankungen in den Renditen hat. Gelegentlich werden diese Cluster auch durch die absoluten Renditen wie in Abbildung 1.8 dargestellt.

Ziel ist es nun die Cluster, die die stetigen Renditen aufweisen, zu nutzen, um den Verlauf der Volatilität der stetigen Renditen zu modellieren.

Für einen Anleger stellt die Prognostizierbarkeit der Volatilität eine Chance dar. Unglück-

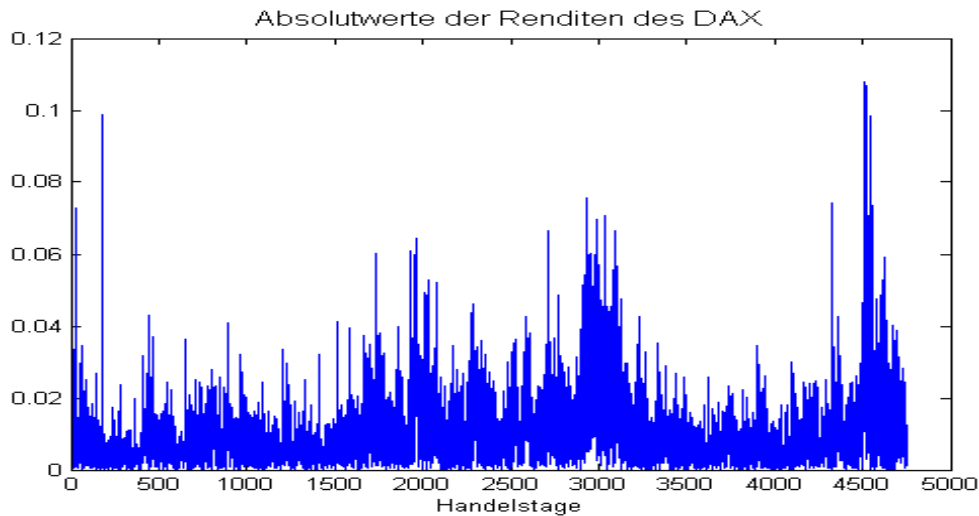


Abbildung 1.8: Darstellung der Absolutwerte der stetigen Renditen des DAX vom 26.11.1990 bis 17.09.2009 (n=4750 Handelstage)

licherweise ermöglicht es der Handel mit Aktien, Renten etc. nicht, in die Volatilität isoliert von der Richtung der Kursentwicklung zu investieren. Für einen Marktteilnehmer, der sich gegen Risiken absichern möchte, ist die Prognose der Volatilität als Maß für das Risiko ebenso von großer Bedeutung. Eine Möglichkeit diese Erkenntnisse umzusetzen, eröffnet der Handel mit derivativen Finanzinstrumenten. Sehr verbreitet ist die Kombination verschiedener Optionen sowohl zu spekulativen Zwecken als auch zur Absicherung, da durch die Kombination sehr individuelle Strategien ermöglicht werden.

## 1.4 Kombinationsstrategien auf Volatilität

In Abschnitt 1.2.1 wurden die Einflussfaktoren auf den Optionspreis behandelt. Wird eine Option erst mal gehandelt, so sind mit dem Kurs des Underlyings, dem Basispreis, dem Zinsniveau sowie der Restlaufzeit vier der fünf Einflussgrößen festgelegt und bekannt. Als einzige einzige unbekannte Größe verbleibt die Volatilität.

Bei Kombinationsstrategien handelt es sich um Kombinationen verschiedener Optionen mit dem Ziel bestimmte Gewinn- und Verlustmuster zu generieren. Dabei werden Long- und Short-Positionen sowie Call- und Put-Optionen kombiniert. Bei der Auswahl der Kombinationen sind die Erwartungen bezüglich Richtung und Ausmaß der Kursbewegungen entscheidend. Da in dieser Arbeit der Fokus auf der Modellierung und Prognose der Volatilität liegt, werden im Folgenden Strategien betrachtet, die sich auf das Ausmaß



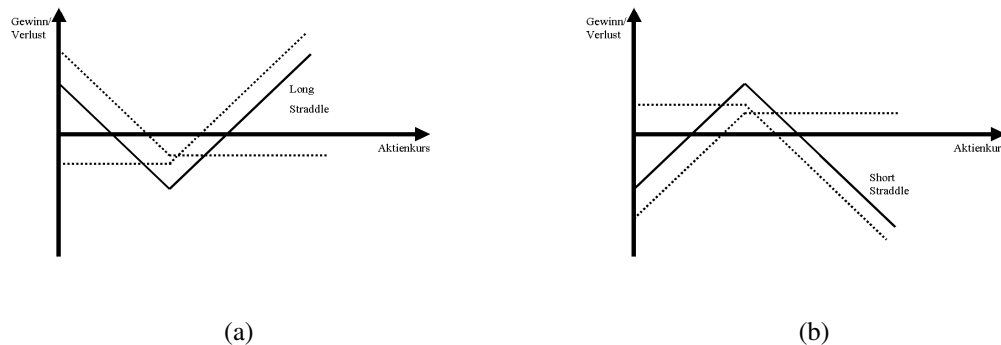


Abbildung 1.9: Gewinn- und Verlustdiagramm eines Long (a) bzw. Short (b) Straddle

der Kursbewegungen konzentrieren. Eine Berücksichtigung der Richtung unterbleibt an dieser Stelle auch, da der Kursverlauf nach der Random-Walk-Theorie ohnehin nicht prognostizierbar ist.

### Straddle

Beim Long Straddle handelt es sich um die Kombination eines Long-Call und eines Long-Put, welche beide denselben Basispreis haben. Bei dieser Kombination profitiert der Investor umso mehr, je stärker sich der Kurs des Underlyings vom Basispreis unterscheidet. Der Ertrag ist unabhängig davon, ob der Kurs der Underlyings steigt oder fällt, da in beiden Fällen jeweils nur die im Geld befindliche Option ausgenutzt wird. Die andere Option verfällt. Entspricht der Kurs des Underlyings dem Basispreis, so sind beide Optionen wertlos. Die Investition in einen Long Straddle erscheint also sinnvoll in einem volatilen Markt. Der Short Straddle ist die entsprechende Gegenposition zum Long Straddle und besteht im Verkauf einer Call- und einer Put-Option zur gleichen Basis. Diese Strategie entspricht der Erwartung geringer Kursschwankungen bzw. geringer Volatilität.

Die Abbildung 1.9 zeigt die Gewinn- und Verlustdiagramme der Straddle-Kombinationen. Aus dem Schaubild wird deutlich, dass Kombinationen verschiedener Long- bzw. Short-Positionen den maximalen Verlust vergrößern, da mehrere Optionen gekauft werden müssen. Aus dem gleichen Grund verringert sich der mögliche Gewinn gegenüber der Investition in eine einzelne Option. Dies entspricht der üblichen Annahme, dass höhere erwartete Erträge mit mehr Risiko verbunden sind.

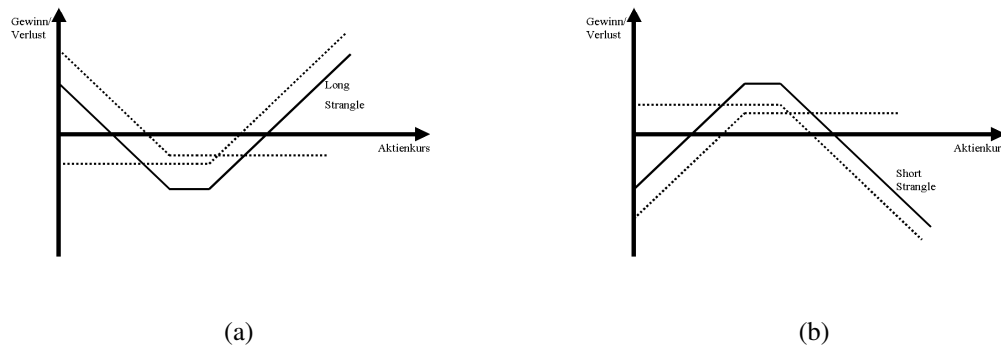


Abbildung 1.10: Gewinn- und Verlustdiagramm eines Long (a) bzw. Short (b) Strangle

### Strangle

Der Long Strangle ist der Strategie des Long Straddle sehr ähnlich. Der wichtige Unterschied ist, dass der Call und der Put beim Strangle nicht denselben Basispreis haben. Dadurch wird das Gewinn- und Verlustdiagramm dahingehend verändert, dass der Bereich für den Kurs des Underlyings, der zu dem maximalen Verlust führt, vergrößert wird. Ansonsten entspricht der Strangle dem Straddle. Der Investor erwartet auch hier große Kursschwankungen. Analog verhält sich der Short Straddle nur ebenfalls mit voneinander abweichenden Basispreisen, das Grundkonzept bleibt aber gleich.

Die Gewinn- und Verlustdiagramme der Strangle-Kombinationen können Abbildung 1.10 entnommen werden.

### Strip

Dies ist eine kombinierte Optionsstrategie, die der Straddle-Variante ähnelt. Im Unterschied zum Straddle wird aber bei einem Strip ein Call mit zwei Puts verbunden. Ein Long Strip ist ein gleichzeitiger Kauf eines Calls und zweier Puts auf denselben Basiswert, bei gleichem Verfalldatum und identischem Basispreis. Der Short Strip dagegen ist ein gleichzeitiger Verkauf einer Call- und zwei Put-Optionen auf denselben Basiswert bei gleichem Verfalldatum und identischem Basispreis.

Die Abbildung 1.11 zeigt die Gewinn- und Verlustdiagramme der Strip-Kombinationen. Es ist leicht ersichtlich, dass diese Kombination gewählt wird, um an einer negativen Entwicklung des Underlyings stärker zu profitieren, als an einer positiven.

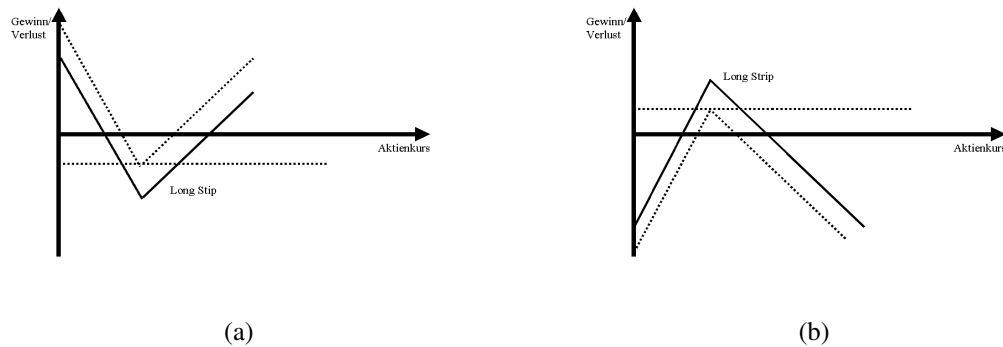


Abbildung 1.11: Gewinn- und Verlustdiagramm eines Long (a) bzw. Short (b) Strip

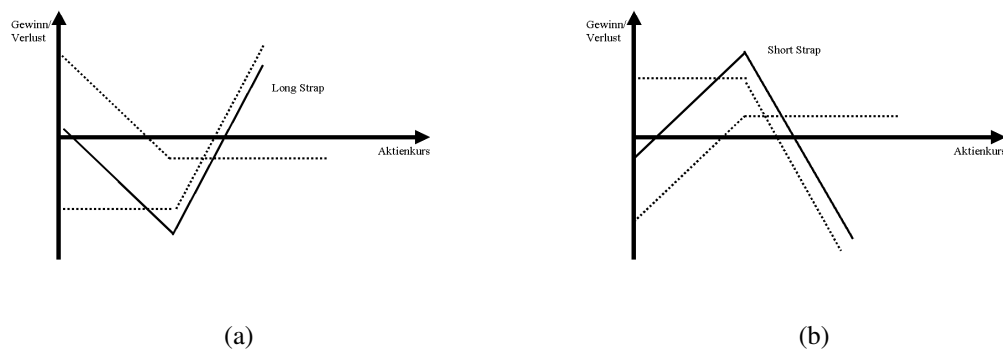


Abbildung 1.12: Gewinn- und Verlustdiagramm eines Long (a) bzw. Short (b) Strap

### Strap

Auch der Strap ähnelt dem Straddle. Im Unterschied dazu werden bei einem Strap zwei Calls mit einem Put kombiniert. Der Long Strap ist der gleichzeitige Kauf zweier Calls und eines Puts auf denselben Basiswert bei gleichem Verfallsdatum und identischem Basispreis. Der Short Strap ist der gleichzeitige Verkauf von zwei Call-Optionen und einer Put-Option auf denselben Basiswert bei gleichem Verfallsdatum und gleichem Basispreis. Die Abbildung 1.12 zeigt die Gewinn- und Verlustdiagramme der Strap-Kombinationen. Beide Strategien, d.h. Strip und Strap, dienen dazu die Winkel des Verlaufs im Gewinn- und Verlustdiagramm anzupassen. Die Asymmetrie der Gewinn- und Verlustdiagramme zeigt auch, dass die Erwartungen bezüglich der Kursentwicklung bei dieser Kombinationsstrategie durchaus eine Rolle spielen. Zentraler Bestandteil der Strategie ist aber weiterhin die Volatilität. Bei diesen Kominationsstrategien ist das Verhältnis von Calls

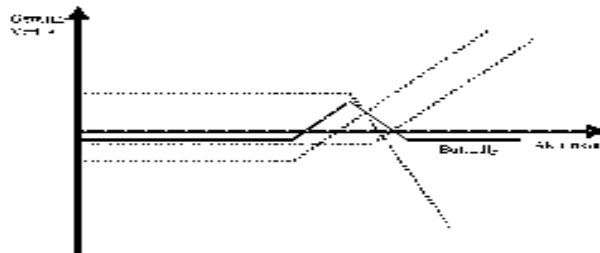


Abbildung 1.13: Gewinn- und Verlustdiagramm eines Butterfly-Spread

und Puts (hier: 1:2 bzw. 2:1) auch variierbar, so dass sich eine Vielzahl an Gewinn- und Verlustdiagrammen generieren lässt.

### Butterfly-Spread

Spread-Strategien unterscheiden sich von den bisher dargestellten Kombinationen durch die Tatsache, dass nicht Put- und Call-Positionen, sondern Long- und Short-Positionen mit verschiedenen Basispreisen kombiniert werden. Den Butterfly-Spread gibt es auf Basis von Puts und auf Basis von Calls, das Gewinn- und Verlustdiagramm unterscheidet sich dabei nicht. Man benötigt für den Butterfly-Spread drei verschiedene Optionen mit drei verschiedenen Basispreisen. Man kauft jeweils einen Call (Put) mit dem niedrigsten und höchsten Basispreis und verkauft zwei Calls (Puts) mit dem mittleren Basispreis. Abbildung 1.13 zeigt das Gewinn- und Verlustdiagramm eines Butterfly-Spread auf Basis von Calls. Man kann sehen, dass die Strategie darauf abzielt, sehr geringe Schwankungen des Basiswertes auszunutzen. Der Butterfly-Spread auf Basis von Puts ergibt dasselbe Gewinn- und Verlustdiagramm.

Da als einzige unbekanntes Größe die Volatilität bleibt, wird deutlich, dass es für einen Investor von besonderem Interesse ist, die Entwicklung der Volatilität möglichst gut zu prognostizieren.

## Kapitel 2

# GARCH-Modelle und deren Erweiterungen

In diesem Kapitel werden einige Vertreter der Familie der GARCH-Modelle vorgestellt. Diese Modelle ermöglichen es die Abhängigkeitsstrukturen in den bedingten Varianzen von Zeitreihen  $Var(X_t | \mathcal{F}_{t-1})$  zu modellieren. Es werden zunächst in den Abschnitten 2.1 und 2.2 die Grundmodelle von Engle und Bollerslev aufgeführt, die geeignet sind bedingte Heteroskedastie zu modellieren. In Abschnitt 2.4 werden einige Erweiterungen der Modelle gezeigt, die es erlauben, die bedingten Varianzen in Zeitreihen sehr viel flexibler zu betrachten. Schließlich werden in Abschnitt 2.5 multivariate Modelle vorgestellt, welche die Zusammenhänge von mehreren Zeitreihen mehrdimensional erfassen.

Univariate Modelle werden verwendet, wenn es nicht das Ziel ist, Zusammenhänge zwischen unterschiedlichen Prozessen zu berücksichtigen, sondern eine Zeitreihe isoliert betrachtet werden soll. Werden die Entwicklungen der Aktienkurse von Unternehmen betrachtet, so mag z.B. die Entwicklung von anderen Unternehmen derselben Branche durchaus bei der Erklärung hilfreich sein und eine multivariate Analyse ist angebracht. Wenn aber beispielsweise Indizes oder Wechselkurse betrachtet werden, so ist es durchaus plausibel, diese als univariate Zeitreihe zu untersuchen.

Das GARCH-Modell bildet die Grundlage für eine breite Palette an Erweiterungen, die jeweils spezielle Formen der bedingten Heteroskedastie modellieren. So können fraktionale integrierte GARCH (FIGARCH) -Modelle im Vergleich zu GARCH-Modellen auch einen längerfristigen Einfluss der sog. Schocks oder Störungen berücksichtigen, vgl. dazu Baillie, Bollerslev und Mikkelsen (1996) [5]. Unter Schocks versteht man die Abweichung zwischen der beobachteten Rendite zu Zeitpunkt  $t$  und der erwarteten Rendite

ausgehend vom Zeitpunkt  $t - 1$ . Ist der Einfluss der Schocks permanent, so spricht man auch von Persistenz. In diesem Fall kommen integrierte GARCH(IGARCH)-Modelle zum Einsatz. In Abschnitt 2.4.2 wird ein Modell vorgestellt, in welchem diverse Modelle genestet sind. Unter anderem bieten manche von ihnen die Möglichkeit Asymmetrie zu modellieren, so zum Beispiel die in Abschnitt 2.4.3 beschriebenen exponentiellen GARCH (EGARCH)-Modelle. Die Abschnitte 2.4.7 bis 2.4.6 befassen sich noch mit weiteren Modellen. Eine exzellente Übersicht zu verschiedenen Erweiterungen der GARCH-Modelle findet man in Tsay (2002) [91]. Einen Überblick über die Eigenschaften verschiedener Modelle geben Bollerslev, Engle und Nelson (1994) [19] oder Bera und Higgins (1993) [9]. Eine Übersicht über die Literatur zu GARCH-Modellen findet man in Bollerslev, Chou und Kroner (1992) [18].

## 2.1 ARCH-Modell

Mit den von Engle 1982 [34] eingeführten ARCH (Autoregressive Conditional Heteroscedastic)-Modellen ist es möglich, die bedingte Varianz in Zeitreihen und ökonomischen Modellen über die Zeit variieren zu lassen. Die bedingte Varianz ist dabei eine Funktion der vergangenen Störterme, während die unbedingte Varianz konstant bleibt. Es sei  $r_t$  ein reellwertiger, stochastischer Prozess mit diskretem Zeitverlauf, und  $\mathcal{F}_{t-1}$  sei das Informationsset ( $\sigma$ -Feld) aller Informationen zum Zeitpunkt  $t$ . Beispielsweise betrachten wir die Log-Renditen aus Abschnitt 1.2.2. Werden diese um ihren bedingten Erwartungswert  $\mu_t$  bereinigt, wird häufig davon ausgegangen, dass sie sich gut durch einen Martingaldifferenzenprozess darstellen lassen.

$$z_t = r_t - E(r_t | \mathcal{F}_{t-1}) = R_t - \mu_t \quad t = 1, \dots, T. \quad (2.1)$$

Daher können die Ergebnisse für  $z_t$  auf  $r_t$  übertragen werden. Gleichung (2.1) bezeichnet man auch als Mittelwertgleichung. Das univariate ARCH-Modell nach Engle ist durch die beiden Gleichungen:

$$z_t | \mathcal{F}_{t-1} = \sigma_t \varepsilon_t \quad (2.2)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 z_{t-1}^2 \quad \text{mit} \quad \alpha_0 > 0, \alpha_1 > 0. \quad (2.3)$$

gegeben. Die  $\{\varepsilon_t\}$  sind unabhängig und identisch verteilt und folgen einem White-Noise-Prozess mit  $E(\varepsilon_t) = 0$  und  $Var(\varepsilon_t) = 1$ .  $\alpha_0 > 0$  und  $\alpha_1 > 0$  stellen sicher, dass die

bedingte Varianz  $\sigma_t^2$  in Gleichung (2.3) positiv ist, diese Gleichung bezeichnet man als Varianzgleichung. Für den bedingten Erwartungswert von  $z_t$  gilt:

$$E(z_t | \mathcal{F}_{t-1}) = E(\sigma_t \varepsilon_t | \mathcal{F}_{t-1}) = \sigma_t E(\varepsilon_t | \mathcal{F}_{t-1}) = 0,$$

da  $\sigma_t^2$  entsprechend (2.3) nur von  $z_{t-1}$  abhängig ist und  $z_{t-1}$  in  $\mathcal{F}_{t-1}$  enthalten ist.

Aufgrund des Gesetzes des iterierten Erwartungswerts gilt auch für den unbedingten Erwartungswert:

$$E(z_t) = E(E(z_t | \mathcal{F}_{t-1})) = 0.$$

Für die bedingte Varianz gilt:

$$\begin{aligned} \text{Var}(z_t | \mathcal{F}_{t-1}) &= \text{Var}(z_t | z_{t-1}) = E(z_t^2 | z_{t-1}) - E(z_t | z_{t-1})^2 \\ &= E(\varepsilon_t^2 \sigma_t^2 | \varepsilon_{t-1}) - E(\varepsilon_t \sigma_t | z_{t-1})^2 \\ &= \sigma_t^2 E(u_t^2 | z_{t-1}) - (\sigma_t E(\varepsilon_t | z_{t-1}))^2 \\ &= \sigma_t^2. \end{aligned}$$

Ausführlich werden die Eigenschaften eines ARCH-Modells in Bera und Higgins (1993) [9] besprochen.

## 2.2 GARCH-Modell

GARCH (Generalized Autoregressive Conditional Heteroscedastic)-Modelle sind eine Erweiterung der ARCH-Modelle aus Abschnitt 2.1, die 1986 durch Bollerslev in die Literatur eingeführt wurden, vgl. Bollerslev (1986) [14] und (1987) [15]. Die bedingten Varianzen werden dabei im Gegensatz zum ARCH-Modell nicht nur aus der Vergangenheit des Prozesses erklärt, sondern berücksichtigen darüber hinaus noch die eigene Vergangenheit.

Es wird angenommen die Mittelwertgleichung habe die Form

$$r_t = z_t,$$

wobei  $r_t$  der beobachtete Prozess, also beispielsweise die Log-Renditen einer Aktie, ist. Das GARCH(p,q)-Modell (Bollerslev 1986) [14] wird durch die folgende Formulierung gegeben:

$$z_t | \mathcal{F}_{t-1} = \sqrt{\sigma_t^2} \varepsilon_t, \quad t = 1, \dots, T, \quad (2.4)$$

mit der Varianzgleichung

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \cdot z_{t-i}^2 + \sum_{j=1}^q \beta_j \cdot \sigma_{t-j}^2, \quad (2.5)$$

wobei

$$\varepsilon \stackrel{u.i.v.}{\sim} \mathcal{P}_{0,1}.$$

$\varepsilon_t$  ist also unabhängig, identisch verteilt.  $\mathcal{P}_{0,1}$  bezeichnet eine Verteilung mit  $E(\varepsilon_t) = 0$  und  $Var(\varepsilon_t) = 1$  für  $t = 1, \dots, T$ .  $\sigma_t^2$  ist die bedingte Varianz des Prozesses  $z_t$  zum Zeitpunkt  $t$ .  $p$  und  $q$  sind ganze Zahlen mit  $p > 0$ ,  $q \geq 0$ , sie geben die Lag-Ordnungen wieder.

Um sicher zu stellen, dass die Varianzen nur positive Werte annehmen, werden folgende Restriktionen angenommen:

$$\alpha_0 > 0 \text{ und } \alpha_i \geq 0, \text{ für } i = 1, \dots, p \quad \text{und} \quad \beta_j \geq 0, \text{ für } j = 1, \dots, q.$$

Diese Restriktionen sind hinreichend, aber nicht notwendig. In Gleichung (2.5) wird angenommen, dass  $z_t = \sigma_t = 0$  für  $t < 0$ .

Die Bedingung

$$\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$$

sichert die Varianzstationarität. Nelson und Cao (1992) [80] nennen weniger restriktive Bedingungen. Offensichtlich ergibt sich das ARCH-Modell als Spezialfall eines GARCH-Modells für  $q = 0$ .

Mit Hilfe der dargestellten GARCH-Modellfamilie werden flexiblere Lag-Strukturen erreicht als mit ARCH-Modellen. Im Vergleich zu ARCH-Modellen kommt man bei GARCH-Modellen meist mit einer niedrigeren Ordnung aus. D.h., während es bei ARCH-Modellen oft zu höheren Ordnungen kommen kann, benötigt man bei GARCH-Modellen selten eine Ordnung, welche höher als  $p = q = 1$  ist. Higgins und Bera (1993) [9] kommen zu dem Schluss, dass Ordnungen bei denen entweder  $p = 2$  oder  $q = 2$  ist, kaum überstiegen werden müssen. Daher soll dem GARCH(1,1)-Modell besondere Aufmerksamkeit gewidmet werden. Die Varianzgleichung eines GARCH(1,1)-Modells hat folgende Form:

$$\sigma_t^2 = \alpha_0 + \alpha_1 z_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad t = 1, \dots, T. \quad (2.6)$$

Unter der Annahme, dass  $\varepsilon_t$  aus Gleichung (2.4) normalverteilt ist, lautet die Likelihood-



funktion für eine Stichprobe von  $T$  Beobachtungen  $z = (z_1, \dots, z_T)$  dann:

$$L(\theta; z) = (2\pi)^{-\frac{T}{2}} \prod_{t=1}^T \left\{ (\sigma_t^2)^{-1/2} \exp\left(-\frac{z_t^2}{2\sigma_t^2}\right) \right\}, \quad (2.7)$$

wobei sich die  $\sigma_t^2$  aus Gleichung (2.6) ergeben. Mit Hinblick auf die im folgenden Abschnitt beschriebene Maximum-Likelihood-Schätzung der Parameter sei noch die logarithmierte Likelihoodfunktion erwähnt, sie hat folgende Form:

$$\ln L(\theta; z) = -T \ln(\sqrt{2\pi}) - \frac{1}{2} \sum_{t=1}^T \ln(\sigma_t^2) - \frac{1}{2} \sum_{t=1}^T \frac{z_t^2}{\sigma_t^2}.$$

Wie in Abschnitt 1.3 gezeigt, ist die Annahme der Normalverteilung der Renditen häufig nicht gerechtfertigt, daher werden auch in GARCH-Modellen häufig andere Verteilungen für  $\varepsilon_t$  unterstellt. Abschnitt 2.3 beschäftigt sich mit ihnen und nennt die daraus resultierenden Loglikelihoodfunktionen.

## 2.2.1 Schätzung des Modells mit der (Quasi-)Maximum-Likelihood-Methode

Die verbreitetsten Methoden der Schätzung von GARCH-Modellen sind die Maximum-Likelihood-Schätzung (ML-Schätzung) und die Quasi-Maximum-Likelihood-Schätzung (QML-Schätzung), vgl. dazu Bera und Higgins (1993) [9]. Dabei besteht der Unterschied der beiden Methoden lediglich darin, dass bei der QML-Schätzung die zugrundeliegende Verteilung nicht bekannt sein muss. Es wird also eine Verteilung verwendet, die die wahre Verteilung u.U. fehlspezifiziert.

Für die Schätzung sind in beiden Fällen die partiellen Ableitungen nach den Parametern sowie die Hesse-Matrix zu bestimmen, was aufgrund der rekursiven Struktur der bedingten Varianzen sehr schwierig ist. Eine analytische Lösung ist deshalb meist nicht verfügbar, vgl. dazu beispielsweise Fiorentini et al. (1996) [39], daher werden i.d.R. numerische Optimierungsverfahren verwendet. Der gängige Algorithmus ist der von Berndt, Hall, Hall und Hausman (1974) [11] (BHHH), der im Vergleich zu dem Verfahren von Newton und Raphson mit einer Näherung der Hesse-Matrix auskommt. Eine ähnliche Näherung der Hesse-Matrix wird in dem Algorithmus von Broyden, Fletcher, Goldfarb und Shanno benutzt.

Daneben gibt es die auf der ARMA-Darstellung des GARCH-Modells basierenden Schätzmethoden nach Art der Generalized Method of Moments (GMM). Diese werden

von Baillie und Chung (2001) [3] vorgeschlagen und werden Minimum Distance Estimator (MDE) genannt, vgl. dazu auch Storti (2006) [84] oder Kristensen und Linton (2006) [63].

## 2.2.2 Test auf GARCH-Effekte

Der Lagrange-Multiplikator- (LM-) Test bietet eine Möglichkeit auf ARCH-Effekte zu testen. Dieser wurde von Engle in seinem Artikel von 1982 [34] vorgeschlagen. Hierbei wird die Nullhypothese  $H_0$  getestet, dass in der Zeitreihe  $X$  keine ARCH-Effekte vorliegen, indem man eine Regression mit den quadrierten Werten zum Zeitpunkt  $t$  als endogener und den verzögerten quadrierten Werten als exogener Größe, also:

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_q X_{t-q}^2 + v_t, \quad t = 1, \dots, T,$$

schätzt. Anschließend wird für diese Regression das multiple Bestimmtheitsmaß  $R^2$  berechnet. Bei Gültigkeit der Nullhypothese ist der Ausdruck  $T \cdot R^2$  asymptotisch  $\chi^2$ -verteilt mit  $q$  Freiheitsgraden. Die Nullhypothese wird abgelehnt, falls der Ausdruck  $T \cdot R^2$  den  $(1 - \alpha)\%$ -Quantilwert der  $\chi_{(q)}^2$ -Verteilung übersteigt. Bera und Higgins (1993) [9] zeigen, dass der Test auf ARCH-Effekte identisch ist mit einem Test auf GARCH-Effekte. Demos und Sentana (1998) [30] vergleichen dieses Verfahren mit Alternativen wie dem Likelihood-Verhältnis-Test und dem Wald-Test.

## 2.2.3 Prognose mit GARCH-Modellen

Aus den bereits genannten Gründen besteht bei der Betrachtung der GARCH-Modelle großes Interesse an der Prognose der zukünftigen Werte für die Volatilität, also den bedingten Varianzen. Unter Verwendung der Mittelwertgleichung (2.4) mit  $\sigma_t^2 = \alpha_0 + \alpha_1 z_{t-1}^2 + \beta_1 \sigma_{t-1}^2$  mit  $t = 1, \dots, T$  erhält man eine  $h$ -Schritt Prognose für die bedingte Varianz mittels:

$$\begin{aligned} \hat{\sigma}_{t+h}^2 &= \hat{E}(\sigma_{t+h}^2 \mid \Psi_t) = \hat{\alpha}_0 + \hat{\alpha}_1 E(\varepsilon_{t+h-1}^2) + \dots + \hat{\alpha}_q E(\varepsilon_{t+h-q}^2) \\ &\quad + \hat{\beta}_1 E(\sigma_{t+h-1}^2) + \dots + \hat{\beta}_q E(\sigma_{t+h-p}^2). \end{aligned}$$

Aus Gründen der Lesbarkeit wird auf die Bedingung von  $\Psi_t$  in allen Erwartungswerten verzichtet. Die bedingten Erwartungswerte werden dabei folgendermaßen gesetzt:

$$E(\varepsilon_{t+i}^2) = \begin{cases} E(\sigma_{t+i}^2), & i > 0, \\ \varepsilon_{t+i}^2, & i \leq 0, \end{cases}$$

$$E(\sigma_{t+i}^2) = \frac{\hat{\alpha}_0}{1 - \sum_{j=1}^q \hat{\alpha}_j - \sum_{j=1}^p \hat{\beta}_j} \quad \text{für } i \leq -t + q.$$

Für ein GARCH(1,1)-Modell für eine 1-Schritt-Prognose:

$$\begin{aligned} E_T[\sigma_{T+1}^2] &= \alpha_0 + \alpha_1 E_T[z_T^2] + \beta_1 E_T[\sigma_T^2] \\ &= \alpha_0 + \alpha_1 z_T^2 + \beta_1 \sigma_T^2 \end{aligned}$$

Für h=2 ergibt sich:

$$\begin{aligned} E_T[\sigma_{T+2}^2] &= \alpha_0 + \alpha_1 E_T[z_{T+1}^2] + \beta_1 E_T[\sigma_{T+1}^2] \\ &= \alpha_0 + (\alpha_1 + \beta_1) E_T[\sigma_{T+1}^2] \end{aligned}$$

mit

$$E_T[z_{T+1}^2] = E_T[\varepsilon_{T+1}^2 \sigma_{T+1}^2] = E_T[\sigma_T^2]$$

Allgemein erhält man die h-Schritt-Prognose durch:

$$\begin{aligned} E_T[\sigma_{T+k}^2] &= \alpha_0 + \alpha_1 E_T[z_{T+k-1}^2] + \beta_1 E_T[\sigma_{T+1}^2] \\ &= \alpha_0 \sum_{i=0}^{k-1} (\alpha_1 + \beta_1)^i + (\alpha_1 + \beta_1)^{k-1} (\alpha_1 \varepsilon_T^2 + \beta_1 \sigma_T^2) \end{aligned}$$

Auf die Prognose zukünftiger Werte mit Hilfe eines GARCH-Modells wird in Abschnitt 4.5 noch detaillierter eingegangen.

## 2.3 Alternative Residualverteilungen

Mittels bekannter Tests wie beispielsweise von Jarque und Bera, Kolmogorow und Smirnow sowie grafischen Verfahren kann die Annahme der Normalverteilung der Residuen leicht getestet werden. Neben der allgemein verwendeten Normalverteilung eignen sich diverse andere Verteilungen ebenfalls als Residualverteilung. In Abschnitt 1.3 wurde gezeigt, dass die Normalverteilungsannahme für die Renditen nicht ohne weiteres unterstellt werden darf. Daher werden die beiden gängigsten Alternativen dazu in den Abschnitten 2.3.1 und 2.3.2 dargestellt. Diese beiden Alternativen sind i.d.R. auch in der gängigen Software implementiert und häufig besser geeignet die Verteilung der Renditen nachzubilden.

### 2.3.1 t-Verteilung

Die Leptokurtosis, also die Spitzgipfligkeit, von Finanzmarktdaten kann z.B. durch die  $t$ -Verteilung in das Modell eingebracht werden. Bollerslev (1987) [15] und Bollerslev und Baillie (1989) [4] verwenden für den Fehlerprozess  $\varepsilon_t$  aus Gleichung (2.4) die standardisierte  $t$ -Verteilung

$$z_t | \Psi_{t-1} = \sqrt{\sigma_t^2} \varepsilon_t, \quad t = 1, \dots, T,$$

mit  $\varepsilon_t \stackrel{u.i.v.}{\sim} t_n(0, 1)$ .

Im Zusammenhang mit dem Bayes-Ansatz verwenden Bauwens und Lubrano (1998) [7] ebenfalls die  $t$ -Verteilung.

Die Likelihoodfunktion einer Stichprobe mit  $T$  Beobachtungen für  $Z$  kann geschrieben werden als:

$$L_T(z|\theta) = \prod_{t=1}^T \left\{ \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) [(\nu-2) \cdot \sigma_t^2]^{1/2}} \left(1 + \frac{Z_t^2}{(\nu-2) \cdot \sigma_t^2}\right)^{-(\nu+1)/2} \right\}, \quad (2.8)$$

wobei  $\nu > 2$  die Anzahl der Freiheitsgrade der  $t$ -Verteilung kennzeichnet und  $\Gamma(\cdot)$  die Gammafunktion darstellt. GARCH-Modelle, die die  $t$ -Verteilung verwenden, werden im Folgenden mit  $\text{GARCH}_t$  bezeichnet.

Damit ergibt sich die Log-Likelihoodfunktion der  $t$ -Verteilung:

$$\begin{aligned} \ln L_T(z|\theta) &= \sum_{t=0}^T \ln \left[ \Gamma\left(\frac{\nu+1}{2}\right) \right] - \left[ \Gamma\left(\frac{\nu}{2}\right) \right] - \frac{1}{2} \cdot \ln[\pi(\nu-2)] \\ &\quad - \frac{1}{2} \sum_{t=0}^T \left[ \ln(\sigma_t^2) + (1+\nu) \cdot \ln \left( 1 + \frac{Z_t^2}{(\nu-2) \cdot \sigma_t^2} \right) \right]. \end{aligned}$$

Für die Wölbung  $\kappa$  der  $t$ -Verteilung gilt unter der Bedingung  $\nu > 4$ :

$$\kappa = 3 \cdot \frac{\nu-2}{\nu-4},$$

vgl. dazu beispielsweise Johnson, Kotz und Balakrishnan (1995) [61]. Daran lässt sich erkennen, dass die Wölbung der  $t$ -Verteilung größer ist als drei und somit größer als bei der Normalverteilung. Ebenso gilt, dass die  $t$ -Verteilung für  $n \rightarrow \infty$  gegen die Standardnormalverteilung strebt.

Abbildung 2.1 zeigt die Anpassung einer  $t$ -Verteilung mit 9 FHG an die Renditen des DAX aus Abschnitt 1.3. Die Anpassung ist augenscheinlich sehr gut. Die  $t$ -Verteilung bildet die Renditen des DAX im Vergleich zur Normalverteilung in Abbildung 1.5 deutlich besser ab. Die Schätzung der Anzahl der Freiheitsgrade wird dann in Abschnitt 4.3.3 gezeigt.

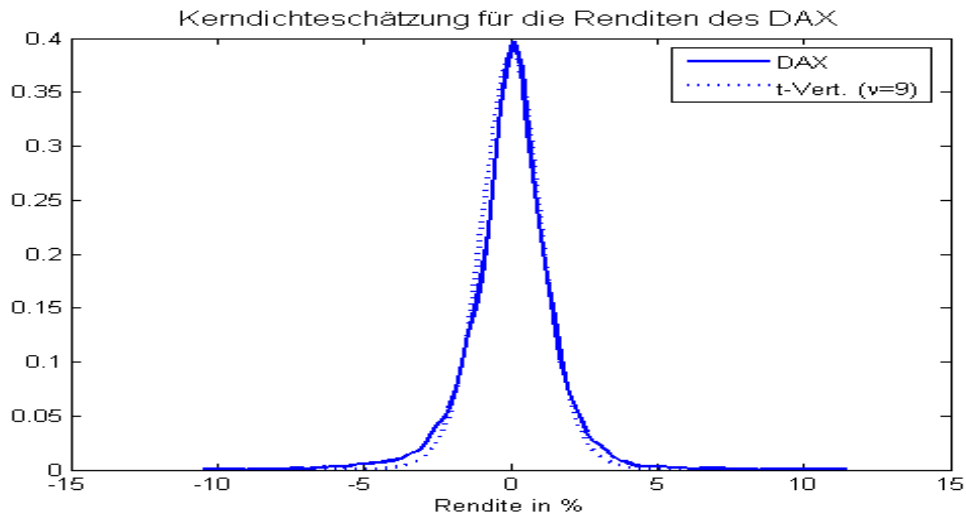


Abbildung 2.1: Angepasste t-Verteilung und Dichteschätzung

### 2.3.2 GED-Verteilung

Die Verwendung der generalisierten Fehlerverteilung oder auch Generalized Error Distribution, kurz *GED*-Verteilung, wurde von Nelson 1991 im Zusammenhang mit dem von ihm entwickelten EGARCH-Modell vorgeschlagen, welches in Abschnitt 2.4.3 vorgestellt wird. Für  $\varepsilon_t$  wird angenommen:

$$\varepsilon_t \stackrel{u.i.v.}{\sim} GED(0, 1).$$

Die *GED*-Verteilung hat folgende Dichtefunktion:

$$f(x) = \frac{\nu \cdot \exp\left\{-\frac{1}{2}\left|\frac{x}{\lambda}\right|^\nu\right\}}{\lambda \cdot 2^{1+\frac{1}{\nu}}\Gamma\left(\frac{1}{\nu}\right)}, \quad -\infty < x < +\infty, \quad 0 < \nu < +\infty,$$

mit

$$\lambda = \left[ 2^{-\frac{2}{\nu}} \Gamma\left(\frac{1}{\nu}\right) / \Gamma\left(\frac{3}{\nu}\right) \right]^{1/2}.$$

Der Parameter  $\nu$  bestimmt die Form der Ränder. Für Werte von  $\nu < 2$  besitzt die *GED*-Verteilung schwerere Ränder als die Normalverteilung, entsprechend bedeuten Werte  $\nu > 2$ , dass weniger Wahrscheinlichkeitsmasse in den Rändern steckt. Als Spezialfall erhält man für  $\nu = 2$  die Standardnormalverteilung, siehe dazu Malmsten und Teräsvirta (2004) [71].

Die *GED*-Verteilung beinhaltet noch weitere Spezialfälle, die sich in Abhängigkeit vom Parameter  $\nu$  ergeben:

- für  $\nu = 1$  : Standard-Laplace-Verteilung

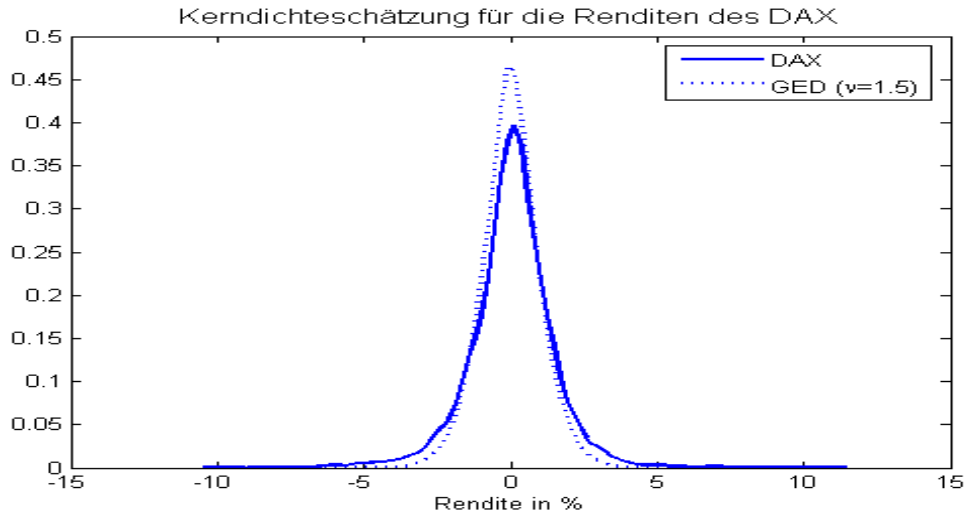


Abbildung 2.2: Angepasste GED-Verteilung und Dichteschätzung

- für  $\nu \rightarrow +1$  : Standard-Gleichverteilung.

Bei Verwendung der  $z_t$  und  $\sigma_t^2$  ist die Likelihoodfunktion gegeben durch:

$$L(\theta; y) = \left[ \frac{\nu}{\lambda 2^{1+(1/\nu)} \Gamma(\frac{1}{\nu})} \right]^T \times \prod_{t=1}^T \left\{ (\sigma_t^2)^{-1/2} \exp \left( -\frac{1}{2} \left| \frac{z_t}{\sigma_t \lambda} \right|^\nu \right) \right\}, \quad (2.9)$$

wobei wiederum gilt:

$$\lambda = \left[ 2^{(-2/\nu)} \frac{\Gamma(1/\nu)}{\Gamma(3/\nu)} \right]^{1/2}.$$

Die logarithmierte Likelihoodfunktion lautet:

$$\ln L(\theta; y) = T \cdot \ln \left( \frac{\nu}{\lambda \cdot 2^{(1+1/\nu)} \cdot \Gamma(\frac{1}{\nu})} \right) - \frac{1}{2} \cdot \sum_{t=1}^T \ln(\sigma_t^2) - \frac{1}{2} \cdot \sum_{t=1}^T \left| \frac{z_t}{\lambda \sigma_t} \right|^\nu,$$

diese kann alternativ auch folgendermaßen formuliert werden

$$\begin{aligned} \ln L(\theta; y) = & T \cdot \ln \left( \frac{\nu}{\lambda} \right) - T \cdot \left( 1 + \frac{1}{\nu} \right) \cdot \ln(2) - T \cdot \ln \left( \Gamma \left( \frac{1}{\nu} \right) \right) \\ & - \frac{1}{2} \sum \left| \frac{Z_t}{\lambda \cdot \sigma_t} \right|^\nu - \frac{1}{2} \sum \cdot \ln(\sigma_t^2). \end{aligned}$$

Für die Renditen des DAX aus Abschnitt 1.3 zeigt Abbildung 2.2 eine angepasste *GED*-Verteilung mit  $\nu = 1.5$ . Auch hier sei bezüglich der Schätzung des Parameters  $\nu$  auf Abschnitt 4.3.3 verwiesen. Auch die *GED*-Verteilung ergibt eine akzeptable Anpassung an die Verteilung der Renditen, welche augenscheinlich besser zu den Daten passt, als die Normalverteilung. Modelle, die eine *GED*-Verteilung unterstellen werden im Folgenden

mit  $\text{GARCH}_{GED}$  bezeichnet.

Auch wenn die  $t$ -Verteilung und die  $GED$ -Verteilung in der Praxis und auch in der Literatur sehr weit verbreitet sind, kommen noch diverse andere Verteilungen für die Residuen in Frage. So wurden u.a. die Laplace Verteilung, Logistische Verteilung, Hyperbolische Sekant, EGB2 Verteilung, NEF-GHS Verteilung, Verallgemeinerte hyperbolische Verteilung verwendet.

## 2.4 Erweiterungen des GARCH-Modells

Seit der Einführung der ARCH-Modelle im Jahre 1982 wurde eine wahre Flut an Weiterentwicklungen dieses Modells produziert. Unzählige Modelle versuchen die bedingte Volatilität in Zeitreihen präziser zu modellieren und zu prognostizieren. Eine sehr umfangreiche Sammlung der bestehenden GARCH-Erweiterungen hat Bollerslev in seinem Beitrag zur Festschrift zu Ehren von Robert F. Engle zusammengetragen [17]. Diese zeigt, dass die Fülle an Abkürzungen selbst schon wieder unübersichtlich ist. Er listet über 130 Akronyme und Abkürzungen auf, die seit dem Artikel von Engle im Zusammenhang mit dem ARCH-Modell entwickelt wurden. Um diesbezüglich eine gewisse Übersichtlichkeit zu bewahren, halten sich die in den folgenden Abschnitten dargestellten GARCH-Erweiterungen hinsichtlich der Abkürzungen an die von Bollerslev verwendeten Bezeichnungen.

Ein wichtiger Aspekt bei der Wahl der Modellvariante ist die Berücksichtigung und die Behandlung der Asymmetrie der Renditen, d.h. der Tatsache, dass negative Schocks einen größeren Einfluss auf die Kursschwankungen haben, als positive Schocks. Dieses Phänomen untersuchte Black 1976 [13], seitdem wird es allgemein als „Leverage-Effekt“ bezeichnet.

Neben den Modellen, die immer speziellere Formen der Heteroskedastie modellieren, gibt es noch einige Modelle, die versuchen allgemeinere Formen zu modellieren. Dazu gehören das Aug-GARCH-Modell von Duan (1997) [33] und das H-GARCH-Modell von Hentschel (1995) [55].

### 2.4.1 Erweiterungen der Mittelwertgleichung

Das Standard GARCH-Modell aus Abschnitt 2.2 lässt sich sowohl bezüglich der Mittelwertgleichung als auch der Varianzgleichung variieren. Bevor diverse Erweiterungen für

die Varianzgleichung dargestellt werden, soll in diesem Abschnitt ein kurzer Überblick über alternative Mittelwertgleichungen gegeben werden.

### 2.4.1.1 ARMA-GARCH

Ein ARMA-GARCH-Modell kombiniert die Eigenschaften der bekannten autoregressiven moving average (ARMA) Modelle mit GARCH-Effekten, d.h. für die Zeitreihe wird ein ARMA-Modell unterstellt und für die Residuen ein GARCH-Modell.

### 2.4.1.2 GARCH-in-Mean

Das GARCH-in-Mean-Modell von Engle, Lilien und Robbins (1987) [36] ermöglicht es Renditen in Abhängigkeit der bedingten Varianzen zu modellieren,

$$z_t = c + \gamma_t \cdot \sigma_{t-1} + u_t.$$

Dabei ist  $z_t$  die Rendite, sie wird ausgedrückt durch eine Konstante  $c$ , einen zeitabhängigen Teil und einen heteroskedastischen Term  $u_t$ , für den wiederum ein GARCH-Prozess unterstellt wird. Dies ist nur ein sehr kurzer Überblick über mögliche Varianten der Mittelwertgleichung. Soweit nicht anders angegeben wird im Folgenden die Mittelwertgleichung (2.1), also  $z_t = \sigma_t \cdot \varepsilon_t$ .

## 2.4.2 H-GARCH-Modell nach Hentschel

Das Modell von Hentschel ist eine Erweiterung der GARCH-Modelle. Es nestet eine Vielzahl symmetrischer und asymmetrischer Modelle. Das Modell dient in diesem Abschnitt dazu, einen Überblick über eine breite Auswahl an verallgemeinerten GARCH-Modellen zu bekommen. In den weiteren Abschnitten folgen dann diverse Spezifikationen der Varianzgleichung, die dann zu speziellen GARCH-Modellen führen.

Die Varianzgleichung eines verallgemeinerten H-GARCH(1,1)-Modells lautet:

$$\frac{\sigma_t^\lambda - 1}{\lambda} = \omega' + \alpha^* \cdot \sigma_{t-1}^\lambda \cdot f^\nu(\varepsilon_t) + \beta_1 \cdot \frac{\sigma_{t-1}^\lambda - 1}{\lambda}, \quad (2.10)$$

$$\text{mit } f(\varepsilon_t) = |\varepsilon_t - b| - c \cdot (\varepsilon_t - b). \quad (2.11)$$



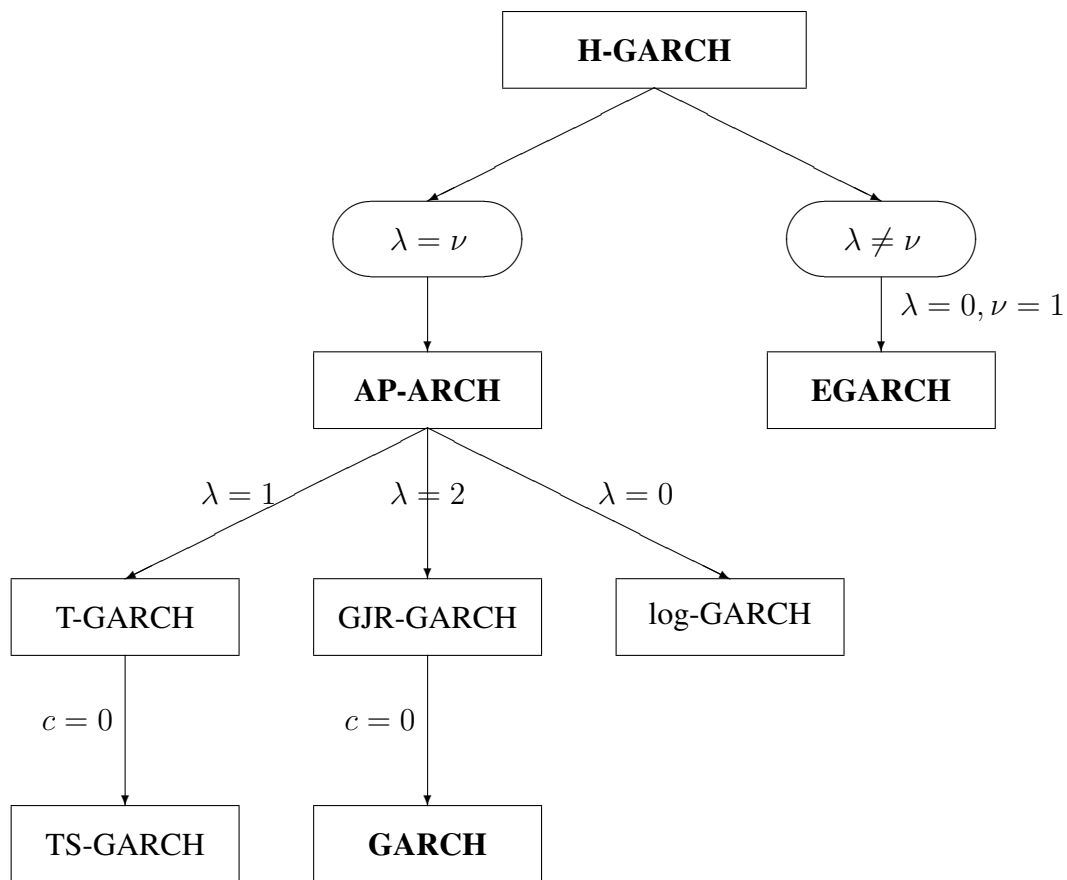


Abbildung 2.3: H-GARCH und genestete Modelle

Mit  $\alpha_0 = \omega' \cdot \lambda + 1 - \beta_1$  und  $\alpha_1 = \alpha^* \cdot \lambda$  lässt sich die Gleichung in die bekannte Form

$$\sigma_t^\lambda = \alpha_0 + \alpha_1 \cdot \sigma_{t-1}^\lambda \cdot f^\nu(\varepsilon_t) + \beta_1 \cdot \sigma_{t-1}^\lambda$$

überführen.

Die geeignete Wahl der Parameter  $\lambda$ ,  $\nu$ ,  $b$  und  $c$  erlaubt es unterschiedliche Varianzgleichungen zu spezifizieren und ermöglicht somit verschiedene Erweiterungen des GARCH-Modells. Dabei stehen sowohl symmetrische als auch asymmetrische Modelle zur Verfügung. Tabelle 2.1 listet die genesteten Modelle des H-GARCH-Modells auf, dabei bedeutet das Wort „frei“, dass der Parameter keinen Beschränkungen unterliegt.

Abbildung 2.3 gibt einen Überblick über die Verzweigungen der erweiterten GARCH-Modelle und ihren Beziehungen untereinander.

In den folgenden Abschnitten werden einige der genesteten Modelle näher dargestellt. Dabei wird sowohl die Notation verwendet, welche sich aus dem H-GARCH-Modell ergibt, als auch alternative Versionen, beispielsweise solche, wie sie in der Software Mat-

| $\lambda$ | $\nu$     | $b$  | $c$          | Modell                |
|-----------|-----------|------|--------------|-----------------------|
| 0         | 1         | 0    | frei         | EGARCH                |
| 1         | 1         | 0    | $ c  \leq 1$ | TGARCH                |
| 1         | 1         | frei | $ c  \leq 1$ | Abs. Value GARCH      |
| 2         | 2         | 0    | 0            | GARCH                 |
| 2         | 2         | frei | 0            | Nonlinear-asym. GARCH |
| 2         | 2         | 0    | frei         | GJR-GARCH             |
| frei      | $\lambda$ | 0    | 0            | Nonlinear ARCH        |
| frei      | $\lambda$ | 0    | $ c  \leq 1$ | AP-ARCH               |

Tabelle 2.1: Genestete Modelle in der Übersicht

lab und dazugehörigen Toolboxen verwendet werden. Analog zu der Schreibweise beim H-GARCH(1,1)-Modell wird im Folgenden auf die Bezeichnung der Ordnung des Parameters  $c$  bei der Bezeichnung des jeweiligen Modells verzichtet.

### 2.4.3 EGARCH-Modell nach Nelson

Eine sehr weit verbreitete Variante des GARCH-Modells ist das exponentielle GARCH (EGARCH)-Modell von Nelson (1991) [79]. Malmsten und Teräsvirta (2002) [70] oder Malmsten (2004) [69] behandeln intensiv dieses Modell. Es gehört zu den asymmetrischen GARCH-Modellen und bietet eine Möglichkeit den genannten Leverage-Effekt zu berücksichtigen. Bisher wurden sämtliche Ausschläge mit ihrem Quadrat  $\varepsilon_t^2$  berücksichtigt. Man erhält das Modell, indem man in den Gleichungen (2.10) und (2.11) die Parameter  $b = \lambda = 0$  und  $\nu = 1$  setzt.

Es wird der gleiche Fehlerprozess unterstellt wie beim GARCH-Modell (siehe Gleichung (2.4)):

$$z_t \mid \Psi_{t-1} = \sqrt{\sigma_t^2} \varepsilon_t \quad t = 1, \dots, T. \quad (2.12)$$

Die Varianzgleichung eines EGARCH(p,q)-Modells lautet:

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{j=1}^q \beta_j \cdot \ln(\sigma_{t-j}^2) + \sum_{k=1}^p [c_k \varepsilon_{t-k} + \alpha_k (|\varepsilon_{t-k}| - E(|\varepsilon_{t-k}|))], \quad (2.13)$$

wobei  $\sigma_t = z_t = 0$  für  $t < 0$ . Analog zum GARCH-Modell handelt es sich bei den  $\varepsilon_t$  um eine Folge identisch, unabhängig verteilter Zufallsvariablen mit  $E = 0$  und  $Var = 1$ . Dabei ist analog zum GARCH-Modell meist eine Form anzutreffen, bei der  $p = q = 1$ .

Die Varianzgleichung eines EGARCH(1,1)-Modells lautet

$$\ln(\sigma_t^2) = \alpha_0 + \beta_1 \cdot \ln(\sigma_{t-1}^2) + c \cdot \varepsilon_{t-1} + \alpha_1(|\varepsilon_{t-1}| - E(|\varepsilon_{t-1}|)). \quad (2.14)$$

Der Erwartungswert von  $|\varepsilon_{t-k}|$  ist in Abhängigkeit der Residualverteilung durch die nachstehenden Ausdrücke gegeben.

Für die Normalverteilung gilt:

$$E|\varepsilon_{t-k}| = \sqrt{2/\pi}.$$

Für die  $t$ -Verteilung mit  $\nu$  Freiheitsgraden gilt:

$$E|\varepsilon_{t-k}| = \sqrt{\frac{\nu-2}{\pi}} \cdot \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})},$$

und für die GED-Verteilung erhält man:

$$E|\varepsilon_{t-k}| = \frac{\Gamma(\frac{2}{\nu})}{[\Gamma(\frac{1}{\nu}) \cdot \Gamma(\frac{3}{\nu})]^{1/2}}.$$

Des Weiteren trifft Nelson die Annahme, dass  $\varepsilon_t$  in Gleichung (2.12) einer *GED*-Verteilung folgt, mit Erwartungswert 0, Varianz  $\sigma_t^2$  und  $\nu = 2$ , woraus sich die Normalverteilung ergibt.

Eine Besonderheit des EGARCH-Modells ist die Formulierung der bedingten Varianzen durch Logarithmen, welche dazu führt, dass in der Varianzgleichung keine Restriktionen notwendig sind, um positive Varianzen zu garantieren. Daher können die Parameter in einem EGARCH-Modell auch negative Werte annehmen.

#### 2.4.4 GJR-GARCH-Modell nach Glosten, Jagannathan und Runkle

Das Modell von Glosten, Jagannathan und Runkle (1993) [49] erhält man, wenn man in den Gleichungen (2.10) und (2.11) die Parameter  $b = 0$  und  $\nu = \lambda = 2$  setzt. Damit ergibt sich nach der Schreibweise von Hentschel:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 [(1 + c^2) \varepsilon_t^2 - 2c|\varepsilon_t| \varepsilon_t] + \beta_1 \sigma_{t-1}^2 \quad t = 1, \dots, T,$$

mit  $\alpha_1 = \alpha^* \cdot 2$ . Die Asymmetrie kann man in folgender Umformung leicht erkennen:

$$\sigma_t^2 = \begin{cases} \alpha_0 + \alpha_1(1 + c)^2 \varepsilon_t^2 \sigma_{t-1}^2 + \beta_1 \sigma_{t-1}^2, & \varepsilon < 0, \\ \alpha_0 + \alpha_1(1 - c)^2 \varepsilon_t^2 \sigma_{t-1}^2 + \beta_1 \sigma_{t-1}^2, & \varepsilon \geq 0, \end{cases}$$

Glosten, Jagannathan und Runkle spezifizieren das GJR(1,1)-Modell in einer alternativen Schreibweise wie folgt:

$$\sigma_t^2 = \alpha_0 + [\alpha'_1 z_t^2 - c' \cdot (z_t^+)^2] + \beta_1 \sigma_{t-1}^2 \quad \text{mit } t = 1, \dots, T,$$

Beide Schreibweisen lassen sich ineinander überführen, die Parameterwerte erhalten wir durch folgende Umformungen:  $\alpha'_1 = \alpha_1 \cdot (1 - c)^2$  und  $c' = \alpha_1 \cdot (1 - c) - \alpha_1 \cdot (1 + c)^2$ .

### 2.4.5 AP-ARCH-Modell nach Ding, Granger und Engle

Setzt man den Parameter des H-GARCH-Modells  $b = 0$ , führt dies zu dem asymmetrischen Power-ARCH-Modell von Ding, Granger und Engle [32]. Damit erhält man die Varianzgleichung nach Hentschel:

$$\sigma_t^\lambda = \alpha_0 + \alpha_1 \cdot \lambda \cdot (|z_t| - c \cdot z_t)^\lambda + \beta_1 \cdot \sigma_{t-1}^\lambda \quad t = 1, \dots, T,$$

alternativ dazu kann für das AP-ARCH-Modell folgende Schreibweise verwendet werden:

$$\sigma_t^\lambda = \alpha_0 + \alpha'_1 \cdot (|z_t| - c \cdot z_t)^\lambda + \beta_1 \cdot \sigma_{t-1}^\lambda \quad t = 1, \dots, T$$

mit  $\alpha'_1 = \alpha_1 \cdot \lambda$ . Die Besonderheit dieser Varianzgleichung liegt offensichtlich in der Verwendung des Parameters  $\lambda$ . Das AP-ARCH-Modell beinhaltet für  $c = 0$  auch das nichtlineare (G)ARCH-Modell von Higgins und Bera (1992) [56].

### 2.4.6 TS-GARCH-Modell nach Taylor und Schwert

Folgende Festlegung der Parameter in den Gleichungen (2.10) und (2.11):  $\lambda = \nu = 1$ ,  $|c| \leq 1$  und ein beliebiger Wert für  $b$ , charakterisiert das TS-GARCH-Modell von Taylor (1986) [89] und Schwert (1989) [83]. Das Modell lässt sich sehr leicht vereinfachen, indem man  $b = c = 0$  setzt. Dies führt zu folgender Varianzgleichung:

$$\sigma_t = \alpha_0 + \alpha_i \cdot |z_{t-1}| + \beta_j \sigma_{t-j} \quad t = 1, \dots, T,$$

Es werden also im Gegensatz zu den bisher vorgestellten Modellen absolute Beträge verwendet. Es ist leicht zu sehen, dass es sich hierbei um eine Varianzgleichung handelt, die der Gleichung (2.5) aus dem Standard GARCH(1,1) ähnelt. Es werden lediglich die

Standardabweichung anstatt der Varianzen betrachtet. Wegen Jenens's Ungleichung ist der Einfluss großer Schocks bei diese Betrachtung kleiner als bei den Standard GARCH-Modellen nach Bollerslev.

### 2.4.7 TGARCH-Modell nach Zakoian

Das Threshold-GARCH-Modell von Zakoian (1991) [97] ergibt sich aus dem H-GARCH-Modell wenn die Parameter in den Gleichungen (2.10) und (2.11) wie folgt gewählt werden  $\lambda = \nu = 1$ ,  $b = 0$  und ein beliebiger Wert für  $c$ .

Wie bei dem TS-GARCH-Modell werden die bedingten Standardabweichungen modelliert. Damit erhält man als Varianzgleichung eines T-GARCH(1,1)-Modells

$$\begin{aligned}\sigma_t &= \alpha_0 + \alpha_1[|Z_t| - c \cdot Z_t] + \beta_1\sigma_{t-1} \\ &= \alpha_0 + \alpha_1 \cdot (1 - c)Z_t^+ - \alpha_1 \cdot (1 + c) \cdot Z_t^-] + \beta_1\sigma_{t-1}\end{aligned}$$

dabei ist  $z_t^+ = \max(z - t, 0)$  und  $z_t^- = \max(Z_t, 0)$ .

Daneben kann

$$\sigma_t = \alpha_0 + \alpha_1^+ z_t^+ - \alpha_1^- z_t^- + \beta_1\sigma_{t-1} \quad t = 1, \dots, T,$$

als Schreibweise für die Varianzgleichung gewählt werden. Offensichtlich lassen sich die Parameterwerte der beiden Varianten einfach ineinander überführen:  $\alpha_1^+ = \alpha_1 \cdot (1 - c)$  und  $\alpha_1^- = \alpha_1 \cdot (1 + c)$ . Diese Schreibweise entspricht

$$\sigma_t = \alpha_0 + \alpha_1' \cdot |z_t| + c' \cdot |z_t| \cdot I(z_t < 0) + \beta_1\sigma_{t-1} \quad t = 1, \dots, T.$$

In diesem Fall können die Parameter durch  $\alpha_1' = \alpha_1 \cdot (1 - c)$  und  $c' = 2 \cdot \alpha \cdot c$  überführt werden.

## 2.5 Multivariate GARCH-Modelle

Abbildung 2.4 zeigt den Verlauf des DAX, des Dow Jones und des japanischen Nikkei 225 in dem Zeitraum vom 26.11.1990 bis 17.09.2009. Dabei ist zu berücksichtigen, dass es in den verschiedenen Ländern in diesem Zeitraum unterschiedlich viele Handelstage gab, daher sind die Zeitreihen minimal gegeneinander verschoben. Es fällt auf, dass sich die Indizes in Teilabschnitten ähnlich entwickeln. Besonders gut ist dies bei den Zeitreihen des Dax und des Dow Jones zu sehen. Für den Nikkei gilt dies ca. ab dem 3000.

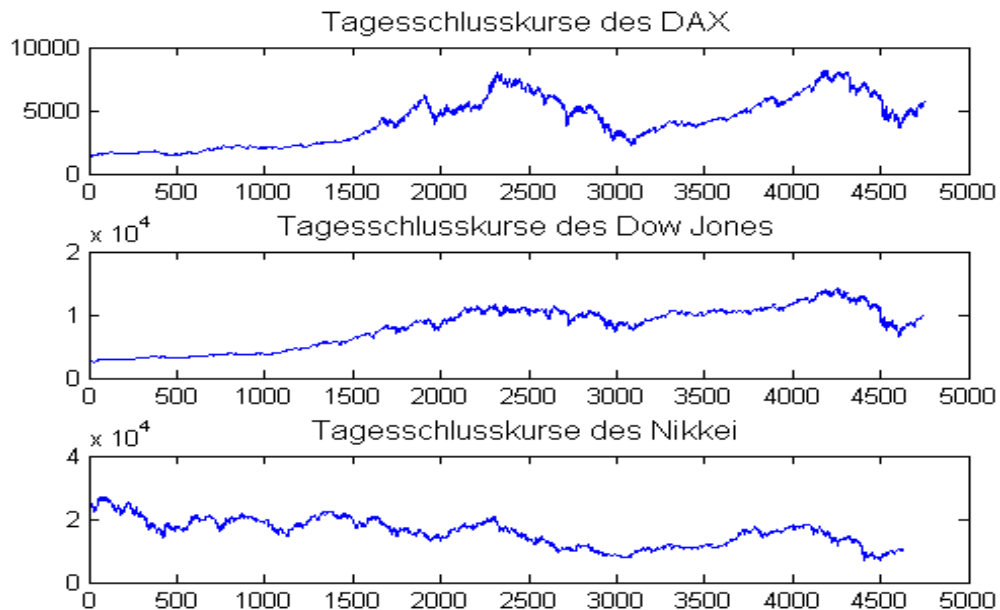


Abbildung 2.4: Tagesschlusskurse vom 26.11.1990 bis 17.09.2009 für DAX, Dow Jones und Nikkei

Wert dieser Zeitreihe, also um die Jahrtausendwende. Dies ist angesichts des globalen Zusammenspiels der Finanzmärkte auch nicht verwunderlich. Auch die stetigen Renditen in Abbildung 2.5 deuten darauf hin, dass die Märkte über die Zeit voneinander abhängig sind. Dies legt nahe, die kontemporären Korrelationen von Zeitreihen für die Analyse zu berücksichtigen. Ebenso verhält es sich mit Unternehmen derselben Branche. Auch gegensätzliche Entwicklungen z.B. beim Goldpreis kommen in Betracht.

Im Folgenden werden zwei GARCH-Modelle vorgestellt, die geeignet sind Zeitreihen multivariat zu modellieren. Es existieren diverse weitere Modelle, die deutlich komplexere Möglichkeiten der Modellierung bieten. Allerdings werden hier nur die Modelle vorgestellt, die in dieser Arbeit durch die in Kapitel 3 vorgestellten Bayes-Schätzungen geschätzt werden. Die vorgestellten Modelle haben ebenfalls den Vorteil, dass die Zeitreihen univariat mit unterschiedlichen GARCH-Varianten modelliert werden können. Einen sehr guten Überblick über verschiedene multivariate Modelle findet man bei Bauwens, Laurent und Rombouts [6]. Auch das Buch von Gouriéroux (1997) [50] bietet einen guten Zugang zu den multivariaten GARCH-Modellen.

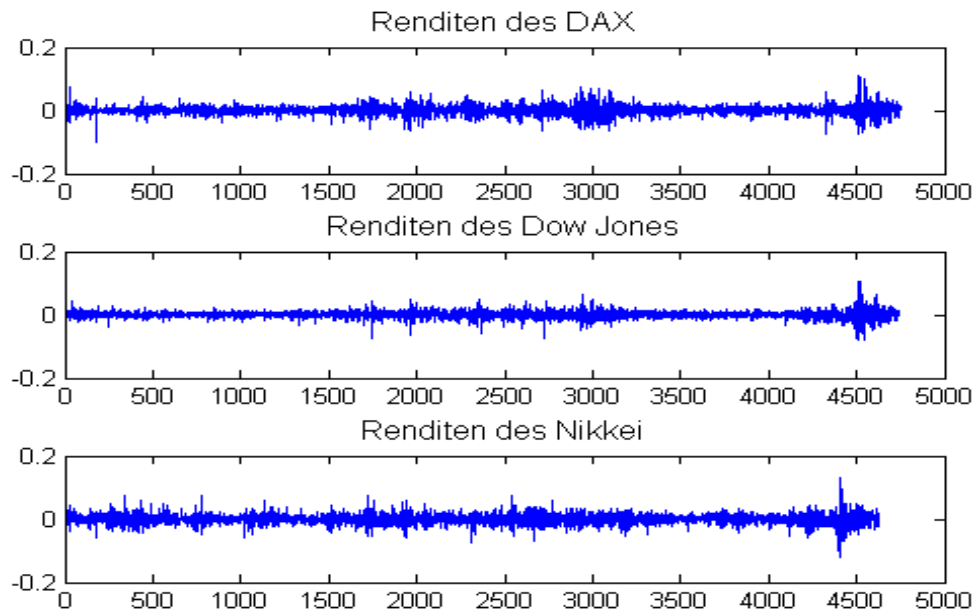


Abbildung 2.5: Renditen vom 26.11.1990 bis 17.09.2009 für DAX, Dow Jones und Nikkei

### 2.5.1 Constant-Conditional-Correlation-GARCH-Modell

Das Constant-Conditional-Correlation-GARCH-Modell (CCC-GARCH) wurde von Bollerslev (1990) [16] vorgeschlagen. Es bietet eine sehr einfache Möglichkeit der Modellierung multivariater Zeitreihen. Dabei werden für die  $N$  univariaten Zeitreihen jeweils GARCH-Prozesse verwendet und zwischen den einzelnen Prozessen werden untereinander konstante bedingte Korrelationen unterstellt. Das CCC-Modell bietet aber generell die Möglichkeit die univariaten Zeitreihen durch andere Varianten des GARCH-Modells, wie z.B. den in Abschnitt 2.4 erwähnten, zu modellieren.

Die zentrale Annahme im CCC-Modell ist, dass die bedingten Korrelationen über die Zeit konstant sind, d.h.  $\rho_{ij,t} = \rho_{ij}$ . Für die bedingten Kovarianzen gilt dann:

$$\sigma_{ij,t} = \rho_{ij} \sqrt{\sigma_{i,t}^2 \cdot \sigma_{j,t}^2} \quad \text{mit } j = 1, \dots, N \quad \text{und} \quad i = j + 1, \dots, N,$$

bzw. für die bedingten Korrelationen zwischen  $z_{i,t}$  und  $z_{j,t}$  gilt:

$$\text{Korr}(z_{i,t}, z_{j,t} \mid \mathcal{F}_{t-1}) = \frac{\sigma_{ij,t}}{\sqrt{\sigma_{i,t}^2 \cdot \sigma_{j,t}^2}} = \frac{\rho_{ij} \sqrt{\sigma_{i,t}^2 \cdot \sigma_{j,t}^2}}{\sqrt{\sigma_{i,t}^2 \cdot \sigma_{j,t}^2}} = \rho_{ij}.$$

Bollerslev formuliert die bedingte Varianz-Kovarianz-Matrix folgendermaßen:

$$\Sigma_t = D_t \cdot R \cdot D_t, \quad (2.15)$$

dabei ist  $R$  die symmetrische, positiv definite,  $N \times N$  Korrelationsmatrix mit den Elementen  $|\rho_{ij}| \leq 1$  für  $i \neq j$  und  $\rho_{ij} = 1$  für alle  $i = 1, \dots, N$ . Die  $N \times N$  Diagonal-Matrix

$$D_t = \begin{pmatrix} \sigma_{11,t} & 0 & \cdots & \cdots \\ 0 & \sigma_{22,t} & 0 & \cdots \\ \vdots & 0 & \ddots & \\ \vdots & \vdots & & \sigma_{nn,t} \end{pmatrix} \quad (2.16)$$

beinhaltet für den Zeitpunkt  $t$  die bedingten Standardabweichungen der einzelnen Prozesse. Für jede der  $N$  univariaten Zeitreihen wird z.B. entsprechend Gleichung (2.5) ein GARCH-Modell verwendet<sup>1</sup>:

$$\sigma_{ii,t}^2 = \alpha_{i0} + \sum_{p=1}^P \alpha_{ip} \cdot z_{i,t-p}^2 + \sum_{q=1}^Q \beta_{iq} \cdot \sigma_{ii,t-q}^2 \quad \forall \quad i = 1, \dots, N. \quad (2.17)$$

Das CCC-Modell bietet die Möglichkeit die bedingten Varianzen der univariaten Zeitreihen flexibel zu modellieren. Ein weiterer Vorteil ist die geringe Anzahl zu schätzender Parameter, so sind für  $N$  Zeitreihen lediglich  $\left[ \frac{N(N-1)}{2} \right] + (1 + P + Q) \cdot N$  Parameter zu schätzen. Bollerslev zeigt auch Möglichkeiten auf, die ML-Schätzung der Parameter deutlich zu vereinfachen.

Die Annahme, die empirischen Korrelationen verschiedener Kapitalmarktzeitreihen seien auch über einen längeren Zeitraum konstant, ist allerdings sehr restriktiv. Überprüft werden kann sie beispielsweise anhand der Tests von Tse (2000) [92], Bera und Kim (2002) [10] oder Nakatani und Teräsvirta (2009) [77], wobei der letztgenannte Test auf das Extended Constant Correlation (ECCC-) GARCH-Modell von Jeantheau (1998) [59] zurück greift, welches eine Erweiterung des Ansatzes von Bollerslev darstellt und dieses beinhaltet.

Der folgende Abschnitt stellt ein weiteres Modell vor, bei dem die bedingten Korrelationen nicht als konstant angenommen werden.

<sup>1</sup>Es gelten die Parameterrestriktionen für jedes der GARCH-Modelle aus Abschnitt 2.2.



## 2.5.2 Dynamic-Conditional-Correlation-GARCH-Modell

Für das Dynamic-Conditional-Correlation-GARCH-Modell (DCC-GARCH) gibt es in der Literatur unterschiedliche Varianten. Zwei sehr verbreitete Versionen stammen von Tse und Tsui (2000) [93] bzw. Engle (2000a) [35] und Engle und Sheppard (2001) [37]. Beide verallgemeinern das CCC-Modell, modellieren aber die bedingte Korrelationsmatrix  $R_t$  unterschiedlich. Daher werden die beiden Modelle gelegentlich mit  $DCC_T$  und  $DCC_E$  bezeichnet. Da sich diese Arbeit ausschließlich auf das Modell nach Engle und Sheppard bezieht, wird im Folgenden auf die Unterscheidung verzichtet.

Auch in diesem Modell besteht die Möglichkeit zunächst die Parameter der univariaten GARCH-Prozesse zu schätzen. Dies kann auch hier wie in Gleichung (2.17) geschrieben werden. Andere GARCH-Varianten sind ebenfalls möglich. Die bedingten Varianzen der univariaten GARCH-Prozesse  $\sigma_{i,t}^2$  werden dann zur Standardisierung der  $z_{i,t}$  und zum Aufbau der Diagonalmatrix der bedingten Standardabweichungen  $D_t = \text{diag}[\sigma_{ii,t}]$  für  $i = 1, \dots, N$ , analog zu Gleichung (2.16), genutzt.

Im Gegensatz zum CCC-GARCH wird im DCC-Modell die konstante bedingte Korrelationsmatrix in (2.15) als zeitlich variabel modelliert. Für die bedingte Varianz-Kovarianz-Matrix ergibt sich dann:

$$\Sigma_t = D_t \cdot R_t \cdot D_t, \quad (2.18)$$

Engle verwendet folgende Schreibweise zur Darstellung der Korrelationsmatrix  $R_t$ :

$$R_t = Q_t^{*-1} \cdot Q_t \cdot Q_t^{*-1}, \quad (2.19)$$

und schlägt dabei in Anlehnung an ein univariates GARCH(1,1)-Modell folgende Schätzung der Korrelationsmatrizen vor. Zuerst werden die Matrizen  $Q_t$  geschätzt:

$$Q_t = \tilde{Q} \left( 1 - \sum_{w=1}^W \alpha_w^* - \sum_{v=1}^V \beta_v^* \right) + \sum_{w=1}^W \alpha_w^* \cdot (\tilde{z}_{t-w} \tilde{z}'_{t-w}) + \sum_{v=1}^V \beta_v^* \cdot Q_{t-v}, \quad (2.20)$$

hier ist  $\tilde{Q}$  die unbedingte Kovarianzmatrix der standardisierten Residuen  $z_{i,t}/\sigma_{ii,t}$  und  $\alpha^*$  und  $\beta^*$  sind skalare Parameter, vergleichbar mit den Parametern eines GARCH-Modells<sup>2</sup>. Aus den  $Q_t$  wird dann die bedingte Korrelationsmatrix entsprechend Gleichung (2.19) durch

<sup>2</sup>Daher gelten die Parameterrestriktionen aus Abschnitt 2.2 auch hier für  $\alpha^*$  und  $\beta^*$ .

$$R_t = \begin{pmatrix} \frac{1}{\sqrt{q_{11,t}}} & 0 & \cdots & \cdots \\ 0 & \frac{1}{\sqrt{q_{22,t}}} & 0 & \cdots \\ \vdots & 0 & \ddots & \\ \vdots & & & \frac{1}{\sqrt{q_{nn,t}}} \end{pmatrix} Q_t \begin{pmatrix} \frac{1}{\sqrt{q_{11,t}}} & 0 & \cdots & \cdots \\ 0 & \frac{1}{\sqrt{q_{22,t}}} & 0 & \cdots \\ \vdots & 0 & \ddots & \\ \vdots & & & \frac{1}{\sqrt{q_{nn,t}}} \end{pmatrix} \quad (2.21)$$

berechnet. Es ergibt sich eine Korrelationsmatrix mit Einsen auf der Hauptdiagonalen. Für  $\alpha^* + \beta^* = 1$  ergibt sich in (2.18) die exponentielle Glättung. Die Anzahl zu schätzender Parameter erhöht sich im Vergleich zum CCC-Modell beim DCC-Modell nur gering auf  $\left[ \frac{N(N-1)}{2} \right] + (1 + P + Q) \cdot N + (W + V)$ . Das DCC-Modell bietet zusätzliche Flexibilität gegenüber dem CCC-Modell und hält die Anzahl zu schätzender Parameter dennoch gering. Aber auch hier werden die Parameter  $\alpha^*$  und  $\beta^*$  konstant gehalten, d.h. die bedingten Korrelationen sind zwar zeitlich variabel, unterliegen aber über den gesamten Zeitraum demselben Prozess.

## Kapitel 3

# Bayes-Ansatz und Markov Chain Monte Carlo-Methoden

Nachdem in Kapitel 2 GARCH-Modelle inklusive deren Schätzung durch die Maximum-Likelihood-Methode angesprochen wurden, sollen in diesem Kapitel die Grundzüge dargelegt werden, mit denen eine Betrachtung der Modelle aus bayesianischer Sicht möglich ist.

Die Anwendungsgebiete des Bayes-Ansatzes werden parallel mit der fortschreitenden Steigerung der Leistungsfähigkeit von Computern immer vielfältiger und seit geraumer Zeit wird er auch auf GARCH-Modelle angewendet. In Abschnitt 3.1 werden die Grundzüge der Bayes-Statistik erläutert und mit den Markov Chain Monte Carlo-Verfahren aus Abschnitt 3.2 eine Möglichkeit vorgestellt, mit der man die Schätzung der Modellparameter durchführen kann.

### 3.1 Bayes-Ansatz

Unzählige Bücher und Artikel bieten einen guten Zugang zum Bayes-Ansatz. Einige herausragende Bücher sind die Werke von Carlin und Louis (2000) [22], Gelman et al. (2003) [43] und der Artikel von Geweke und McCausland (2001) [46]. Bauwens und Lubrano (2002) [8] zeigen, wie der Bayes-Ansatz verwendet werden kann, um Optionen zu bepreisen.

### 3.1.1 Einführung

In der frequentistischen Statistik betrachtet man den Datenvektor  $Y$  als Zufallsvariable mit dessen Hilfe man den, die Daten beschreibenden „wahren“ Parametervektor, zu schätzen versucht. In dieser Betrachtungsweise des Parametervektors und des Datenvektors unterscheidet sich der Bayes-Ansatz grundlegend. In der Bayes-Statistik wird der unbekannte Parametervektor  $\Theta = (\Theta_1, \Theta_2, \dots, \Theta_p)'$  selbst als Zufallsvariable aufgefasst, welcher eine eigene Verteilung besitzt. Es werden also sowohl die Beobachtungen als auch die Parameter als Zufallsvariablen betrachtet.  $Y$  bezeichnet die Daten und  $f_{Y,\Theta}(\theta, y)$  die gemeinsame Verteilung der Zufallsvariablen. Diese lässt sich darstellen als

$$f_{\Theta,Y}(\theta, y) = f_{Y|\Theta}(y | \theta) f_{\Theta}(\theta).$$

Der Bayes-Statistiker hat die Möglichkeit Informationen über die Verteilung des Parametervektors, über die er vor Bekanntwerden der Daten verfügt, über die a priori Dichte  $f_{\Theta}(\theta)$  mit in die Analyse einfließen zu lassen, vgl. dazu Abschnitt 3.1.2. Schlussfolgerungen über den Parametervektor  $\theta = (\theta_1, \dots, \theta_n)$  bedingt auf die Daten  $Y = y$  werden über die a posteriori Dichte  $f_{\Theta|Y}(\theta | y)$  gemacht. Diese erhält man unter Verwendung des Bayes-Theorems

$$f_{\Theta|Y}(\theta | y) = \frac{f_{\Theta,Y}(\theta, y)}{f_Y(y)} = \frac{f_{\Theta}(\theta) f_{Y|\Theta}(y | \theta)}{f_Y(y)}. \quad (3.1)$$

Der Ausdruck im Nenner  $f_Y(y) = \int f_{\Theta}(\theta) f_{Y|\Theta}(y | \theta)$  ist eine Normierungskonstante, die sicher stellt, dass es sich bei der a posteriori Verteilung um eine Dichtefunktion handelt. Der Term  $f_{Y|\Theta}(y | \theta)$  wird als Funktion von  $\Theta$  betrachtet und stellt im frequentistischen Sinne die Likelihoodfunktion  $L(\theta; y)$  dar. Die a posteriori Verteilung lautet vor Bekanntwerden der Daten  $f_{\Theta|Y}(\theta | y) = c \cdot L(y | \theta) \cdot f_{\Theta}(\theta)$ , mit der Normierungskonstante  $c$ , der Likelihoodfunktion  $L(y | \theta)$ , und der a priori Dichte  $f_{\Theta}(\theta)$ . Sobald die Daten vorliegen, nimmt die Normierungskonstante einen festen Wert an. Somit ist

$$c^{-1} = \begin{cases} \int_{\Omega} f_{Y|\Theta}(y | \theta) f_{\Theta}(\theta) d\theta & \text{falls } \Theta \text{ stetig,} \\ \sum_{\Omega} f_{Y|\Theta}(y | \theta) f_{\Theta}(\theta) & \text{falls } \Theta \text{ diskret.} \end{cases}$$

Damit kann man die a posteriori Verteilung schreiben, als

$$f_{\Theta|Y}(\theta | y) \propto L(y | \theta) \cdot f_{\Theta}(\theta). \quad (3.2)$$

Dabei wird in der praktischen Anwendung die Existenz der Normierungskonstante unterstellt. Die Normierungskonstante ist aber insbesondere bei der Modellselektion von ent-

scheidender Bedeutung, da sie die so genannte Modellwahrscheinlichkeit darstellt. Dies wird in Abschnitt 3.3 erläutert.

In der Praxis stellt die exakte Bestimmung von  $f_{\Theta|Y}(\theta | y)$  oft ein schwieriges bis unlösbares Problem dar. Für die Bestimmung der interessierenden a posteriori Verteilung haben sich daher Markov Chain Monte Carlo-(MCMC-)Methoden als hilfreiche Simulationsmethoden etabliert.

### 3.1.2 Die a priori Verteilung

Wie im vorangegangenen Abschnitt beschrieben, wird die Unsicherheit bezüglich des als Zufallsvariable angesehenen Parametervektors  $\theta$  in Form der *a priori* Verteilung ausgedrückt. In den Gleichungen (3.1) und (3.2) wird der Einfluss der a priori Verteilung deutlich. Daher ist besondere Sorgfalt bei der Auswahl einer geeigneten a priori Verteilung nötig. Der Bayes-Statistiker kann sein vorhandenes Wissen über den Parametervektor  $\theta$  vor Kenntnis der Daten über die a priori einbringen und muss dabei allerdings vermeiden die Ergebnisse über die a priori fälschlicherweise zu manipulieren. Deshalb ist es in der Praxis, auch im Zusammenhang mit den in Kapitel 2 angesprochenen GARCH-Modellen, eher eine Herausforderung eine geeignete a priori zu finden, die zum Ausdruck bringt, dass ausdrücklich kein Wissen über  $\theta$  im Vorhinein vorlag und die Ergebnisse somit frei von Manipulation sind. Daraus ergeben sich unterschiedliche Ansätze der Bestimmung einer als geeignet angesehenen a priori. Wird auf die Verwendung einer a priori verzichtet, so wird aus Gleichung (3.2) deutlich, dass die a posteriori Verteilung der Likelihoodverteilung entspricht.

Es lassen sich *informative* und *uninformative* a priori Verteilungen, also solche mit und ohne Beitrag zusätzlicher Informationen, unterscheiden. Bei der Verwendung von informativen a priori Verteilungen setzt sich der Anwender immer auch der möglichen Kritik der Subjektivität aus. Daher wird in der Praxis oft auf die Berücksichtigung von vorab verfügbaren Informationen verzichtet. Man beraubt sich dabei allerdings einem der Hauptvorteile der Bayes-Analyse. Meist werden deshalb uninformative a priori Verteilungen benutzt. Dabei unterscheidet man zwischen diffusen und unechten a priori Verteilungen. Während unechte a priori Verteilungen nur Konstanten sind, die keine Dichtefunktion ergeben, folgen diffuse oder vage a priori Verteilungen prinzipiell einer Dichtefunktion. Allerdings weisen sie i.d.R. eine so hohe Streuung auf, dass sie beinahe als lokal gleichverteilt gelten können.

### 3.1.3 Die a posteriori Prädiktivverteilung

Basierend auf den beobachteten Daten kann die a posteriori Prädiktivverteilung verwendet werden, um zukünftige Werte vorherzusagen, vgl. dazu Gelman et al. (2003) [43] oder Carlin und Louis (2000) [22]. Im Zusammenhang mit der Zeitreihenanalyse ist sie die Basis für Prognosen zukünftiger Werte. Dazu werden die vergangenen und beobachteten Werte  $Y_1, \dots, Y_n$  herangezogen und die Verteilung der zukünftigen Werte gegeben die beobachteten Daten  $f_{Y_{n+1}, \dots, Y_{n+m} | Y_1, \dots, Y_n}(y_{n+1}, \dots, y_{n+m} | y_1, \dots, y_n)$  wird als a posteriori Prädiktivverteilung bezeichnet. Sie ergibt sich gemäß

$$\begin{aligned} & f_{Y_{n+1}, \dots, Y_{n+m} | Y_1, \dots, Y_n}(y_{n+1}, \dots, y_{n+m} | y_1, \dots, y_n) \\ &= \int_{\Omega} f_{Y_{n+1}, \dots, Y_{n+m} | \Theta, Y_1, \dots, Y_n}(y_{n+1}, \dots, y_{n+m} | \theta, y_1, \dots, y_n) f_{\Theta | Y_1, \dots, Y_n}(\theta | y_1, \dots, y_n) d\theta \\ &= \int_{\Omega} f_{Y_{n+1}, \dots, Y_{n+m} | \Theta}(y_{n+1}, \dots, y_{n+m} | \theta) f_{\Theta | Y_1, \dots, Y_n}(\theta | y_1, \dots, y_n) d\theta \\ &= \int_{\Omega} \prod_{i=n+1}^m f_{Y_i | \Theta}(y_i | \theta) f_{\Theta | Y_1, \dots, Y_n}(\theta | y_1, \dots, y_n) d\theta. \end{aligned}$$

Dabei wird vereinfachend unterstellt, dass bedingt auf  $\Theta = \theta$  die  $Y_i$  stochastisch unabhängig sind. Generell kann die Qualität der a posteriori Prädiktivverteilung und damit des zugrunde liegenden Modells durch einen Vergleich mit der empirischen Verteilung neuer Werte überprüft werden. Dies unterscheidet die a posteriori Prädiktivverteilung von der a posteriori Verteilung von  $\Theta$  gegeben die Daten, bei der es keine vergleichbare Möglichkeit gibt, da der Parametervektor grundsätzlich nicht beobachtbar ist.

## 3.2 Markov Chain Monte Carlo-Verfahren

Im Zentrum der Bayes-Analyse steht die Bestimmung der a posteriori Verteilung. Da die zum Teil hochdimensionale a posteriori Verteilung der Parameter in der Praxis oft nicht exakt berechenbar ist, werden numerische oder asymptotische Verfahren verwendet, um die gesuchten Verteilungen zu simulieren; einen guten Überblick über mögliche Verfahren geben beispielsweise Evans und Swartz (1995) [38] oder Gilks et al. (1996) [48].

### Monte Carlo-Simulation

Monte Carlo(MC)-Verfahren bieten u.a. eine Möglichkeit Integrale der Form

$$\int g(y) \cdot \pi(y) dy = E_{\pi}[g(Y)]$$

zu berechnen. In der Bayes-Statistik steht  $\pi$  hier für die in Gleichung (3.1) dargestellte a posteriori Verteilung. Es werden voneinander unabhängige Zufallszahlen  $Y = Y_1, \dots, Y_n$  aus einer Verteilung mit der Dichte  $\pi(y)$  generiert. Es ergibt sich als Schätzung:  $E_\pi[g(Y)] \approx \frac{1}{n} \sum_{i=1}^n g(Y_i)$ . Das Starke Gesetz der Großen Zahlen sichert bei Unabhängigkeit der  $Y$ , dass der empirische Mittelwert fast sicher gegen den wahren Wert  $E[g(Y)]$  konvergiert. Das Erzeugen unabhängiger Zufallszüge aus der Dichte  $\pi(y)$  ist allerdings nicht immer möglich. So wird es bei steigender Anzahl von Parametern zunehmend schwerer eine geeignete Dichte  $\pi(y)$  zu finden, aus der unabhängige Werte für die Parameter gezogen werden können.

Allerdings kann man grob gesprochen sagen, dass die Unabhängigkeit nicht notwendig für die Konvergenz ist. Es würde auch genügen, wenn man  $Y_i$ 's durch einen Prozess erzeugen kann, der Werte aus dem gesamten Wertebereich von  $\pi(y)$  im richtigen Verhältnis liefert.

Die Markov Chain Monte Carlo-Verfahren stellen eine Verallgemeinerung der MC-Simulationen dar. Sie verbinden die Monte Carlo-Verfahren zur numerischen Berechnung von Integralen mit der Theorie der Markovketten. Der entscheidende Unterschied besteht in der Tatsache, dass nun voneinander abhängige Zufallszahlen gezogen werden.

### 3.2.1 Theorie der Markovketten

Markov Chain Monte Carlo-(MCMC-)Methoden bieten eine gute Möglichkeit für die Bestimmung selbst hochdimensionaler a posteriori Verteilungen. Sie sind benannt nach dem russischen Mathematiker Andrei Andrejewitsch Markow.

Dieser Abschnitt befasst sich mit den Eigenschaften der Markovketten und gibt einen Einblick in die dazugehörige Theorie. Grimmett und Stirzaker (1992) [52] und Gamermann (2006) [41] befassen sich intensiv mit der Theorie von Markovketten. Weitere Erläuterungen zu Markovketten und deren Anwendung in der Praxis finden sich in Gilks, Richardson und Spiegelhalter (1996) [48]. Ebenfalls zur Standardliteratur im Zusammenhang mit Monte Carlo-Methoden gehört Robert und Casella (2004) [81].

#### 3.2.1.1 Definition und Eigenschaften einer Markovkette

Man unterscheidet Markovketten in diskreter und in stetiger Zeit. Markovketten in stetiger Zeit werden häufig als Markov-Prozess bezeichnet. In der gängigen Literatur werden die Eigenschaften der Markovketten meist anhand von diskreten, endlichen Zustands-

räumen erklärt, daher beschränken sich die folgenden Ausführungen ebenfalls auf diskrete Zustandsräume. Alle Aussagen können jedoch auf stetige Zustandsräume übertragen werden. Eine ausführliche Darstellung zu Markovketten mit stetigem Zustandsraum findet man bei Tierney (1994) [90].

Markovketten sind stochastische Prozesse, also Verläufe in der Zeit, für die die Markoveigenschaft gilt: "Der zukünftige Prozessverlauf hängt bei bekanntem gegenwärtigem Wert nicht vom vorangegangenen Prozessverlauf ab", siehe Fahrmeir (1981).

$X^{(t)}$  sei ein stochastischer Prozess mit  $t \in \mathbb{N}_0$ , dann ist  $X^{(t)}$  eine Markovkette, falls für alle Zeitpunkte  $t = 0, \dots, n$  und eine beliebige Teilmenge  $A$  aus der  $\sigma$ -Algebra des Zustandsraums  $E$ :

$$Pr[X^{(t)} \in A \mid X^{(0)}, X^{(1)}, \dots, X^{(t-1)}] = Pr[X^{(t)} \in A \mid X^{(t-1)}]$$

gilt.

Eine Markovkette wird als homogen bezeichnet, wenn die bedingten Wahrscheinlichkeiten unabhängig vom Zeitpunkt  $t$  sind, also wenn gilt:

$$Pr[X^{(t+1)} = y \mid X^{(t)} = x] = Pr[X^{(1)} = y \mid X^{(0)} = x] \quad t = 1, \dots, n.$$

Befindet sich die Kette zum Zeitpunkt  $t \geq 0$  im Zustand  $x \in \mathbb{R}^k$ , so gibt

$$P(x, A) = Pr[X^{(t+1)} \in A \mid X^{(t)} = x] \quad (3.3)$$

die Wahrscheinlichkeit an, sich zum Zeitpunkt  $t + 1$  in der Menge  $A \in E$  zu befinden.

Für den Fall eines diskreten Zustandsraums  $E$  ergibt sich die Übergangswahrscheinlichkeit

$$P(x, y) = Pr[X^{(t+1)} = y \mid X^{(t)} = x] = P_{xy}$$

analog zu (3.3). Diese Übergangswahrscheinlichkeiten lassen sich nun zusammenfassen zu einer Übergangsmatrix  $P$ . Für die Übergangswahrscheinlichkeiten gilt:  $P(x, y) \geq 0 \forall x, y \in E$ , sowie  $\sum_{y \in E} P(x, y) = 1 \forall x \in E$ . Daraus ergibt sich die Verteilungsfunktion  $P(x, \cdot)$  der Markovkette von  $x \in E$  ausgehend.

Die Zeit, welche die Markovkette ausgehend vom Zustand  $i$  benötigt, um zum Zustand  $i$  zurückzukehren, wird durch  $\tau_{xx} = \min_{t > 0} \{X^{(t)} = x \mid X^{(0)} = x\}$  gemessen.

Die Wahrscheinlichkeit, dass der Zustand  $y$ , ausgehend vom Zustand  $x$ , in exakt  $n$  Schritten erreicht wird, lässt sich formulieren, als:

$$P^n(x, y) = Pr[X^{(n)} = y \mid X^{(0)} = x].$$



Man beachte dabei die Homogenität der Kette. Mittels der mit Hilfe der Chapman-Kolmogorov Gleichung kann man nun

$$P^{t+n}(x, y) = \sum_k P^t(x, k)P^n(k, y)$$

definieren. Damit lässt sich also auch die Verteilung der Kette nach  $t$  Schritten bestimmen, im diskreten Fall ergibt sie sich als Vektor der  $t$ -ten Zeile der Übergangsmatrix. Die Verteilung  $\pi^t$  der Markovkette nach  $t$  Iterationen ist gegeben durch

$$\begin{aligned} \pi^t(y) &= Pr[X^{(t)} = y] \\ &= \sum_{x \in E} Pr[X^{(t)} = y \mid X^{(0)} = x] Pr[X^{(0)} = x] \\ &= \sum_{x \in E} P^t(x, y) \pi^0(x). \end{aligned}$$

Oder in Matrixschreibweise  $\pi^n = \pi^0 P^n$ . Dies verdeutlicht auch, dass die Suche nach der stationären Verteilung der Bestimmung der Übergangswahrscheinlichkeiten entspricht. Von höchstem Interesse ist nun das asymptotische Verhalten der Übergangsmatrix für  $t \rightarrow \infty$ , insbesondere ob die Kette nach ausreichender Anzahl von Kettengliedern einen Zustand erreicht, der stationär ist. Die Verteilung  $\pi$  heißt stationär oder invariant, falls sie

$$\sum_{x \in E} \pi(x) P(x, dy) = \Pi(dy), \text{ bzw. } \int_{-\infty}^{\infty} \pi(x) P(x, dy) dx = \Pi(dy),$$

für alle  $y \in E$  erfüllt, in Matrixschreibweise ergibt sich  $\Pi = \Pi \cdot P$ . Sobald die Verteilung der Markovkette "Π" entspricht, besitzt jeder nachfolgende Zustand der Kette die stationäre Verteilung Π.

Damit die Verteilung von  $X^t$  gegen eine stationäre Verteilung konvergiert, muss die Markovkette zusätzlich noch drei Bedingungen erfüllen. Sie muss irreduzibel, aperiodisch und positiv rekurrent sein. Diese drei Bedingungen werden im Folgenden kurz erläutert.

### Irreduzibilität

Ist der Zustand  $y$  von  $x$  aus erreichbar, und gilt dies auch umgekehrt, so spricht man von zwei kommunizierenden Zuständen, oder die Zustände  $y$  und  $x$  sind kommunizierend. Sind sowohl die Zustände  $x$  und  $y$ , als auch die Zustände  $y$  und  $z$  gegenseitig kommunizierend, dann sind auch die Zustände  $x$  und  $z$  kommunizierend, also gegenseitig erreichbar. Eine homogene Markovkette  $X^{(t)}$  heißt irreduzibel, wenn jeder Punkt von

jedem Punkt aus mit einer positiven Wahrscheinlichkeit erreicht werden kann, also für alle  $x, y \in E$  ein  $t > 0$  existiert, für das  $P_{xy}(t) > 0$  gilt. Man kann also auch sagen eine Markovkette ist irreduzibel, wenn alle Zustände kommunizierend sind.

### Aperiodizität

Als aperiodisch bezeichnet man eine Markovkette, wenn kein wiederkehrendes Muster besteht, mit dem die Markovkette bestimmte Zustände oder Zustandsmengen annimmt. Das bedeutet man kann für ein  $x$  zeigen, dass für den größten gemeinsamen Teiler:  $ggT_{t>0}\{P_{xx}(t) > 0\} = 1$  gilt. Der größte gemeinsame Teiler wird gelegentlich auch Periode genannt. Ist die Periode  $ggT > 1$ , so nennt man die Kette periodisch. Ist eine Markovkette periodisch, so ist hierbei nicht entscheidend mit welchem Muster die Kette bestimmte Zustände annehmen kann, sondern die Tatsache, dass sie zu bestimmten Zeitpunkten bestimmte Zustände nicht annehmen kann. Die Eigenschaft der Aperiodizität ist notwendig für die Konvergenz der Übergangswahrscheinlichkeiten. Es kann gezeigt werden, dass die verschiedenen Zustände einer irreduziblen Markovkette alle die gleiche Periode haben.

### Rekurrenz und Transienz

Die Markovkette  $X^{(t)}$  heißt positiv rekurrent, sofern jeder mögliche Zustand mit einer positiven Wahrscheinlichkeit unendlich oft erreicht werden kann, und damit

$$E[\tau_{ii}] < \infty \quad (3.4)$$

ist. Erfüllt eine Kette die Bedingung:  $Pr[\tau_{ii} < \infty] = \sum_t f_i(t) = 1$ , so nennt man sie rekurrent. Eine rekurrente Kette, welche die Bedingung (3.4) nicht erfüllt, heißt null rekurrent. Kehrt eine Markovkette mit einer positiven Wahrscheinlichkeit nicht zum Zustand  $y$  zurück, so nennt man sie transient. Dies ist bei so genannten absorbierenden Zuständen der Fall. Erreicht die Kette einen absorbierenden Zustand, so behält die Kette diesen Zustand bei, es gilt  $P(y, y) = 1$ . Ein alternativer Weg eine positiv rekurrente Kette zu beschreiben, ergibt sich über

$$\sum_{x \in E} \pi(x) P^t(x, y) = \Pi(y).$$

Da man in der Bayes Statistik die stationäre Verteilung  $\pi$  vorgibt, genügt der Nachweis der Irreduzibilität, um sicherzustellen, dass die Markovkette positiv rekurrent ist. Eine

Markovkette, die aperiodisch und positiv rekurrent ist, nennt man auch ergodisch. Für eine irreduzible Markovkette mit stationärer Verteilung  $\Pi$  und eine reellwertige Funktion  $g$  mit  $E_\pi(|g(Y)|) < \infty$  gilt der Ergodensatz:

$$Pr \left[ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n g(Y_t) = E_\pi(g(Y)) \right] = 1.$$

Damit der in Abschnitt 3.2.2 beschriebene Metropolis-Hastings-Algorithmus zu einer stationären Verteilung führt, muss die Markovketten noch die Umkehrbarkeitsbedingung oder Detailed Balance erfüllen. Diese Bedingung verlangt, dass die Gleichgewichtswahrscheinlichkeit einer Bewegung ausgehend vom Zustand  $x$  hin zum Zustand  $y$  der Wahrscheinlichkeit sich ausgehend vom Zustand  $y$  zum Zustand  $x$  zu bewegen, entspricht, also  $\pi(x)p(x, y) = \pi(y)p(y, x)$ . Einen deutschsprachigen Überblick kann man beispielsweise bei Gesslbauer (1997) [45] oder bei Biller (1999) [12] finden.

Um zu überprüfen, ob die Kette gegen die stationäre Verteilung konvergiert, benötigt man noch ein weiteres Maß. Die Distanz zwischen zwei Verteilungen  $\nu_1$  und  $\nu_2$  wird gemessen durch:

$$\|\nu_1 - \nu_2\| = 2 \sup_{A \subseteq E} |\nu_1(A) - \nu_2(A)|.$$

Erfüllt eine Markovkette die in Abschnitt 3.2.1.1 erläuterten Bedingungen, so ist sie gegen die stationäre Verteilung konvergiert, wenn für  $n \rightarrow \infty$

$$\|P^n(x, \cdot) - \pi\| \rightarrow 0$$

gilt. Im stetigen Fall muss berücksichtigt werden, dass  $P(x, \{y\}) = 0$ .

$$P^{(t)}(x, A) = \int_{\mathbb{R}^k} P^{(t-1)}(x, dy) P(y, A)$$

ist die t-te Iteration des Übergangskerns. Allgemein erhält man eine stationäre Verteilung durch die Bedingung

$$\Pi(A) = \int_{\mathbb{R}^k} P(x, A) \pi(x) dx. \quad (3.5)$$

Für die Suche von geeigneten Übergangswahrscheinlichkeiten wählt man eine Funktion  $p(x, y)$  so, dass für  $x \in \mathbb{R}^k$  und  $y \in \mathbb{R}^k$

$$P(x, dy) = p(x, y)dy + r(x)I_{(x \in dy)}$$

gilt und  $p(x, x) = 0$  ist.  $I$  ist die Indikatorfunktion, d.h.  $I_{(x \in dy)} = 1$ , falls  $x \in dy$  und 0 sonst. Die Funktion  $r(x) = 1 - \int_{\mathbb{R}^k} p(x, y) dy$  bezeichnet die Wahrscheinlichkeit, dass die Markovkette in einem Schritt im Punkt  $x$  verharrt. Erfüllt  $p(x, y)$  außerdem die Umkehrbedingung (3.2.1.1), dann ist  $\pi(\cdot)$  die stationäre Dichtefunktion von  $P(x, \cdot)$ . Um dies zu zeigen, unterzieht man Gleichung (3.5) einer genaueren Betrachtung:

$$\begin{aligned} \int P(x, A)\pi(x) dx &= \int \left[ \int_A p(x, y) dy \right] \pi(x) dx + \int r(x)I_{x \in A}\pi(x) dx \\ &= \int_A \left[ \int p(x, y)\pi(x) dx \right] dy + \int_A r(x)\pi(x) dx \\ &= \int_A \left[ \int p(y, x)\pi(y) dx \right] dy + \int_A r(x)\pi(x) dx \\ &= \int_A \int (1 - r(y))\pi(y) dy + \int_A r(x)\pi(x) dx \\ &= \int_A \pi(y) dy. \end{aligned}$$

Hierbei wird auch die Bedeutung der Bedingung der Detailed Balance deutlich. Man sieht also, dass sich  $\pi$  als stationäre Dichte der Übergangswahrscheinlichkeit ergibt.

### 3.2.1.2 Vorgehensweise

Im Gegensatz zu dem Grundgedanken der Markovtheorie, bei dem nach der stationären Verteilung  $\Pi$  der Markovkette gesucht wird, interessiert den Bayes-Statistiker diejenige Markovkette, deren stationäre Dichte die gesuchte a posteriori Verteilung ist. Gesucht ist also derjenige Übergangskern, der ausgehend von einem Startwert, die Markovkette mit der a posteriori Verteilung der betrachteten Parameter liefert.

Nachdem die Eigenschaften einer Markovkette in Abschnitt 3.2.1.1 dargestellt wurden, soll in diesem Abschnitt gezeigt werden, wie man eine solche Markovkette erzeugt. Um zu einer ausreichend großen Stichprobe aus der gesuchten a posteriori Verteilung zu gelangen, soll eine Markovkette  $X^t$  erzeugt werden, welche als stationäre Verteilung die gesuchte a posteriori Verteilung des Parametervektors  $\theta$  besitzt. Dabei entspricht die Anzahl der Komponenten von  $X_k^t$  den  $k$  Komponenten des Parametervektors  $\theta_k$ . Die weitere Vorgehensweise basiert auf der Konstruktion einer irreduziblen und aperiodischen Markovkette  $X^t$  mit den Kettengliedern  $X^1, X^2, \dots, X^t$ . Wenn  $t$  jetzt hinreichend

groß gewählt wird, so liefert die Markovkette, unter Berücksichtigung der Phase bis zum Erreichen der stationären Verteilung, eine ausreichend große Stichprobe, um sie mit herkömmlichen Mitteln zu untersuchen.

Im Allgemeinen werden die zu Beginn erzeugten Glieder der Markovkette noch nicht nach der stationären Verteilung  $\Pi$  verteilt sein. Es wird daher eine gewisse Anzahl von Gliedern bei der Schätzung vernachlässigt. Man bezeichnet dies als burn-in Phase, in der die Markovkette in die stationäre Verteilung einschwingt. Die Abschnitte 3.2.4 und 3.2.5 befassen sich noch näher mit diesem Thema.

Es existieren nun verschiedene Möglichkeiten einen geeigneten Übergangskern zu konstruieren. Dabei ist besonders der Metropolis-Hastings-Algorithmus zu nennen, dieser wird in Abschnitt 3.2.2 beschrieben. Aus diesem Algorithmus ergeben sich viele weitere Möglichkeiten zur Bestimmung eines geeigneten Übergangskerns als Spezialfälle.

### 3.2.2 Der Metropolis-Hastings-Algorithmus

Mit Hilfe des in diesem Abschnitt vorgestellten Metropolis-Hastings-Algorithmus (MH-Algorithmus), soll der gesuchte Übergangskern ermittelt werden. Dieser basiert auf dem von Metropolis, Rosenbluth, Rosenbluth, Teller und Teller (1953) [74] vorgestellten Algorithmus, der durch Hastings (1970) [54] verallgemeinert wurde. Einen guten Einblick in den Metropolis-Hastings-Algorithmus geben auch Chib und Greenberg (1995) [23].

Beim Metropolis-Hastings-Algorithmus verwendet man eine Dichte  $q(x, y)$ , um von einem Punkt  $x$  ausgehend einen Punkt  $y$  zu generieren. Idealerweise würde die verwendete Vorschlagsdichte die Detailed Balance Eigenschaft bereits erfüllen. Dies ist jedoch meist nicht der Fall, so dass man bspw. auf Ungleichungen der Form:

$$\pi(x)q(x, y) > \pi(y)q(y, x) \quad (3.6)$$

trifft. In diesem Fall würde die Kette mit einer höheren Wahrscheinlichkeit von Punkt  $x$  zu Punkt  $y$  springen als umgekehrt. Damit gelangt man nicht mehr zu der gesuchten stationären Verteilung. Aus diesem Grund wird eine Wahrscheinlichkeit eingeführt, mit der der Schritt unternommen wird. Man spricht hierbei von der Akzeptanzwahrscheinlichkeit  $\alpha(x, y)$ . Dies bedeutet für die Kette, dass die Kette weiterhin in dem letzten Zustand verharrt, falls der Schritt abgelehnt wird. In diesem Punkt unterscheiden sich die Verfahren ohne Akzeptanzwahrscheinlichkeit, hier werden solange Kandidaten gezo-

gen, bis ein neuer Kandidat akzeptiert wird. Die Kette bleibt also nie im gleichen Zustand.

Betrachtet man nun die Ungleichung (3.6), so wird der Schritt von  $y$  nach  $x$  nicht oft genug gemacht. Die Wahrscheinlichkeit für diesen Schritt wird nun maximiert, d.h man setzt  $\alpha(y, x) = 1$ . Fordert man zusätzlich, dass die Detailed Balance Bedingung erfüllt wird, so wird die Wahrscheinlichkeit  $\alpha(x, y)$  durch

$$\begin{aligned}\pi(x)q(x, y)\alpha(x, y) &= \pi(y)q(y, x)\alpha(y, x) \\ &= \pi(y)q(y, x)\end{aligned}$$

beschrieben. Die Akzeptanzwahrscheinlichkeit ergibt sich also als:

$$\alpha(x, y) = \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}.$$

Hierbei kürzt sich offensichtlich die Normierungskonstante der a posteriori Verteilung raus. Dies ist der Grund, warum in der Praxis oft vernachlässigt wird, die Existenz der Normierungskonstante zu überprüfen.

Die Übergangswahrscheinlichkeit der simulierten Kette lässt sich folgendermaßen beschreiben:

$$p_{yx} = q(x | y)\alpha(x | y) + I(x = y)r(y).$$

Dabei ist die Indikatorfunktion und  $r(y)$  die Wahrscheinlichkeit im Zustand  $y$  zu bleiben, wenn ein Kandidat  $x \neq y$  gezogen wurde. Aus der Definition der Akzeptanzwahrscheinlichkeit folgt die Eigenschaft

$$\pi(y)\alpha(x | y) = \pi(x)\alpha(y | x)$$

und, wegen der Symmetrie der Vorschlagsdichte

$$\pi(y)q(x | y)\alpha(x | y) = \pi(x)q(y | x)\alpha(y | x).$$

Daraus ergibt sich die Eigenschaft

$$\pi(y)p_{yx} = \pi(x)p_{xy},$$

die Detailed Balance genannt wird. Durch Summation über  $y$  auf beiden Seiten erhält man

$$\sum_y \pi(y) p_{yx} = \pi(x) \text{ für alle } x,$$

oder in Matrixschreibweise

$$\Pi \cdot P = \Pi.$$

Um zu zeigen, dass  $P^{(t)}$  wirklich gegen  $\Pi$  konvergiert, müssen noch Irreduzierbarkeit und Aperiodizität gezeigt werden. Diese hängen jedoch von der Wahl der Vorschlagsdichte ab. In der praktischen Anwendung ist diese Frage jedoch unproblematisch.

D.h. man erzeugt eine Markovkette der Länge  $N$ , indem man einen Kandidaten aus der Vorschlagsdichte  $q(X_t, \cdot)$  zieht, und die Akzeptanzwahrscheinlichkeit  $\alpha(x, y) = \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}$  für diesen Wert mit dem Wert aus der Gleichverteilung  $(0,1)$  vergleicht. Ist der Wert für die Akzeptanzwahrscheinlichkeit größer 1, ist also der Kandidat  $Y$  "wahrscheinlicher" als der momentane Zustand  $X^t$ , dann wird der Wert mit Sicherheit angenommen. Ist der Quotient kleiner 1, so wird er mit einem Wert aus der Gleichverteilung verglichen und anschließend akzeptiert, sofern der Quotient den Wert aus der Gleichverteilung übersteigt.

### Metropolis-Hastings-Algorithmus

```

Initialisiere  $X^0$ ; setze  $t=0$ 
Solange  $t < N$ 
{
  Erzeuge  $Y$  aus  $q(X^t, Y)$ 
  Erzeuge  $U$  aus  $U(0,1)$ 
  Falls  $U \leq \alpha(X^t, Y)$  setze  $X^{t+1} = Y$ 
  Sonst setze  $X^{t+1} = X^t$ 
  Setze  $t=t+1$ 
}

```

### 3.2.3 Spezialfälle

Für den Metropolis-Hastings-Algorithmus existieren unterschiedliche Varianten, die sich hinsichtlich der Wahl der Vorschlagsdichte unterscheiden. Eine Übersicht über die

unterschiedliche Wahl der Vorschlagsdichte findet sich beispielsweise in Geweke und Tanizaki (2003) [47]. Im Folgenden werden die wichtigsten Varianten kurz erläutert, insbesondere der Metropolis-Algorithmus, der in den Simulationen in Kapitel 4 verwendet wird.

### Metropolis-Algorithmus

In der praktischen Anwendung spielt der häufig verwendete Metropolis-Algorithmus (vgl. Metropolis *et al.*, 1953 [74]) eine große Rolle. Er ist ein Spezialfall des Metropolis-Hastings-Algorithmus und ergibt sich, wenn die Vorschlagsdichte symmetrisch gewählt wird, d.h.  $q(x, y) = q(y, x)$ . Dies ist sehr häufig der Fall, da in sehr vielen Anwendungen die Kandidaten aus einer Normalverteilung oder aber aus einer t-Verteilung gezogen werden. Durch die Symmetrie vereinfacht sich der Term der Akzeptanzwahrscheinlichkeit zu

$$\alpha(X, Y) = \min \left\{ \frac{\pi(Y)}{\pi(X)}, 1 \right\}.$$

In dieser Arbeit werden die Kandidaten immer aus einer Normalverteilung gezogen, somit handelt es sich stets um den Metropolis-Algorithmus. Sprachlich wird häufig nicht zwischen Metropolis-Algorithmus und Metropolis-Hastings-Algorithmus unterschieden.

### Random-Walk-Metropolis-Algorithmus

Der Random-Walk-Metropolis-Algorithmus wiederum ist ein Spezialfall des Metropolis-Algorithmus. Er ergibt sich, wenn die beim Metropolis-Algorithmus verwendete symmetrische Vorschlagsdichte für die Differenz des letzten Kettenglieds zum Kandidaten verwendet wird. Wenn also  $q(x, y) = q(|x - y|)$ , dann spricht man vom Random-Walk-Metropolis-Algorithmus. Es wird also nicht der Kandidat selbst gezogen, sondern seine Abweichung zum letzten Kettenglied.

### Aufbau der Vorschlagsdichte

Werden alle einzelnen Parameter  $\theta_i$ , für  $i = 1, 2, \dots, k$  des Parametervektors  $\theta$  in jeder Iteration in einem Schritt gemeinsam erneuert, so spricht man auch von einem „simultaneous-vector-update“. Dafür benötigt man eine k-dimensionale Vorschlagsdichte aus der der Kandidat für  $\theta$  gezogen wird.

Alternativ dazu besteht die Möglichkeit den Parametervektor  $\theta$  zu gruppieren und in



$h \leq k$  einzelne Blöcke  $\{\theta_{.1}, \theta_{.2}, \dots, \theta_{.h}\}$  aufzuteilen. Dann können die Kandidaten blockweise aus unterschiedlichen Vorschlagsdichten gezogen werden. Bei jeder Iteration wird für jeden Block der Wert einzeln mit einem Metropolis-Hastings-Schritt aktualisiert, man benötigt also  $h$  unabhängige Metropolis-Schritte. Man kann nun bspw.  $h$  Random-Walk-Ketten verwenden und für jeden der  $h$  Blöcke die Kandidaten aus einer entsprechenden Vorschlagsdichte  $N(0, \Sigma)$  ziehen.

Außerdem kann für jeden Parameter  $\theta_i$ , für  $i = 1, 2, \dots, k$  eine eigene Vorschlagsdichte gewählt werden. Man spricht dann vom Einzelkomponenten-Metropolis-Hastings-Algorithmus von Metropolis et al. (1953) [74] oder auch vom „single-component-update“. Weitere Erläuterungen zu diesem Algorithmus findet man bei Tierney (1994) [90] oder auch bei Chib und Greenberg (1995) [23].

Der Vorteil des „simultaneous-vector-update“-Verfahrens besteht klar in der Möglichkeit Abhängigkeitsstrukturen zwischen den einzelnen Parametern abzubilden und für die Erzeugung neuer Kandidaten zu berücksichtigen. Dies führt dazu, dass die burn-in Phase verkürzt werden kann und die Anzahl „ungünstiger“ Kandidaten sinkt, was die Akzeptanzrate erhöht und die Rechenzeit senkt, vgl. dazu Vrontos et al. (2000) [94].

Wird der Kandidat bspw. aus einer multivariaten Normalverteilung gezogen und werden dabei keine Kovarianzen berücksichtigt, so entspricht dieses Vorgehen dem Vorgehen beim „single-component-update“-Verfahren. Ist die Abhängigkeitsstruktur der Parameter unbekannt, so kann auf diese Weise eine Markovkette erzeugt werden, mit deren Hilfe man eine geeignete Kovarianzmatrix  $\Sigma$  schätzen kann, mit der dann neue Kandidaten erzeugt werden.

### Independence-Sampler

Wird die Vorschlagsdichte beim Metropolis-Hastings-Algorithmus so gewählt, dass der neue Zustand unabhängig vom letzten Zustand ist, also  $q(Y, X) = q(Y)$ , erhält man den sog. Independence-Sampler von Tierney (1994) [90]. Hierbei muss die Akzeptanzwahrscheinlichkeit

$$\alpha(X, Y) = \min \left\{ 1, \frac{\pi(Y)q(Y)}{\pi(X)q(X)} \right\}$$

berücksichtigt werden. Beim Independence-Sampler ist die Wahl der Vorschlagsdichte besonders entscheidend für die Güte der Simulation.

### 3.2.4 Mischung und Konvergenz

Wurde eine Markovkette  $X^t, t = 1, \dots, N$  mit den beschriebenen Methoden erzeugt, so sind für die Beurteilung der Qualität der erzeugten Kette die Begriffe Mischung und Konvergenz von großer Bedeutung. Um die a posteriori Verteilung auswerten zu können, benötigt man eine ausreichend große Anzahl von Kettengliedern nach Erreichen des stationären Zustands, d.h. nachdem die Markovkette gegen die stationäre Verteilung konvergiert ist. Die Phase bis die Markovkette gegen die stationäre Verteilung konvergiert ist, nennt man, wie oben schon genannt, burn-in Phase. Sie ist entscheidend für die Wahl der Gesamtlänge der Kette. Für die Auswertung der Markovkette werden die  $N$  Kettenglieder abzüglich der burn-in Phase berücksichtigt.  $N$  muss also ausreichend groß gewählt werden, so dass die Genauigkeit des Schätzers  $\bar{f}_N$  ein bestimmtes Niveau erreicht.

Darüber hinaus ist die Mischung einer konvergenten Kette entscheidend. Bei Methoden zur Bildung der Markovkette, welche eine Akzeptanzwahrscheinlichkeit verwenden, gibt es Züge für die Kettenglieder, die abgelehnt werden. Geschieht dies sehr oft, d.h. verweilt die Kette sehr oft bzw. sehr lange in dem gleichen Zustand, so ist die Mischung der Kette schlecht. Dies ist gleichbedeutend mit einem niedrigen Anteil akzeptierter Kandidaten und somit einer niedrigen Akzeptanzrate. Ist die Akzeptanzrate hoch, jedoch aufeinander folgende Zustände  $X^t$  und  $X^{t+1}$  stark korreliert, spricht man ebenfalls von einer schlechten Mischung. Um eine schlechte Mischung zu vermeiden, ist damit sowohl auf die Wahl der Vorschlagsdichte  $q(x,y)$ , als auch auf die Parametrisierung der Ausgangsdaten zu achten.

Die Konvergenz einer Markovkette ist unabhängig von ihrer Mischung zu betrachten, obwohl eine schlechte Mischung die Konvergenzgeschwindigkeit negativ beeinflusst.

Ein besonderes Problem bezüglich der Konvergenz stellen jene stationäre Verteilungen dar, die mehrere Modi aufweisen. Von entsprechenden konvergenten Markovketten erwartet man, dass deren Zustände nicht nur um einen einzigen Modus konzentriert sind, sondern jeder Extremwert der stationären Verteilung regelmäßig von der Kette besucht wird.

### 3.2.5 Konvergenzanalyse

Man unterscheidet zwischen qualitativen und quantitativen Analysemethoden. Einen guten Überblick über Konvergenzanalysetechniken (wie bspw. den Test auf Konvergenz von Heidelberger und Welch) liefern Cowles und Carlin (1996) [28].

Im Allgemeinen wird am Anfang der Simulation der Startwert einen großen Einfluss haben, sofern der Startwert nicht schon aus der stationären Verteilung stammt. Bei mehrdimensionalen Markovketten sind dabei alle Startwerte sowie die jeweiligen Ketten zu berücksichtigen. Befindet sich die Kette nicht schon zu Beginn der Simulation in der stationären Verteilung, so werden einige Glieder  $X^0, X^1, \dots, X^m$  benötigt, bis die Markovkette die stationäre Verteilung erreicht. Die Konvergenzanalyse hat nun die Aufgabe zu ermitteln, wie groß die burn-in Phase zu wählen ist, d.h. wie viele Iterationen bei den Berechnungen der empirischen Erwartungswerte unberücksichtigt bleiben. Dazu muss berücksichtigt werden, dass die Genauigkeit der Analyse mit der Anzahl der verwendeten Iterationen steigt, so dass die Konvergenzanalyse gleichzeitig einen Hinweis darauf gibt, wie viele Iterationen für eine Berechnung mit einer bestimmten Genauigkeit notwendig sind. Mit der steigenden Leistungsfähigkeit von Computern, die die Erzeugung langer Ketten begünstigen, besteht die Möglichkeit die burn-in Phase großzügig zu wählen. Damit sinkt die Notwendigkeit einer exakten Bestimmung der Länge der burn-in Phase.

Oft erfolgt die Analyse nur durch eine optische Inspektion der Kette. Der Abschnitt 4.3.1 beschäftigt sich unter anderem mit der burn-in Phase. Es zeigt sich, dass die in dieser Arbeit durchgeführten MCMC-Schätzungen hinsichtlich der Konvergenz völlig unproblematisch sind. Eine optische Inspektion der Markovkette hat sich für die in dieser Arbeit betrachteten Markovketten als absolut ausreichend erwiesen. Daher werden hier weitere Möglichkeiten der Analyse des Konvergenzverhaltens nicht thematisiert.

### 3.3 Modellselektion

In der Statistik erfolgt eine Modellselektion, also die Wahl des "besten" Modells, meist mittels Größen wie dem Akaike Informationskriterium (AIC) oder dem Bayes Informationskriterium (BIC). Gesucht wird das Modell, welches die Zeitreihe am geeignetsten modelliert. Daher ist es notwendig sich auf ein "bestes" Modell festzulegen. Diese Einschränkung erlaubt es nicht die Vorteile verschiedener Modelle zu kombinieren. In Abschnitt 3.4 wird mittels Model Averaging ein Möglichkeit beschrieben diesem Problem zu begegnen.

$M$  sei ein Set konkurrierender Modelle für gegebene Werte der Daten  $y$ . Das Modell  $m \in M$  hat einen unbekannt Parametervektor  $\theta_m \in \Theta_m$ . Es ist wichtig zu berücksichtigen, dass die Dimensionen der verschiedenen Parametervektoren von Modell zu Modell variieren können. Die a posteriori Wahrscheinlichkeit von Modell  $m$  ist gegeben durch

$$\pi(m | Y) = \frac{\pi(m) \int_{\Theta_m} \pi(y | m, \theta_m) \pi(\theta_m | m) d\theta_m}{\sum_{m \in \mathcal{M}} \pi(m) \int_{\Theta_m} \pi(y | m, \theta_m) \pi(\theta_m | m) d\theta_m},$$

wobei  $\pi(y | m, \theta_m)$  die Likelihood gegeben Modell  $m$  ist, der Parametervektor  $\theta_m$ ,  $\pi(m)$  ist die a priori Wahrscheinlichkeit für das Modell  $m$ , und  $\pi(\theta_m)$  ist die a priori Verteilung für den Parametervektor  $\theta_m$  gegeben das Modell  $m$ .

Die sog. Modell-Likelihood

$$\pi(Y | m_i) = \int_{\Theta_{m_i}} \pi(y | m_i, \theta_{m_i}) \pi(\theta_{m_i} | m_i) d\theta_{m_i}$$

ist offensichtlich die Normierungskonstante für das jeweilige Modell.

Der Bayes-Ansatz ermöglicht es eine Entscheidungen bezüglich der Modellselektion mit Hilfe des sog. Bayes-Faktors zu treffen, welcher in Abschnitt 3.3.1 erläutert wird. Mit dessen Hilfe erfolgt ein Vergleich des Modells  $m_i$  gegen das Modell  $m_j$ .

### 3.3.1 Bayes-Faktor

In diesem Abschnitt wird das Konzept des Bayes-Faktors (BF) gezeigt, mit dem in der Bayes-Statistik Modellselektion betrieben werden kann. Der Bayes-Faktor hat folgende Form:

$$BF = \frac{\pi(m_i | y) \pi(m_j)}{\pi(m_j | y) \pi(m_i)} = \frac{\int_{\Theta_{m_i}} \pi(y | m_i, \theta_{m_i}) \pi(\theta_{m_i} | m_i) d\theta_{m_i}}{\int_{\Theta_{m_j}} \pi(y | m_j, \theta_{m_j}) \pi(\theta_{m_j} | m_j) d\theta_{m_j}}.$$

Gegenüber dem AIC und dem BIC hat der BF laut Kass und Raftery (1995) den Vorteil, dass er nicht nur asymptotisch gilt und keinen subjektiven Teil enthält.

Mit Hilfe der Modell-Likelihoods können die a posteriori Wahrscheinlichkeiten für die Modelle durch

$$\pi(m_i | Y) = \frac{p(Y | m_i) \cdot \pi(m_i)}{\sum_{k=1}^n p(Y | m_k) \cdot \pi(m_k)}$$

berechnet werden. Im Folgenden werden einige Möglichkeiten demonstriert, um den Ausdruck  $p(Y | m_i)$  zu schätzen, vgl. dazu Miazhynskaia et al. (2006) [75].

### 3.3.2 Harmonic Mean Estimator

Die einfachste Möglichkeit  $\pi(Y | m_i)$  zu schätzen, stellt sich die Berechnung des arithmetischen Mittels von Zufallszügen aus der Likelihood-Verteilung dar,

$$\hat{\pi}(Y | m_i) = \frac{1}{G} \sum_{g=1}^G f(y | \theta^{(g)}, m_i),$$

wobei  $\theta^{(g)}$  aus den a priori Verteilungen  $\pi(\theta | m_i)$  stammt. Allerdings ist dieser Schätzer sehr ineffizient, da viele Werte in der Summe nahe 0 sind.

Der Harmonic Mean Estimator berechnet das harmonische Mittel von Zufallszügen aus der Likelihood-Verteilung,

$$\hat{\pi}_{HM}(Y | m_i) = \left[ \frac{1}{M} \sum_{m=1}^M \frac{1}{f(y | \theta^{(m_i)}, m_i)} \right]^{-1},$$

wobei  $\theta^{(m_i)}$  jetzt aus den a posteriori Verteilungen  $\pi(\theta | Y, m_i)$  stammt. Eine gravierende Einschränkung dieser Methode ist, dass, ebenso wie beim arithmetischen Mittel, die a posteriori Verteilung Werte nahe 0 annimmt.

### 3.3.3 Reciprocal Importance Estimator

Von Gelfand und Dey (1994) [42] stammt der Vorschlag des Reciprocal Importance Estimator.

$$\hat{\pi}_{RI}(Y | m_i) = \frac{1}{M} \sum_{m=1}^M \frac{h(\theta^{m_i})}{f(y | \theta^{(m_i)})\pi(\theta^{(m_i)})},$$

wobei  $\theta^{(m_i)}$  aus den a posteriori Verteilungen stammt. Der Schätzer reagiert sensibel auf die Funktion  $h$ , sie sollte der a posteriori Verteilung möglichst ähnlich sein. Gelfand und Dey empfehlen die Verwendung einer multivariaten Normalverteilung oder multivariaten  $t$ -Verteilung, wobei Mittelwert und Kovarianz aus der a posteriori Verteilung geschätzt werden.

### 3.3.4 Chib' Candidate's Estimator

Ein eleganter Vorschlag zur Bestimmung der Modell-Likelihood stammt von Chib (1995) [24]. Die Methode verwendet Ergebnisse aus einem Gibbs-Sampler und wurde 2001 für die Anwendung in Verbindung mit dem Metropolis-Hasting-Algorithmus von Chib und Jeliazkov [25] erweitert. Als Ausgangspunkt dient der Zusammenhang

$$\pi(Y | m_i) = \frac{f(Y | \theta, m_i) \cdot \pi(\theta | m_i)}{\pi(\theta | Y, m_i)}.$$

Da nur der Ausdruck  $\pi(\theta | Y, M_i)$  unbekannt ist, nutzt man die Tatsache aus, dass der Zusammenhang für jeden Wert für  $\theta$  gilt. Als günstig erscheint dabei die Verwendung des Modus der a posteriori Verteilung. Die Berechnungen sind sehr einfach für den Fall, dass  $\theta$  in einem Block simuliert wird. Chib und Jeliazkov behandeln ebenfalls die Möglichkeit  $\theta$  in mehrere Blöcke zu unterteilen.

Anstatt für jedes Modell die Modell-Likelihood separat zu berechnen, kann der Indikator des Modells auch in den Simulationsalgorithmus integriert werden, dies geschieht z.B. bei der Reversible-Jump Markov Chain Monte Carlo-Methode.

### 3.3.5 Reversible-Jump MCMC

Auf dem Metropolis-Hastings Ansatz beruhend, stellte Green (1995) [51] die Reversible-Jump MCMC-Methode vor, um damit die gemeinsame a posteriori  $\pi(m, \theta_m | y)$  zu generieren.

Bei der Reversible-Jump MCMC-Methode bewegt sich die konstruierte Markovkette innerhalb und zwischen den Modellen, so dass die Wahrscheinlichkeit eines Besuchs eines gegebenen Modells die gesuchte  $\pi(m | y)$  ist. Angenommen der momentane Zustand der Markovkette zum Zeitpunkt  $t$  sei  $(m, \theta_m)$ , wobei  $\theta_m$  die Dimension  $d(\theta_m)$  hat, dann wird ein Sprung zum Zeitpunkt  $t + 1$  zu einem neuen Modell  $m'$  mit der Wahrscheinlichkeit  $j(m, m')$  und dem entsprechenden Parametervektor  $\theta'_{m'}$  vorgeschlagen. Dann wird ein Vektor  $u$  aus einer spezifizierten Vorschlagsdichte  $q(u | \theta_m, m, m')$  generiert, und  $(\theta'_{m'}, u') = g_{m,m'}(\theta_m, u)$  gesetzt für eine spezifizierte, invertierbare Funktion  $g_{m,m'}$ , so dass  $g'_{m',m} = g_{m,m'}^{-1}$ . Beachte, dass  $d(\theta_m) + d(u) = d(\theta'_{m'}) + d(u')$ . Green zeigte, dass wenn der Sprung als nächste Realisation der Markovkette mit der Wahrscheinlichkeit  $\alpha = \min\{1, r\}$  akzeptiert wird, wobei  $r = \frac{\pi(y|m',\theta'_{m'})\pi(\theta'_{m'}|m')\pi(m')j(m',m)q(u'|\theta'_{m'},m',m)}{\pi(y|m,\theta_m)\pi(\theta_m|m)\pi(m)j(m,m')q(u|\theta_m,m,m')}$  mit  $J = \partial(\theta'_{m'}, u')/\partial(\theta_m, u)$  dem Jacobian Term der Transformation, dann erfüllt die Kette die Bedingung der Detailed Balance und hat als stationäre Verteilung  $\pi(m, \theta_m | y)$ .

## 3.4 Theorie des Model Averaging

Neben der Modellierung der bedingten Varianzen ist die Prognose der zukünftigen Volatilität von besonderem Interesse bei der praktischen Anwendung der in Kapitel 2 beschriebenen GARCH-Modelle. Der Anwender ist dabei mit dem Problem konfrontiert sich auf ein einzelnes Modell aus einer Auswahl verschiedener konkurrierender Modelle

festlegen zu müssen. Wie oben beschrieben, existieren verschiedene Möglichkeiten, das "beste" Modell zu bestimmen, also das einzelne Modell, welches bedingten Varianzen am besten modelliert und somit zu den besten Prognosewerten führt.

Bei diesem Vorgehen wird unterstellt, dass es ein Modell gibt, welches allein zu den besten Prognosewerten führt. Es wird dabei vernachlässigt, dass andere Modelle helfen könnten die bedingten Varianzen zu modellieren, dass bspw. einzelne Modelle es ermöglichen könnten bestimmte Aspekte, wie z.B. die Asymmetrie, mit in die Untersuchung einfließen zu lassen. Das Prinzip des Model Averaging versucht die Unsicherheit bezüglich des "wahren" Modells mit in die Schlussfolgerungen einfließen zu lassen. Bei der Untersuchung jedes einzelnen Modells wird automatisch die jeweilige a posteriori Wahrscheinlichkeit mitbestimmt. Anstatt das einzelne "beste" Modell auszuwählen und dann Schlussfolgerungen zu treffen, unter der Annahme das gewählte Modell sei das "wahre" Modell, werden beim Model Averaging Ansatz mehrere Modelle verwendet, die eine gewichtete Prognose liefern. Angenommen man ist interessiert an  $\sigma_{T+1}^2$ , also der prognostizierten Volatilität zum Zeitpunkt T+1, dann ist die a posteriori Verteilung gegeben die Daten  $y$   $\pi(\sigma_{T+1}^2 | y) = \sum_{m \in \mathcal{M}} \pi(\sigma_{T+1}^2 | m, y) \pi(m | y)$ , was ein Durchschnitt der a posteriori Prädiktivverteilungen unter jedem Modell ist, gewichtet mit seiner a posteriori Modellwahrscheinlichkeit. Ausdrücklich soll darauf hingewiesen werden, dass das Model Averaging ausschließlich für Prognosen verwendet wird.

Zuerst wird für ein gegebenes Modell  $m$  die a posteriori Verteilung von  $\pi(\sigma_{T+1}^2 | m, y)$  ermittelt, indem man für jedes Kettenglied in  $\theta$  die Varianzen  $\sigma_0^2, \sigma_1^2, \dots, \sigma_{T+1}^2$  berechnet. Dann erhält man eine Stichprobe von  $\pi(\sigma_{T+1}^2 | y)$ , indem man jeden ermittelten Wert unter Modell  $m$  mit der Wahrscheinlichkeit  $\pi(m | y)$  akzeptiert. Also, wird jeder Wert von  $\pi(\sigma_{T+1}^2 | y)$  errechnet, indem man alle Züge von  $\pi(\sigma_{T+1}^2 | m, y)$  mit  $\pi(m | y)$  gewichtet. Vrontos et al. [94] betonen, dass diese Methode ausschließlich zur Prognose verwendet werden sollte und nicht, um den datengenerierenden Prozess zu verstehen.

## Kapitel 4

# Ergebnisse der Simulationsstudien

In diesem Kapitel werden die Ergebnisse von Simulationen präsentiert, welche die in Kapitel 2 vorgestellten GARCH-Modelle mit dem Bayes-Ansatz aus Kapitel 3 verbinden. In Abschnitt 4.3 wird zunächst anhand eines simulierten Datensatzes die Durchführung der Bayes-Schätzung eines GARCH(1,1)-Modells durch MCMC-Methoden demonstriert. Abschnitt 4.1 erläutert dazu detailliert das Vorgehen bei dieser MCMC-Schätzung und den weiteren MCMC-Simulationen in diesem Kapitel. Nachdem durch diese Simulation die grundsätzliche Vorgehensweise veranschaulicht wurde, befasst sich Abschnitt 4.3.1 mit wichtigen Aspekten im Zusammenhang mit der MCMC-Schätzung von GARCH-Modellen. Abschnitt 4.3.2 beschäftigt sich mit der Annahme der asymptotischen Normalverteilung der ML-Schätzer und stellt den Bezug zu der a posteriori Verteilung mittels MCMC-Schätzung her. Dabei steht die Frage im Vordergrund, inwieweit die a posteriori Verteilung geeignet ist, die Verteilung der Parameter, insbesondere bei niedrigen Stichprobenumfängen, abzubilden. Alternative Verteilungsannahmen werden in Abschnitt 4.3.3 behandelt und Abschnitt 4.4 demonstriert die MCMC-Schätzung an den in Abschnitt 2.4 vorgestellten Erweiterungen des GARCH-Modells. Die Prognose mittels der Prädikativverteilung wird in Abschnitt 4.5 gezeigt und Abschnitt 4.6 zeigt die Anwendung der MCMC-Schätzung auf multivariate Datensätze. Im letzten Abschnitt dieses Kapitels wird die MCMC-Schätzung an realen Datensätzen durchgeführt und Prognosen mittels Model Averaging gezeigt.

Sämtliche Berechnungszeiten beziehen sich, soweit angegeben, auf einen Personal Computer mit zweikerniger CPU mit 3GHz und 4GB RAM. Das Betriebssystem ist Microsoft Windows XP.



## 4.1 Ablauf der Schätzung

Es existiert eine breite Palette von Vorschlägen für die Anwendung von MCMC-Methoden auf GARCH-Modelle. Asai (2006) [1] gibt eine kleine Übersicht über bekannte Ansätze mitsamt einer Bewertung. Er behandelt den „Griddy Gibbs sampler“ von Bauwens und Lubrano (1998) [7], den Ansatz von Nakatsuma (2000) mit Verwendung des Metropolis-Hastings-Algorithmus, den „Adaptive rejection Metropolis sampling“ Ansatz von Kim et al. (1998) [62] und den „Acceptance-rejection“ Ansatz von Mitsui und Watanabe (2003) [76]. Auch Takaishi baut mit seinen Arbeiten 2006 [85] und 2008 [86] darauf auf. Die Studie von Asai kommt zu dem Schluss, dass der Acceptance-Rejection/Metropolis-Hastings-Algorithmus den anderen Vorschlägen überlegen ist. So auch dem verbreiteten Ansatz von Nakatsuma und Tsurumi (1996) [78] für ARMA-GARCH-Modelle. Der Nachteil dieses Ansatzes ist, dass die ARMA-Darstellung des GARCH-Modells verfügbar sein muss, was nicht für jede Modellvariante ohne weiteres möglich ist.

Im Folgenden wird die konkrete Vorgehensweise der durchgeführten Simulationen in dieser Arbeit dargestellt.

### Spezifikationen

Für die Erzeugung einer Markovkette, deren stationäre Verteilung die jeweilige gesuchte a posteriori Verteilung ist, werden zunächst Startwerte für die Markovketten sowie Vorschlagsdichten für die Ziehung neuer Kandidaten benötigt. Dabei kommen hier Normalverteilungen zum Einsatz, so dass es sich um symmetrische Vorschlagsdichten handelt. Damit entspricht das Vorgehen, dem in Abschnitt 3.2.3 beschriebenen Metropolis-Algorithmus. Es werden keine Kovarianzen berücksichtigt, d.h. die Kandidaten werden unabhängig voneinander gezogen. Die Kovarianzmatrix  $\Sigma$  enthält also nur auf der Hauptdiagonalen von Null abweichende Werte.

### Ziehung der Kandidaten des ersten Laufs

Aus der Vorschlagsdichte wird dann ein Kandidat für den Parametervektor  $\theta$  gezogen. Dabei enthält  $\theta$  im Falle eines GARCH(1,1)-Modells die für die Varianzgleichung relevanten Parameter ( $\alpha_0$ ,  $\alpha_1$  und  $\beta_1$ ) sowie gegebenenfalls einen Parameter  $\nu$  für den Scharparameter der Verteilung der Residuen.

Mit diesem Kandidat werden die Werte von  $\sigma_t^2$  aus der Varianzgleichung des jeweiligen

GARCH-Modells bestimmt. Dann wird der Wert der Log-Likelihood für diesen Kandidaten berechnet. Anschließend wird zusätzlich die a priori Verteilung berücksichtigt und somit die a posteriori Verteilung bestimmt. Entsprechend dem Metropolis-Algorithmus wird der Quotient der a posteriori Verteilung des Kandidaten und dem Wert der a posteriori Verteilung des letzten Wertes für  $\theta$  berechnet und mit dem Zufallszug einer Gleichverteilung zwischen 0 und 1 verglichen. Ist der Quotient größer als der Wert aus der Gleichverteilung, so wird der Kandidat akzeptiert, andernfalls verbleibt die Markovkette in ihrem letzten Zustand. Ein neuer Kandidat wird aus der Normalverteilung mit unveränderter Varianz aber gegebenenfalls mit dem akzeptierten Kandidaten als neuem Erwartungswert generiert. Auf diese Weise wird eine Kette der gewünschten Länge erzeugt. Das bis hier beschriebene Vorgehen kennzeichnet den ersten Lauf der Simulation.

### **Berücksichtigung der burn-in Phase**

Für die weiteren Berechnungen werden nur die Werte der zweiten Hälfte der bisher erzeugten Markovkette verwendet. D.h. die erste Hälfte wird vollständig verworfen, um dem Einfluss der Startwerte Rechnung zu tragen. Die burn-in Phase beträgt somit immer gerade  $\frac{1}{2}$  Länge der Markovkette des ersten Laufs.

### **Der zweite Lauf**

In Abhängigkeit der Fragestellung ist es durchaus möglich, die Schätzungen nach dem ersten Lauf abzubrechen. Für die Bestimmung der Punktschätzer der Parameter ist der erste Lauf in vielen Fällen ausreichend, falls die Akzeptanzrate nicht zu klein ist; siehe dazu Abschnitt 4.3.1. Allerdings werden in den durchgeführten Simulationen, in Abhängigkeit von den Vorschlagsdichten, sehr niedrige Akzeptanzraten erzielt. Dies ist hier durchaus beabsichtigt, denn aus den erzeugten Markovketten werden die Startwerte und eine geeignete Kovarianzmatrix der Vorschlagsdichte für den zweiten Lauf ermittelt. Die Ergebnisse des ersten Laufs werden anschließend nicht weiter verwendet. Die niedrige Akzeptanzrate des ersten Laufs resultiert aus der Verwendung relativ großer Varianzen für die Vorschlagsdichten des ersten Laufs, dies hat den Vorteil, dass wenig Feingefühl bei der Wahl der Varianzen notwendig ist und die burn-in Phase verkürzt wird.

Der Vorteil geringer Akzeptanzraten ( $< 10\%$  akzeptierte Kandidaten) im ersten Lauf sind meist günstige Schätzungen der Varianzen und Kovarianzen für den zweiten Lauf. Die Ergebnisse liefern meist auch sehr günstige Startwerte. Anschließend werden die Markovketten analog zum Vorgehen im ersten Lauf durchgeführt, allerdings unter Be-

rücksichtigung der Kovarianzen.

Sind die Varianzen der Vorschlagsdichten zu hoch, kommt es zu einer zu hohen Streuung der Kandidaten, was zu einer niedrigen Akzeptanzrate führt. In der Literatur wird daher häufig eine Konstante verwendet, um die Kovarianzmatrix zu manipulieren und damit eine bestimmte Höhe der Akzeptanzrate zu erreichen. Die Erfahrungen aus den Simulationen zeigen, dass eine Manipulation der Kovarianzmatrix bei dem beschriebenen Vorgehen in der Regel nicht notwendig ist.

Die Akzeptanzrate ist eine entscheidende Größe bei der Beurteilung der Qualität der erzeugten Ketten. Daher wird eine „optimale“ Akzeptanzrate im zweiten Lauf angestrebt. Die Angaben zur Höhe dieser „optimalen“ Akzeptanzrate in der Literatur sind sehr uneinheitlich. Während häufig 50% angestrebt werden, empfehlen Gelman et al. (1995) [44] Raten zwischen 25% und 40%. Auch Carlin und Louis (2000) [22] geben diese Spanne an. Daher werden in dieser Arbeit Akzeptanzraten in diesem Bereich angestrebt, wobei dies ausdrücklich nur für den zweiten Lauf der Simulation gilt.

### **Varianten**

Es ist möglich, das Verfahren auszuweiten und iterativ immer neue Kovarianzmatrizen zu erzeugen, vgl. Takaishi (2009) [88]. Allerdings hat es sich gezeigt, dass weitere Läufe nicht notwendig waren, da sie zu keiner weiteren Verbesserung geführt haben. Eine besonders rechenintensive Variante des beschriebenen Vorgehens ist die Neuberechnung der Kovarianzmatrix nach jedem Schritt im zweiten Lauf. Diese Variante wird in Abschnitt 4.3.1 ebenfalls behandelt.

## **4.2 Schätzer**

Die aus dem Metropolis-Hastings-Algorithmus resultierende a posteriori Verteilung wird mit dem Mittelwert, dem Median und dem Modus analysiert. Mit dem Modus erhält man den so genannten Maximum-A-Posteriori-Schätzer (MAP), vergleiche dazu beispielsweise Carlin und Louis (2000) [22]. Die Autoren weisen darauf hin, dass bei der Wahl einer flachen a priori Verteilung der Modus der a posteriori Verteilung dem Schätzer der Maximum-Likelihood-Methode entspricht, dies wird auch an den Ergebnissen der folgenden Simulationen sichtbar. Der MAP erfordert keine Berechnung aus der a posteriori Verteilung, da für die Parameter jedes Zugs die Werte der Log-Likelihood und der a posteriori Verteilung ohnehin berechnet werden. Somit muss nur diejenige Parameterkonstel-

lation ausgewählt werden, die zum Maximum der a posteriori Verteilung geführt hat.

Daneben wird für jede Simulation der Mittelwert und der Median der a posteriori Verteilung der einzelnen Parameter berechnet. Selbstverständlich sind die beiden Schätzer identisch, falls die a posteriori Verteilung symmetrisch ist.

Neben den Ergebnissen der MCMC-Schätzung werden auch die der ML-Schätzungen präsentiert. Zur Berechnung der ML-Schätzer wird die Software Matlab in der Version 7.0.4.365 (R14) mit Service Pack 2 in Verbindung mit der kommerziellen GARCH-Toolbox der Firma Mathworks verwendet (im Folgenden ML-Matlab). Damit lassen sich folgende Modellvarianten schätzen: GARCH, EGARCH und GJR-GARCH. Als Residualverteilung kann man zwischen der Normalverteilung und der  $t$ -Verteilung wählen. Des weiteren werden die nicht kommerziellen Toolboxen UCSD\_Garch (im Folgenden: ML-UCSD) und Oxford MFE (im Folgenden ML-MFE) verwendet<sup>1</sup>. Sie orientieren sich stark an dem Modell von Hentschel aus Abschnitt 2.4.2 und beinhalten eine umfangreiche Auswahl an GARCH-Modellvarianten. Außerdem enthalten beide als Residualverteilung neben der Normalverteilung und der  $t$ -Verteilung die  $GED$ -Verteilung, vgl. dazu Abschnitt 2.3. Die MFE-Toolbox stellt eine Weiterentwicklung der UCSD-Toolbox dar und sollte bevorzugt werden. Allerdings sind bislang keine multivariaten Anwendungen in der MFE-Toolbox verfügbar.

Da für die MCMC-Schätzungen in den folgenden Abschnitten die jeweiligen Berechnungszeiten angegeben sind, sei erwähnt, dass der Zeitaufwand für sämtliche durchgeführten ML-Schätzungen unerheblich war.

---

<sup>1</sup>Beide Toolboxen stammen von Kevin Sheppard von der Oxford Universität und sind über [http://www.kevinsheppard.com/wiki/UCSD\\_GARCH](http://www.kevinsheppard.com/wiki/UCSD_GARCH) verfügbar.

## 4.3 ML- und MCMC-Schätzung eines GARCH(1,1)-Modells

Dieser Abschnitt demonstriert das in Abschnitt 4.1 erläuterte Vorgehen exemplarisch für ein GARCH(1,1)-Modell anhand eines konkreten Datensatzes.

### Erzeugung der Daten

Mit den Gleichungen (2.4) und (2.6) aus Abschnitt 2.2:

$$z_t = \sigma_t \cdot \varepsilon_t, \quad t = 1, \dots, T,$$

mit

$$\sigma_t^2 = \alpha_0 + \alpha_1 z_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad t = 1, \dots, T,$$

werden Daten aus einem GARCH(1,1)-Modell erzeugt. Dabei stammen die  $\varepsilon_t$  aus einer Normalverteilung. Für die Demonstration der MCMC-Schätzung wird ein Datensatz der Länge  $t = 5000$  mit den folgenden Parametern erzeugt:

$$\alpha_0 = 0.1, \alpha_1 = 0.1 \text{ und } \beta_1 = 0.8.$$

Die Wahl der Parameterwerte orientiert sich an Ergebnissen in Abschnitt 4.7 für Zeitreihen aus dem Finanzmarktbereich. Der Datensatz wird entsprechend der Verteilung und der Länge mit „y5000garch\_norm“ bezeichnet.

### Für die Simulation notwendige Spezifikationen

#### Startwerte

Für die MCMC-Simulation des GARCH(1,1)-Modells in diesem Abschnitt wurden folgende Startwerte verwendet:

$$\alpha_0 : 0.05, \alpha_1 : 0.005, \beta_1 : 0.9 \text{ und } \sigma_0^2 : 1.0615.$$

In dieser Simulation wird  $\sigma_0^2$  als zu schätzender Parameter aufgefasst, dies ist nicht notwendig und unterscheidet sich von dem Vorgehen bei der ML-Schätzung. Die Schätzung von  $\sigma_0^2$  wird in Abschnitt 4.3.1 nochmals thematisiert. Der Startwert von  $\sigma_0^2$  entspricht dem Schätzer für die Varianz der Daten. Die Bedeutung der Startwerte wird ebenfalls im folgenden Abschnitt noch diskutiert.

### Vorschlagsdichte

Die Kandidaten werden aus einer multivariaten Normalverteilung mit folgender Kovarianzmatrix gezogen:

$$\begin{aligned} \Sigma &= \begin{pmatrix} \text{Var}(\alpha_0) & \text{Cov}(\alpha_0, \alpha_1) & \text{Cov}(\alpha_0, \beta_1) & \text{Cov}(\alpha_0, \sigma_0^2) \\ \text{Cov}(\alpha_1, \alpha_0) & \text{Var}(\alpha_1) & \text{Cov}(\alpha_1, \beta_1) & \text{Cov}(\alpha_1, \sigma_0^2) \\ \text{Cov}(\beta_1, \alpha_0) & \text{Cov}(\beta_1, \alpha_1) & \text{Var}(\beta_1) & \text{Cov}(\beta_1, \sigma_0^2) \\ \text{Cov}(\sigma_0^2, \alpha_0) & \text{Cov}(\sigma_0^2, \alpha_1) & \text{Cov}(\sigma_0^2, \beta_1) & \text{Var}(\sigma_0^2) \end{pmatrix} \\ &= \begin{pmatrix} 0.001 & 0 & 0 & 0 \\ 0 & 0.001 & 0 & 0 \\ 0 & 0 & 0.001 & 0 \\ 0 & 0 & 0 & 0.5 \end{pmatrix}. \end{aligned} \quad (4.1)$$

Bei der Wahl der Vorschlagsdichte ist es wichtig zu bemerken, dass sie hier nicht speziell auf diesen konkreten Datensatz abgestimmt wird, sondern vielmehr für ein breites Spektrum der in diesem Kapitel durchgeführten Simulationen gültig sein soll. Wird in den folgenden Abschnitten eine andere Vorschlagsdichte verwendet, so wird diese angegeben.

### A priori Verteilungen

Die folgenden Gleichverteilungen wurden als uninformative a priori Verteilungen für die Parameter  $\alpha_1$  und  $\beta_1$  angenommen:

$$f(\alpha_1) = \begin{cases} 1 & \text{für } 0 \leq \alpha_1 \leq 1 \\ 0 & \text{sonst} \end{cases}, \quad f(\beta_1) = \begin{cases} 1 & \text{für } 0 \leq \beta_1 \leq 1 \\ 0 & \text{sonst.} \end{cases}$$

Für den Parameter  $\alpha_0$  wurde folgende Log-Normalverteilung als a priori Verteilungen angenommen:

$$\alpha_0 \sim \log N(-2.0, 5),$$

dies entspricht dem Vorschlag von Miashynskaia (2006) [75]. Für den Parameter  $\sigma_0^2$  wurde Vrotos et al. [94] folgend

$$f(\sigma_0^2) = \frac{1}{\sigma_0^2}$$

verwendet.

| Schätzung       | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $\sigma_0^2$ | Log-L   |
|-----------------|------------------|------------------|-----------------|--------------|---------|
| MCMC-Modus      | 0.1184           | 0.0859           | 0.8031          | 0.9892       | -7163.1 |
| MCMC-Mittelwert | 0.1285           | 0.0908           | 0.7883          | 2.1217       |         |
| MCMC-Median     | 0.1263           | 0.0901           | 0.7907          | 1.7145       |         |

Tabelle 4.1: GARCH-Modell mit y5000garch\_norm

### Ergebnisse

Mit den oben genannten Spezifikationen werden für die Daten die beiden Läufe durchgeführt. Tabelle 4.1 zeigt die Ergebnisse der Berechnungen für den ersten Lauf der Simulation für den Datensatz y5000garch\_norm. Dabei wurden 100000 Kettenglieder erzeugt. Unter Berücksichtigung der burn-in Phase wurden aus den letzten 50000 Werten des ersten Laufs die angegebenen Ergebnisse berechnet und Startwerte sowie die Kovarianzmatrix der Vorschlagsdichte für den zweiten Lauf bestimmt.

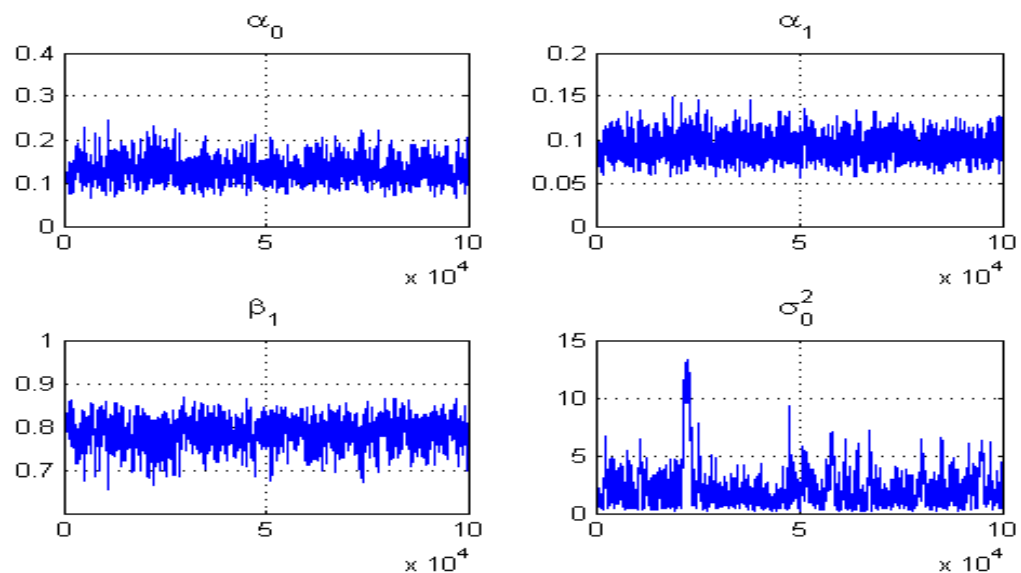


Abbildung 4.1: Markovketten des ersten Laufs für y5000garch\_norm

| Schätzung       | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $\sigma_0^2 : 0.9588$ | Log-L   |
|-----------------|------------------|------------------|-----------------|-----------------------|---------|
| ML-Matlab       | 0.1200           | 0.0883           | 0.7984          | 1.0301                | -7162.7 |
| ML-UCSD         | 0.1204           | 0.0884           | 0.7980          | 1.0363                | -7163.3 |
| ML-MFE          | 0.1199           | 0.0882           | 0.7986          | 1.3127                | -7163.1 |
| MCMC-Modus      | 0.1154           | 0.0877           | 0.8032          | 0.9288                | -7163.1 |
| MCMC-Mittelwert | 0.1279           | 0.0911           | 0.7886          | 2.7165                | -7163.2 |
| MCMC-Median     | 0.1258           | 0.0906           | 0.7905          | 2.0612                | -7163.1 |

Tabelle 4.2: GARCH-Modell mit y5000garch\_norm

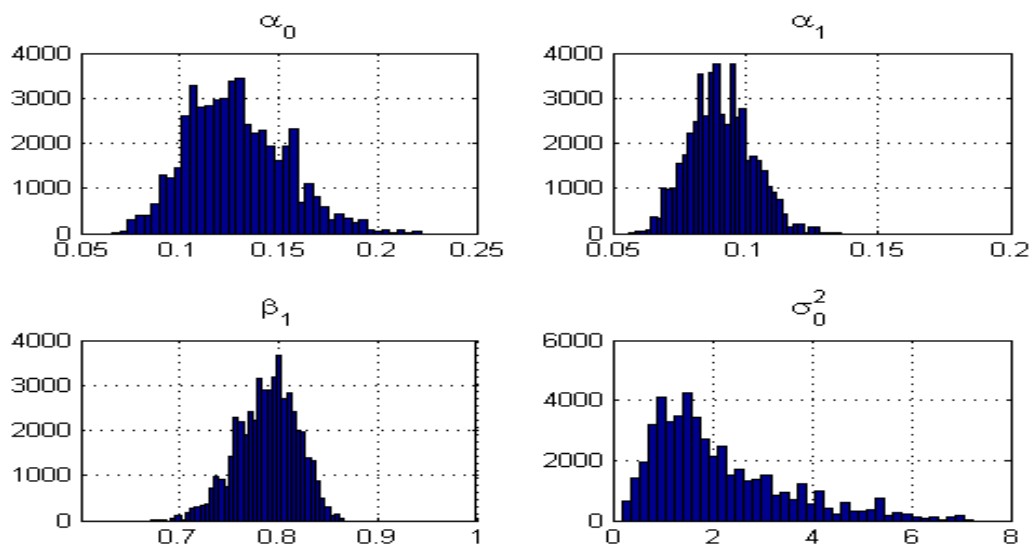


Abbildung 4.2: Histogramme des ersten Laufs für y5000garch\_norm

In dem zweiten Lauf werden dann 200000 Werte erzeugt. In Tabelle 4.2 sind die Ergebnisse des zweiten Laufs angegeben, zusätzlich enthält sie die ML-Schätzer der anfangs genannten Matlab Routinen.

Es zeigt sich bereits deutlich an den Markovketten des ersten Laufs (Abbildung 4.1), dass sie einen stationären Zustand erreicht haben. Die Ergebnisse des ersten Laufs ermöglichen schon eine Auswertung der a posteriori Verteilung und deuten bspw. auf leichte Schiefe der a posteriori Verteilungen des Parameter  $\beta_1$  hin, vgl. Abbildung 4.2. Die Akzeptanzrate des ersten Laufs ist allerdings sehr niedrig und liegt bei ca. 4%. Dies ist entsprechend den Ausführungen in Abschnitt 4.1 unproblematisch und führt zu günstigen Varianzen für die Vorschlagsdichten des zweiten Laufs. Die Dauer der Simulation der 100000 Werte des



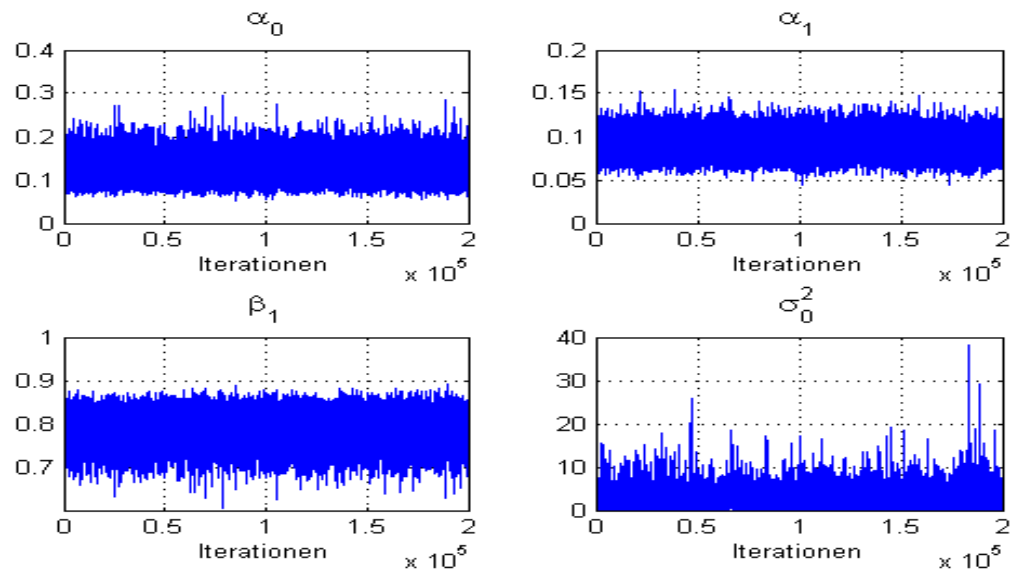


Abbildung 4.3: Markovketten des zweiten Laufs für y5000garch\_norm

ersten Laufs betrug unter zwei Minuten.

Die Abbildungen 4.3 und 4.4 zeigen die Markovketten und Histogramme der 200000 Werte des zweiten Laufs. Die Akzeptanzrate ist erkennbar höher als im ersten Lauf und betrug ca. 40%. Der zweite Lauf dauerte ca. 4 Minuten, so dass die Gesamtdauer der Simulation ca. sechs Minuten betrug.

Tabelle 4.2 kann man entnehmen, dass die MCMC-Schätzer für Modus, Mittelwert und Median nahe beieinander liegen. Das deutet auf minimale Asymmetrie hin. Dies kann man auch Abbildung 4.4 entnehmen. Der Modus der a posteriori Verteilung kommt dem ML-Schätzer aber sehr nahe, was an den angesprochenen flachen a priori Verteilungen liegt. Außerdem ist für den MCMC-Mittelwert und MCMC-Median jeweils die Log-Likelihood angegeben, die sich mit diesen Parametern ergibt. Auch der Wert der Log-Likelihood ist bei allen angegebenen Schätzern nahezu identisch. Die Tabelle enthält einen Wert für  $\sigma_0^2$ , dieser Wert stellt den initialen Wert der  $\sigma_t^2$  dar. Für die verschiedenen ML-Schätzungen ist der für die jeweilige Schätzung verwendete Wert angegeben. Da es sich um simulierte Daten handelt, kann der wahre Wert für  $\sigma_0^2$  angegeben werden, hier betrug er 0.9588.

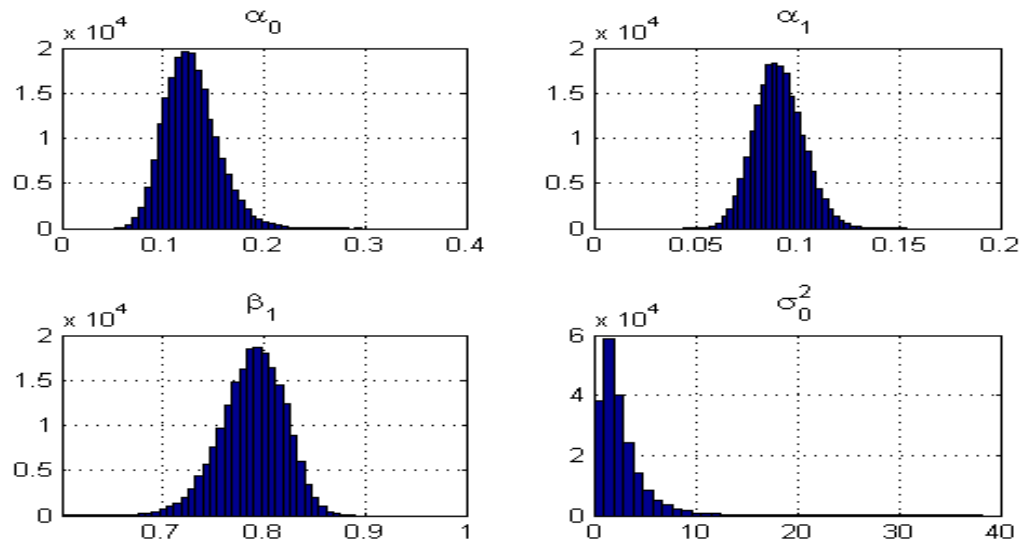


Abbildung 4.4: Histogramme des zweiten Laufs für y5000garch\_norm

### 4.3.1 Wichtige Aspekte bei der Simulation

In diesem Abschnitt werden Besonderheiten und Probleme bei der Anwendung des Bayes-Ansatzes auf GARCH-Modelle diskutiert und mögliche Lösungsvorschläge dargestellt.

#### Einfluss der Startwerte

Probleme bei der Simulation können, ebenso wie bei der ML-Schätzung, durch die Verwendung sehr ungünstiger Startwerte entstehen. Allerdings lässt sich die Phase des Einschwingens sehr leicht an den Markovketten erkennen. Wenn beispielsweise folgende Werte als Startwerte für die zu schätzenden Parameter verwendet werden:

$$\alpha_0 : 1, \alpha_1 : 0.8, \beta_1 : 0.1 \text{ und } \sigma_0^2 : 20,$$

welche verglichen mit den ML-MFE-Schätzwerten der GARCH-Parameter von

$$\alpha_0 : 0.1199, \alpha_1 : 0.0882 \text{ und } \beta_1 : 0.7986,$$

als extrem ungünstig angesehen werden können, ergeben sich die Markovketten in den Abbildungen 4.5 und 4.6.

Die Startwerte wirken sich direkt auf die Länge der burn-in Phase aus, wobei die Wahl sehr günstiger Startwerte die burn-in Phase selbstverständlich verkürzt. Ebenso können

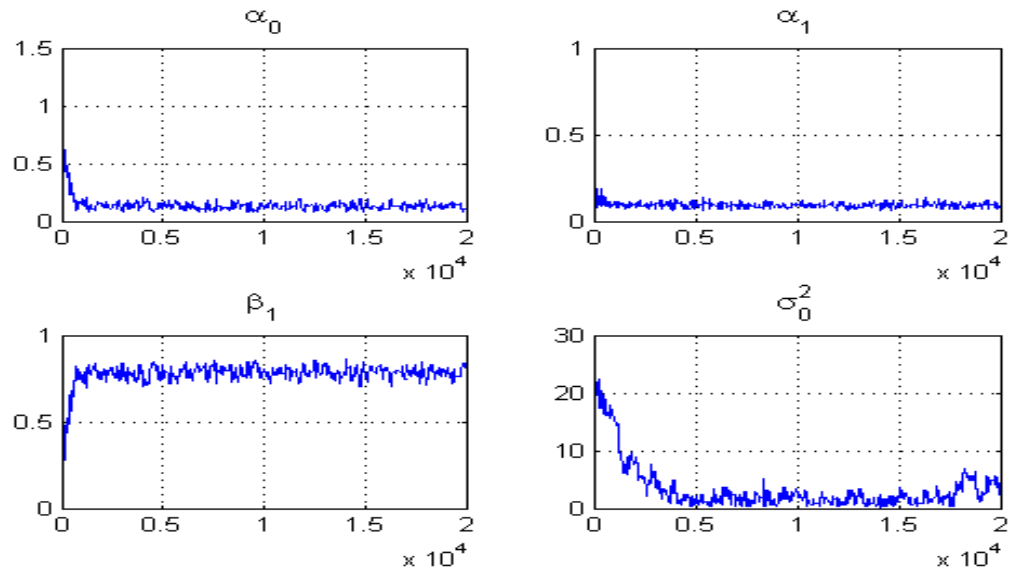


Abbildung 4.5: Markovketten des ersten Laufs bei ungünstigen Startwerten

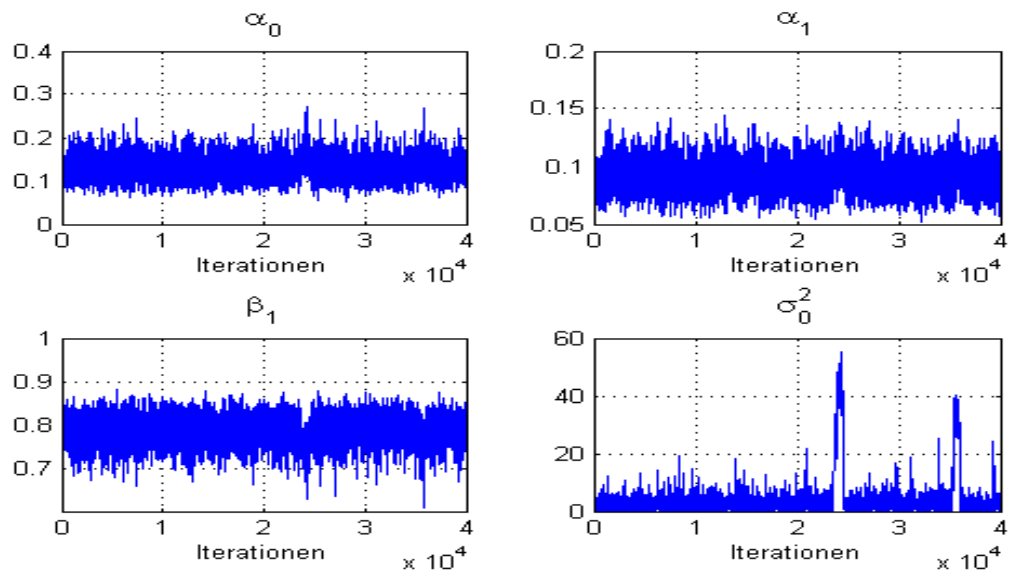


Abbildung 4.6: Markovketten des zweiten Laufs bei ungünstigen Startwerten im ersten Lauf

| Schätzung              | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $\sigma_0^2$ |
|------------------------|------------------|------------------|-----------------|--------------|
| MCMC-Modus Lauf 1      | 0.1188           | 0.0918           | 0.7957          | 0.7769       |
| MCMC-Mittelwert Lauf 1 | 0.1285           | 0.0922           | 0.7867          | 2.1064       |
| MCMC-Median Lauf 1     | 0.1286           | 0.0921           | 0.7869          | 1.7831       |
| MCMC-Modus Lauf 2      | 0.1127           | 0.0862           | 0.8069          | 1.0745       |
| MCMC-Mittelwert Lauf 2 | 0.1290           | 0.0913           | 0.7871          | 4.1413       |
| MCMC-Median Lauf 2     | 0.1267           | 0.0906           | 0.7894          | 2.2312       |

Tabelle 4.3: Ergebnisse bei ungünstigen Startwerten

aus ungünstigen Startwerten tendenziell schwächere Akzeptanzraten resultieren, selbst im zweiten Lauf. Man kann allerdings an dem Beispiel deutlich erkennen, dass die Kette sehr schnell die Startwerte verlässt und sich den gesuchten Werten nähert. In dem Beispiel wurde eine sehr kurze Länge von 20000 Iterationen des ersten Laufs gewählt, um die Auswirkung deutlicher darzustellen.

Tabelle 4.3 enthält die Ergebnisse der beiden Läufe und zeigt deutlich, dass die letzten 10000 Werte des ersten Laufs bereits passable Ergebnisse liefern und die MCMC-Punktschätzer für Modus und Mittelwert der a posteriori Verteilung denen der 40000 Werte des zweiten Laufs bereits sehr ähnlich sind.

Neben der Wahl geeigneter Startwerte kann man dem Problem begegnen, indem man die Varianzen anpasst. D.h. man wählt größere Varianzen für die Vorschlagsdichten, so werden die Bereiche für die einzelnen Parameter schneller durchlaufen und die burn-in Phase verkürzt. Erst das Zusammenspiel von ungeeigneten Startwerten und zu kleinen Varianzen der Vorschlagsdichten führt zu Markovketten, die nicht innerhalb der vorgegeben Anzahl an Iterationen konvergieren. Dies zeigt Abbildung 4.7, dabei wurde  $\sigma_0^2$  aus  $N(\sigma_{0,t-1}^2, 0.004)$  gezogen. Man kann nun erkennen, dass einzig die Markovkette für  $\sigma_0^2$  nicht konvergiert. Dies beeinträchtigt die Ergebnisse für den zweiten Lauf allerdings nicht, so dass die Markovketten und die daraus resultierenden MCMC-Schätzer den oben genannten stark ähneln. Dazu sei angemerkt, dass weitere Simulationen gezeigt haben, dass die Kombination ungeeigneter Startwerte und zu kleiner Varianzen der Vorschlagsdichte eines GARCH-Parameters zwar auch die anderen GARCH-Parameter stärker beeinflusst als bei  $\sigma_0^2$ . Allerdings wurde die burn-in Phase dadurch nur unwesentlich verlängert.

Generell bietet es sich an für die Startwerte entweder die ML-Schätzer zu verwenden

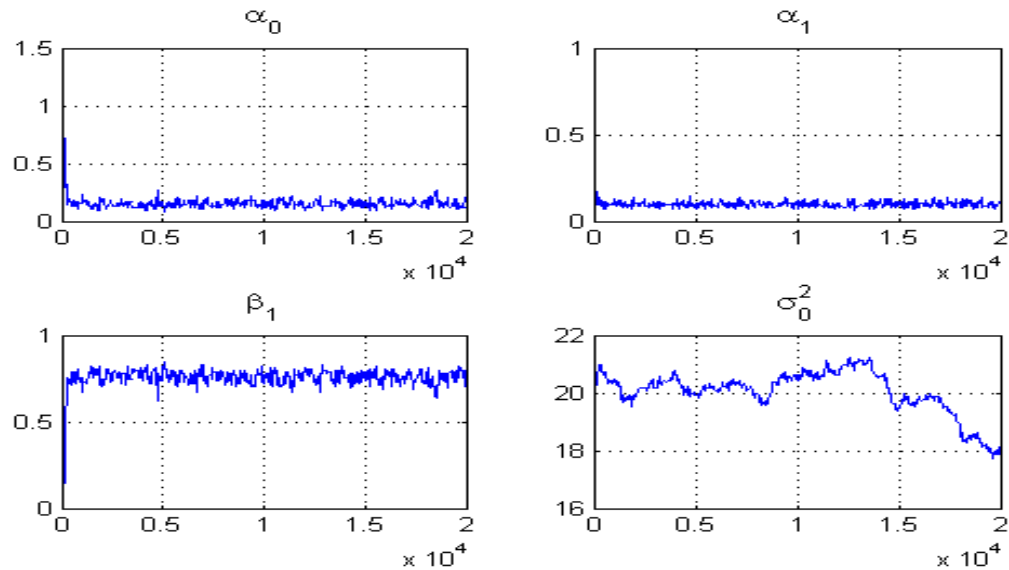


Abbildung 4.7: Folge ungünstiger Startwerte und Vorschlagsdichten

oder auf die Ergebnisse eines Testlaufs zurückzugreifen.

### Varianzen der Vorschlagsdichten

Wie oben schon erwähnt, spielt die Wahl der Vorschlagsdichten eine entscheidende Rolle. In den Simulationen werden ausschließlich Normalverteilungen verwendet. Vgl. dazu Takaishi (2009) [87] und [88], dort stammen die Kandidaten aus einer multivariaten  $t$ -Verteilung. Die Erwartungswerte der Vorschlagsdichten ergeben sich nach Festlegung der Startwerte aus der Simulation selbst, da immer der letzte Zustand der Markovkette als Erwartungswert dient. Die Varianzen werden im ersten Lauf vorgegeben. Diese Vorgabe ist entscheidend für das Ergebnis.

Ein einfaches Verfahren die gesuchte Varianzmatrix  $\Sigma$  zu finden, ist selbstverständlich die Varianzmatrix der ML-Schätzer zu verwenden. Dabei kommt sowohl die robust geschätzte Kovarianzmatrix nach Bollerslev und Wooldridge (1992) [20], als auch die nicht robust geschätzte Matrix in Betracht.

Durch das in Abschnitt 4.1 beschriebene Vorgehen, wird die Kovarianzmatrix für den zweiten Lauf bestimmt, aber für den ersten Lauf wird sie frei gewählt. Wird die Varianz hierbei zu klein gewählt, so werden tendenziell Kandidaten vorgeschlagen, die sehr nahe am momentanen Zustand der Markovkette sind. Damit bewegen sich die Ketten sehr langsam weg von ihren Startwerten. Außerdem ist der Einfluss einzelner Parameter sehr groß, so dass die Ketten daran gehindert werden können alle Parameter zu finden. Die

Anzahl akzeptierter Kandidaten ist sehr groß und die Ketten verweilen evtl. in lokalen Maxima.

Bei zu großen Varianzen der Vorschlagsdichten springen die Kandidaten sehr viel, so dass nur sehr wenige von ihnen akzeptiert werden. Dies führt zu relativ kleinen Werten der Kovarianzmatrix für den zweiten Lauf aufgrund der Phasen, in denen die Markovketten ihren Zustand nicht ändern. Die Kovarianzmatrix aus Gleichung (4.1) kann selbstverständlich auch durch Multiplikation mit einer Konstanten variiert werden, dies ist in vielen Fällen hilfreich, um eine geeignete Kovarianzmatrix zu finden.

Zusätzlich erschwert wird die Suche nach der geeigneten Varianz der Vorschlagsdichte durch die Tatsache, dass unterschiedliche a priori Verteilungen verschiedene Varianzen notwendig machen. Insbesondere gilt dies für die Varianz der Vorschlagsdichte von  $\sigma_0^2$ , dies wird später noch thematisiert.

### **Behandlung der Stationaritätsbedingung**

Das in Abschnitt 4.3 beschriebene Verfahren zur Schätzung eines GARCH(1,1)-Modells wird mit Berücksichtigung der Stationaritätsbedingung ( $\alpha_1 + \beta_1 < 1$ ) durchgeführt, d.h. jeder erzeugte Vektor mit Kandidaten wird daraufhin überprüft, ob die Stationaritätsbedingung eingehalten wird. Ein Kandidatenvektor, der die Bedingung nicht erfüllt, wird sofort verworfen, ohne die Akzeptanzwahrscheinlichkeit des Kandidaten zu prüfen. Ein solcher Kandidat senkt somit die Akzeptanzrate der MCMC-Schätzung. Dies entspricht einem Vorschlag von Gelman. Alternativ wird vorgeschlagen die Einhaltung der Stationaritätsbedingung durch die a priori Verteilung zu steuern, dies führt aber zu einem größeren Rechenaufwand.

Die Schätzung kann auch ohne Prüfung der Stationaritätsbedingung durchgeführt werden. Vrontos et al. (2000) [95] vergleichen Verfahren mit und ohne Berücksichtigung der Stationaritätsbedingung im Zusammenhang mit multivariaten GARCH-Modellen und zeigen, dass beide Varianten bereits nach 3000 Iterationen gleiche Ergebnisse liefern. Es muss aber betont werden, dass das Vernachlässigen der Überprüfung der Stationaritätsbedingung durchaus dazu führen kann, dass die Vorschlagsdichten angepasst werden müssen. Diverse eigene Simulationen haben gezeigt, dass die MCMC-Schätzungen durch zum Teil deutlich niedrigere Varianzen der Vorschlagsdichten zu denselben Ergebnissen führen. Auch im Folgenden wird bei den MCMC-Simulationen die Stationaritätsbedingung stets berücksichtigt.

### Notwendigkeit des zweiten Laufs

Wie in Abschnitt 4.1 erwähnt, ist es möglich, die Simulation nach dem ersten Lauf abzubrechen und ausschließlich die bis dahin gewonnenen Informationen zu nutzen. Bei dieser Entscheidung ist der relevante Faktor in aller Regel die Berechnungszeit. Allerdings zeigen diverse Simulationen, dass es günstiger ist, eher den ersten Lauf deutlich zu verkürzen und den zweiten Lauf entsprechend zu wählen, da die Markovketten im zweiten Lauf fast immer vorzuziehen sind. Ausnahmen ergeben sich nur bei hohen Akzeptanzraten im ersten Lauf.

Wie bereits Abbildung 4.5 gezeigt hat, sind die burn-in Phasen zwar zu berücksichtigen, aber meist überschaubar. In dem Beispiel wäre eine Markovkette der Länge 10000 ausreichend gewesen, um mit den 5000 verbleibenden Werten Startwerte und die Kovarianzmatrix für den zweiten Lauf zu berechnen. Ein kleines Beispiel soll dies verdeutlichen: Mit den folgenden Varianzen der Vorschlagsdichten wird der erste Lauf für 100000 Iterationen betrachtet:

$$\alpha_0 : 0.001, \alpha_1 : 0.001, \beta_1 : 0.001 \text{ und } \sigma_0^2 : 0.5.$$

Damit wird eine Akzeptanzrate von unter 5% erreicht, die Berechnungszeit beträgt ca. eine Minute. Wird bei gleichen Voraussetzungen eine zweistufige Simulation durchgeführt mit 20000 Iterationen für den ersten Lauf und 80000 Iterationen für den Zweiten, so dauern die Berechnungen beider Läufe insgesamt ebenfalls ca. eine Minute, aber die Akzeptanzrate innerhalb des zweiten Laufs liegt bei ca. 40%.

Natürlich kann die Akzeptanzrate des ersten Laufs durch sorgfältigere Auswahl der Vorschlagsdichten in den gewünschten Bereich zwischen 20% und 40% verändert werden, allerdings ist die Suche nach geeigneten Vorschlagsdichten ohne ausreichende Vorkenntnis häufig sehr aufwändig, die Durchführung des zweiten Laufs der Simulation reduziert diesen Aufwand deutlich. Entscheidend ist, dass die Vorschlagsdichte die a posteriori Verteilung möglichst gut nachbildet, dies erhöht die Akzeptanzrate deutlich. Ein weiterer entscheidender Vorteil ist die Berücksichtigung der Kovarianzen in dem zweiten Lauf.

### Schätzung von $\sigma_0^2$

Im Gegensatz zu der ML-Schätzung ist bei der Bayes-Schätzung ein Wert für  $\sigma_0^2$  zu schätzen, vgl. Vrontos et al. [94]. Allerdings wird in der Literatur meist ein konstanter Wert für  $\sigma_0^2$  angenommen. Als Werte kommen beispielsweise der Lageparameter  $a_0$  oder die unbedingte Varianz  $\frac{\alpha_0}{1-\alpha_1-\beta_1}$  in Frage, vgl. dazu Ardia (2006) [29]. Wird  $\sigma_0^2$  als eigenstän-

| Schätzung           | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $\sigma_0^2 : 0.9588$ | Log-L   |
|---------------------|------------------|------------------|-----------------|-----------------------|---------|
| ML-MFE              | 0.1199           | 0.0882           | 0.7986          | 1.3127                | -7163.1 |
| (1) MCMC-Modus      | 0.1251           | 0.0908           | 0.7915          | 0.1000                | -7171.9 |
| (1) MCMC-Mittelwert | 0.1362           | 0.0948           | 0.7774          | 0.1000                | -7171.9 |
| (1) MCMC-Median     | 0.1340           | 0.0943           | 0.7794          | 0.1000                | -7171.9 |
| (2) MCMC-Modus      | 0.1152           | 0.0867           | 0.8046          | 1.0000                | -7163.4 |
| (2) MCMC-Mittelwert | 0.1264           | 0.0911           | 0.7901          | 1.0000                | -7163.4 |
| (2) MCMC-Median     | 0.1242           | 0.0905           | 0.7921          | 1.0000                | -7163.4 |
| (3) MCMC-Modus      | 0.1155           | 0.0882           | 0.8030          | 1.0179                | -7163.1 |
| (3) MCMC-Mittelwert | 0.1284           | 0.0912           | 0.7878          | 2.9479                | -7163.3 |
| (3) MCMC-Median     | 0.1264           | 0.0909           | 0.7898          | 2.1539                | -7163.1 |

Tabelle 4.4: Einfluss der Schätzung von  $\sigma_0^2$ 

diger Parameter mitgeschätzt, so hat das erheblichen Einfluss auf die Durchführung der Schätzung. Zum einen erhöht jeder zu schätzende Parameter die benötigte Rechenzeit, zum anderen senkt er die Akzeptanzrate. Außerdem erschwert er das Abstimmen der Varianzen der Vorschlagsdichten. Die Schätzung von  $\sigma_0^2$  hat damit auch großen Einfluss auf die restlichen Schätzer. Die a posteriori Verteilung von  $\sigma_0^2$  ist extrem schief, daher weichen Mittelwert und Median stark von einander ab. Dies beeinflusst auch die Ergebnisse der anderen Schätzer. Selbstverständlich nimmt der Einfluss des initialen Werts von  $\sigma_0^2$  mit steigender Länge der Daten ab.

Am Beispiel der Berechnungen zu den Daten y5000garch\_norm wird der Einfluss der Schätzung von  $\sigma_0^2$  demonstriert. Tabelle 4.4 zeigt die Ergebnisse der Simulationen mit verschiedenen Varianten der Behandlung von  $\sigma_0^2$ . Die Ergebnisse sind dabei entsprechend folgender Auflistung geordnet:

- (1) konstanter Wert:  $\alpha_0$ , hier 0.05,
- (2) konstanter Wert:  $\frac{\alpha_0}{1-\alpha_1-\beta_1}$ , also der Schätzer für die unbedingte Varianz,
- (3)  $\sigma_0^2$  wird als eigenständiger Parameter geschätzt.

Dabei wird deutlich, dass der Einfluss der Schätzung von  $\sigma_0^2$  sehr stark ist. Die Ergebnisse der Schätzer und des Maximalwerts der Log-Likelihood zeigen, dass die Varianten (2) und (3) sehr ähnlich sind und sich deutlich von Variante (1) unterscheiden. Dies liegt



daran, dass die unbedingte Varianz in diesem Fall näher an dem geschätzten Wert von  $\sigma_0^2$  liegt, als  $\alpha_0$ . Dies ist natürlich von je nach Datensatz unterschiedlich. Allerdings hat sich gezeigt, dass  $\alpha_0$  in aller Regel zu schlechteren Ergebnissen führt. Selbstverständlich kann anstatt  $\frac{\alpha_0}{1-\alpha_1-\beta_1}$  auch  $\widehat{Var}(z_t)$  berechnet werden. Da es sich um simulierte Daten handelt ist es in diesem Fall möglich den tatsächlichen Wert von  $\sigma_0^2$  zu beobachten. Hier lautete der Wert: 0.9588. Der Wert, den die MFE-Toolbox als initialen Wert für  $\sigma_t$  verwendet, ist ebenfalls in der Tabelle angegeben.

Alternativ wurde noch versucht  $\alpha_0$  nicht konstant zu halten, sondern den Wert der jeweiligen Iteration von  $\alpha_0$  zu verwenden, allerdings verzerrt dieses Vorgehen die Ergebnisse zu stark.

### A priori Verteilungen

Grundsätzlich kommen sehr unterschiedliche a priori Verteilungen für die betrachteten Modelle in Frage. Für die Parameter des GARCH-Modells können beispielsweise Konstanten verwendet werden. Wie in Abschnitt 3.1.2 erwähnt, handelt es sich dann um unechte a priori Verteilungen. Zur Erinnerung sei erwähnt, dass in Abschnitt 4.3 Gleichverteilungen für  $\alpha_1$  und  $\beta_1$  verwendet wurden, sowie eine Log-Normalverteilung für  $\alpha - 0$  und  $f(\sigma_0^2) = \frac{1}{\sigma_0^2}$ . Vrontos et al. verwenden  $f(\alpha_0) = \frac{1}{\alpha_0}$ . Sie ähnelt der hier verwendeten a priori Verteilung, allerdings strebt sie für  $\alpha_0$  gegen null, gegen unendlich. Miazhyńska et al. (2006) [75] schlagen Log-Normalverteilungen vor, z.B.  $\alpha_1 \sim \log N(-2.0, 5)$ , und  $\beta_1 \sim \log N(-0.2, 5)$ . Diese erwiesen sich als günstig für große Stichprobenumfänge, da hier die Daten die a priori Verteilungen deutlich überlagern. Allerdings führen sie bei kleinen Stichproben dazu, dass die Parameter  $\alpha_1$  und  $\beta_1$  gegen Null streben. Die Wahl der a priori Verteilung ist auch abhängig von dem Ziel der MCMC-Schätzung. Wie bereits erwähnt führt die Wahl flacher a priori Verteilung dazu, dass der MAP dem ML-Schätzer entspricht. Soll also nur ein Schätzer für  $\theta$  gefunden werden, der die Likelihood maximiert, so können für  $\alpha_1$  und  $\beta_1$  Gleichverteilungen zwischen 0 und 1 verwendet werden und eine Konstante für  $\alpha_0$ . Allerdings besteht bei kleinen Stichproben die Gefahr, dass die a priori Verteilung die a posteriori Verteilung überlagert, da die Annahme einer Gleichverteilung zwischen 0 und 1 auch eine Art der Information darstellt.

Bislang wurde ausschließlich auf die Parameter der Varianzfunktion des GARCH-Modells abgestellt, da von einer Normalverteilung der Residuen ausgegangen wurde. Im Vorgriff auf die Verwendung alternativer Verteilungen soll noch die a priori Verteilung

für den Scharparameter der Verteilung angesprochen werden. Von besonderer Bedeutung ist dabei der Parameter  $\nu$  der  $t$ -Verteilung. Bauwens und Lubrano (1998) [7] zeigen, dass die a posteriori Dichte nicht "proper" ist, wenn man eine flache a priori Verteilung im Bereich  $(0, \infty)$  verwendet. Miazhyńska et al. (2006) [75] schlagen die Exponentialverteilung  $\nu \sim \text{Exp}(0.1)$  vor, dabei ist zu beachten, dass in Matlab der Parameter  $\lambda$  der Exponentialverteilung invers verwendet wird ( $f(x|\lambda) = \frac{1}{\lambda}e^{-\frac{x}{\lambda}}$ ), daher  $\nu \sim \text{Exp}(10)$ . Diese Verteilung wird auch hier im Folgenden benutzt. Alternativ schlagen Vrontos et al. (2000) [94]  $\frac{1}{n-2}$  als priori Verteilung für den Parameter  $\nu$  vor.

### Auswahl der Werte zur Analyse der a posteriori Verteilung

Die Länge der simulierten Ketten hat nach Berücksichtigung der burn-in Phase nur eine kleine Bedeutung für die Schätzer der Parameter. Der Zeitaufwand spielt mit zunehmender Rechenleistung der Computer eine immer untergeordnetere Rolle. Wichtig ist die Frage, ob sämtliche Werte der Markovkette in die Analyse aufgenommen werden sollten. Dabei spielt die Autokorrelation der Kettenglieder eine wichtige Rolle. Daher wird in der Literatur häufig geraten nur jeden lag-ten Zug zu verwenden, damit in den verwendeten Zügen keine Autokorrelation mehr vorhanden ist. Dies führt allerdings zu einem enormen Verlust an Informationen. Immerhin 50% der Werte bleiben unberücksichtigt bei Vernachlässigung nur jedes zweiten Wertes, bzw. entsprechend 80% bei Berücksichtigung jedes fünften Zugs. Die steigende Rechenleistung ermöglicht es allerdings dennoch für jede Variante ausreichend lange Markovketten zu erzeugen.

Eine weitere denkbare Alternative wäre die ausschliessliche Berücksichtigung der akzeptierten Kandidaten. Dies führt je nach Simulation zu einem Verlust von 60%-80%, entsprechend der angestrebten Akzeptanzrate von 20%-40%. Allerdings verfälscht dies die Ergebnisse deutlich, da nun die Kandidaten nicht mehr im korrekten Verhältnis berücksichtigt werden.

Tabelle 4.5 zeigt die Ergebnisse der Schätzung eines GARCH(1,1)-Modells für den Datensatz „y5000garch\_norm“. Die Spezifikationen entsprechen den oben genannten, damit werden für den zweiten Lauf wieder 200000 Werte bestimmt. Bei den angegebenen Ergebnissen werden bei (1) alle Werte (200000) des zweiten Laufs für die Analyse a posteriori Verteilung verwendet, bei (2) wird nur jeder 10. Zug verwendet, also 20000 Werte. Zusätzlich werden bei (3) der Mittelwert und der Median der akzeptierten Kandidaten angegeben. Bei dieser Schätzung wurden 82870 Kandidaten akzeptiert.

| Schätzung           | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $\sigma_0^2$ |
|---------------------|------------------|------------------|-----------------|--------------|
| (1) MCMC-Mittelwert | 0.1143           | 0.1040           | 0.7725          | 2.1991       |
| (1) MCMC-Median     | 0.1117           | 0.1033           | 0.7756          | 1.6161       |
| (2) MCMC-Mittelwert | 0.1143           | 0.1040           | 0.7725          | 2.199        |
| (2) MCMC-Median     | 0.1118           | 0.1033           | 0.7753          | 1.6177       |
| (3) MCMC-Mittelwert | 0.1169           | 0.1050           | 0.7687          | 2.2797       |
| (3) MCMC-Median     | 0.1145           | 0.1045           | 0.7714          | 1.6927       |

Tabelle 4.5: Auswahl unterschiedlicher Werte für die Analyse der a posteriori Verteilung:  
 (1) alle Werte, (2) jeder 10. Wert und (3) akzeptierte Kandidaten

Es wird deutlich, dass (1) und (2) zu nahezu identischen Ergebnissen führen, d.h. es ist für die Parameterschätzer unerheblich, ob neun von zehn Zügen verworfen werden oder nicht. Allerdings weichen die Schätzer bei Verwendung der akzeptierten Kandidaten erwartungsgemäß ab. Im Folgenden werden stets die Ergebnisse bei Berücksichtigung jedes 10. Zugs angegeben. Davon ausgenommen ist der MAP, dieser wird stets aus allen verfügbaren Werten ermittelt.

### A posteriori Verteilung versus Log-Likelihood

Abbildung 4.8 zeigt den Verlauf der a posteriori Verteilung und der Log-Likelihood für die Kandidaten, d.h. den entsprechenden Wert für jede Parameterkonstellation in der Markovkette. Man kann erkennen, wie die a priori Verteilung die Kette beeinflusst. Dabei kommt es zu einer Niveauverschiebung, aber auch zu einer Reduktion extremer Werte, da diese in der a priori Verteilung eine niedrige Wahrscheinlichkeit haben.

### Neuberechnung der Kovarianzmatrix in jedem Schritt

Takaishi (2009) [87] hat eine "adaptive Vorschlagsdichte" vorgeschlagen. Dazu werden bei einer Simulation in bestimmten Intervallen die gewonnenen Daten dazu verwendet, die Vorschlagsdichten anzupassen. Der Autor demonstriert wie dieses Vorgehen die Autokorrelation innerhalb der Markovkette reduziert. Dabei werden die Kandidaten wie schon erwähnt aus einer multivariaten  $t$ -Verteilung gezogen.

Es folgt ein Beispiel für einen GARCH(1,1)-Datensatz der Länge 5000, der wie oben beschrieben erzeugt wurde. Die Kandidaten werden wie oben beschrieben aus einer multivariaten Normalverteilung gezogen, aber die Vorschlagsdichte wird nicht wie in Abschnitt

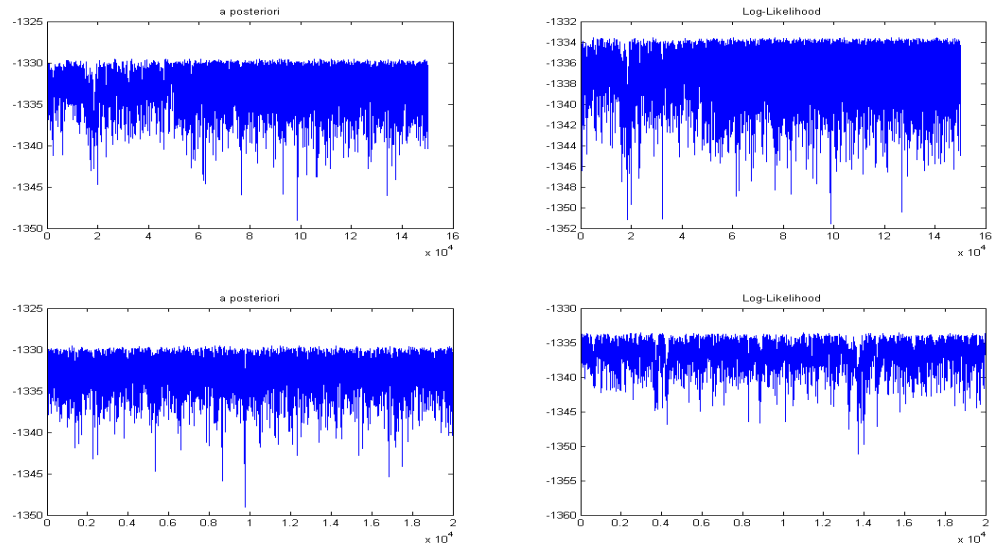


Abbildung 4.8: Vergleich der a posteriori Verteilung und der Log-Likelihood

4.1 beschrieben nur einmalig angepasst, sondern in dem zweiten Lauf erfolgt die Anpassung der Vorschlagsdichte nach jedem Zug. Alle sonstigen Spezifikationen entsprechen den oben genannten. Dies führt selbstverständlich zu einer erheblich höheren Berechnungsdauer. Gezeigt werden soll das Ausmaß an Veränderung der Kovarianzmatrix  $\Sigma$  durch dieses Vorgehen.

Für den ersten Lauf werden die Kandidaten aus einer multivariaten Normalverteilung mit folgender Kovarianzmatrix gezogen<sup>2</sup>:

$$\Sigma = \begin{pmatrix} \text{Var}(\alpha_1) & \text{Cov}(\alpha_0, \alpha_1) & \text{Cov}(\alpha_0, \beta_1) \\ \text{Cov}(\alpha_1, \alpha_0) & \text{Var}(\alpha_1) & \text{Cov}(\alpha_1, \beta_1) \\ \text{Cov}(\beta_1, \alpha_0) & \text{Cov}(\beta_1, \alpha_1) & \text{Var}(\beta_1) \end{pmatrix} = \begin{pmatrix} 0.001 & 0 & 0 \\ 0 & 0.001 & 0 \\ 0 & 0 & 0.001 \end{pmatrix}.$$

Nachdem mit dem oben beschriebenen Vorgehen der erste Lauf der Simulation durchgeführt wurde, ergibt sich beispielsweise folgende Kovarianzmatrix<sup>3</sup>:

$$\Sigma = \begin{pmatrix} 0.0008 & 0.0002 & -0.0011 \\ 0.0002 & 0.0002 & -0.0004 \\ -0.0011 & -0.0004 & 0.0017 \end{pmatrix}.$$

<sup>2</sup>Aus Gründen der Übersichtlichkeit und Vergleichbarkeit mit Ergebnissen der ML-Schätzung werden nur die GARCH-Parameter betrachtet.

<sup>3</sup>Die Akzeptanzrate des ersten Laufs lag bei knapp 5%.

Diese dient als Kovarianzmatrix der Vorschlagsdichte für den ersten Kandidaten des zweiten Laufs, danach wird nach jedem Zug anhand der bis dahin verfügbaren Markovketten eine neue Kovarianzmatrix für den nächsten Zug bestimmt. Damit ergibt sich am Ende des zweiten Laufs die folgende Kovarianzmatrix für die Parameter:

$$\Sigma = \begin{pmatrix} 0.0007 & 0.0002 & -0.0010 \\ 0.0002 & 0.0002 & -0.0004 \\ -0.0010 & -0.0004 & 0.0015 \end{pmatrix}.$$

Es zeigt sich, dass die Kovarianzmatrix, die mit Hilfe der Werte aus dem ersten Lauf der Simulation bestimmt wurde, durch die weiteren Anpassungen kaum weiter verändert wird. Da die Kovarianzmatrix nahezu unverändert bleibt, verändert sich die Akzeptanzrate im Vergleich zur einmaligen Anpassung der Kovarianzmatrix nicht, sie beträgt knapp 43%. Alle weiteren Ergebnisse entsprechen den oben genannten Resultaten für diesen Datensatz, allerdings dauerten die Berechnungen ca. drei Stunden. Aufgrund des Aufwands durch die Neuberechnung der Kovarianzmatrix, wird im Folgenden stets eine einmalige Anpassung der Kovarianzmatrix verwendet. Der Vergleich mit Kovarianzmatrizen, die aus der ML-Schätzungen resultieren, wird im Anschluss behandelt.

Zu den Korrelationen sei noch angemerkt, dass zwischen den GARCH-Parametern sehr hohe Korrelationen zu beobachten waren. So beträgt die Korrelation zwischen  $\alpha_1$  und  $\beta_1$  -0.8 und die Korrelation zwischen  $\alpha_0$  und  $\beta_1$  häufig sogar -0.95.

### Vergleich der Kovarianzmatrix der ML-Schätzung und der Bayes-Schätzung

Die verwendete MFE-Toolbox von Sheppard berechnet sowohl die einfache Kovarianzmatrix, sowie die robuste Kovarianzmatrix nach Bollerslev und Wooldridge (1992) [20]. Bei der MCMC-Schätzung kann die Varianz der Schätzer aus der a posteriori Verteilung der Parameter geschätzt werden. Für den verwendeten Beispiel-Datensatz ergeben sich für die ML- und die MCMC-Schätzung die Werte aus Tabelle 4.6.

### 4.3.2 Asymptotische Normalverteilung der ML-Schätzer

Die in Abschnitt 2.2.1 angesprochene ML-Schätzung der Parameter eines GARCH(1,1)-Modells führt zu asymptotisch normalverteilten Schätzern, dies wird in Beiträgen von Bollerslev und Wooldridge (1992) [20], Lumsdaine (1996) [68] oder Lee und Hansen

| Schätzung             | $\alpha_0$ | $\alpha_1$ | $\beta_1$ |
|-----------------------|------------|------------|-----------|
| ML - Varianz          | 0.0005     | 0.0002     | 0.0011    |
| ML - Varianz (robust) | 0.0006     | 0.0002     | 0.0014    |
| MCMC                  | 0.0007     | 0.0002     | 0.0015    |

Tabelle 4.6: Vergleich der Schätzer für die Varianzen (ML versus MCMC)

(1994) [67] speziell für GARCH-Modelle behandelt und bewiesen. Dies gilt ebenso für die Quasi-ML-Schätzung. Allerdings bleibt dabei offen, ab wann die Asymptotik greift, d.h. wie viele Beobachtungen notwendig sind, bis die Verteilungen der Schätzer tatsächlich Normalverteilungen folgen. Hwang und Pereira (2006) [58] beschäftigen sich mit den Eigenschaften von kleinen Stichproben in GARCH-Modellen, den so genannten Small Sample Eigenschaften. Sie kommen zu dem Schluss, dass für GARCH-Modelle mindestens 500 Beobachtungen notwendig sind, da die Verzerrungen und "Convergence errors", also das Problem nicht konvergierender Schätzer, sonst zu groß ist. Allerdings stellt sich die Frage, ob die geforderten 500 Beobachtungen in Einklang zu bringen sind mit der Berücksichtigung von Strukturbrüchen, analog zu den Überlegungen in der Regressionsrechnung. Also ob die Parameter über die gesamte Zeitreihe konstant bleiben. Immerhin entsprechen die 500 Beobachtungen bei der Verwendung von Tagesschlusskursen etwa zwei Jahren. Eine Betrachtung der Zeitreihe des DAX ist geeignet zu klären, ob die Parameter eines GARCH-Modells über einen langen Zeitraum beibehalten werden können. Daher wurden für die Zeitreihe der Tagesschlusskurse des DAX folgende Berechnungen durchgeführt. Es wurden die Renditen von 4000 aufeinander folgenden Handelstagen bis einschließlich 17.09.2009 in vier Zeitreihen mit jeweils 1000 Handelstagen zerlegt. Für diese vier Zeiträume wurde jeweils ein  $GARCH_t$ -Modell angepasst. Tabelle 4.7 zeigt die Ergebnisse für die vier Zeitspannen.

Schon bei dieser einfachen Aufteilung der Zeitreihe ergeben sich Abweichungen in den GARCH-Parametern und insbesondere für den Parameter  $\nu$  der Freiheitsgrade der  $t$ -Verteilung. Weitere Berechnungen haben gezeigt, dass auch für andere Unterteilungen der Zeitreihe Schwankungen in GARCH-Parametern vorhanden sind. Die Schwankungen in den Freiheitsgraden waren noch sehr viel stärker ausgeprägt. Insgesamt fällt auf, dass die Summe der Parameter  $\alpha_1$  und  $\beta_1$  in allen Bereichen nahe 1 ist, d.h. es liegt durchgängig eine hohe Persistenz vor.

Die Eigenschaften kleiner Stichproben bei Verwendung von Hochfrequenzdaten wird z.B.

| Teilabschnitt | $\alpha_0$ | $\alpha_1$ | $\beta_1$ | $\nu$   |
|---------------|------------|------------|-----------|---------|
| 1             | 0.0131     | 0.0603     | 0.9301    | 9.9331  |
| 2             | 0.0757     | 0.1000     | 0.8721    | 30.8353 |
| 3             | 0.0082     | 0.0625     | 0.9333    | 28.2490 |
| 4             | 0.0202     | 0.0973     | 0.8973    | 7.6130  |

Tabelle 4.7: Vergleich der GARCH-Parameter für die Renditen des DAX für vier Zeiträume mit jeweils  $n = 1000$

von Galbraith und Zinde-Walsh (2000) [40] untersucht. Interessant in diesem Zusammenhang sind die Ergebnisse von Hwang und Pereira (2006) [58], welche zeigen, dass bei einer Zeitreihe von 2000 Daten mit konstanten ARCH- und GARCH-Parametern schon eine Phase mit nur 20 Daten mit starker Persistenz die Persistenz der gesamten Zeitreihe erheblich beeinflussen. Lastrapes (1989) [65], Lastrapes und Koray (1990) [66] und Lamoureux und Lastrapes (1990) [64] beschäftigen sich mit der Schätzung von Persistenz in langen Zeitreihen. Es wird die Möglichkeit eines Strukturbruchs im Lageparameter, aber auch in den ARCH- und GARCH-Parametern in Betracht gezogen. Markov-Switching-Modelle bieten eine Möglichkeit den unterschiedlichen Phasen in Zeitreihen auch im GARCH-Modell Rechnung zu tragen, vgl. dazu bspw. Cai (1994) [21], Haas et al. (2004) [53] oder Marcucci (2005) [73].

Zur Demonstration der Eigenschaft der asymptotischen Normalverteilung der ML-Schätzer wird eine weitere Simulationsstudie durchgeführt. Dazu werden für 5 Varianten jeweils 1000 Datensätze mit den jeweiligen Parametern erzeugt und geschätzt. Dabei wurden folgende Werte für die Parameter verwendet:

Variante 1:  $\alpha_0 = 0.2$ ,  $\alpha_1 = 0.2$  und  $\beta_1 = 0.6$ ,

Variante 2:  $\alpha_0 = 0.1$ ,  $\alpha_1 = 0.1$  und  $\beta_1 = 0.8$ ,

Variante 3:  $\alpha_0 = 0.05$ ,  $\alpha_1 = 0.05$  und  $\beta_1 = 0.94$ ,

Variante 4:  $\alpha_0 = 0.4$ ,  $\alpha_1 = 0.49$  und  $\beta_1 = 0.49$ ,

Variante 5:  $\alpha_0 = 0.1$ ,  $\alpha_1 = 0.3$  und  $\beta_1 = 0.3$ .

Dabei ist zu beachten, dass sich bei den ersten beiden Varianten die Parameterwerte  $\alpha_1$  und  $\beta_1$  zu 0.8 bzw. 0.9 aufsummieren und somit eine mäßige bis hohe Persistenz vorliegt. Bei den beiden anderen Varianten ergibt die Summe nahezu 1, damit liegt eine sehr hohe

Persistenz vor. Insbesondere die Varianten 2 und 3 orientieren sich an den Ergebnissen für den DAX aus Tabelle 4.7. In Variante 5 ergibt sich eine niedrige Persistenz.

Zur Simulation sowie zur Schätzung wird die oben erwähnte MFE-Toolbox von Sheppard verwendet. Dies wird jeweils für sechs verschiedene Stichprobenumfänge durchgeführt:  $n = 100, 250, 500, 1000, 2500$  und  $5000$ . Tatsächlich werden jeweils 3000 Daten mehr erzeugt, aber die ersten 3000 Daten werden verworfen, um den Einfluss der Startwerte zu minimieren. Dabei sei erwähnt, dass dies bei ca. 250 Handelstagen pro Jahr den Tageschlusskursen von ca. 5 Monaten sowie 1, 2, 4, 10 und 20 Jahren entspricht.



| Variante 1                 | n=100            |                  |                 | n=250            |                  |                 | n=500                  |                        |                        |
|----------------------------|------------------|------------------|-----------------|------------------|------------------|-----------------|------------------------|------------------------|------------------------|
|                            | $\alpha_0 : 0.2$ | $\alpha_1 : 0.2$ | $\beta_1 : 0.6$ | $\alpha_0 : 0.2$ | $\alpha_1 : 0.2$ | $\beta_1 : 0.6$ | $\alpha_0 : 0.2$       | $\alpha_1 : 0.2$       | $\beta_1 : 0.6$        |
| Mittelwert (1)             | 0.2672           | 0.1840           | 0.5396          | 0.2386           | 0.1956           | 0.5604          | 0.2192                 | 0.1937                 | 0.5842                 |
| Varianz robust (3)         | 182.8503         | 10.2513          | 339.0424        | 0.0293           | 0.0125           | 0.0566          | 0.0177                 | 0.0044                 | 0.0329                 |
| Varianz der $\theta_i$ (2) | 0.0472           | 0.0211           | 0.0873          | 0.0247           | 0.0086           | 0.0445          | 0.0108                 | 0.0035                 | 0.0195                 |
|                            | n=1000           |                  |                 | n=2500           |                  |                 | n=5000                 |                        |                        |
| Mittelwert (1)             | 0.2096           | 0.1987           | 0.5896          | 0.2033           | 0.1989           | 0.5975          | 0.2021                 | 0.2000                 | 0.5977                 |
| Varianz robust (3)         | 0.0046           | 0.0019           | 0.0088          | 0.0015           | 0.0007           | 0.0029          | 0.0007                 | 0.0004                 | 0.0014                 |
| Varianz der $\theta_i$ (2) | 0.0039           | 0.0017           | 0.0073          | 0.0014           | 0.0007           | 0.0027          | 0.0014                 | 0.0007                 | 0.0027                 |
| Variante 2                 | n=100            |                  |                 | n=250            |                  |                 | n=500                  |                        |                        |
|                            | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $\alpha_0 : 0.1$       | $\alpha_1 : 0.1$       | $\beta_1 : 0.8$        |
| Mittelwert (1)             | 0.2216           | 0.0901           | 0.6771          | 0.1603           | 0.0948           | 0.7408          | 0.1299                 | 0.0962                 | 0.7713                 |
| Varianz robust (3)         | 0.6147           | 0.2107           | 1.4474          | 0.1893           | 0.0285           | 0.4071          | 0.0183                 | 0.0024                 | 0.0299                 |
| Varianz der $\theta_i$ (2) | 0.0622           | 0.0104           | 0.0988          | 0.0277           | 0.0042           | 0.0430          | 0.0109                 | 0.0018                 | 0.0179                 |
|                            | n=1000           |                  |                 | n=2500           |                  |                 | n=5000                 |                        |                        |
| Mittelwert (1)             | 0.1133           | 0.1004           | 0.7856          | 0.1049           | 0.0998           | 0.7950          | 0.1029                 | 0.1001                 | 0.7970                 |
| Varianz robust (3)         | 0.0042           | 0.0010           | 0.0075          | 0.0010           | 0.0003           | 0.0020          | $0.4100 \cdot 10^{-3}$ | $0.1627 \cdot 10^{-3}$ | $0.8191 \cdot 10^{-3}$ |
| Varianz der $\theta_i$ (2) | 0.0038           | 0.0009           | 0.0066          | 0.0009           | 0.0003           | 0.0016          | $0.4185 \cdot 10^{-3}$ | $0.1581 \cdot 10^{-3}$ | $0.8226 \cdot 10^{-3}$ |

Tabelle 4.8: Mittelwerte und Varianzen der ML-Schätzer bei unterschiedlichen Stichprobenumfängen (mäßige Persistenz)

| Variante 3                 | n=100             |                   | n=100            |                        | n=250                  |                        | n=250                  |                        | n=500                  |                        | n=500                  |                        |
|----------------------------|-------------------|-------------------|------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
|                            | $\alpha_0 : 0.05$ | $\alpha_1 : 0.05$ | $\beta_1 : 0.94$ | $\alpha_0 : 0.05$      | $\alpha_1 : 0.05$      | $\beta_1 : 0.94$       | $\alpha_0 : 0.05$      | $\alpha_1 : 0.05$      | $\beta_1 : 0.94$       | $\alpha_0 : 0.05$      | $\alpha_1 : 0.05$      | $\beta_1 : 0.94$       |
| Mittelwert (1)             | 1.1109            | 0.0460            | 0.7214           | 0.5313                 | 0.0469                 | 0.8338                 | 0.1880                 | 0.0460                 | 0.8338                 | 0.0460                 | 0.0460                 | 0.9113                 |
| Varianz robust (3)         | 268.5534          | 0.7701            | 7.0918           | 2.4749                 | 0.0689                 | 0.2836                 | 0.2107                 | 0.0013                 | 0.2836                 | 0.0013                 | 0.0013                 | 0.0124                 |
| Varianz der $\theta_i$ (2) | 3.2764            | 0.0059            | 0.1110           | 1.0599                 | 0.0021                 | 0.0569                 | 0.1830                 | 0.0005                 | 0.0569                 | 0.0005                 | 0.0005                 | 0.0123                 |
|                            | n=1000            |                   | n=1000           |                        | n=2500                 |                        | n=2500                 |                        | n=5000                 |                        | n=5000                 |                        |
| Mittelwert (1)             | 0.0872            | 0.0485            | 0.9323           | 0.0612                 | 0.0492                 | 0.9381                 | 0.0556                 | 0.0499                 | 0.9381                 | 0.0499                 | 0.0499                 | 0.9388                 |
| Varianz robust (3)         | 0.0047            | 0.0002            | 0.0007           | $0.6853 \cdot 10^{-3}$ | $0.0777 \cdot 10^{-3}$ | $0.1470 \cdot 10^{-3}$ | $0.2504 \cdot 10^{-3}$ | $0.0379 \cdot 10^{-3}$ | $0.1470 \cdot 10^{-3}$ | $0.0379 \cdot 10^{-3}$ | $0.0379 \cdot 10^{-3}$ | $0.0617 \cdot 10^{-3}$ |
| Varianz der $\theta_i$ (2) | 0.0125            | 0.0002            | 0.0014           | $0.6718 \cdot 10^{-3}$ | $0.0719 \cdot 10^{-3}$ | $0.1247 \cdot 10^{-3}$ | $0.2720 \cdot 10^{-3}$ | $0.0384 \cdot 10^{-3}$ | $0.1247 \cdot 10^{-3}$ | $0.0384 \cdot 10^{-3}$ | $0.0384 \cdot 10^{-3}$ | $0.0612 \cdot 10^{-3}$ |
| Variante 4                 | n=100             |                   | n=100            |                        | n=250                  |                        | n=250                  |                        | n=500                  |                        | n=500                  |                        |
|                            | $\alpha_0 : 0.4$  | $\alpha_1 : 0.49$ | $\beta_1 : 0.49$ | $\alpha_0 : 0.4$       | $\alpha_1 : 0.49$      | $\beta_1 : 0.49$       | $\alpha_0 : 0.4$       | $\alpha_1 : 0.49$      | $\beta_1 : 0.49$       | $\alpha_0 : 0.4$       | $\alpha_1 : 0.49$      | $\beta_1 : 0.49$       |
| Mittelwert (1)             | 0.5860            | 0.4258            | 0.4640           | 0.4752                 | 0.4626                 | 0.4771                 | 0.4369                 | 0.4706                 | 0.4771                 | 0.4369                 | 0.4706                 | 0.4874                 |
| Varianz robust (3)         | 0.1880            | 0.0344            | 0.0363           | 0.0399                 | 0.0124                 | 0.0094                 | 0.0158                 | 0.0062                 | 0.0094                 | 0.0158                 | 0.0062                 | 0.0041                 |
| Varianz der $\theta_i$ (2) | 0.1752            | 0.0305            | 0.0348           | 0.0416                 | 0.0111                 | 0.0106                 | 0.0153                 | 0.0050                 | 0.0106                 | 0.0153                 | 0.0050                 | 0.0042                 |
|                            | n=1000            |                   | n=1000           |                        | n=2500                 |                        | n=2500                 |                        | n=5000                 |                        | n=5000                 |                        |
| Mittelwert (1)             | 0.4169            | 0.4801            | 0.4875           | 0.4078                 | 0.4860                 | 0.4893                 | 0.4029                 | 0.4885                 | 0.4893                 | 0.4029                 | 0.4885                 | 0.4897                 |
| Varianz robust (3)         | 0.0068            | 0.0031            | 0.0019           | 0.0026                 | 0.0013                 | 0.0007                 | 0.0012                 | 0.0006                 | 0.0007                 | 0.0012                 | 0.0006                 | 0.0004                 |
| Varianz der $\theta_i$ (2) | 0.0064            | 0.0025            | 0.0018           | 0.0025                 | 0.0011                 | 0.0007                 | 0.0012                 | 0.0006                 | 0.0007                 | 0.0012                 | 0.0006                 | 0.0004                 |

Tabelle 4.9: Mittelwerte und Varianzen der ML-Schätzer bei unterschiedlichen Stichprobenumfängen (hohe Persistenz)

### Interpretation der Resultate

Die Tabellen 4.8 und 4.9 zeigen die Mittelwerte und Varianzen für die verschiedenen Stichprobenumfänge  $n$  für die Varianten 1 bis 4. Die Ergebnisse von Variante 5 befinden sich im Anhang.

Aufgrund von Konvergenzproblemen kann es bei der Schätzung der Varianzen mit der MFE-Toolbox zu negativen Werten kommen. Daher sind in der Tabelle für die Varianzen nur die robusten Schätzer nach Bollerslev und Wooldridge angegeben, in Form der Mittelwerte der Elemente der Hauptdiagonale der Varianz-Kovarianzmatrix. Hier tritt dieses Problem nicht auf. Es lässt sich deutlich ablesen, dass sich der Mittelwert der Schätzer (1)

$$\bar{\theta}_i = \widehat{E}(\widehat{\theta}^{ML}) = \frac{1}{1000} \sum_{i=1}^{1000} \widehat{\theta}_i^{ML}$$

dem "wahren" Wert mit steigendem Stichprobenumfang annähert, ebenso reduziert sich die Varianz der Schätzer (2)

$$S^2(\widehat{\theta}) = \frac{1}{1000 - 1} \sum_{k=1}^{1000} (\widehat{\theta}_i^k - \bar{\theta}_i)^2.$$

Man kann auch deutlich erkennen, dass die robuste Schätzung der Varianz nach Bollerslev und Wooldridge (3)

$$\widehat{E}(\widehat{Var}(\widehat{\theta}^{ML}))$$

die Varianz der Schätzer bei niedrigen Stichprobenumfängen zum Teil deutlich überschätzt, und dass dies mit steigendem  $n$  deutlich abnimmt. Allerdings erkennt man auch, dass die Konvergenz nicht für alle drei Parameter gleichermaßen gut ist. Während  $\alpha_1$  bereits bei kleinen Stichprobenumfängen relativ kleine Verzerrungen aufweist, bleibt die Verzerrung bei den Parametern  $\alpha_0$  und  $\beta_1$  länger und deutlicher bestehen. Dies wird im Folgenden auch grafisch veranschaulicht.

Nach den Ergebnissen in den Tabellen 4.8 und 4.9 wird der Parameter  $\alpha_1$  durchgängig gut geschätzt, selbst bei kleinen Stichprobenumfängen sind die Verzerrungen relativ klein. Einzig in Variante 4 ist die Verzerrung bei einem Stichprobenumfang von 100 auffällig. Wobei für alle Varianten gilt, dass bei einem Stichprobenumfang von 100 starke Verzerrungen und auch sehr schlechte Schätzer für die Varianzen aller Parameter zu beobachten sind.

Der Parameter  $\beta_1$  weist umso stärkere Verzerrungen auf, je mehr er sich dem Wert 1 nähert. Aufgrund der Stationaritätsbedingung  $\alpha_1 + \beta_1 < 1$  ist der Parameter dort beschränkt. Interessanterweise ist der Parameter  $\alpha_1$  davon nicht betroffen und in Variante 3

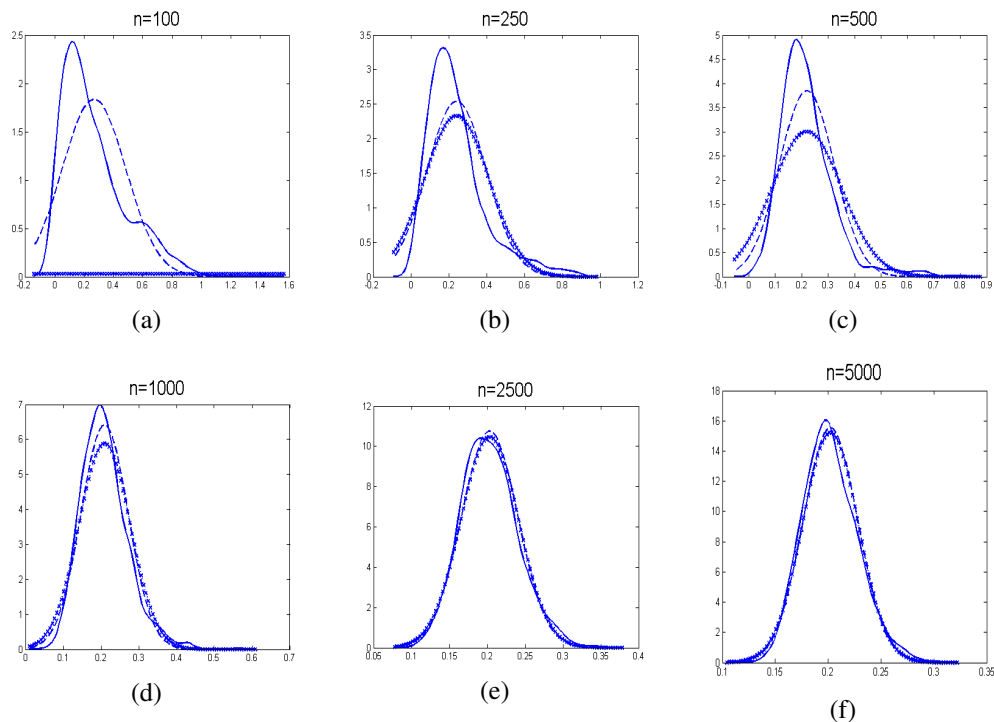


Abbildung 4.9: Verteilung des ML-Schätzers des Parameters  $\alpha_0$ : 0.2 bei Variante 1, bei unterschiedlichen Stichprobenumfängen

(mit  $\alpha_1 = 0.05$ ) wird der Parameter bereits bei einem Stichprobenumfang von 100 gut getroffen. Vielmehr ist der Parameter  $\alpha_0$  betroffen. Seine Verzerrung scheint umso stärker, je näher der Parameter  $\beta_1$  an 1 ist. Dies steht im Zusammenhang mit der bereits erwähnten extrem hohen negativen Korrelation der Parameter  $\alpha_0$  und  $\beta_1$ .

Die Abbildungen 4.9, 4.10 und 4.11 zeigen für die drei Parameter  $\alpha_0$ ,  $\alpha_1$  und  $\beta_1$ , jeweils die aus der Simulation resultierenden Verteilungen der ML-Parameterschätzer der 1000 Replikationen für die unterschiedlichen Stichprobenumfänge in Variante 1. Die durchgezogene Linie stellt eine Kerndichteschätzung der Verteilung der 1000 ML-Parameterschätzer dar, '- -' ist eine angepasste Normalverteilung mit der aus den 1000 Parameterschätzern geschätzten Varianz und  $\bar{\theta}_i$  als Erwartungswert, und die gepunktete Linie stellt die angepasste Normalverteilung dar, die sich mit dem Durchschnitt der robusten Varianzschätzer von Bollerslev und Wooldridge ergibt.

Auffallend ist bei der Betrachtung der Schaubilder, dass bei einem Stichprobenumfang von 100 die angepassten Normalverteilungen deutlich von der Verteilung der 1000 ML-Schätzwerte abweichen. Insbesondere ist der Durchschnitt der robust geschätzten Varianz deutlich zu groß, was zu einer nahezu flachen Verteilung führt. Ansonsten ist zu sehen,

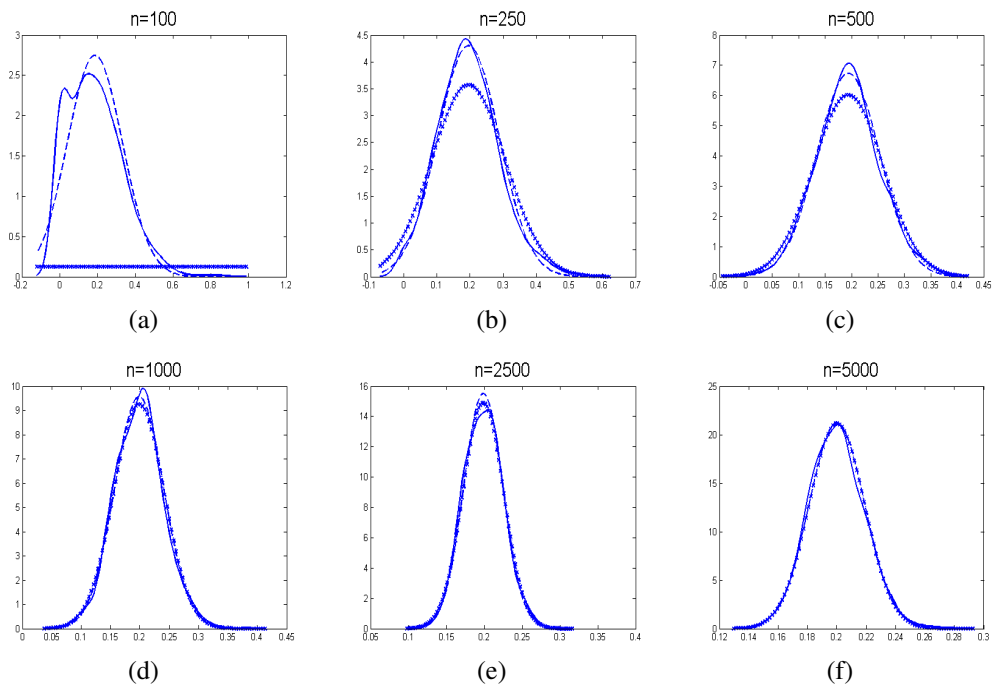


Abbildung 4.10: Verteilung des ML-Schätzers des Parameters  $\alpha_1: 0.2$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

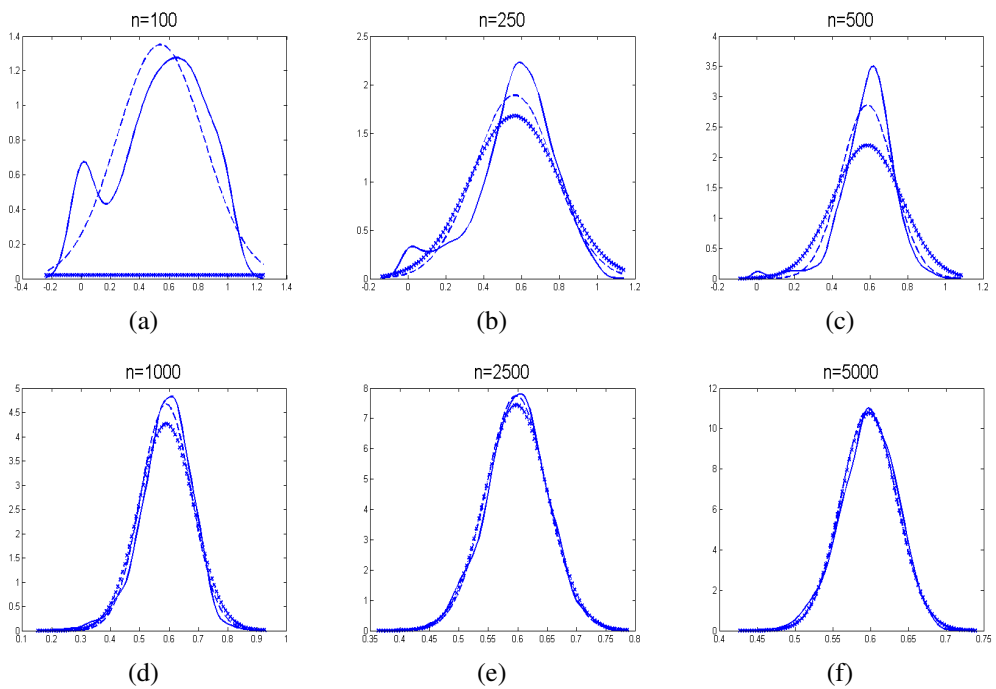


Abbildung 4.11: Verteilung des ML-Schätzers des Parameters  $\beta_1: 0.6$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

dass die Abweichungen mit steigendem Stichprobenumfang erwartungsgemäß abnehmen, allerdings wird deutlich, dass dies für die Parameter unterschiedlich gut geschieht. So weisen die Parameter  $\alpha_0$  und  $\beta_1$  bei einem Stichprobenumfang von 500 stärkere Abweichungen auf, als der Parameter  $\alpha_1$ .

Tendenziell unterschätzt die ML-Schätzung die Varianz der GARCH-Parameter, dies wird durch die robuste Schätzung nach Bollerslev und Wooldridge kompensiert. Umso erstaunlicher ist die Tatsache, dass insbesondere bei kleinen Stichprobenumfängen der Durchschnitt der robust geschätzten Varianz hier deutlich zu groß erscheint. Abbildung 4.12 zeigt die robust geschätzten Varianzen der 1000 Replikationen für die Stichprobenumfänge  $n = 100, 250$  und  $500$ . Es wird deutlich, dass die angepassten Normalverteilungen insbesondere bei einem Stichprobenumfang von  $n = 100$  aufgrund von extremen Ausreißern eine zu große Varianz haben. Diese extremen Ausreißer sind bei den Parametern  $\alpha_0$  und  $\beta_1$  deutlich stärker als bei dem Parameter  $\alpha_1$ . Der Einfluss der Ausreißer nimmt mit steigendem Stichprobenumfang ab. Bei den grafischen Darstellungen der anderen Stichprobenumfänge setzt sich die Entwicklung fort, sie sind im Anhang A.1 abgebildet.

Die Ergebnisse der Variante 3 (mit  $\alpha_0 = 0.05$ ,  $\alpha_1 = 0.05$  und  $\beta_1 = 0.94$ ,) sind in den Abbildungen 4.13, 4.14 und 4.15 grafisch dargestellt. Auffallend ist bei der Betrachtung der Schaubilder insbesondere die Schiefe der Verteilung der Parameter  $\alpha_0$  und  $\beta_1$ , was aus den Ergebnissen den Tabellen 4.8 und 4.9 bereits zu erkennen ist. Die Asymmetrie nimmt mit steigendem  $n$  ab. Es ist wichtig die bereits in Abschnitt 4.3.1 angesprochene hohe negative Korrelation der Parameter  $\alpha_0$  und  $\beta_1$  zu beachten, da diese in der Vorschlagsdichte beachtet wird, und sich so auf die a posteriori Verteilung auswirkt.

Es wurden für alle Parameter Jarque-Bera und Shapiro-Wilk Tests auf Normalverteilung durchgeführt. Die Annahme der Normalverteilung wird beinahe für alle Verteilungen verworfen, außer für  $\alpha_1$  bei den Stichprobenumfängen  $n = 2500$  und  $5000$ . Die Kerndichteschätzungen verfälschen den Eindruck allerdings, da hierbei entweder Bereiche angenommen werden können, die der Parameter aufgrund der Restriktionen nicht annehmen kann, oder die Dichtefunktion an den Rändern gegen unendlich strebt.

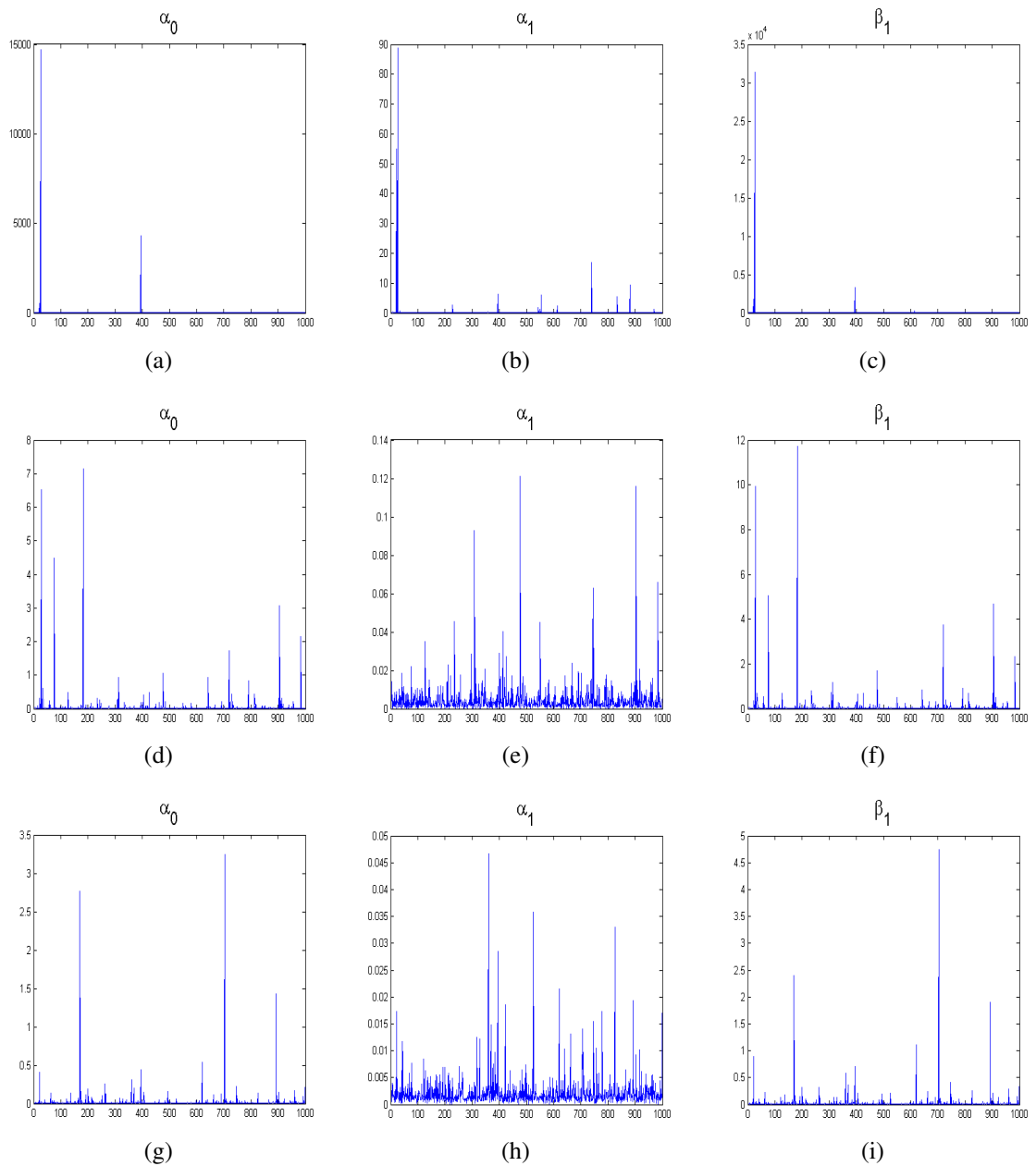


Abbildung 4.12: Grafische Darstellung der robusten Schätzer für die Varianz der 1000 Replikationen für die Stichprobenumfänge  $n=100$ ,  $250$  und  $500$  bei Varianz 2

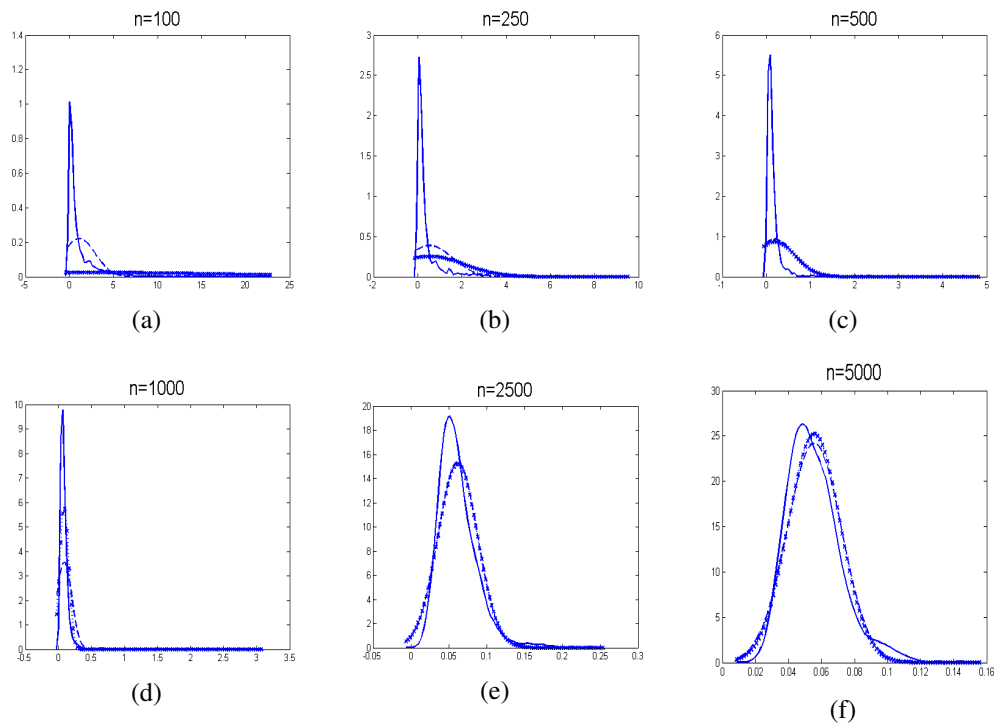


Abbildung 4.13: Verteilung des ML-Schätzers des Parameters  $\alpha_0 : 0.05$  bei Variante 3, bei unterschiedlichen Stichprobenumfängen

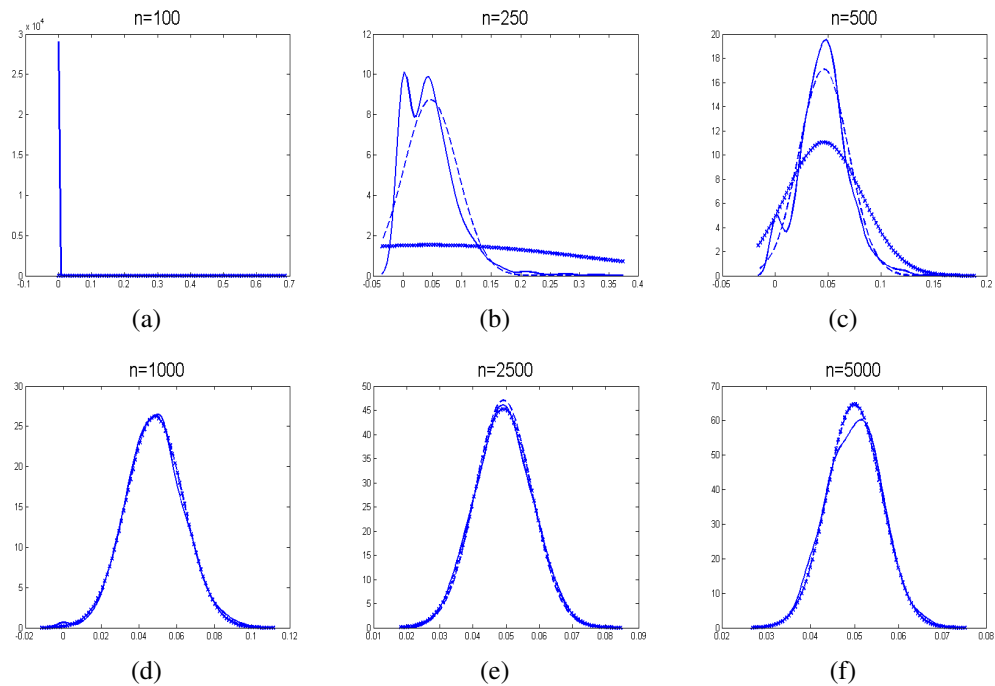


Abbildung 4.14: Verteilung des ML-Schätzers des Parameters  $\alpha_1 : 0.05$  bei Variante 3, bei unterschiedlichen Stichprobenumfängen



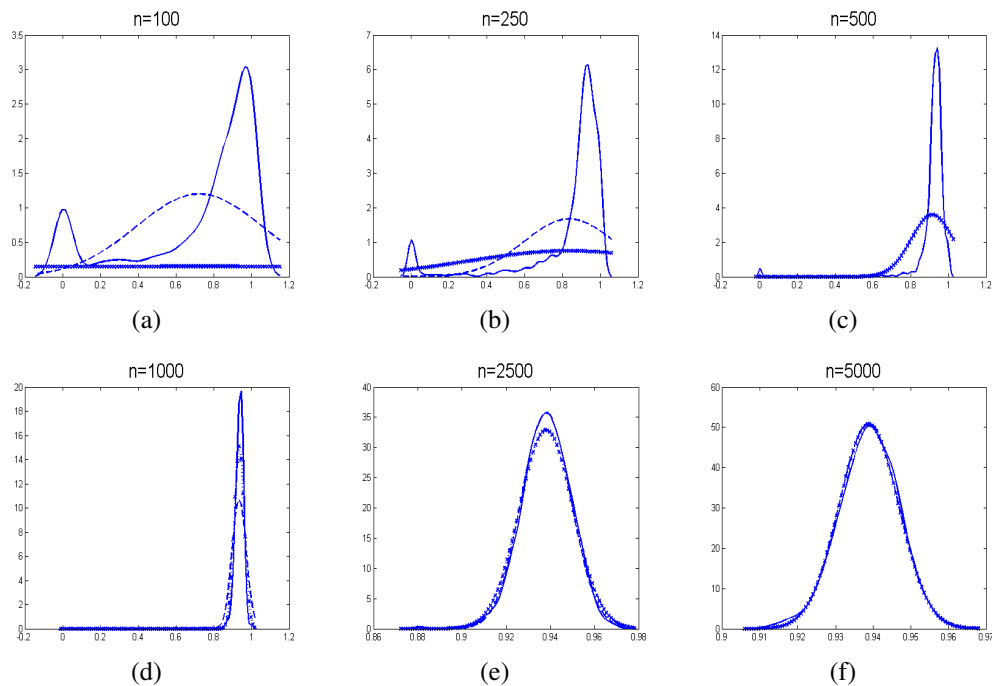


Abbildung 4.15: Verteilung des ML-Schätzers des Parameters  $\beta_1 : 0.94$  bei Variante 3, bei unterschiedlichen Stichprobenumfängen

### 4.3.3 Vergleich der MCMC-Schätzung und der ML-Schätzer

Um die bisherigen Ergebnisse der ML- und der MCMC-Schätzung zu vergleichen, werden noch sechs weitere einzelne Datensätze mit Stichprobenumfang  $n = 100, 250, 500, 1000, 2500$  und  $5000$  mit den oben genannten Parametern

$$\alpha_0 = 0.1, \alpha_1 = 0.1 \text{ und } \beta_1 = 0.8$$

erzeugt. Dabei sei ausdrücklich erwähnt, dass die Auswahl der Datensätze nicht zufällig geschehen ist. So wurden für die MCMC-Schätzungen speziell Datensätze ausgewählt, deren ML-Schätzer nicht zu stark von den oben genannten Parametern abwichen. Dies schränkt speziell bei niedrigen Stichprobenumfängen die Auswahl ein.

Zunächst soll anhand dieser ausgewählten Datensätze verglichen werden, wie gut die MCMC-Schätzung dieser Datensätze und die daraus resultierenden a posteriori Verteilungen gegenüber den entsprechenden Verteilungen der ML-Schätzer der 1000 Replikationen aus dem vorigen Abschnitt abschneiden.

Alle MCMC-Simulationen wurden wie in Abschnitt 4.1 beschrieben durchgeführt.

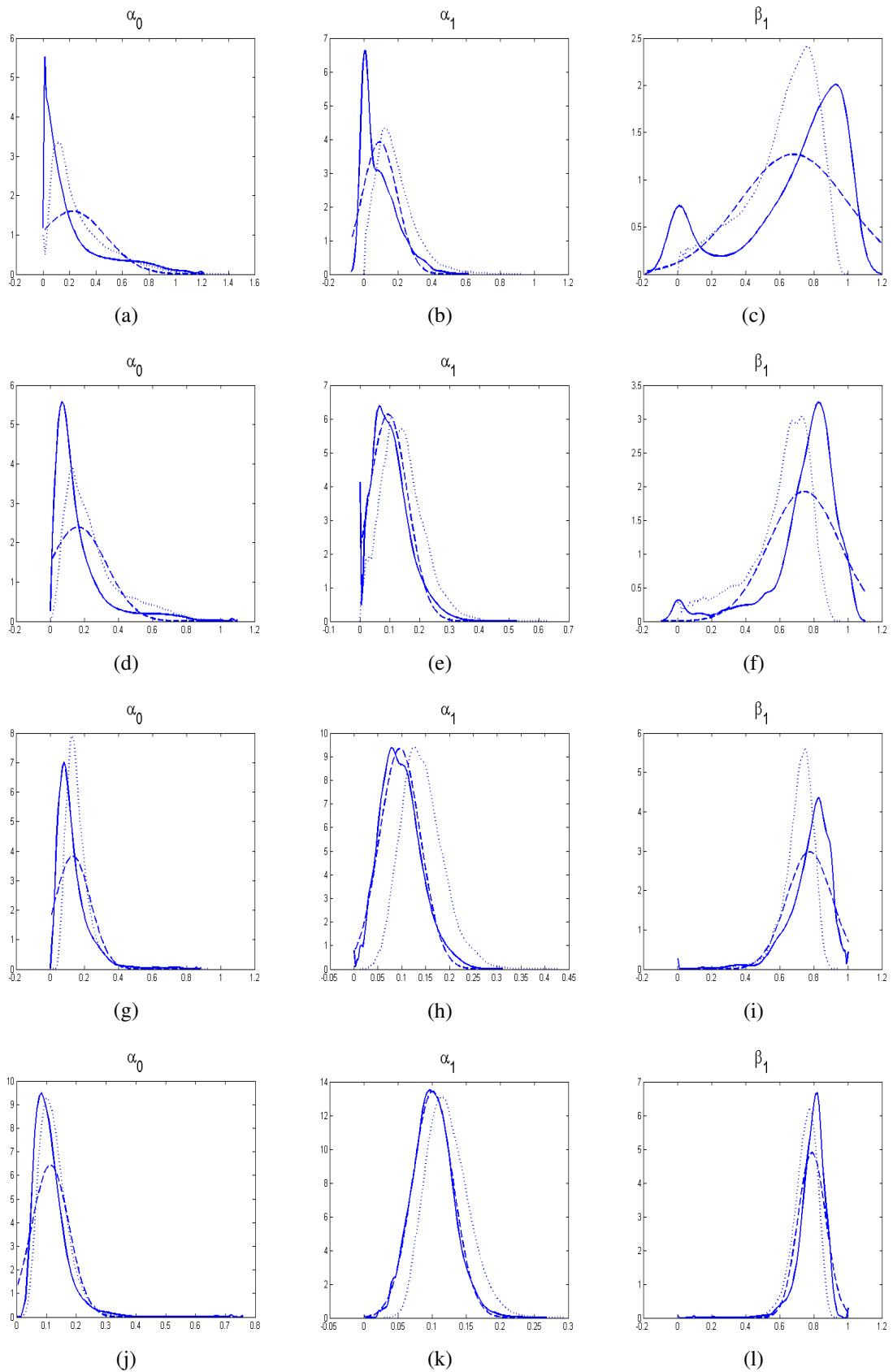


Abbildung 4.16: Randverteilungen der ML-Schätzer bei unterschiedlichen Stichprobenumfängen im Vergleich zu der a posteriori Verteilung der MCMC-Schätzung eines einzelnen Datensatzes ( $n=100, 250, 500$  und  $1000$ )

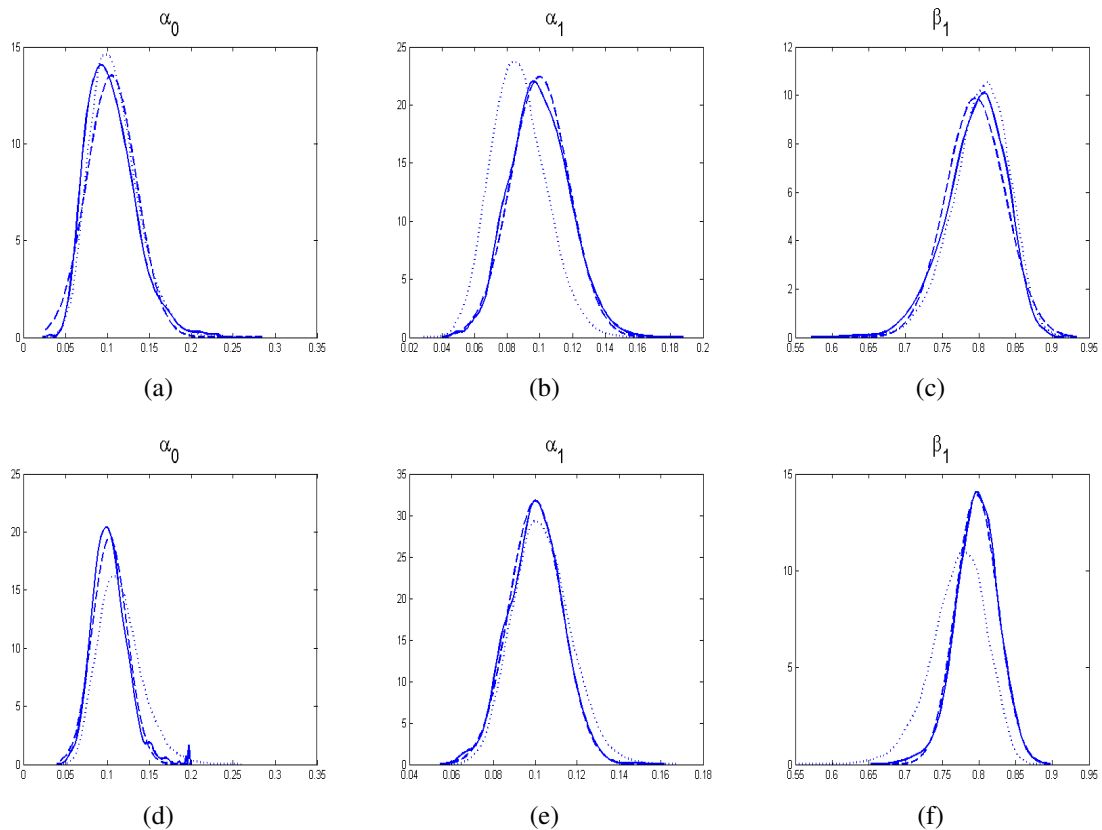


Abbildung 4.17: Randverteilungen der ML-Schätzer bei unterschiedlichen Stichprobenumfängen im Vergleich zu der a posteriori Verteilung der MCMC-Schätzung eines einzelnen Datensatzes ( $n=2500$  und  $5000$ )

In den beiden Abbildungen 4.16 und 4.17 sind die Ergebnisse der 1000 Replikationen und die a posteriori Verteilung der MCMC-Schätzung für einen Datensatz abgebildet. Die durchgezogene Linie stellt die Kerndichteschätzung der Verteilung dar, ' - ' ist eine angepasste Normalverteilung mit der aus den 1000 ML-Parameterschätzern geschätzten Varianz und  $\bar{\theta}_i$  als Erwartungswert, und die gepunktete Linie ist die Kerndichteschätzung der Werte der a posteriori Verteilung.

### Kritische Anmerkungen

An dieser Stelle müssen einige Aspekte dieses Abschnitts näher erläutert werden.

Die verwendete a priori Verteilung von  $\alpha_0$ :  $\alpha_0 \sim \log N(-2.0, 5)$ , ist nicht flach. Bei diversen Simulationen unterschiedlicher Datensätze mit Stichprobenumfängen von 1000 und mehr hat sie sich als sehr günstig erwiesen. Auch bei einem Stichprobenumfang von 500 ist sie meist gut geeignet. Bei den Stichprobenumfängen 100 und 250 kann sie

allerdings die Daten überlagern, weil ihr Einfluss auf die a posteriori Verteilung, relativ betrachtet, zu groß ist. Dies führt in Einzelfällen dazu, dass der MAP-Schätzer für  $\alpha_0$  gegen Null strebt, was auch die Schätzergebnisse der anderen Parameter beeinflusst. Ist dies der Fall, so lässt sich das leicht an der a posteriori Verteilung erkennen, da sie bei den Parametern  $\alpha_0$  und  $\beta_1$  dann sehr flach ist und nahezu den gesamten Bereich zwischen 0 und 1 gleichermaßen abdeckt und geht meist mit einer sehr hohen Akzeptanzrate einher. Daher ist es zur Bestimmung eines MAP-Schätzers, der mit dem ML-Schätzer verglichen werden soll, günstig eine flache a priori Verteilung für  $\alpha_0$  zu wählen. Allerdings wird die a posteriori Verteilung dadurch wie beschrieben beeinflusst. Auf der anderen Seite führt die Verwendung von  $\frac{1}{\alpha_0}$  oder  $\alpha_0 \sim \log N(-2, 5)$  als a priori Verteilung für  $\alpha_0$  zu einer a posteriori Verteilung, die der Verteilung der 1000 ML-Schätzer ähnlicher ist, und damit auch zu besseren Schätzern für Mittelwert und Median der a posteriori Verteilung. Für den Vergleich der a posteriori Verteilung mit der Verteilung der ML-Schätzer werden daher keine flachen a priori Verteilungen benutzt.

Die für den Parameter  $\alpha_0$  beschriebenen Erkenntnisse lassen sich ebenso auf den Parameter  $\sigma_0^2$  übertragen. Auch hier gilt, dass in kleinen Stichproben der Schätzwert von  $\sigma_0^2$  alle anderen Parameterschätzer überlagern kann. Ebenso kann die a priori Verteilung von  $\sigma_0^2$  der Parameter gegen 0 streben lassen. Daher ist es notwendig sich in kleinen Stichproben die Daten und die Schätzergebnisse detailliert anzusehen. Dies erschwert die Durchführung einer Simulationsstudie, analog zu der für die ML-Schätzung, erheblich.

### **Zusammenfassung**

Die Ergebnisse in diesem Abschnitt bestätigen, dass sowohl ML-Schätzung als auch MCMC-Schätzung mit steigendem Stichprobenumfang erwartungsgemäß bessere Ergebnisse liefern. Die Frage, ab welchem Stichprobenumfang von der asymptotischen Normalverteilung der ML-Schätzer bei GARCH-Modellen ausgegangen werden kann, kann anhand dieser kurzen Simulationsstudie nicht abschließend geklärt werden. Die Empfehlung von Hwang und Pereira, Stichproben ab einer Länge von 500 zu verwenden, wird gestützt. Allerdings zeigen die Abbildungen A.3 bis A.5 im Anhang A.2, dass selbst bei  $n = 1000$  die Verteilungen von  $\alpha_0$  und  $\beta_1$  nicht normalverteilt sind.

Die a posteriori Verteilungen einer einmaligen MCMC-Schätzung bilden die Verteilungen der ML-Parameter bereits sehr gut wieder. Insbesondere gilt dies für kleine Stichprobenumfänge, bei denen die Verteilungen der ML-Schätzer deutlich nicht normalverteilt sind.

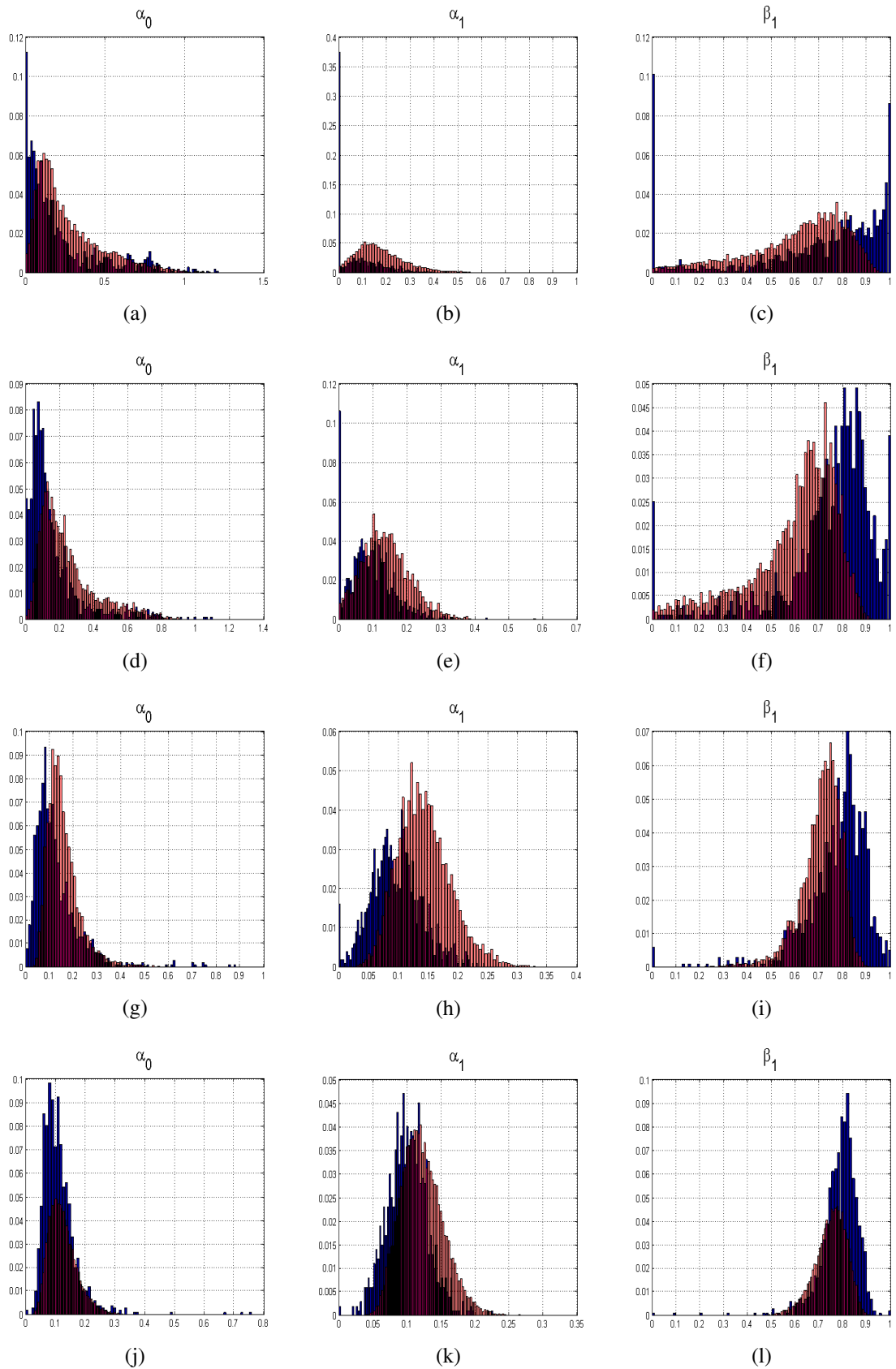


Abbildung 4.18: Vergleich der Histogramme der Verteilung der ML-Schätzer und der a posteriori Verteilung auf Basis eines Datensatzes ( $n=100, 250, 500$  und  $1000$ )

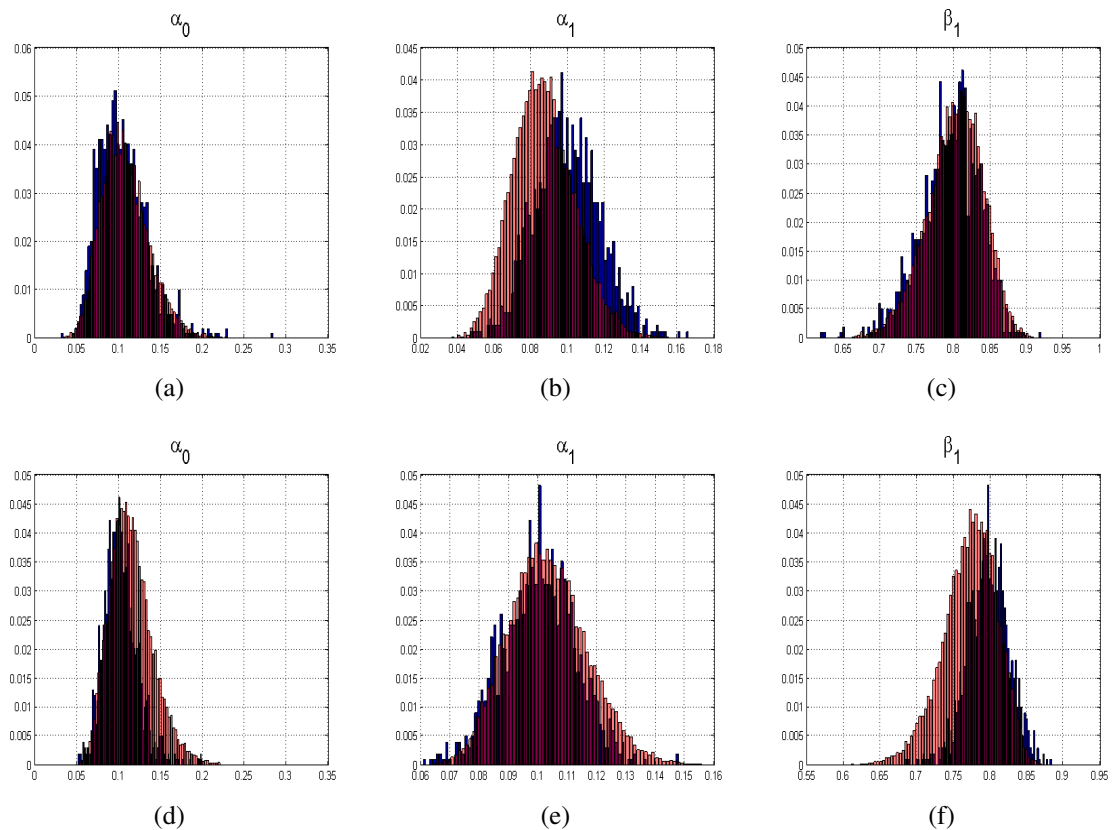


Abbildung 4.19: Vergleich der Histogramme der Verteilung der ML-Schätzer und der a posteriori Verteilung auf Basis eines Datensatzes ( $n=2500$  und  $5000$ )

### Simulationsstudie zur MCMC-Schätzung

In einer weiteren Simulationsstudie werden die Ergebnisse beider Schätzmethode bei 1000 Replikationen verglichen. Für die Studie werden mit der Routine „garchsim“ aus der GARCH-Toolbox 1000 Datensätze mit den verschiedenen Stichprobenumfängen erzeugt<sup>4</sup>. Anschließend erfolgt die Schätzung der Parameter durch die ML-Methode mit der „tarch“-Routine aus der MFE-Toolbox. Die MCMC-Schätzung erfolgt wie oben beschrieben. Zusätzlich werden die Schätzer für die Varianz der Parameterschätzer bei beiden Schätzmethode berechnet. Dies wird für die Stichprobenumfänge  $n = 100, 250, 500, 1000, 2500$  und  $5000$  gemacht.

### Interpretation der Ergebnisse

Die Tabellen 4.10 bis 4.13 zeigen die Ergebnisse der Simulationsstudie für Stichprobe-

<sup>4</sup>In dieser Studie wird ein Einfluss der Toolbox durch Simulation und Schätzung mittels der gleichen Toolbox ausgeschlossen.

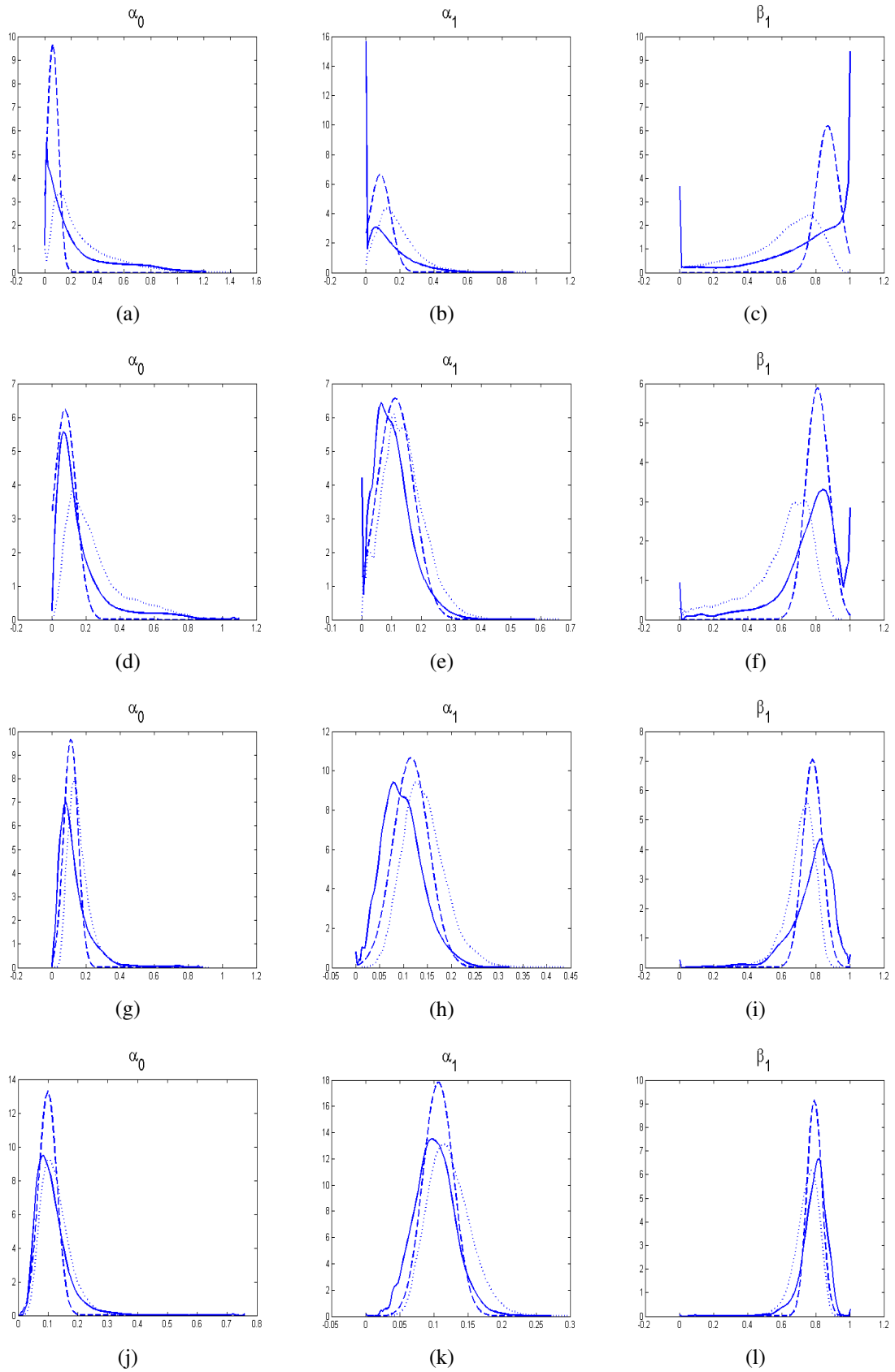


Abbildung 4.20: Vergleich der Randverteilungen der ML-Schätzung und einer MCMC-Schätzung an einem Datensatz ( $n=100, 250, 500$  und  $1000$ )

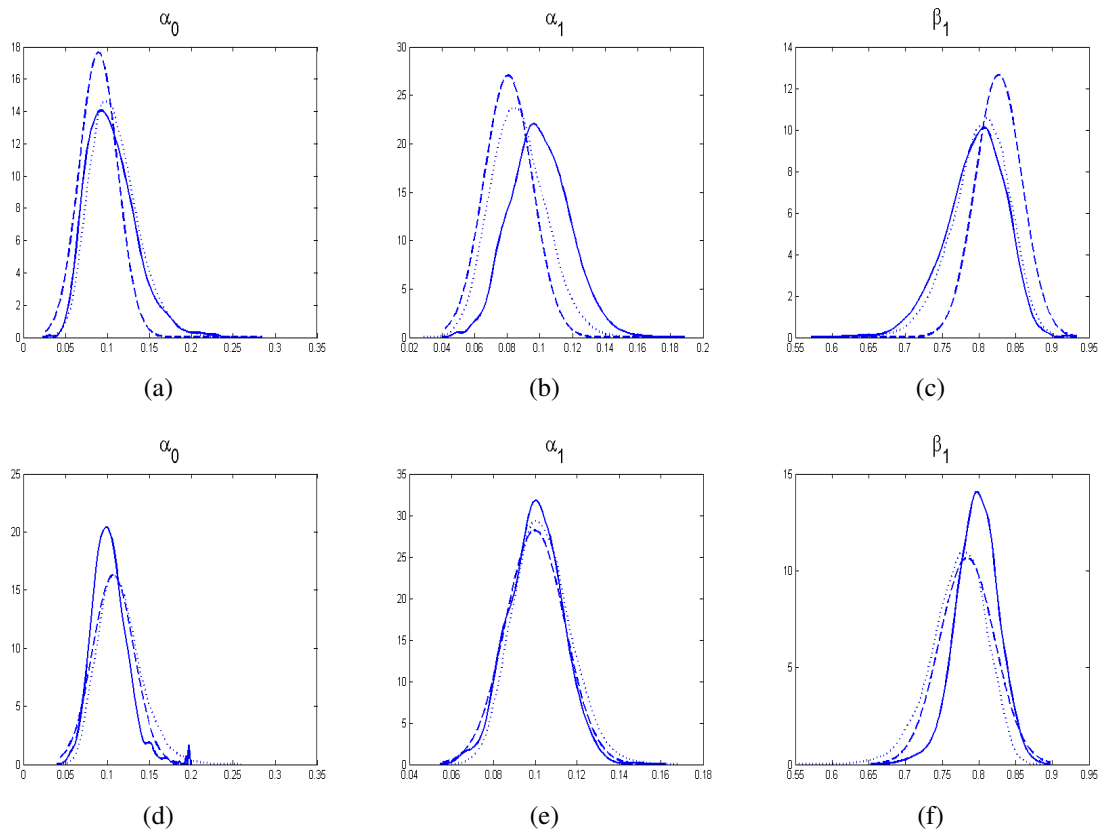


Abbildung 4.21: Vergleich der Randverteilungen der ML-Schätzung und einer MCMC-Schätzung an einem Datensatz ( $n=2500$  und  $5000$ )

numfänge bis 1000. Der Vergleich der beiden Schätzmethode erfolgt über den mean square error (MSE), die geschätzte Varianz der Parameterschätzer und die Verzerrung (Bias).

Besonders auffällig ist der extrem niedrige Bias der MAP-Schätzer bei einem Stichprobenumfang von  $n=100$ . Dies liegt an der Häufigkeit mit der  $\alpha_0$  nahe 0 ist. Der Bias für Mittelwert und Median sind sehr hoch. Auffallend gut ist die Schätzung der Varianz der Schätzer über die Varianz der a posteriori Verteilung. Dies ist schon bei dem Vergleich der MCMC-Schätzung anhand eines Datensatzes aufgefallen.

Die hier präsentierten Ergebnisse sind extrem unzuverlässig. Die 1000 MCMC-Simulationen wurden in unterschiedlichen Varianten wiederholt. Dabei wurden unterschiedliche a priori Verteilungen für  $\alpha_0$  und  $\sigma_0^2$  verwendet oder die Simulation mit unterschiedlichen Startwerten für  $\sigma_0^2$  gestartet.  $\sigma_0^2$  wurde als Parameter geschätzt oder nicht. Es zeigt sich, dass keine Konstellation gefunden werden konnte, welche zu durchgängig guten Ergebnissen für alle 1000 Datensätze und alle Stichprobenumfänge führt. In großen



| Schätzer        |         | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ |
|-----------------|---------|------------------|------------------|-----------------|
| ML              | Bias    | 0.1254           | -0.0091          | -0.1251         |
|                 | Varianz | 4.7341           | 0.1284           | 4.0539          |
|                 | MSE     | 4.8148           | 0.1388           | 4.1654          |
| MCMC-MAP        | Bias    | 0.0039           | -0.0062          | 0.0001          |
| MCMC-Mittelwert | Bias    | 0.3174           | 0.0987           | -0.3829         |
|                 | Varianz | 0.0482           | 0.0146           | 0.0440          |
|                 | MSE     | 0.1606           | 0.0270           | 0.2110          |
| MCMC-Median     | Bias    | 0.2982           | 0.0789           | -0.3876         |

Tabelle 4.10: Stichprobenumfang: 100

| Schätzer        |         | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ |
|-----------------|---------|------------------|------------------|-----------------|
| ML              | Bias    | 0.0577           | -0.0065          | -0.0551         |
|                 | Varianz | 0.0404           | 0.0110           | 0.0683          |
|                 | MSE     | 0.0686           | 0.0150           | 0.1106          |
| MCMC-MAP        | Bias    | 0.0142           | -0.0115          | -0.0050         |
| MCMC-Mittelwert | Bias    | 0.2411           | 0.0503           | -0.2861         |
|                 | Varianz | 0.0331           | 0.0054           | 0.0372          |
|                 | MSE     | 0.1110           | 0.0111           | 0.1376          |
| MCMC-Median     | Bias    | 0.2190           | 0.0404           | -0.2743         |

Tabelle 4.11: Stichprobenumfang: 250

Stichproben führen beide Schätzmethoden zu ähnlichen Punktschätzern und Schätzern für die Varianz der Parameter, wie schon in Abschnitt 4.3 beschrieben. Hierfür sind relativ allgemeine Spezifikationen ausreichend. In kleinen Stichproben hängen die Ergebnisse stark von den Spezifikationen ab. Durch die Verwendung ausschließlich flacher a priori Verteilungen erhält man für den MCMC-MAP die ML-Schätzwerte, allerdings beeinflusst dies die Schätzung für die Varianz, die dann meist überschätzt wird.

Bislang wurden die Simulationen bei Annahme der Normalverteilung durchgeführt. In Abschnitt 2.3 wurden bereits andere Verteilungsannahmen diskutiert. Daher werden in diesem Abschnitt für das GARCH(1,1)-Modell Simulationen mit der  $t$ - bzw.  $GED$ -Verteilung durchgeführt und besonders auf die Ergebnisse für den Parameter  $\nu$  der Ver-

| Schätzer        |         | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ |
|-----------------|---------|------------------|------------------|-----------------|
| ML              | Bias    | 0.0323           | -0.0006          | -0.0331         |
|                 | Varianz | 0.0113           | 0.0021           | 0.0184          |
|                 | MSE     | 0.0238           | 0.0039           | 0.0370          |
| MCMC-MAP        | Bias    | 0.0195           | -0.0032          | -0.0165         |
| MCMC-Mittelwert | Bias    | 0.1483           | 0.0330           | -0.1802         |
|                 | Varianz | 0.0190           | 0.0024           | 0.0239          |
|                 | MSE     | 0.0560           | 0.0051           | 0.0734          |
| MCMC-Median     | Bias    | 0.1269           | 0.0276           | -0.1626         |

Tabelle 4.12: Stichprobenumfang: 500

| Schätzer        |         | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ |
|-----------------|---------|------------------|------------------|-----------------|
| ML              | Bias    | 0.0426           | 0.0012           | -0.0395         |
|                 | Varianz | 0.0089           | 0.0011           | 0.0120          |
|                 | MSE     | 0.0181           | 0.0018           | 0.0227          |
| MCMC-MAP        | Bias    | 0.0014           | -0.0091          | 0.0095          |
| MCMC-Mittelwert | Bias    | 0.1069           | 0.0194           | -0.1194         |
|                 | Varianz | 0.0091           | 0.0012           | 0.0116          |
|                 | MSE     | 0.0395           | 0.0022           | 0.0487          |
| MCMC-Median     | Bias    | 0.0980           | 0.0167           | -0.1125         |

Tabelle 4.13: Stichprobenumfang: 1000

teilung geachtet.

### Annahme einer $t$ -Verteilung

Zunächst werden Simulationen für ein GARCH-Modell mit  $t$ -verteilten Residuen durchgeführt. Der zu schätzende Parametervektor schließt also nun den Parameter der Verteilung mit ein und lautet  $\theta = (\alpha_0, \alpha_1, \beta_1, \sigma_0^2, \nu)$ . Die Daten (y5000garch\_t) wurden mit den Parameterwerten:

$$\alpha_0 = 0.1, \alpha_1 = 0.1, \beta_1 = 0.8 \text{ und } \nu = 6$$

erzeugt. Die Spezifikationen der GARCH-Parameter entsprechen den genannten Angaben und für den Parameter  $\nu$  wird der Startwert auf 5 gesetzt und die Varianz der Vorschlags-

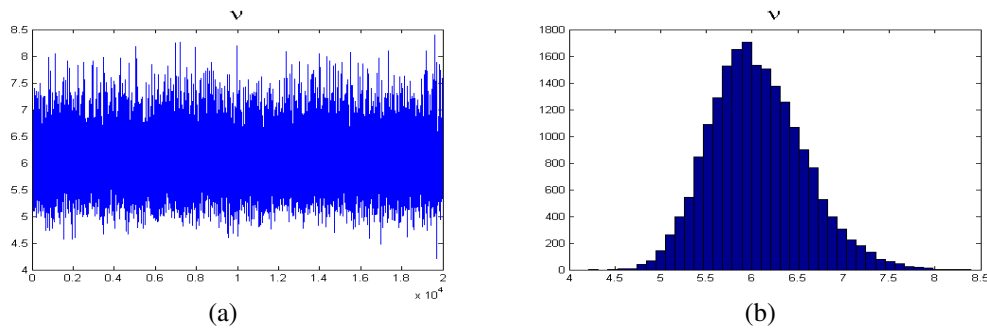


Abbildung 4.22: Markovkette und Histogramm für die Anzahl der Freiheitsgrade einer  $t$ -Verteilung

dichte von  $\nu$  beträgt 0.3.

| Schätzung       | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $\sigma_0^2$ | $\nu : 6$ | Log-L      |
|-----------------|------------------|------------------|-----------------|--------------|-----------|------------|
| ML-Matlab (t)   | 0.0896           | 0.1187           | 0.8068          | —            | 3.7171    | -6504.9    |
| ML-MFE (t)      | 0.0973           | 0.1027           | 0.8024          | —            | 6.0397    | -6795.4    |
| MCMC-MAP        | 0.1007           | 0.1053           | 0.7974          | 0.3119       | 6.0060    | -6775.0141 |
| MCMC-Mittelwert | 0.1030           | 0.1066           | 0.7942          | 2.1834       | 6.0601    |            |
| MCMC-Median     | 0.1017           | 0.1059           | 0.7954          | 1.5727       | 6.0196    |            |

Tabelle 4.14: GARCH $_t$ -Modell mit y5000garch\_t

Die Darstellung der Markovketten und Histogramme für die Parameter  $\alpha_0$ ,  $\alpha_1$  und  $\beta_1$  entsprechen denen aus den vorigen Abschnitten und werden daher nicht nochmals aufgeführt. Die Ergebnisse für die Anzahl der Freiheitsgrade der  $t$ -Verteilung stellt Abbildung 4.23 dar.

Analog zu dem vorigen Abschnitt wird eine Simulation durchgeführt, bei der jeweils 1000 Datensätze der Länge  $n=100, 250, 500, 1000, 2500$  und  $5000$  erzeugt und mit der ML-Methode geschätzt werden. Abbildung 4.23 zeigt die Ergebnisse der 1000 Replikationen für den Parameter  $\nu$  der  $t$ -Verteilung. Den Abbildungen aus Abschnitt 4.3.2 entsprechend, stellt die durchgezogene Linie eine Kerndichteschätzung der Verteilung des Parameters  $\nu$  dar, '- -' ist eine angepasste Normalverteilung mit der aus den 1000 Parameterschätzern geschätzten Varianz und  $\bar{\theta}_i$  als Erwartungswert, und die gepunktete Linie stellt die angepasste Normalverteilung dar, die sich mit dem Durchschnitt der robusten Varianzschätzer von Bollerslev und Wooldridge ergibt. Es ist klar zu erkennen,

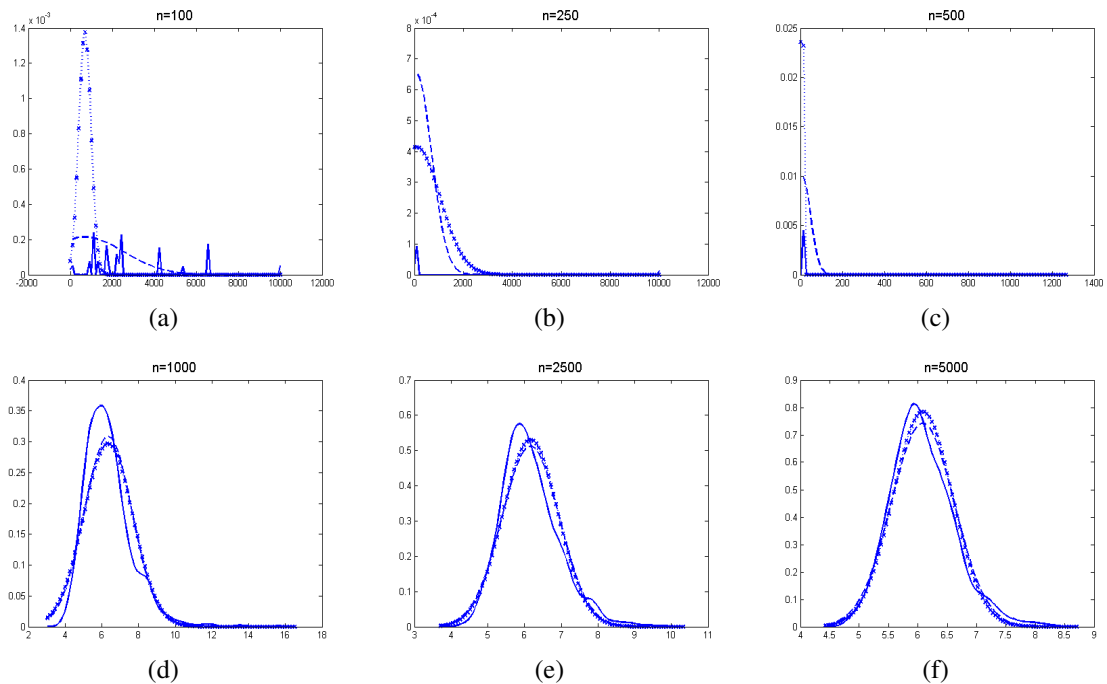


Abbildung 4.23: Verteilung des ML-Schätzers für die Anzahl der Freiheitsgrade einer  $t$ -Verteilung bei unterschiedlichen Stichprobenumfängen

dass es in den Ergebnissen einen Sprung gibt zwischen dem Stichprobenumfang  $n = 500$  und  $n = 1000$ . Bei einem Stichprobenumfang von  $n = 1000$  bilden die angepassten Normalverteilungen die Verteilung der 1000 Werte von  $\nu$  akzeptabel wieder.

### Annahme einer $GED$ -Verteilung

Die Berechnungen für ein  $\text{GARCH}_{GED}(1,1)$ -Modell mit  $GED$ -verteilten Residuen erfolgen analog. Die Daten (y5000garch\_ged) wurden mit den Parameterwerten:

$$\alpha_0 = 0.1, \alpha_1 = 0.1, \beta_1 = 0.8 \text{ und } \nu = 1.5$$

erzeugt. Die Spezifikationen der GARCH-Parameter entsprechen den genannten Angaben und für den Parameter  $\nu$  wird der Startwert auf 2 gesetzt und die Varianz der Vorschlagsdichte von  $\nu$  beträgt 0.005.

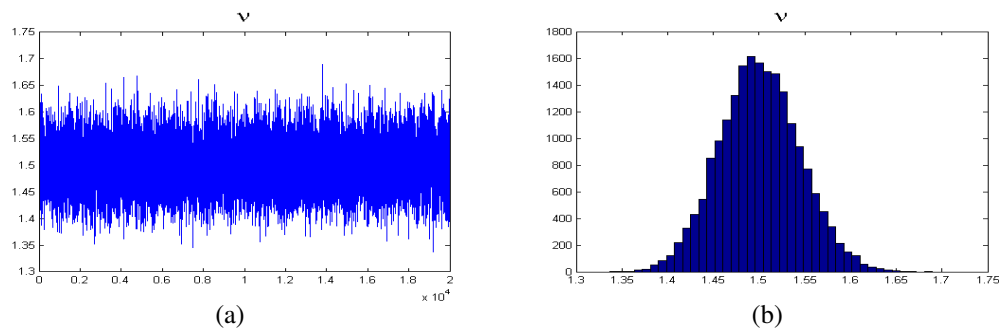


Abbildung 4.24: Markovkette und Histogramm für den Parameter  $\nu$  einer  $GED$ -Verteilung

| Schätzung       | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $\sigma_0^2$ | $\nu : 1.5$ | Log-L      |
|-----------------|------------------|------------------|-----------------|--------------|-------------|------------|
| ML-MFE (GED)    | 0.0881           | 0.0904           | 0.8168          | —            | 1.5020      | -6783.2    |
| MCMC-MAP        | 0.0721           | 0.0708           | 0.8091          | 0.0107       | 1.5067      | -6754.3521 |
| MCMC-Mittelwert | 0.0737           | 0.0726           | 0.8043          | 0.3698       | 1.5006      |            |
| MCMC-Median     | 0.0722           | 0.0722           | 0.8063          | 0.1980       | 1.4999      |            |

Tabelle 4.15:  $GARCH_{GED}$ -Modell mit y5000garch\_ged

Wiederum werden 1000 Replikationen durchgeführt, bei denen jeweils Datensätze der Länge  $n=100, 250, 500, 1000, 2500$  und  $5000$  erzeugt und geschätzt werden. Die Ergebnisse für den Scharparameter  $\nu$  der  $GED$ -Verteilung zeigt Abbildung 4.25. Man kann sehen, dass die Ergebnisse für den Schätzer  $\nu$  im Vergleich zu den GARCH-Parametern recht früh gegen die Normalverteilung konvergieren. Sehr interessant ist auch der Vergleich mit Abbildung 4.23, bei der  $GED$ -Verteilung ist die Annahme der Normalverteilung des Parameterschätzers für  $\nu$  augenscheinlich erheblich früher erfüllt als bei der  $t$ -Verteilung. Dies kann bei der Wahl der Verteilung im Rahmen kleiner Stichprobenumfänge durchaus relevant sein. Während der Nachteil der  $GED$ -Verteilung ist, dass sie nicht gleichzeitig Spitzgipfligkeit und schwere Ränder abbilden kann, ist sie in kleinen Stichproben der  $t$ -Verteilung evtl. vorzuziehen.

### Überprüfung der Normalverteilungsannahme

Selbstverständlich kann durch die Schätzung des Parameters  $\nu$  bei beiden Verteilungen auch die Annahme der Normalverteilung überprüft werden. Die folgenden beiden Beispiele zeigen das Verhalten des Parameters  $\nu$ , wenn die Normalverteilung tatsächlich vor-

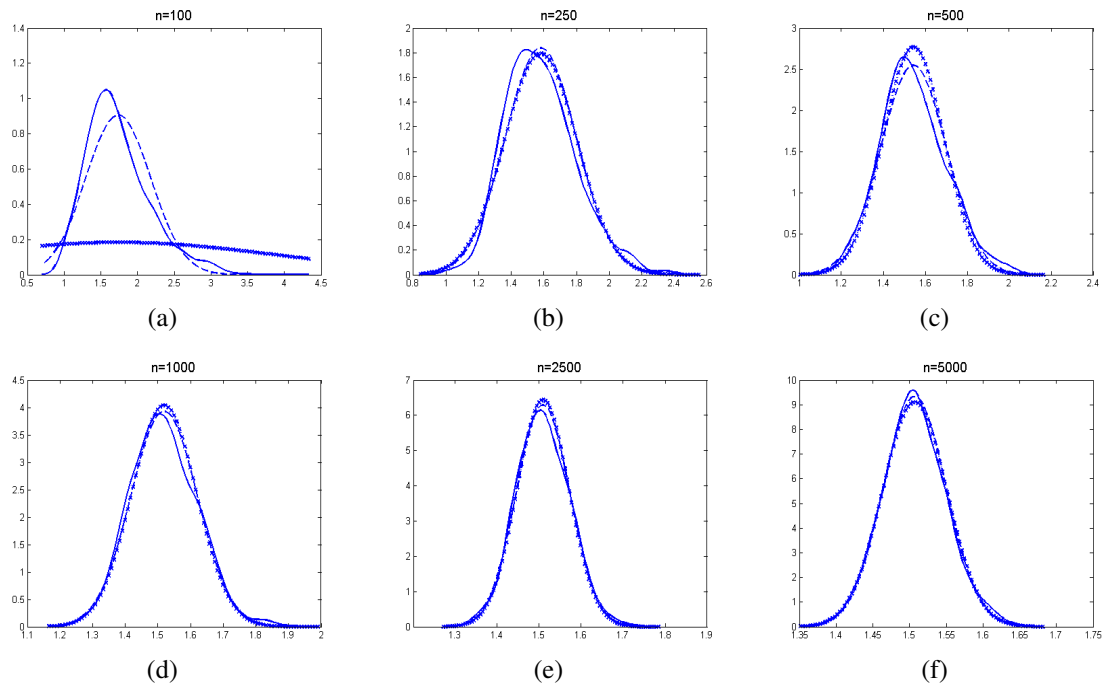


Abbildung 4.25: Verteilung des ML-Schätzers für den Parameter  $\nu$  einer  $GED$ -Verteilung bei unterschiedlichen Stichprobenumfängen

liegt. Wird bei der Schätzung eine  $t$ -Verteilung zugrunde gelegt, strebt die Markovkette schnell gegen sehr große Werte, entsprechend nimmt  $\nu$  bei der Annahme einer  $GED$ -Verteilung erwartungsgemäß Werte um 2 an. Die entsprechenden Ergebnisse zeigen die Tabellen 4.16 und 4.17.

Wie oben bereits erwähnt ist die Wahl der a priori Verteilung für den Parameter  $\nu$  der  $t$ -Verteilung von großer Bedeutung. Verschiedene Simulationen haben gezeigt, dass die Verwendung einer konstanten a priori Verteilung häufig dazu führt, dass die Markovkette von  $\nu$  immer größere Werte annimmt, damit nicht konvergiert, und dass dies die gesamte Schätzung überlagert. Die resultierenden Schätzer für die GARCH-Parameter sind dann unbrauchbar. Daher wird hier der Vorschlag von Miazhyńska et al. (2006) [75]  $\nu \sim \text{Exp}(0.1)$  aufgegriffen. Allerdings schränkt diese a priori Verteilung den Parameter  $\nu$  extrem stark auf den Bereich  $<50$  ein. Daher deuten bei Verwendung dieser a priori Verteilung Werte nahe 50 auf eine Normalverteilung hin.

| Schätzung       | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $\sigma_0^2$ | $\nu$    | Log-L   |
|-----------------|------------------|------------------|-----------------|--------------|----------|---------|
| ML-Matlab (t)   | 0.0896           | 0.1187           | 0.8068          | —            | 3.7171   | -6504.9 |
| ML-MFE (t)      | 0.1087           | 0.0966           | 0.7894          |              | 689.1869 | -6882.4 |
| MCMC-MAP        | 0.1016           | 0.0935           | 0.8004          | 0.4163       | 42.6061  | -6853.7 |
| MCMC-Mittelwert | 0.1192           | 0.1024           | 0.7737          | 1.2977       | 50.2478  |         |
| MCMC-Median     | 0.1165           | 0.1016           | 0.7768          | 0.9727       | 47.6678  |         |

Tabelle 4.16: GARCH<sub>t</sub>-Modell mit normalverteilten Residuen (unter der Annahme einer  $t$ -Verteilung geschätzt)

| Schätzung       | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $\sigma_0^2$ | $\nu$  | Log-L   |
|-----------------|------------------|------------------|-----------------|--------------|--------|---------|
| ML-Matlab (t)   | 0.0896           | 0.1187           | 0.8068          | —            | 3.7171 | -6504.9 |
| ML-MFE (t)      | 0.1084           | 0.0963           | 0.7899          | —            | 2.0352 | -6882.2 |
| MCMC-MAP        | 0.1017           | 0.0942           | 0.8013          | 0.3122       | 2.0193 | -6853.0 |
| MCMC-Mittelwert | 0.1208           | 0.1034           | 0.7744          | 1.2666       | 2.0340 |         |
| MCMC-Median     | 0.1170           | 0.1026           | 0.7781          | 0.9207       | 2.0331 |         |

Tabelle 4.17: GARCH<sub>GED</sub>-Modell mit normalverteilten Residuen (unter der Annahme einer  $GED$ -Verteilung geschätzt)

## 4.4 ML- und MCMC-Simulationen diverser GARCH-Varianten

Nachdem die Schätzung eines GARCH(1,1)-Modells in Abschnitt 4.3 besprochen wurde, erweitert dieser Abschnitt die Anwendung der beschriebenen MCMC-Schätzung auf diverse Verwandte aus der Familie der GARCH-Modelle.

### 4.4.1 Simulationen zu EGARCH-Modellen

Dieser Abschnitt behandelt das EGARCH(1,1)-Modell aus Abschnitt 2.4.3, d.h. die  $\sigma_t^2$  in Gleichung

$$z_t = \sigma_t \cdot \varepsilon_t, \quad t = 1, \dots, T,$$

| Datensatz       | $\alpha_0 : -0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.9$ | $\sigma_0^2 : 0.4413$ | $c : -0.1$ | Log-L   |
|-----------------|-------------------|------------------|-----------------|-----------------------|------------|---------|
| ML-MFE          | -0.0942           | 0.0984           | 0.9065          | 0.5203                | -0.1142    | -4586.9 |
| MCMC-MAP        | -0.0957           | 0.0961           | 0.9048          | 0.8139                | -0.1150    | -4586.9 |
| MCMC-Mittelwert | -0.1025           | 0.1051           | 0.8982          | 4.0686                | -0.1171    | -4588.7 |
| MCMC-Median     | -0.1016           | 0.1046           | 0.8991          | 2.2246                | -0.1170    | -4587.8 |

Tabelle 4.18: ML-Schätzergebnisse für ein EGARCH-Modell mit simulierten Daten

werden mit der Gleichung (2.14):

$$\ln(\sigma_t^2) = \alpha_0 + \beta_1 \cdot \ln(\sigma_{t-1}^2) + c \cdot \varepsilon_{t-1} + \alpha_1(|\varepsilon_{t-1}| - E(|\varepsilon_{t-1}|)),$$

berechnet und die  $\varepsilon_t$  stammen aus einer Normalverteilung.

Das hier beschriebene EGARCH(1,1)-Modell hat den zu schätzenden Parametervektor  $\theta = (\alpha_0, \alpha_1, \beta_1, c, \sigma_0^2)$ .

Wie bereits in Abschnitt 2.4.3 erwähnt, unterliegen bei dem EGARCH(1,1)-Modell die Parameter der Varianzgleichung keinen Restriktionen. Insbesondere für den Parameter  $\alpha_0$  wirkt sich dies auf die Wahl der a priori Verteilung aus. Die Schätzung wird deshalb mit einer Konstanten als a priori für  $\alpha_0$  durchgeführt. Das Fehlen von Parameterrestriktionen hat zur Folge, dass die Parameter auch negative Werte annehmen können, ohne dabei zu negativen Varianzen zu führen. Allerdings haben die durchgeführten Simulationen gezeigt, dass Werte von  $\sigma_t^2$  nahe 0 aufgrund des Logarithmus leicht zu numerischen Problemen führen können. Bei der Schätzung von EGARCH-Modellen mit der MCMC-Methode haben sich kleinere Varianzen der Vorschlagsdichte als günstig erwiesen, um das Ausmaß an numerischen Problemen zu reduzieren. Generell sinkt das Ausmaß der numerischen Probleme mit zunehmendem Stichprobenumfang. Um den Einfluss der fehlenden Parameterrestriktionen zu zeigen wird ein Datensatz mit 5000 Werten mit den Parametern

$$\alpha_0 = -1, \alpha_1 = 0.1, \beta_1 = 0.9, \text{ und } c = -0.1$$

erzeugt.

Die Ergebnisse der ML- und der MCMC-Schätzung für diesen Datensatz zeigt Tabelle 4.18. Es zeigt sich, dass in dieser großen Stichprobe beide Schätzmethoden die vorgegebenen Parameter gut treffen und die Schätzer sowie die Werte der Log-Likelihood nahezu übereinstimmen.

Um die a posteriori Verteilungen der MCMC-Schätzungen für ein EGARCH-Modell mit denen der ML-Schätzung vergleichen zu können, werden 1000 Datensätze mit  $n = 100$ ,



250, 500, 1000, 2500 und 5000 erzeugt, analog zu dem Vorgehen in Abschnitt 4.3.2, allerdings nur für eine Parameterkonstellation. Die Parameter dieser 1000 Datensätze sind

$$\alpha_0 = 0.05, \alpha_1 = 0.1, \beta_1 = 0.9, \text{ und } c = -0.1.$$

Diese Werte orientieren sich an den Parameterschätzern für eine Zeitreihe des DAX. Die Abbildungen 4.26 bis 4.34 zeigen die Kerndichteschätzungen für die Ergebnisse der einzelnen ML-Parameterschätzer. Die durchgezogene Linie stellt eine Kerndichteschätzung der Verteilung der 1000 ML-Parameterschätzer dar, '-' ist eine angepasste Normalverteilung mit der aus den 1000 Parameterschätzern geschätzten Varianz und  $\bar{\theta}_i$  als Erwartungswert, und die gepunktete Linie stellt die angepasste Normalverteilung dar, die sich mit dem Durchschnitt der robusten Varianzschätzer von Bollerslev und Wooldridge ergibt.

Für die Stichprobenumfänge 250, 500 und 1000 werden einzelne Datensätze generiert, wiederum erfolgt die Auswahl der Datensätze selektiv. Es werden drei Datensätze verwendet, welche die vorgegebenen Parameter gut repräsentieren.

Bei den Ergebnissen in Tabelle 4.19 fällt auf, dass die MCMC-MAP-Schätzer und die ML-Schätzer für die drei Datensätze sehr ähnliche Werte liefern, auch für den Wert von  $\sigma_0^2$ , die a posteriori Mittelwerte und Mediane weichen deutlich ab, und deuten auf starke Asymmetrie hin, die auch in den Histogrammen der Werte der a posteriori Verteilung erkennbar ist. Sehr drastisch fallen die Ergebnisse für  $\sigma_0^2$  aus, hierfür werden extrem große Varianzen aus dem ersten Lauf geschätzt und dies führt zu enormen Sprüngen in der Markovkette. Abbildung 4.30 stellt die Histogramme der 1000 ML-Parameterschätzer gegen die a posteriori der einzelnen Datensätze für die einzelnen Parameter dar. Die a posteriori Verteilungen sind für die drei Stichprobenumfänge spaltenweise angeordnet und bilden die Verteilung der 1000 ML-Schätzer in allen Fällen gut ab. Es ist natürlich wichtig dabei zu beachten, dass hier ein einzelner Datensatz für die Simulation verwendet wurde. Wiederholungen dieser Simulation haben gezeigt, dass Datensätze, die die Parameter gut abbilden auch zu guten a posteriori Verteilungen führen, selbst in kleinen Stichproben.

#### 4.4.2 Simulationen zu GJR-GARCH-Modellen

Am Beispiel des GJR-GARCH-Modells aus Abschnitt 2.4.4 wird die Berücksichtigung eines Parameters in der Mittelwertgleichung demonstriert. Dabei stellt  $\mu$  in der Gleichung

$$z_t = \mu + \sigma_t \cdot \varepsilon_t, \quad t = 1, \dots, T,$$

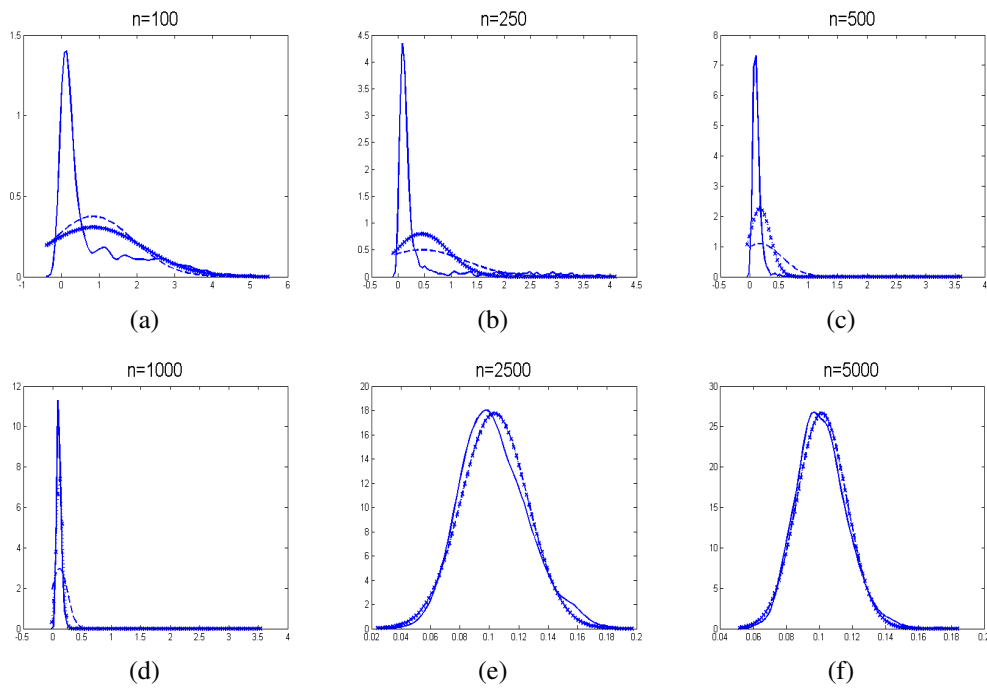


Abbildung 4.26: Verteilung des ML-Schätzers des Parameters  $\alpha_0: 0.1$  bei unterschiedlichen Stichprobenumfängen

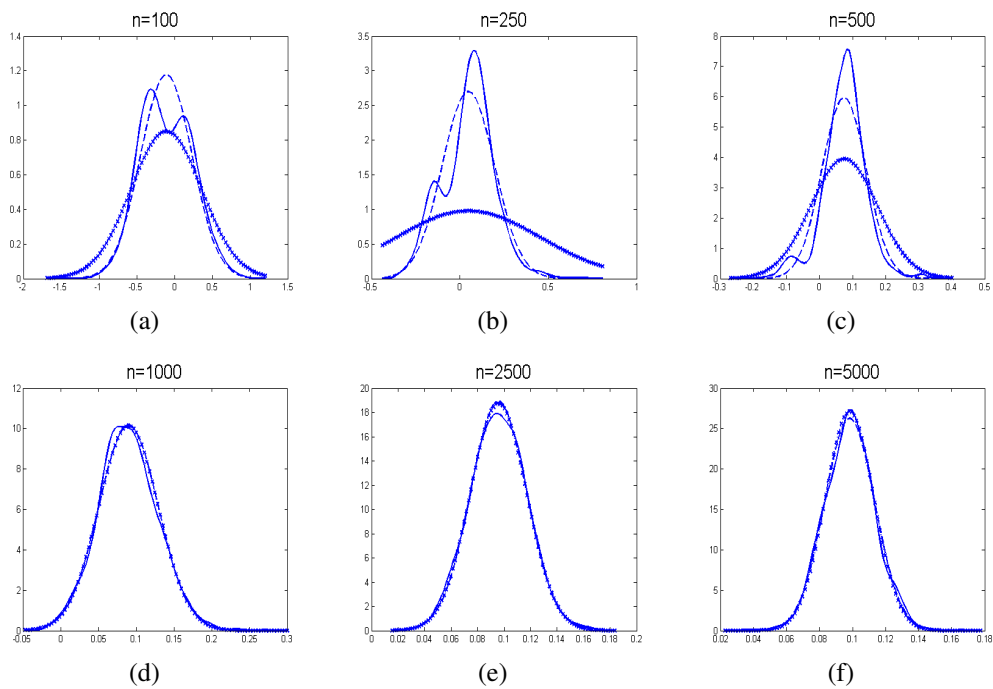


Abbildung 4.27: Verteilung des ML-Schätzers des Parameters  $\alpha_1: 0.1$  bei unterschiedlichen Stichprobenumfängen

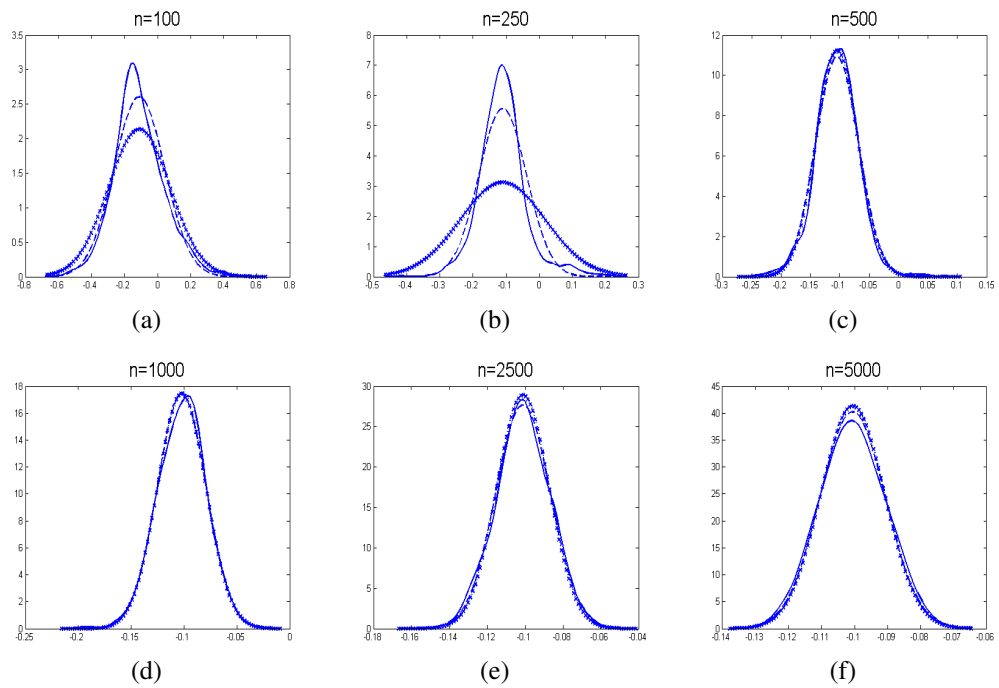


Abbildung 4.28: Verteilung des ML-Schätzers des Parameters  $c$ :  $-0.1$  bei unterschiedlichen Stichprobenumfängen

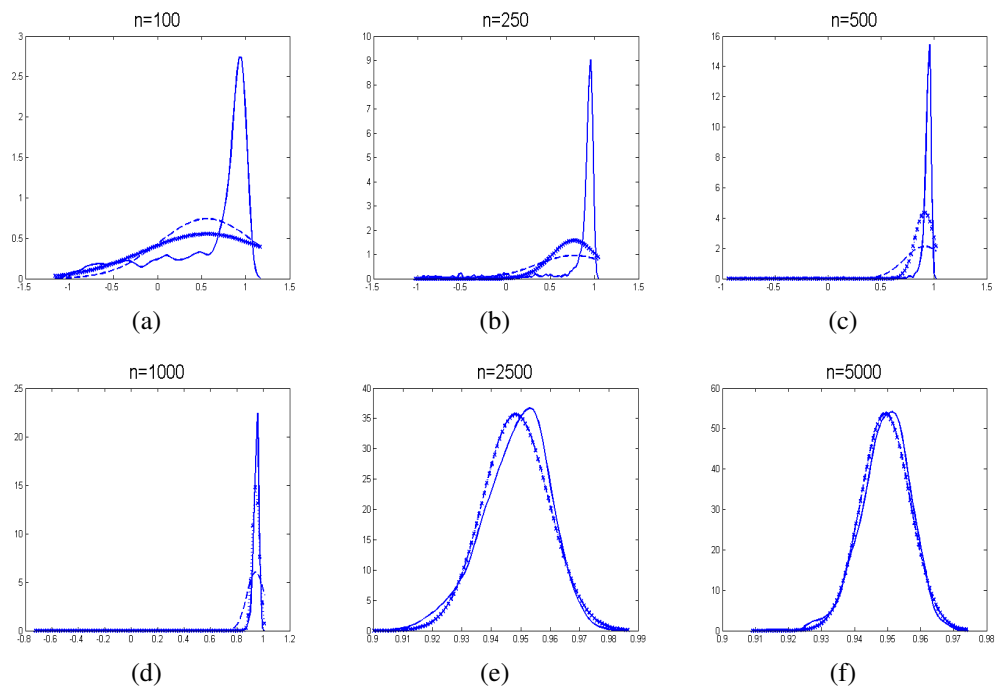


Abbildung 4.29: Verteilung des ML-Schätzers des Parameters  $\beta_1$ :  $0.90$  bei unterschiedlichen Stichprobenumfängen

| Modell            | $\alpha_0$ | $\alpha_1$ | $\beta_1$ | $\sigma_0^2$ | c       | Log-L     |
|-------------------|------------|------------|-----------|--------------|---------|-----------|
| Simulation n=250  | 0.05       | 0.1        | 0.9       | 2.2855       | -0.1    |           |
| ML-MFE            | 0.0206     | 0.0887     | 0.9542    | 1.3160       | -0.1082 | -417.3453 |
| MCMC-Modus        | 0.0206     | 0.0932     | 0.9601    | 1.4699       | -0.1086 | -417.4867 |
| MCMC-Mittelwert   | 0.3702     | 0.2681     | 0.3503    | 160.7346     | -0.1237 | -426.3570 |
| MCMC-Median       | 0.3276     | 0.2665     | 0.4058    | 65.6042      | -0.1226 | -425.5800 |
| Simulation n=500  | 0.05       | 0.1        | 0.9       | 1.2537       | -0.1    |           |
| ML-MFE            | 0.0496     | 0.1119     | 0.9107    | 3.0497       | -0.0850 | -832.9267 |
| MCMC-Modus        | 0.0523     | 0.1011     | 0.9047    | 3.7546       | -0.0853 | -832.8778 |
| MCMC-Mittelwert   | 0.2502     | 0.1866     | 0.5351    | 1153.9074    | -0.1330 | -840.3302 |
| MCMC-Median       | 0.1985     | 0.1818     | 0.6295    | 209.7080     | -0.1310 | -838.1006 |
| Simulation n=1000 | 0.05       | 0.1        | 0.9       | 1.9688       | -0.1    |           |
| ML-MFE            | 0.0469     | 0.1207     | 0.9204    | 1.1499       | -0.1067 | -1687.2   |
| MCMC-Modus        | 0.0485     | 0.1217     | 0.9180    | 1.3323       | -0.1069 | -1687.2   |
| MCMC-Mittelwert   | 0.0687     | 0.1525     | 0.8825    | 9.0617       | -0.1147 | -1690.9   |
| MCMC-Median       | 0.0642     | 0.1497     | 0.8901    | 2.7922       | -0.1140 | -1688.3   |

Tabelle 4.19: EGARCH-Modell mit normalverteilten Residuen

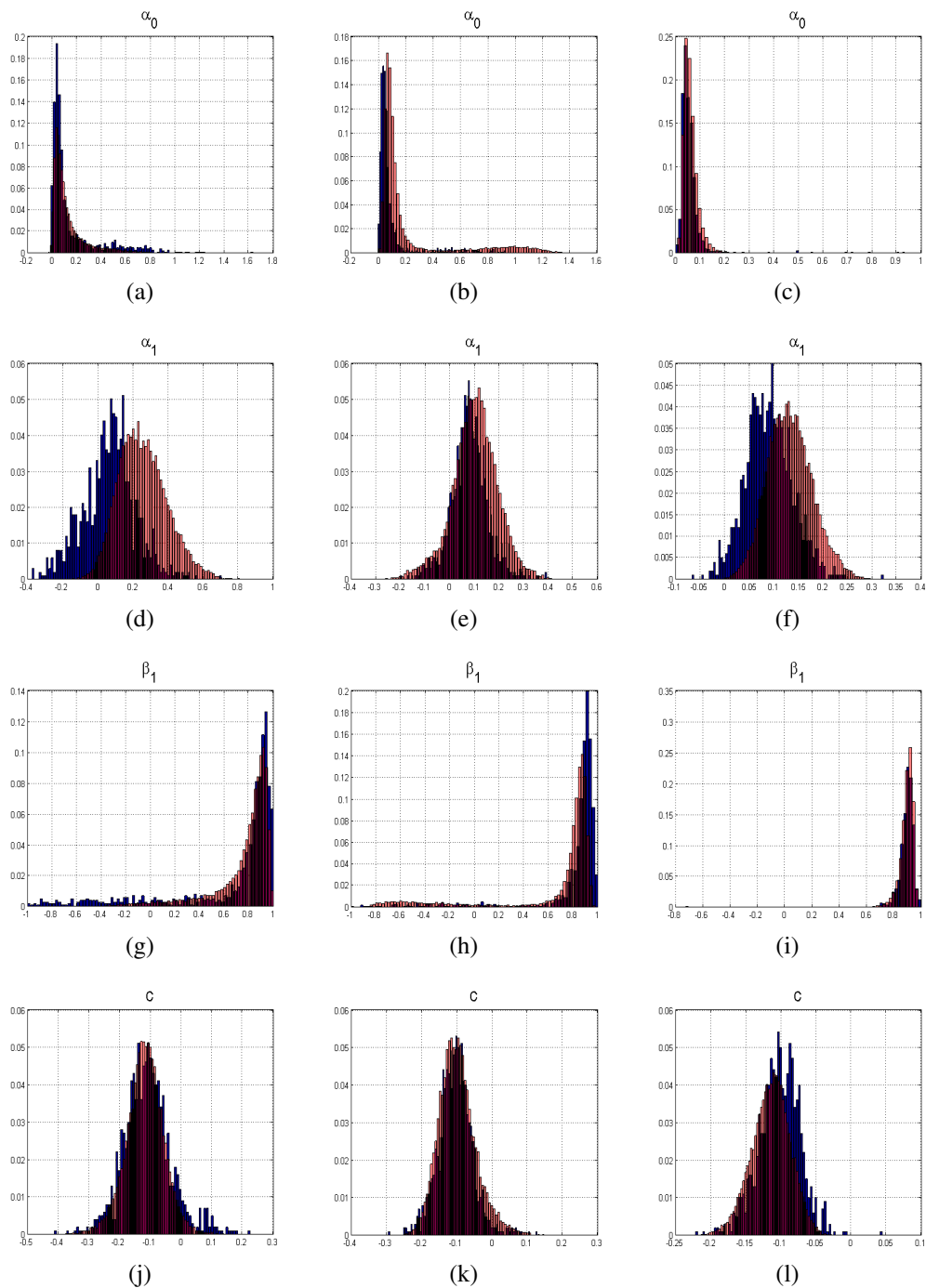


Abbildung 4.30: Verteilung der ML-Schätzer der Parameter bei unterschiedlichen Stichprobenumfängen

einen zu schätzenden Parameter dar, die  $\varepsilon_t$  stammen aus einer Normalverteilung. Mit der Varianzgleichung

$$\sigma_t^2 = \alpha_0 + [\alpha_1 Z_t^2 - c \cdot (Z_t^+)^2] + \beta_1 \sigma_{t-1}^2 \quad \text{mit } t = 1, \dots, T,$$

wird ein Datensatz der Länge  $t=5000$  erzeugt. Dabei wurden die Parameter folgendermaßen gesetzt:

$$\alpha_0 = 0.1, \alpha_1 = 0.1, \beta_1 = 0.8, c = 0.1 \text{ und } \mu = 0.5$$

Anschließend werden die Parameter mit der GARCH-Toolbox und mit der MCMC-Methode geschätzt. Tabelle 4.20 zeigt die Ergebnisse der Schätzungen. Bei dem vorliegenden Datensatz hat sich gezeigt, dass der Wert von  $\sigma_0^2$  völlig unerheblich für die Bestimmung der a posteriori Verteilung war und daher keine eindeutigen Ergebnisse lieferte, daher wird in der Tabelle auf die Präsentation der Ergebnisse von  $\sigma_0^2$  verzichtet. Es zeigt sich, dass die Ergebnisse aller Schätzungen nahezu identisch sind.

| Datensatz       | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $c : 0.1$ | $\mu : 0.5$ | Log-L   |
|-----------------|------------------|------------------|-----------------|-----------|-------------|---------|
| ML-Matlab       | 0.0942           | 0.0980           | 0.8050          | 0.0895    | 0.5095      | -8126.0 |
| MCMC-Modus      | 0.0908           | 0.0995           | 0.8068          | 0.0861    | 0.5101      | -8126.1 |
| MCMC-Mittelwert | 0.0962           | 0.1007           | 0.8016          | 0.0903    | 0.5094      | -8126.1 |
| MCMC-Median     | 0.0955           | 0.1002           | 0.8021          | 0.0900    | 0.5093      | -8126.1 |

Tabelle 4.20: ML-Schätzergebnisse (GJR-GARCH) für y5000gjrgarch

Um die Verteilungen der Parameter darzustellen wird eine Simulationsstudie nach dem Vorbild aus Abschnitt 4.3.2 durchgeführt. Dafür werden 1000 mal Daten aus dem GJR-Modell mit

$$z_t = \sigma_t \cdot \varepsilon_t, \quad t = 1, \dots, T,$$

und der oben genannten Varianzgleichung des GJR-Modells erzeugt mit  $n= 100, 250, 500, 1000, 2500$  und  $5000$ . Dabei werden die Parameter auf

$$\alpha_0 = 0.03, \alpha_1 = 0.03, \beta_1 = 0.9, \text{ und } c = 0.1$$

gesetzt. Diese Werte sind den Ergebnissen für die Schätzung eines GJR(1,1,1)-Modells für den DAX in Abschnitt 4.7 nachempfunden.

Es ist wichtig zu bemerken, dass die Daten mit der MFE-Toolbox erzeugt und anschließend auch mit dieser geschätzt werden. Der Parameter  $c$  ist in der MFE-Toolbox mit

| Modell          | $\alpha_0 : 0.03$ | $\alpha_1 : 0.03$ | $\beta_1 : 0.9$ | $\sigma_0^2 : 1.5786$ | $c : 0.1$ | Log-L   |
|-----------------|-------------------|-------------------|-----------------|-----------------------|-----------|---------|
| ML-MFE (n=250)  | 0.0231            | 0.0493            | 0.8986          | 1.3015                | 0.0800    | -391.4  |
| MCMC-Modus      | 0.0223            | 0.0424            | 0.9000          | 0.8510                | 0.0930    | -391.5  |
| MCMC-Mittelwert | 0.0866            | 0.0999            | 0.8040          | 2.2376                | 0.1103    | -393.4  |
| MCMC-Median     | 0.0740            | 0.0937            | 0.8136          | 1.5583                | 0.1017    | -392.9  |
| ML-MFE (n=500)  | 0.0274            | 0.0249            | 0.9063          | 1.0932                | 0.1091    | -767.6  |
| MCMC-Modus      | 0.0349            | 0.0299            | 0.8966          | 0.4961                | 0.1059    | -767.6  |
| MCMC-Mittelwert | 0.0664            | 0.0578            | 0.8463          | 1.0583                | 0.1184    | -769.1  |
| MCMC-Median     | 0.0559            | 0.0529            | 0.8542          | 0.7707                | 0.1164    | -768.5  |
| ML-MFE (n=1000) | 0.0344            | 0.0238            | 0.9085          | 1.5359                | 0.0937    | -1576.0 |
| MCMC-Modus      | 0.0378            | 0.0273            | 0.9023          | 1.3026                | 0.0965    | -1575.9 |
| MCMC-Mittelwert | 0.0469            | 0.0378            | 0.8860          | 2.2494                | 0.0977    | -1576.6 |
| MCMC-Median     | 0.0447            | 0.0359            | 0.8884          | 1.9098                | 0.0964    | -1576.4 |

Tabelle 4.21: GJR-GARCH-Modell mit normalverteilten Residuen

einem positiven Vorzeichen enthalten. Daher müssen bei einem Vergleich der Ergebnisse der ML-Schätzung mit denen der MCMC-Schätzung die unterschiedlichen Vorzeichen von  $c$  berücksichtigt werden. Die Ergebnisse sind in Tabelle 4.21 aufgelistet.

Die Abbildungen 4.31 bis 4.33 zeigen die grafischen Darstellungen der Verteilungen der Parameter  $\alpha_0$ ,  $\alpha_1$ ,  $c$  und  $\beta_1$ . Da die Stichprobenumfänge 2500 und 5000 unproblematisch sind werden für die Stichprobenumfänge  $n=100, 250, 500, 1000$  einzelne Datensätze erzeugt und mit der MCMC-Methode geschätzt. Die a posteriori Verteilungen werden mit den Verteilungen der Parameterschätzer der 1000 Verteilungen verglichen.

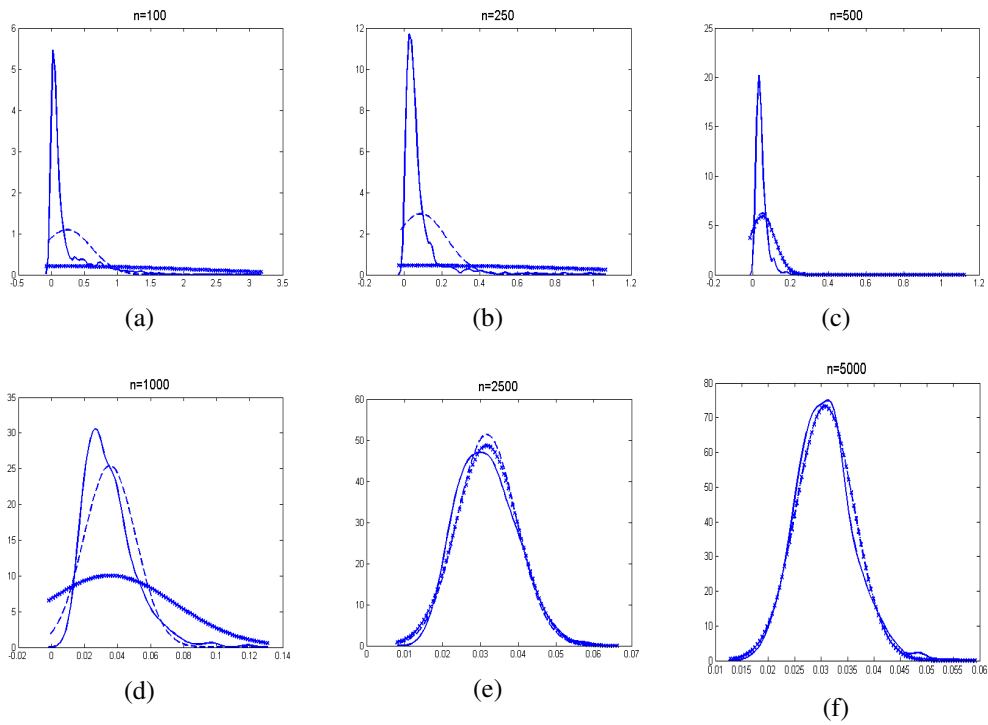


Abbildung 4.31: Verteilung des ML-Schätzers des Parameters  $\alpha_0: 0.03$  bei unterschiedlichen Stichprobenumfängen

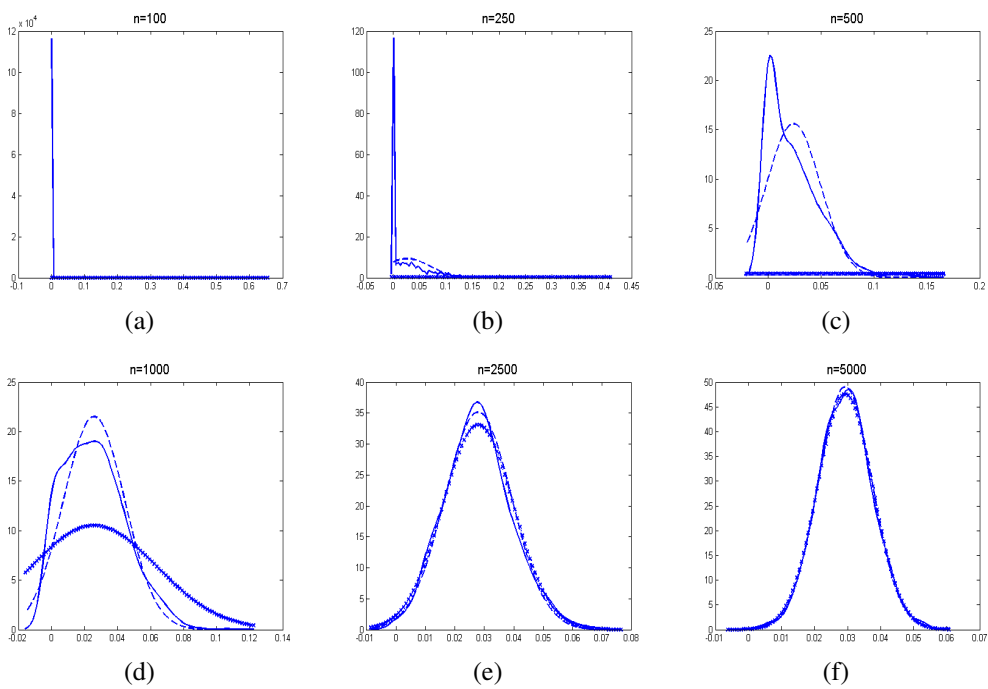


Abbildung 4.32: Verteilung des ML-Schätzers des Parameters  $\alpha_1: 0.03$  bei unterschiedlichen Stichprobenumfängen



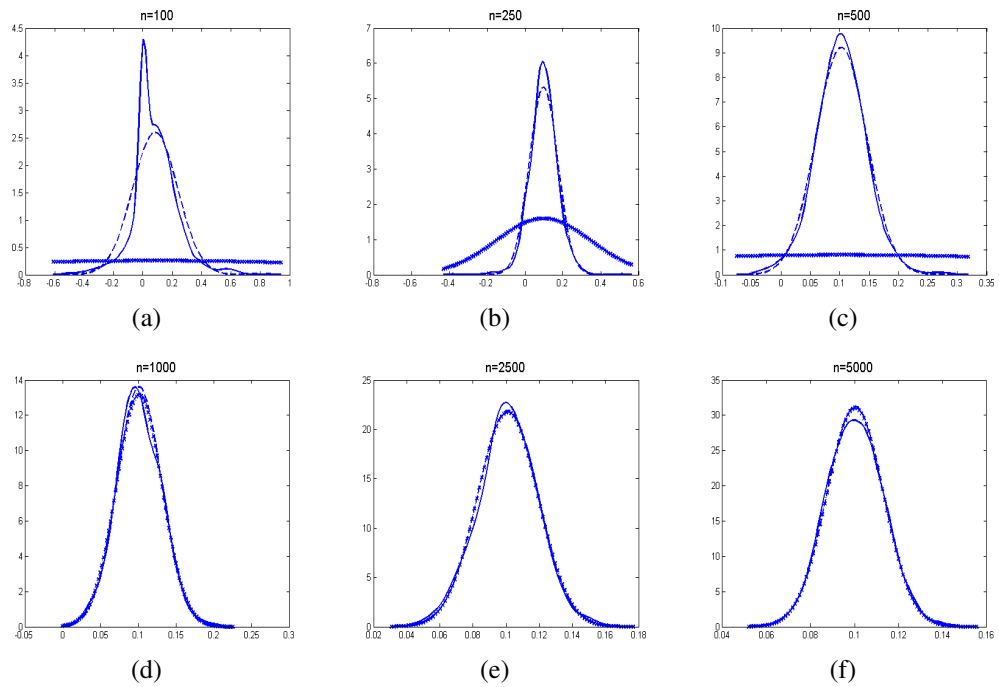


Abbildung 4.33: Verteilung des ML-Schätzers des Parameters  $c$ : 0.1 bei unterschiedlichen Stichprobenumfängen

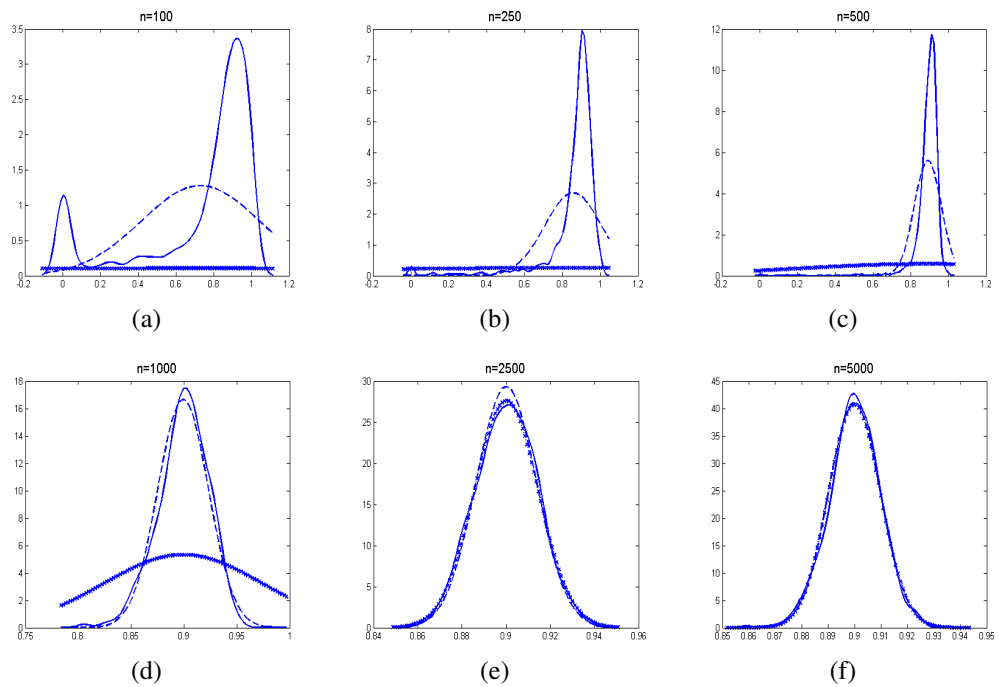


Abbildung 4.34: Verteilung der ML-Schätzers des Parameters  $\beta_1$ : 0.9 bei unterschiedlichen Stichprobenumfängen

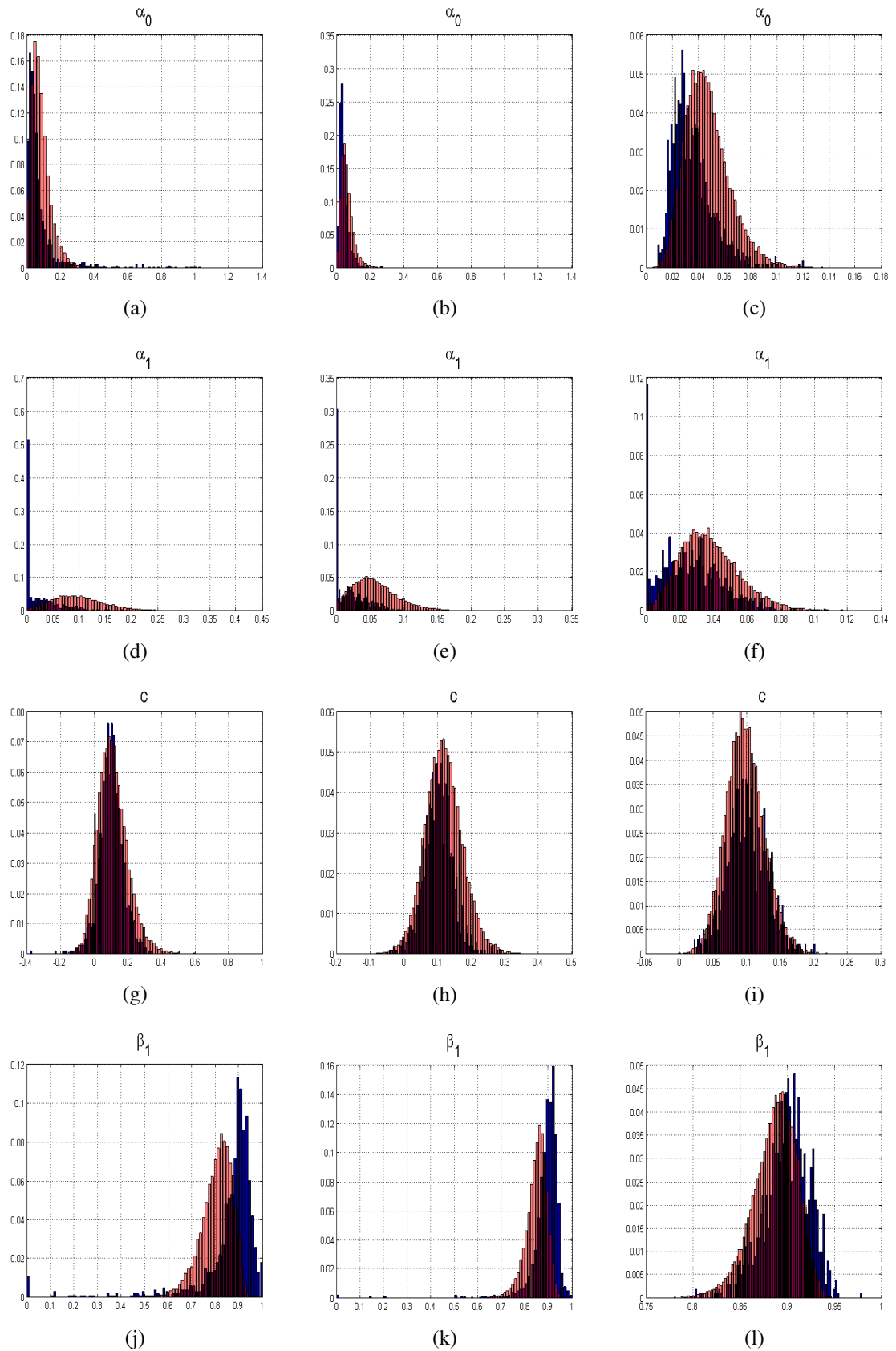


Abbildung 4.35: Vergleich der Histogramme der Verteilung der ML-Parameterschätzer und der a posteriori Verteilung auf Basis eines Datensatzes (Spalte 1:  $n=250$ , Spalte 2:  $n=500$  und Spalte 3:  $n=1000$ )

| $n$  |                          | $\alpha_0 : 0.03$      | $\alpha_1 : 0.03$      | $c : 0.1$              | $\beta_1 : 0.9$        |
|------|--------------------------|------------------------|------------------------|------------------------|------------------------|
| 100  | Mittelwert (1)           | 0.2355                 | 0.0426                 | 0.0832                 | 0.7271                 |
|      | Varianz robust (3)       | 3.8920                 | 6.8897                 | 2.4223                 | 14.2595                |
|      | Varianz der Schätzer (2) | 0.1325                 | 0.0084                 | 0.0236                 | 0.0977                 |
| 250  | Mittelwert (1)           | 0.0866                 | 0.0267                 | 0.0978                 | 0.8541                 |
|      | Varianz robust (3)       | 0.8135                 | 0.8862                 | 0.0623                 | 2.4579                 |
|      | Varianz der Schätzer (2) | 0.0181                 | 0.0018                 | 0.0056                 | 0.0221                 |
| 500  | Mittelwert (1)           | 0.0487                 | 0.0238                 | 0.1024                 | 0.8878                 |
|      | Varianz robust (3)       | 0.0046                 | 1.0237                 | 0.2484                 | 0.4768                 |
|      | Varianz der Schätzer (2) | 0.0041                 | 0.0007                 | 0.0019                 | 0.0051                 |
| 1000 | Mittelwert (1)           | 0.0354                 | 0.0258                 | 0.1006                 | 0.8992                 |
|      | Varianz robust (3)       | 0.0016                 | 0.0014                 | 0.0009                 | 0.0056                 |
|      | Varianz der Schätzer (2) | $0.2463 \cdot 10^{-3}$ | $0.3435 \cdot 10^{-3}$ | $0.8590 \cdot 10^{-3}$ | $0.5746 \cdot 10^{-3}$ |
| 2500 | Mittelwert (1)           | 0.0317                 | 0.0279                 | 0.1010                 | 0.9001                 |
|      | Varianz robust (3)       | $0.0672 \cdot 10^{-3}$ | $0.1454 \cdot 10^{-3}$ | $0.3361 \cdot 10^{-3}$ | $0.2090 \cdot 10^{-3}$ |
|      | Varianz der Schätzer (2) | $0.0604 \cdot 10^{-3}$ | $0.1288 \cdot 10^{-3}$ | $0.3314 \cdot 10^{-3}$ | $0.1855 \cdot 10^{-3}$ |
| 5000 | Mittelwert (1)           | 0.0307                 | 0.0293                 | 0.1003                 | 0.9000                 |
|      | Varianz robust (3)       | $0.0297 \cdot 10^{-3}$ | $0.0702 \cdot 10^{-3}$ | $0.1652 \cdot 10^{-3}$ | $0.0961 \cdot 10^{-3}$ |
|      | Varianz der Schätzer (2) | $0.0296 \cdot 10^{-3}$ | $0.0662 \cdot 10^{-3}$ | $0.1632 \cdot 10^{-3}$ | $0.0941 \cdot 10^{-3}$ |

Tabelle 4.22: Ergebnisse für 1000 Replikationen eines GJR-GARCH(1,1)-Modells mit  $\alpha_0 = 0.03$ ,  $\alpha_1 = 0.03$ ,  $c = 0.1$  und  $\beta_1 = 0.9$  bei unterschiedlichen Stichprobenumfängen  $n$

#### 4.4.3 Simulation zu einem AP-ARCH-Modell

Es folgt die Anwendung der MCMC-Schätzung auf einen Datensatz aus einem AP-ARCH-Modell.

$$z_t = \sigma_t \cdot \varepsilon_t, \quad t = 1, \dots, T,$$

mit

$$\sigma_t^\lambda = \alpha_0 + \alpha_1 \cdot (|Z_t| - c \cdot Z_t)^\lambda + \beta_1 \cdot \sigma_{t-1}^\lambda \quad t = 1, \dots, T,$$

Mit den Parameterwerten

$$\alpha_0 = 0.1, \alpha_1 = 0.1, \beta_1 = 0.8, c = -0.5 \text{ und } \lambda = 1.5$$

wird ein Datensatz der Länge 5000 erzeugt.

Bei der Simulation der Daten ist besonders auffällig, dass einige Versuche notwendig sind, bis ein Datensatz erzeugt werden kann, der die Parameter gut annimmt. Insbesondere für die Werte von  $\lambda$  und  $c$  weisen Simulationen von geringer Länge große Abweichungen zwischen den simulierten Daten und den anschließenden Schätzungen auf.

Die Ergebnisse für den oben erzeugten Datensatz befinden sich in Tabelle 4.23. Aufgrund der Implementierung der Varianzgleichung in der MFE-Toolbox ist das Vorzeichen für den Parameter  $c$  zu beachten.

| Datensatz       | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $\sigma_0^2 : 1.5032$ | $c : -0.5$ | $\lambda : 1.5$ | Log-L   |
|-----------------|------------------|------------------|-----------------|-----------------------|------------|-----------------|---------|
| ML-MFE          | 0.1283           | 0.1089           | 0.7668          | 1.3763                | -0.4888    | 1.4954          | -6884.9 |
| MCMC-Modus      | 0.1280           | 0.1143           | 0.7643          | 1.2323                | 0.4687     | 1.4541          | -6884.2 |
| MCMC-Mittelwert | 0.1340           | 0.1072           | 0.7569          | 4.2067                | 0.4993     | 1.5976          | -6884.4 |
| MCMC-Median     | 0.1325           | 0.1072           | 0.7584          | 3.1150                | 0.4902     | 1.5860          | -6884.4 |

Tabelle 4.23: ML-Schätzergebnisse (AP-ARCH) für y5000aparch

Bei der Simulation von AP-ARCH-Modellen hat es sich für die Akzeptanzrate des zweiten Laufs als günstig erwiesen, relativ große Varianzen für die Vorschlagsdichten des ersten Laufs zu verwenden (dies sorgt für relativ kleine Varianzen im zweiten Lauf). Alternativ kann die Kovarianzmatrix für den zweiten Lauf auch durch eine Konstante verkleinert werden.

Werden Daten geschätzt, die aus einem Modell mit  $\lambda = 1$  stammen, ergibt sich das Problem, dass sehr häufig Werte für  $\lambda$  gezogen werden, die kleiner als 1 sind. Das hat für die Schätzung der Parameter keine Nachteile, allerdings können für die Prädiktivverteilung dann keine plausiblen Werte für die Kandidaten berechnet werden. Daher hat es sich als sinnvoll erwiesen, in solchen Fällen die Restriktion  $\lambda > 1$  zu verwenden.

#### 4.4.4 Simulationen zu TS-GARCH-Modellen

In diesem Abschnitt wird das TS-GARCH-Modell geschätzt. Dazu wird wieder ein Datensatz für

$$z_t = \sigma_t \cdot \varepsilon_t, \quad t = 1, \dots, T,$$

mit

$$\sigma_t = \alpha_0 + \alpha_i \cdot |Z_{t-1}| + \beta_j \sigma_{t-j} \quad t = 1, \dots, T$$

erzeugt, die  $\varepsilon_t$  stammen dabei aus einer Normalverteilung. Anschließend wird der Datensatz mit dem beschriebenen MCMC-Verfahren und der MFE-Toolbox geschätzt. Als Besonderheit wird an diesem Modell verdeutlicht, welchen Einfluss die Kovarianzen der Vorschlagsdichten auf die Akzeptanzrate der Schätzung haben. Dazu wird mit der oben genannten Varianzmatrix der Vorschlagsdichte

$$\Sigma = \begin{pmatrix} 0.001 & 0 & 0 & 0 \\ 0 & 0.001 & 0 & 0 \\ 0 & 0 & 0.001 & 0 \\ 0 & 0 & 0 & 0.5 \end{pmatrix}$$

der erste Lauf durchgeführt. Dabei wird eine Akzeptanzrate von lediglich 3% im ersten Lauf erreicht, wobei die daraus resultierende Kovarianzmatrix wiederum zu guten Ergebnissen im zweiten Lauf führt. Die Akzeptanzrate im zweiten Lauf beträgt 33%. Die Ergebnisse sind in Tabelle 4.24 dargestellt.

Die Schätzung wird wiederholt, wobei für die Kovarianzmatrix der Vorschlagsdichte des ersten Laufs die aus der ersten Schätzung stammenden Varianzen verwendet werden, aber ohne Berücksichtigung der Kovarianzen.

Mit diesen angepassten Varianzen ergibt sich allerdings auch nur eine Akzeptanzrate von ca. 8% im ersten Lauf. Die Akzeptanzrate im zweiten Lauf beträgt. Da die Varianzmatrix der Vorschlagsdichte des zweiten Laufs sich ausschließlich abseits der Hauptdiagonalen verändert, zeigt sich, dass die Qualität der Vorschlagsdichte maßgeblich von der Berücksichtigung der Kovarianzen abhängt.

| Datensatz       | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\beta_1 : 0.8$ | $\sigma_0^2 : 0.6268$ | Log-L   |
|-----------------|------------------|------------------|-----------------|-----------------------|---------|
| ML-MFE          | 0.1097           | 0.1018           | 0.7901          | 0.8052                | -6246.3 |
| MCMC-MAP        | 0.1034           | 0.0987           | 0.7997          | 1.0554                | -6245.4 |
| MCMC-Mittelwert | 0.1174           | 0.1039           | 0.7793          | 4.7874                | -6245.8 |
| MCMC-Median     | 0.1150           | 0.1035           | 0.7825          | 3.0182                | -6245.5 |

Tabelle 4.24: ML-Schätzergebnisse (TS-GARCH) für y1000tsgarch

Auch für das TS-GARCH-Modell wird eine Simulationsstudie zur Verteilung der ML-Schätzer durchgeführt. Aufgrund der Ähnlichkeit des TS-GARCH-Modells und des GARCH-Modells werden wieder die oben besprochenen Varianten verwendet:

Variante 1:  $\alpha_0 = 0.2$ ,  $\alpha_1 = 0.2$  und  $\beta_1 = 0.6$ ,

Variante 2:  $\alpha_0 = 0.1$ ,  $\alpha_1 = 0.1$  und  $\beta_1 = 0.8$ ,

Variante 3:  $\alpha_0 = 0.05$ ,  $\alpha_1 = 0.05$  und  $\beta_1 = 0.94$ ,

Variante 4:  $\alpha_0 = 0.4$ ,  $\alpha_1 = 0.49$  und  $\beta_1 = 0.49$ ,

Variante 5:  $\alpha_0 = 0.1$ ,  $\alpha_1 = 0.3$  und  $\beta_1 = 0.3$ .

Die Ergebnisse für Variante 2 können den Abbildungen 4.36, 4.36 und 4.36 entnommen werden. Weitere Ergebnisse befinden sich im Anhang A.3.

Für die Einzeldatensätze sind die a posteriori Verteilungen in Form der Histogramme in der Abbildungen 4.39 gegeben. Es wird deutlich, dass die a posteriori Verteilungen die Verteilung der ML-Parameterschätzer gut abbilden. Auch die Schiefe der Verteilungen wird erfasst.

| Variante 1                 | n=100               |                     | n=250            |                  | n=500                  |                        |
|----------------------------|---------------------|---------------------|------------------|------------------|------------------------|------------------------|
|                            | $\alpha_0 : 0.2$    | $\alpha_1 : 0.2$    | $\alpha_0 : 0.2$ | $\alpha_1 : 0.2$ | $\alpha_0 : 0.2$       | $\alpha_1 : 0.2$       |
| Mittelwert (1)             | 0.2287              | 0.1721              | 0.2323           | 0.1912           | 0.2245                 | 0.2003                 |
| Varianz robust (3)         | 2.1438              | 22.7263             | 0.0304           | 0.0064           | 0.0122                 | 0.0030                 |
| Varianz der $\theta_i$ (2) | 0.0331              | 0.0167              | 0.0211           | 0.0069           | 0.0098                 | 0.0028                 |
|                            | n=1000              |                     | n=2500           |                  | n=5000                 |                        |
| Mittelwert (1)             | 0.2109              | 0.2006              | 0.2055           | 0.2009           | 0.2021                 | 0.1996                 |
| Varianz robust (3)         | 0.0053              | 0.0014              | 0.0014           | 0.0005           | 0.0007                 | 0.0003                 |
| Varianz der $\theta_i$ (2) | 0.0039              | 0.0014              | 0.0015           | 0.0005           | 0.0006                 | 0.0003                 |
|                            | n=100               |                     | n=250            |                  | n=500                  |                        |
| Variante 2                 | $\alpha_0 : 0.1$    | $\alpha_1 : 0.1$    | $\alpha_0 : 0.1$ | $\alpha_1 : 0.1$ | $\alpha_0 : 0.1$       | $\alpha_1 : 0.1$       |
| Mittelwert (1)             | 0.2088              | 0.0875              | 0.1589           | 0.0977           | 0.1339                 | 0.0979                 |
| Varianz robust (3)         | $0.0912 \cdot 10^3$ | $1.8809 \cdot 10^3$ | 0.0938           | 0.0114           | 0.0585                 | 0.0026                 |
| Varianz der $\theta_i$ (2) | 0.0488              | 0.0095              | 0.0245           | 0.0044           | 0.0130                 | 0.0016                 |
|                            | n=1000              |                     | n=2500           |                  | n=5000                 |                        |
| Mittelwert (1)             | 0.1175              | 0.1002              | 0.1057           | 0.0998           | 0.1022                 | 0.0999                 |
| Varianz robust (3)         | 0.0046              | 0.0008              | 0.0009           | 0.0003           | $0.4002 \cdot 10^{-3}$ | $0.1557 \cdot 10^{-3}$ |
| Varianz der $\theta_i$ (2) | 0.0048              | 0.0009              | 0.0011           | 0.0003           | 0.0005                 | 0.0002                 |
|                            | n=1000              |                     | n=2500           |                  | n=5000                 |                        |
| Mittelwert (1)             | 0.2088              | 0.0875              | 0.1589           | 0.0977           | 0.1339                 | 0.0979                 |
| Varianz robust (3)         | $0.0912 \cdot 10^3$ | $1.8809 \cdot 10^3$ | 0.0938           | 0.0114           | 0.0585                 | 0.0026                 |
| Varianz der $\theta_i$ (2) | 0.0488              | 0.0095              | 0.0245           | 0.0044           | 0.0130                 | 0.0016                 |
|                            | n=1000              |                     | n=2500           |                  | n=5000                 |                        |
| Mittelwert (1)             | 0.1175              | 0.1002              | 0.1057           | 0.0998           | 0.1022                 | 0.0999                 |
| Varianz robust (3)         | 0.0046              | 0.0008              | 0.0009           | 0.0003           | $0.4002 \cdot 10^{-3}$ | $0.1557 \cdot 10^{-3}$ |
| Varianz der $\theta_i$ (2) | 0.0048              | 0.0009              | 0.0011           | 0.0003           | 0.0005                 | 0.0002                 |

Tabelle 4.25: Mittelwerte und Varianzen der ML-Schätzer bei unterschiedlichen Stichprobenumfängen (mäßige Persistenz)

| Variante 3                 | n=100             |                                      | n=250                  |   | n=500                  |   |
|----------------------------|-------------------|--------------------------------------|------------------------|---|------------------------|---|
|                            | $\alpha_0 : 0.05$ | $\alpha_1 : 0.05$   $\beta_1 : 0.94$ | $\alpha_0 : 0.05$      | $\alpha_1 : 0.05$   $\beta_1 : 0.94$            | $\alpha_0 : 0.05$      | $\alpha_1 : 0.05$   $\beta_1 : 0.94$            |
| Mittelwert (1)             | 0.5482            | 0.0433   0.7416                      | 0.3405                 | 0.0488   0.8213                                 | 0.1450                 | 0.0514   0.9001                                 |
| Varianz robust (3)         | 10.7953           | 73.5886   48.7700                    | 1.4055                 | 0.0278   0.3791                                 | 0.0659                 | 0.0221   0.0296                                 |
| Varianz der $\theta_i$ (2) | 0.4721            | 0.0052   0.0904                      | 0.2619                 | 0.0022   0.0503                                 | 0.0496                 | 0.0008   0.0099                                 |
|                            | n=1000            |                                      | n=2500                 |   | n=5000                 |   |
| Mittelwert (1)             | 0.0934            | 0.0503   0.9220                      | 0.0605                 | 0.0506   0.9353                                 | 0.0552                 | 0.0502   0.9377                                 |
| Varianz robust (3)         | 0.1049            | 0.0003   0.0178                      | $0.6672 \cdot 10^{-3}$ | $0.1007 \cdot 10^{-3}$   $0.2579 \cdot 10^{-3}$ | $0.2381 \cdot 10^{-3}$ | $0.0473 \cdot 10^{-3}$   $0.1030 \cdot 10^{-3}$ |
| Varianz der $\theta_i$ (2) | 0.0181            | 0.0003   0.0037                      | $0.5663 \cdot 10^{-3}$ | $0.0995 \cdot 10^{-3}$   $0.2289 \cdot 10^{-3}$ | $0.2233 \cdot 10^{-3}$ | $0.0478 \cdot 10^{-3}$   $0.0957 \cdot 10^{-3}$ |
| Variante 4                 | n=100             |                                      | n=250                  |   | n=500                  |   |
|                            | $\alpha_0 : 0.4$  | $\alpha_1 : 0.49$   $\beta_1 : 0.49$ | $\alpha_0 : 0.4$       | $\alpha_1 : 0.49$   $\beta_1 : 0.49$            | $\alpha_0 : 0.4$       | $\alpha_1 : 0.49$   $\beta_1 : 0.49$            |
| Mittelwert (1)             | 0.5387            | 0.4501   0.4610                      | 0.4551                 | 0.4728   0.4776                                 | 0.4333                 | 0.4844   0.4804                                 |
| Varianz robust (3)         | 0.1004            | 0.0165   0.0243                      | 0.0244                 | 0.0064   0.0075                                 | 0.0103                 | 0.0031   0.0033                                 |
| Varianz der $\theta_i$ (2) | 0.0949            | 0.0205   0.0278                      | 0.0231                 | 0.0062   0.0077                                 | 0.0103                 | 0.0031   0.0036                                 |
|                            | n=1000            |                                      | n=2500                 |   | n=5000                 |   |
| Mittelwert (1)             | 0.4134            | 0.4870   0.4860                      | 0.4064                 | 0.4880   0.4887                                 | 0.4019                 | 0.4889   0.4901                                 |
| Varianz robust (3)         | 0.0045            | 0.0016   0.0016                      | 0.0018                 | 0.0006   0.0006                                 | $0.8598 \cdot 10^{-3}$ | $0.3167 \cdot 10^{-3}$   $0.3016 \cdot 10^{-3}$ |
| Varianz der $\theta_i$ (2) | 0.0039            | 0.0015   0.0015                      | 0.0017                 | 0.0006   0.0006                                 | $0.8649 \cdot 10^{-3}$ | $0.3073 \cdot 10^{-3}$   $0.3084 \cdot 10^{-3}$ |

Tabelle 4.26: Mittelwerte und Varianzen der ML-Schätzer bei unterschiedlichen Stichprobenumfängen (hohe Persistenz)



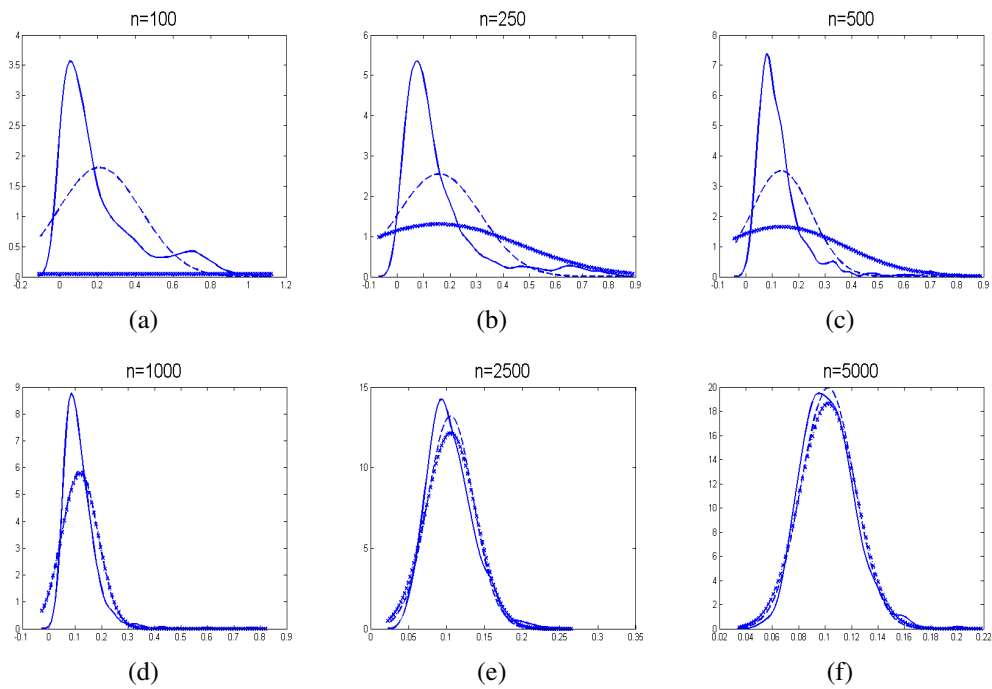


Abbildung 4.36: Verteilung des ML-Schätzers des Parameters bei Variante 1,  $\alpha_0: 0.1$  bei unterschiedlichen Stichprobenumfängen

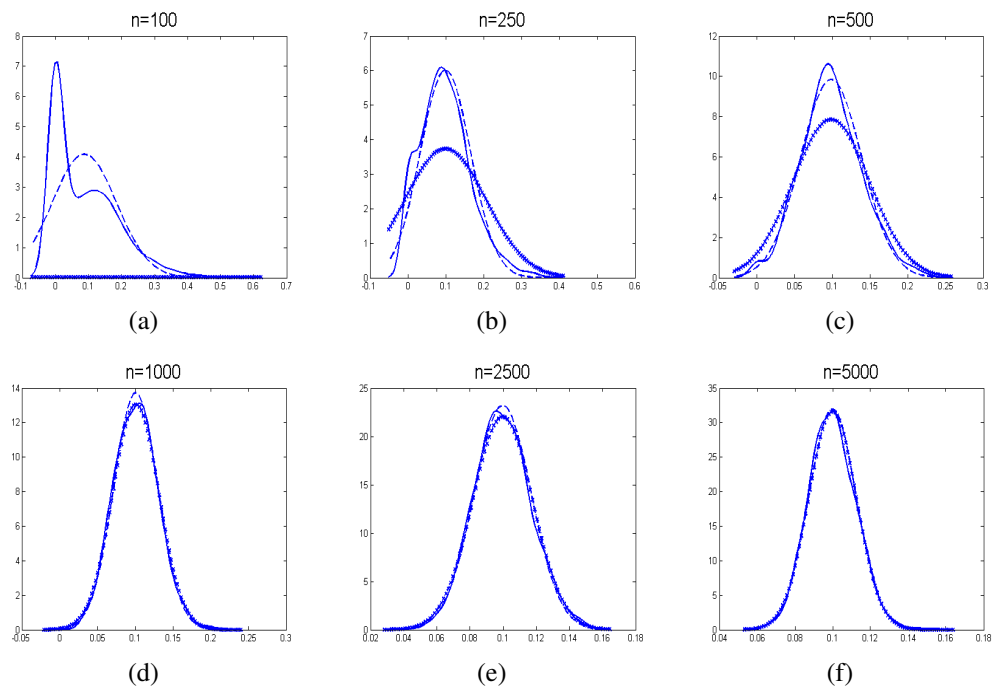


Abbildung 4.37: Verteilung des ML-Schätzers des Parameters bei Variante 1,  $\alpha_1: 0.1$  bei unterschiedlichen Stichprobenumfängen

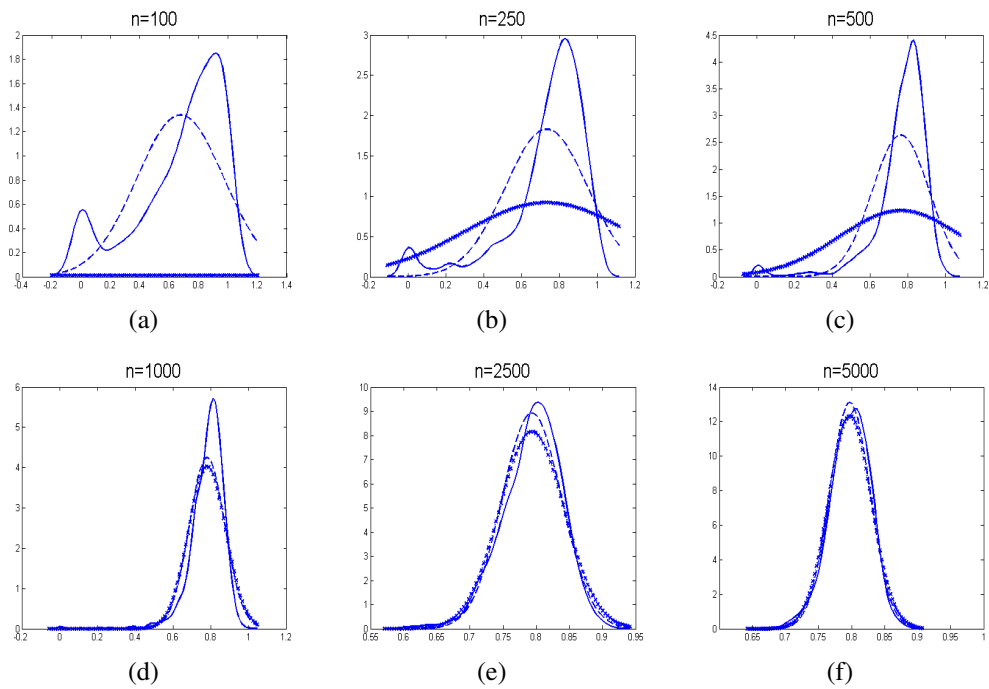


Abbildung 4.38: Verteilung des ML-Schätzers des Parameters bei Variante 1,  $\beta_1: 0.8$  bei unterschiedlichen Stichprobenumfängen

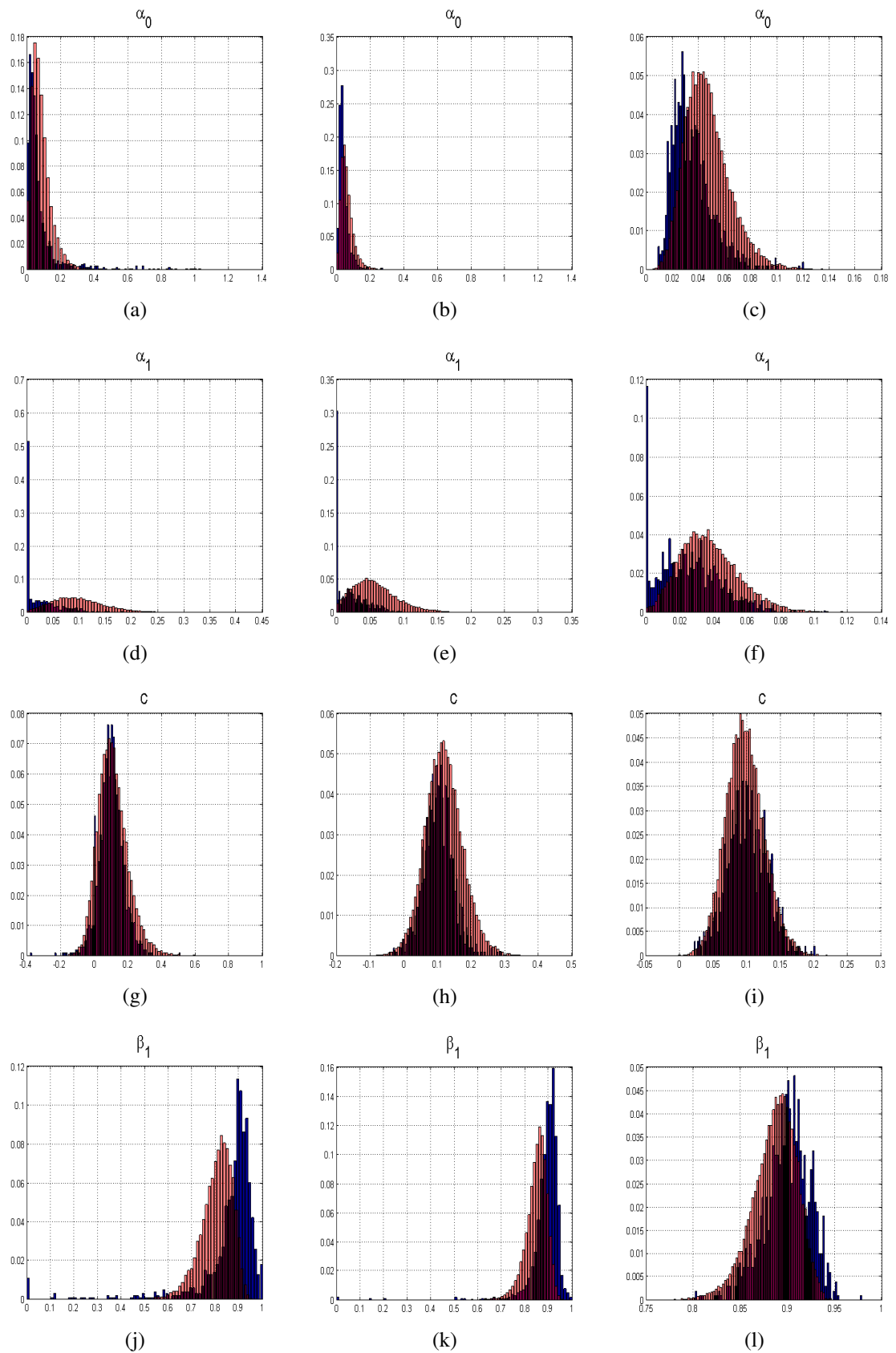


Abbildung 4.39: Vergleich der Histogramme der Verteilung der ML-Schätzer und der a posteriori Verteilung auf Basis eines Datensatzes ( $n=100, 250, 500$  und  $1000$ )

### 4.4.5 Simulation zu einem T-GARCH-Modell

Aus dem T-GARCH-Modell mit der Varianzgleichung

$$\sigma_t = \alpha_0 + \alpha_1 \cdot |Z_t| + c \cdot |Z_t| \cdot I(Z_t < 0) + \beta_1 \sigma_{t-1}$$

für  $t=1, \dots, T$ , wird ein Datensatz mit den Parameterwerten

$$\alpha_0 = 0.05, \alpha_1 = 0.05, \beta_1 = 0.85, \text{ und } c = 0.1$$

erzeugt, diese orientieren sich wiederum an den Werten für die Zeitreihe des DAX. Tabelle 4.27 zeigt die Ergebnisse der ML- und MCMC-Schätzung für die simulierten Daten des T-GARCH-Modells. Es zeigt sich, dass auch beim T-GARCH-Modell die MCMC-Schätzung und die ML-Schätzung bei großen Stichproben zu nahezu identischen Ergebnissen führen.

| Datensatz       | $\alpha_0 : 0.05$ | $\alpha_1 : 0.05$ | $\beta_1 : 0.85$ | $\sigma_0^2 : 0.2928$ | $c : 0.1$ | Log-L   |
|-----------------|-------------------|-------------------|------------------|-----------------------|-----------|---------|
| ML-MFE          | 0.0397            | 0.0457            | 0.8633           | 0.2683                | 0.1096    | -5191.2 |
| MCMC-MAP        | 0.0375            | 0.0462            | 0.8674           | 0.2686                | 0.1070    | -5191.0 |
| MCMC-Mittelwert | 0.0411            | 0.0479            | 0.8592           | 0.7296                | 0.1108    | -5191.2 |
| MCMC-Median     | 0.0407            | 0.0475            | 0.8602           | 0.5444                | 0.1108    | -5191.0 |

Tabelle 4.27: ML-Schätzergebnisse für die simulierten Daten eines T-GARCH(1,1)-Modells

### 4.4.6 Zusammenfassung

Die vorangegangenen Abschnitte zeigen für das GARCH-Modell sowie dessen Erweiterungen, dass sowohl die ML-Schätzung als die MCMC-Schätzung des jeweiligen Modells für große Stichproben vergleichbare Ergebnisse liefern. Anhand von Einzeldatensätzen kann demonstriert werden, dass die MCMC-Schätzung zu a posteriori Verteilungen führt, die bei kleinen Datensätzen in der Lage sind die Verteilung der Parameter akzeptabel abzubilden. Allerdings ist dies stark abhängig von dem jeweiligen Datensatz. Gute Ergebnisse können nur erzielt werden, wenn die Daten die Parameterwerte, mit denen sie erzeugt wurden gut widerspiegeln. In diesem Fall wird u.a. die Schiefe der ML-Parameterschätzer, welche in den Simulationstudien gezeigt werden konnte, durch die a posteriori Verteilung erfasst.

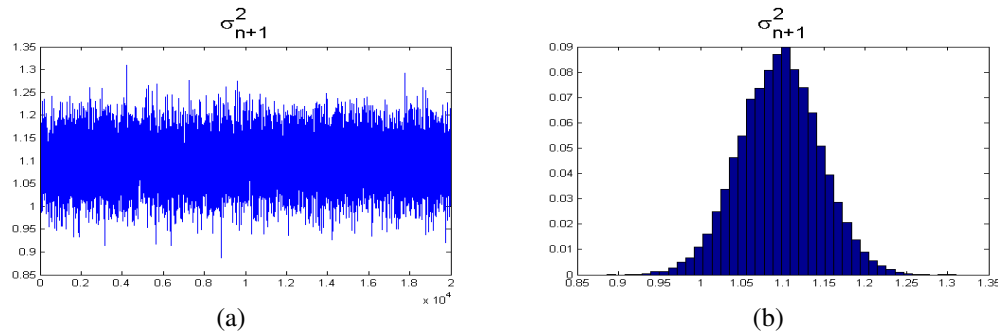


Abbildung 4.40: Prädiktivverteilung

Es zeigt sich, dass in kleinen Stichproben die Wahl der a priori Verteilungen an Bedeutung gewinnt, da diese bei kleinen Stichproben leicht dazu neigen die Daten zu überlagern. Insbesondere trifft dies auf die a priori Verteilung von  $\alpha_0$  zu. Beide verwendeten a priori Verteilungen  $\left(\frac{1}{\alpha_0} \text{ und } \alpha_0 \sim \log N(-2.0, 5)\right)$  neigen dazu die a posteriori Verteilung für kleine Werte des Parameters  $\alpha_0$  in Richtung 0 zu ziehen. Daher sollte in kleinen Stichproben nach geeigneteren a priori Verteilungen gesucht werden.

## 4.5 Prognosen mittels der a posteriori Prädiktivverteilung

Die vergangenen Kapitel hatten die Modellierung der Volatilität zum Inhalt. Wie bereits erwähnt, stellt die Prognose zukünftiger Werte der Volatilität das zentrale Anwendungsgebiet der GARCH-Modelle dar. In diesem Abschnitt werden nun für ausgewählte GARCH-Varianten Prognosen mit Hilfe der a posteriori Prädiktivverteilung gezeigt und mit den Prognosen der ML-Schätzung verglichen. Besonders zu berücksichtigen bei der Prognose mittels MCMC-Schätzung ist, dass für jeden Wert der Markovketten eine Prognose durchgeführt wird. Damit ergibt sich die a posteriori Prädiktivverteilung durch die Prognosen auf Basis der Parameterwerte jedes einzelnen Wertes der Markovketten.

Abbildung 4.40 zeigt die Markovkette und das resultierende Histogramm für die 1-Schritt-Prognosen bei jedem lag-ten Zug und somit die Ergebnisse der a posteriori Verteilung für den Zeitpunkt  $t + 1$ .

Von der Prognose mittels der a posteriori Prädiktivverteilung, bei der für jeden Wert der Markovkette ein Prognosewert berechnet wird, ist die Prognose mit nur einem Punktschätzer zu unterscheiden. Abbildung 4.41 zeigt die Darstellung einer Prognose, bei der

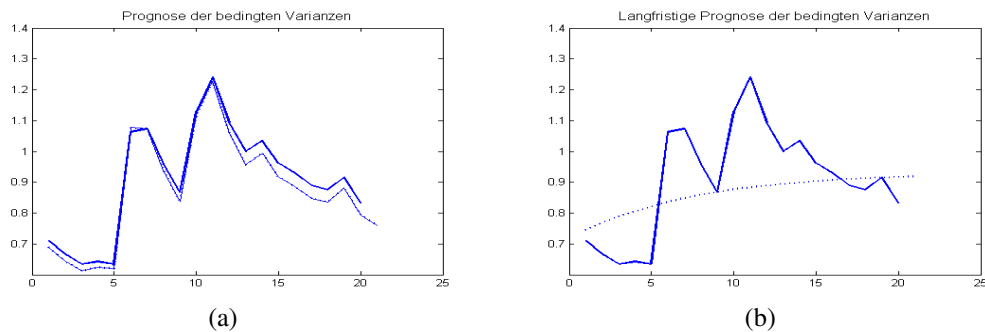


Abbildung 4.41: Prognosewerte für ein GARCH(1,1)-Modell

für den simulierten Datensatz „y5000garch\_norm“ die letzten 20 Werte abgeschnitten wurden. Mit den verbleibenden 4980 Werten wird ein GARCH(1,1)-Modell geschätzt. Mit dem Mittelwert der a posteriori Verteilung wird dann für die 20 abgeschnittenen Werte jeweils eine 1-Schritt-Prognose für die bedingten Varianzen durchgeführt, wobei die Informationen der 20 Werte verwendet wird. Daneben ist eine langfristige Prognose der bedingten Varianzen abgebildet. Es ist bekannt und auch hier deutlich zu sehen, dass die langfristige Prognose gegen die unbedingte Varianz konvergiert.

## 4.6 Simulationen mit multivariaten Zeitreihen

Das in Abschnitt 4.1 erläuterte Vorgehen lässt sich leicht auf multivariate Zeitreihen erweitern. In diesem Abschnitt wird die MCMC-Schätzung auf simulierte und reale multivariate Datensätze angewendet. Es werden die in Abschnitt 2.5 beschriebenen Modelle behandelt, diese haben den Vorteil, dass die Anzahl der zu schätzenden Parameter überschaubar bleibt. Des weiteren ist die Simulation auf den bivariaten Fall begrenzt. Zunächst wird das Verfahren auf ein CCC(1,1)-Modell angewendet.

### Das CCC(1,1)-Modell

Im bivariaten Fall kann die Kovarianzmatrix aus Gleichung (2.15)

$$\Sigma_t = D_t \cdot R \cdot D_t,$$

durch

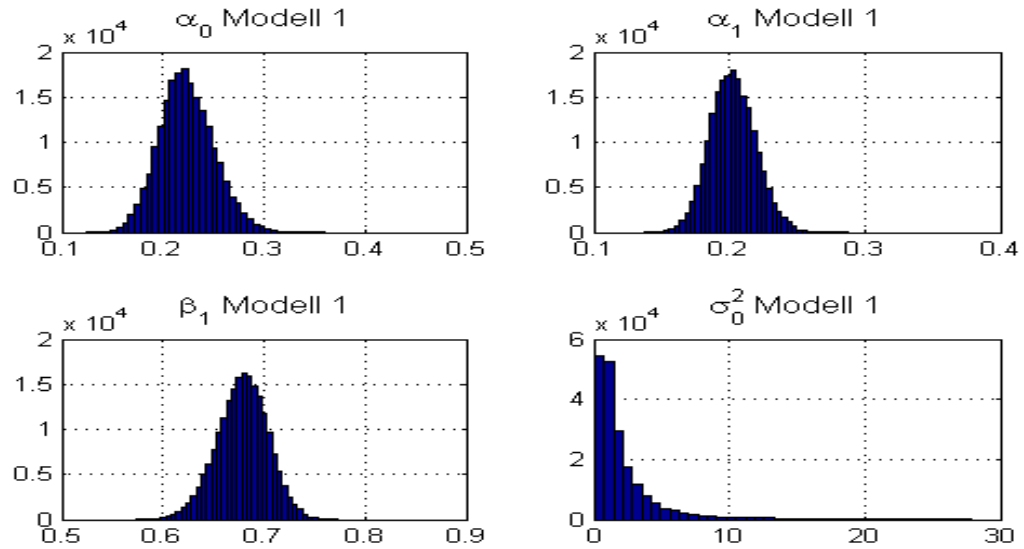


Abbildung 4.42: CCC-GARCH(1,1): Parameter Modell 1

$$\Sigma_t = \begin{bmatrix} \sigma_{11,t} & 0 \\ 0 & \sigma_{22,t} \end{bmatrix} \cdot \begin{bmatrix} 1 & \rho_{12} \\ \rho_{21} & 1 \end{bmatrix} \cdot \begin{bmatrix} \sigma_{11,t} & 0 \\ 0 & \sigma_{22,t} \end{bmatrix}$$

dargestellt werden.

Die beiden GARCH(1,1)-Modelle lauten dabei:

$$\sigma_{11,t}^2 = \alpha_{10} + \alpha_{11}z_{1,t-1}^2 + \beta_{11}\sigma_{11,t-1}^2 \quad t = 1, \dots, T,$$

$$\sigma_{22,t}^2 = \alpha_{20} + \alpha_{12}z_{2,t-1}^2 + \beta_{12}\sigma_{22,t-1}^2 \quad t = 1, \dots, T.$$

Es gelten die Parameterrestriktionen aus dem univariaten Fall.

Der  $(5000 \times 2)$  Datensatz beinhaltet die Zeitreihen der beiden GARCH(1,1)-Modelle mit den Parametern:

$$\text{Modell 1: } \alpha_0 = 0.2, \alpha_1 = 0.2 \text{ und } \beta_1 = 0.7,$$

$$\text{Modell 2: } \alpha_0 = 0.1, \alpha_1 = 0.3 \text{ und } \beta_1 = 0.6,$$

mit  $\rho = 0.5$ . Für die  $2 \times 2$  Korrelationsmatrix  $R$  wird im bivariaten Fall nur ein Wert benötigt. Der Datensatz wurde mit der Routine „cc\_mvlgarch\_simulate“ aus UCSD-Toolbox von Sheppard simuliert. Die Schätzung der Parameter erfolgt mit der Routine „cc\_mvlgarch“ aus derselben Toolbox.

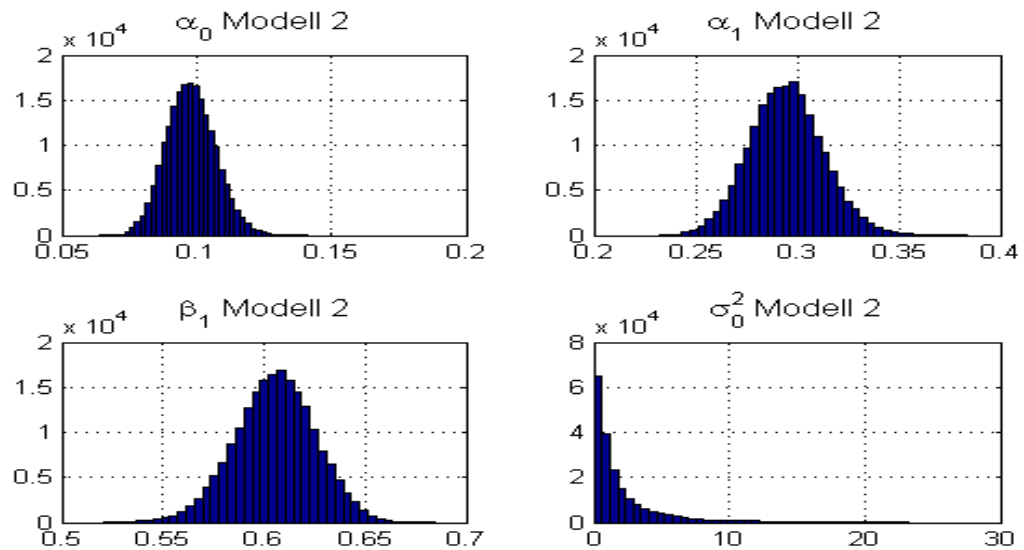


Abbildung 4.43: CCC-GARCH(1,1): Parameter Modell 2

Das Vorgehen im multivariaten Fall erfordert nur geringfügige Anpassungen der oben beschriebenen Methodik. Aufgrund der größeren Anzahl von zu schätzenden Parametern müssen die Vorschlagsdichten modifiziert werden, da die Erhöhung der Anzahl der Parameter die Akzeptanzrate deutlich senkt. Selbstverständlich steigt durch die simultane Schätzung mehrerer GARCH-Modelle und der Korrelationsmatrix  $R$  die Anzahl an Kandidaten, die die Stationaritätsbedingungen nicht erfüllen bzw. zu niedrigen a posteriori Wahrscheinlichkeiten führen. Um dieses Problem zu beheben, ist es ausreichend die Varianzen der Vorschlagsdichten soweit zu verkleinern, bis die Akzeptanzrate des ersten Laufs eine akzeptable Schätzung der Kovarianzmatrix für den zweiten Lauf ermöglicht. Zur weiteren Vereinfachung wurden ausschließlich flache a priori Verteilungen verwendet.

Die Berechnungszeiten steigen bei der Simulation des multivariaten Modells deutlich an, da die Parameter zweier GARCH-Modelle und zusätzlich die Korrelationsmatrix  $R$  geschätzt werden müssen.

Die beiden Abbildungen 4.42 und 4.43 zeigen die Histogramme der Züge für die Parameter des zweiten Laufs beider Modelle. Sie entsprechen den schon beschriebenen Schaubildern. Tabelle 4.28 zeigt die Ergebnisse der ML-Schätzung und MCMC-Schätzung. Dabei wird deutlich, dass sich auch hier beide Methoden hinsichtlich der geschätzten Werte für die Parameter kaum unterscheiden. Abbildung 4.44 zeigt die Ergebnisse des zweiten Laufs für den Parameter  $\rho$ .



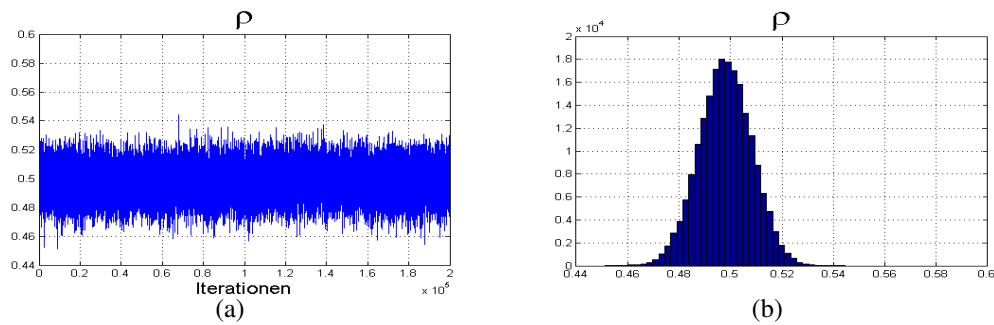


Abbildung 4.44: CCC-GARCH(1,1)-Modell  $\rho$

| Modell | Schätzung       | $\alpha_0$ | $\alpha_1$ | $\beta_1$ | $\rho$ | $\sigma_0^2$ | Log-L  |
|--------|-----------------|------------|------------|-----------|--------|--------------|--------|
| 1 1    | ML-UCSD         | 0.2317     | 0.1976     | 0.6767    | 0.4979 | —            | -13905 |
| 2      | ML-UCSD         | 0.0963     | 0.2950     | 0.6071    | 0.4979 | —            | -13905 |
| 1      | MCMC-MAP        | 0.2209     | 0.1889     | 0.6846    | 0.4959 | 0.7128       | -13903 |
| 1      | MCMC-Mittelwert | 0.2236     | 0.2026     | 0.6786    | 0.4980 | 2.3813       |        |
| 1      | MCMC-Median     | 0.2221     | 0.2022     | 0.6794    | 0.4980 | 1.3288       |        |
| 2      | MCMC-MAP        | 0.0989     | 0.3039     | 0.6010    | 0.4959 | 0.1103       | -13903 |
| 2      | MCMC-Mittelwert | 0.0979     | 0.2949     | 0.6055    | 0.4980 | 2.2794       |        |
| 2      | MCMC-Median     | 0.0977     | 0.2945     | 0.6060    | 0.4980 | 1.0957       |        |

Tabelle 4.28: Ergebnisse der Schätzung der Parameter eines CCC-GARCH(1,1)-Modells

### Das DCC(1,1)-Modell

Eine weitere Anwendung der MCMC-Schätzung wird für das DCC-Modell demonstriert. Auch hier wird ein bivariates Modell verwendet. Wir betrachten nochmal Gleichung (2.18) aus Abschnitt 2.5.2:

$$\Sigma_t = D_t \cdot R_t \cdot D_t,$$

mit der Korrelationsmatrix:

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1}.$$

Für die beiden GARCH(1,1)-Prozesse gilt analog zum vorigen Beispiel:

$$\sigma_{11,t}^2 = \alpha_{10} + \alpha_{11} z_{1,t-1}^2 + \beta_{11} \sigma_{11,t-1}^2 \quad t = 1, \dots, T,$$

| Modell | Schätzung       | $\alpha_0$ | $\alpha_1$ | $\beta_1$ | $\sigma_0^2$ | Log-L      |
|--------|-----------------|------------|------------|-----------|--------------|------------|
| 1      | ML-UCSD         | 0.0614     | 0.1360     | 0.8630    | —            | -9207.6    |
| 2      | ML-UCSD         | 0.1720     | 0.1683     | 0.8317    | —            | -9207.6    |
| 1      | MCMC-MAP        | 0.0940     | 0.1154     | 0.8656    | 36.6652      | -9200.3522 |
| 1      | MCMC-Mittelwert | 0.0931     | 0.1148     | 0.8665    | 16.0644      |            |
| 1      | MCMC-Median     | 0.0921     | 0.1142     | 0.8671    | 13.0701      |            |
| 2      | MCMC-MAP        | 0.2190     | 0.1534     | 0.8323    | 18.2807      | -9200.3522 |
| 2      | MCMC-Mittelwert | 0.2206     | 0.1534     | 0.8322    | 15.9917      |            |
| 2      | MCMC-Median     | 0.2181     | 0.1530     | 0.8323    | 12.5752      |            |

Tabelle 4.29: Ergebnisse der Schätzung der GARCH-Parameter eines DCC-GARCH-Modells

$$\sigma_{22,t}^2 = \alpha_{20} + \alpha_{12}z_{2,t-1}^2 + \beta_{12}\sigma_{22,t-1}^2 \quad t = 1, \dots, T.$$

Wird in Gleichung (2.20) für die Berechnung der  $Q_t$ -Matrizen  $W = V = 1$  gesetzt, erhält man

$$Q_t = \tilde{Q}(1 - \alpha^* - \beta^*) + \alpha^* \tilde{\varepsilon}_{t-1} \tilde{\varepsilon}'_{t-1} + \beta^* Q_{t-1}.$$

Für die Simulation in diesem Abschnitt werden die beiden Zeitreihen der Renditen der Aktienkurse der BASF AG und der Bayer AG aus dem Zeitraum 03.01.2000 bis 20.09.2009 verwendet. Die ML-Schätzung der Parameter wurde mit der Routine „dcc\_mvlgarch“ aus der UCSD-Toolbox durchgeführt.

Die beiden Tabellen 4.29 und 4.30 zeigen die Ergebnisse des zweiten Laufs der MCMC-Schätzung der Parameter. In Tabelle 4.29 sind die ML- und MCMC-Schätzer der beiden univariaten GARCH-Modelle aufgeführt. Tabelle 4.30 beinhaltet die ML- und MCMC-Schätzer für die Parameter  $\alpha^*$  und  $\beta^*$  des Modells zur Schätzung der Korrelationsmatrix  $R_t$  und die entsprechenden Schaubilder der Markovketten und Histogramme zeigen die Abbildungen 4.45 und 4.46. Bei der MCMC-Schätzung erfolgt die Schätzung der GARCH-Parameter beider Modelle simultan, d.h. es werden für beide Modelle gleichzeitig die Kandidaten gezogen und auf die Stationaritätsbedingungen hin untersucht. Anschließend werden Kandidaten für die Gleichung zur Bestimmung der Korrelationsmatrix  $R_t$  gezogen. Die Akzeptanzwahrscheinlichkeit wird dann im Anschluss für sämtliche Parameter bestimmt und mit dem Wert aus der Gleichverteilung verglichen.

| Schätzung       | $\alpha^*$ | $\beta^*$ |
|-----------------|------------|-----------|
| ML-UCSD         | 0.0372     | 0.8958    |
| MCMC-MAP        | 0.0349     | 0.8881    |
| MCMC-Mittelwert | 0.0351     | 0.8886    |
| MCMC-Median     | 0.0346     | 0.8916    |

Tabelle 4.30: Ergebnisse der Schätzung der DCC-Parameter eines DCC-GARCH-Modell

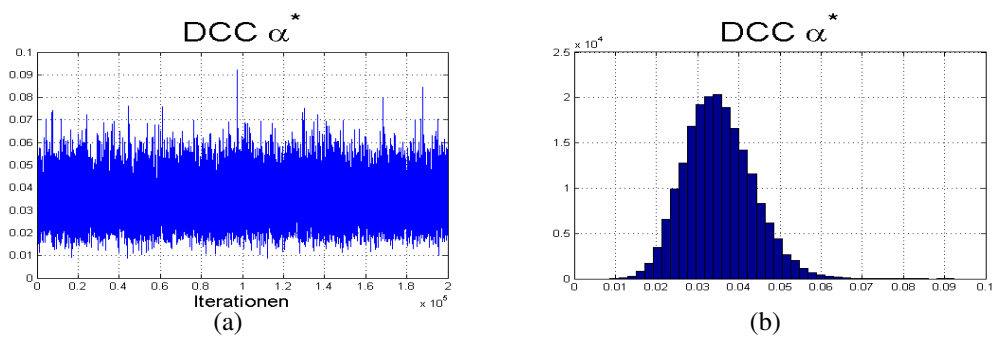


Abbildung 4.45: DCC-GARCH(1,1)-Modell  $\alpha^*$

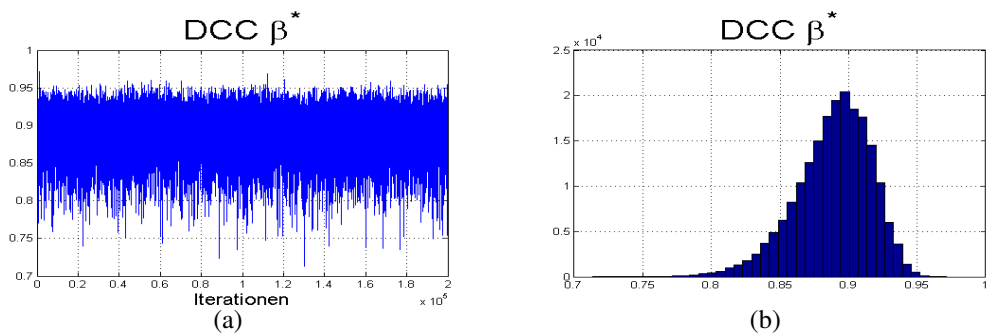


Abbildung 4.46: DCC-GARCH(1,1)-Modell  $\beta^*$

Interessant ist in diesem Zusammenhang eine vergleichende Untersuchung der Eigenschaften der MCMC-Schätzung und der ML-Schätzung im multivariaten Fall. Dies ist allerdings nicht Bestandteil dieser Arbeit.

## 4.7 Anwendung der MCMC-Methodik auf reale univariate Datensätze

Die vergangenen Abschnitte in diesem Kapitel haben diverse Anwendungsbeispiele für das beschriebene MCMC-Verfahren anhand simulierter Daten gezeigt. In diesem Abschnitt werden die Methoden nun auf reale Datensätze angewendet.

### Verwendete Zeitreihen

Es werden die stetigen Renditen des DAX und des Dow Jones in dem Zeitraum vom 26.11.1990 bis 17.09.2009 verwendet. Alle Zeitreihen werden zunächst um ihren Mittelwert bereinigt. Für die Spezifikationen ist es relevant in welcher Form die Renditen verwendet werden. Die angegebenen Spezifikationen sind für die Verwendung der Renditen in Prozent vorgesehen.

Bei realen Daten lässt sich die Annahme der Normalverteilung der Log-Renditen auch anhand der Schätzung der Freiheitsgrade der  $t$ -Verteilung überprüfen. Wie in Abschnitt 2.3.1 gezeigt, sprechen nur sehr hohe Werte für  $\nu$  für die Normalverteilung.

Für die MCMC-Schätzungen werden die oben genannten Startwerte verwendet

$$\alpha_0 : 0.05, \alpha_1 : 0.005, \beta_1 : 0.9 \text{ und } \sigma_0^2 : 0.9253.$$

Die a priori Verteilungen für  $\alpha_1$  und  $\beta_1$  sind:

$$f(\alpha_1) = \begin{cases} 1 & \text{für } 0 \leq \alpha_1 \leq 1 \\ 0 & \text{sonst} \end{cases}, \quad f(\beta) = \begin{cases} 1 & \text{für } 0 \leq \beta \leq 1 \\ 0 & \text{sonst.} \end{cases}$$

Für  $\alpha_0$  und  $\sigma_0^2$  gelten diese a priori Verteilungen:

$$\alpha_0 \sim \log N(-2.0, 5) \text{ bzw. } f(\sigma_0^2) = \frac{1}{\sigma_0^2}.$$

Allerdings hat sich gezeigt, dass die bei den simulierten Datensätzen verwendeten Vorschlagsdichten zu groß waren. Sie führen zwar zu akzeptablen Ergebnissen, doch lässt sich die Performance der Schätzung durch die Verwendung alternativer Vorschlagsdichten weiter verbessern. Dies liegt daran, dass die Kovarianzmatrizen der Parameterschätzer kleinerer Werte annehmen als bei den simulierten Datensätzen.

Es wurden jeweils 100000 Werte im ersten Lauf erzeugt und 200000 Werte im zweiten Lauf. Für die Berechnung der Schätzer wurde stets jeder 10. Wert verwendet. Die Ak-

zeptanzraten lagen für alle Berechnungen im angestrebten Bereich zwischen 25% und 50%.

Es zeigt sich in allen Modellen, dass für alle betrachteten Zeitreihen die Anzahl der Freiheitsgrade bei der  $t$ -Verteilung im einstelligen Bereich geschätzt wird. Analog ergibt sich in allen Schätzungen für die  $GED$ -Verteilung ein  $\nu$  deutlich kleiner als 2.

### **Ergebnisse**

Bei der Betrachtung der Ergebnisse fällt zunächst auf, dass die ML-Schätzung und die MCMC-Schätzung zu sehr ähnlichen Resultaten führen. Über alle Modell-Varianten sind bei der Verwendung der Normal- oder  $t$ -Verteilung nahezu keine Unterschiede zu erkennen.

Unterschiede findet man allerdings jeweils bei der Verwendung der  $GED$ -Verteilung. Beim GARCH- und TGARCH-Modell ist besonders auffällig, dass sich gerade die Werte für  $\nu$ , also den Parameter der  $GED$ -Verteilung und Wert der Log-Likelihood zwischen den Schätzvarianten nicht unterscheiden, sondern lediglich die Werte für  $\alpha_0$  und  $\alpha_1$ . Beim TS-GARCH-Modell dagegen führen die Unterschiede in den Werten der Schätzer für die GARCH-Parameter auch zu Unterschieden in der Log-Likelihood. Interessant sind die Ergebnisse für das EGARCH-Modell. Hierbei fällt insbesondere auf, dass die Vorzeichen bei den Schätzern für  $\alpha_0$  unterschiedlich sind. Bei der Zeitreihe des Dow Jones wirkt sich dies nicht auf den Wert der Log-Likelihood aus, allerdings bei der Betrachtung der Renditen des DAX deutlich. Wiederum sind die Schätzergebnisse für den Parameter  $\nu$  sehr ähnlich.

Bei den Ergebnissen für die AP-ARCH-Modelle ist der Unterschied des Vorzeichens für den Parameter  $c$  zu beachten.

| Schätzer                              | $\alpha_0$ | $\alpha_1$ | $\beta_1$ | $\sigma_0^2$ | $\nu$  | Log-L   |
|---------------------------------------|------------|------------|-----------|--------------|--------|---------|
| GARCH-Modell mit Normalverteilung     |            |            |           |              |        |         |
| MFE ML                                | 0.0314     | 0.0833     | 0.9004    | 2.2392       | —      | -7741.1 |
| MCMC-MAP                              | 0.0311     | 0.0838     | 0.9002    | 1.9061       | —      | -7741.0 |
| MCMC-Mittelwert                       | 0.0320     | 0.0849     | 0.8988    | 3.7016       | —      | -7741.1 |
| MCMC-Median                           | 0.0318     | 0.0846     | 0.8991    | 3.1102       | —      | -7741.0 |
| GARCH-Modell mit $t$ -Verteilung      |            |            |           |              |        |         |
| MFE ML                                | 0.0143     | 0.0819     | 0.9130    | 2.2472       | 8.8031 | -7618.3 |
| MCMC-MAP                              | 0.0136     | 0.0809     | 0.9142    | 2.3340       | 8.9980 | -7618.2 |
| MCMC-Mittelwert                       | 0.0152     | 0.0833     | 0.9108    | 3.6898       | 8.9170 | -7618.3 |
| MCMC-Median                           | 0.0150     | 0.0829     | 0.9111    | 3.1048       | 8.9170 | -7618.2 |
| GARCH-Modell mit $GED$ -Verteilung    |            |            |           |              |        |         |
| MFE ML                                | 0.0205     | 0.0826     | 0.9085    | 2.2450       | 1.4290 | -7647.4 |
| MCMC-MAP                              | 0.0148     | 0.0587     | 0.9108    | 1.3462       | 1.4269 | -7647.3 |
| MCMC-Mittelwert                       | 0.0155     | 0.0623     | 0.9062    | 2.7508       | 1.4293 | -7647.4 |
| MCMC-Median                           | 0.0152     | 0.0620     | 0.9066    | 2.2309       | 1.4289 | -7647.3 |
| TS-GARCH-Modell mit Normalverteilung  |            |            |           |              |        |         |
| MFE ML                                | 0.0271     | 0.0895     | 0.9103    | 1.4171       | —      | -7750.5 |
| MCMC-MAP                              | 0.0267     | 0.0890     | 0.9109    | 1.6416       | —      | -7749.7 |
| MCMC-Mittelwert                       | 0.0294     | 0.0897     | 0.9081    | 2.9880       | —      | -7750.3 |
| MCMC-Median                           | 0.0292     | 0.0895     | 0.9083    | 2.3993       | —      | -7750.3 |
| TS-GARCH-Modell mit $t$ -Verteilung   |            |            |           |              |        |         |
| MFE ML                                | 0.0231     | 0.0856     | 0.9142    | 1.4087       | 9.1250 | -7622.9 |
| MCMC-MAP                              | 0.0232     | 0.0872     | 0.9128    | 1.4806       | 9.0732 | -7621.8 |
| MCMC-Mittelwert                       | 0.0241     | 0.0866     | 0.9124    | 2.9069       | 9.2805 | -7622.9 |
| MCMC-Median                           | 0.0240     | 0.0863     | 0.9127    | 2.3755       | 9.1908 | -7623.0 |
| TS-GARCH-Modell mit $GED$ -Verteilung |            |            |           |              |        |         |
| MFE ML                                | 0.0249     | 0.0871     | 0.9127    | 1.4124       | 1.4285 | -7650.8 |
| MCMC-MAP                              | 0.0135     | 0.0755     | 0.9202    | 1.6428       | 1.4269 | -7646.9 |
| MCMC-Mittelwert                       | 0.0150     | 0.0771     | 0.9172    | 3.2265       | 1.4293 | -7647.1 |
| MCMC-Median                           | 0.0148     | 0.0768     | 0.9175    | 2.5219       | 1.4289 | -7647.0 |

Tabelle 4.31: Schätzergebnisse verschiedener GARCH(1,1) und TS-GARCH(1,1)-Modelle für um den Mittelwert bereinigte, stetige Renditen des DAX

| Schätzer                               | $\alpha_0$ | $\alpha_1$ | $\beta_1$ | $\sigma_0^2$ | $c$    | $\nu$  | Log-L   |
|--|------------|------------|-----------|--------------|--------|--------|---------|
| T-GARCH-Modell mit Normalverteilung    |            |            |           |              |        |        |         |
| MFE ML                                 | 0.0262     | 0.0252     | 0.9279    | 1.4092       | 0.0835 | —      | -7689.6 |
| MCMC-MAP                               | 0.0250     | 0.0239     | 0.9304    | 2.7425       | 0.0823 | —      | -7688.4 |
| MCMC-Mittelwert                        | 0.0266     | 0.0260     | 0.9270    | 4.2754       | 0.0836 | —      | -7688.5 |
| MCMC-Median                            | 0.0264     | 0.0257     | 0.9274    | 3.6379       | 0.0834 | —      | -7688.4 |
| T-GARCH-Modell mit $t$ -Verteilung     |            |            |           |              |        |        |         |
| MFE ML                                 | 0.0221     | 0.0321     | 0.9221    | 1.4093       | 0.0911 | 9.8443 | -7571.4 |
| MCMC-MAP                               | 0.0180     | 0.0339     | 0.9258    | 2.2569       | 0.0885 | 9.2693 | -7569.6 |
| MCMC-Mittelwert                        | 0.0188     | 0.0371     | 0.9225    | 4.0923       | 0.0892 | 9.6863 | -7569.7 |
| MCMC-Median                            | 0.0187     | 0.0368     | 0.9227    | 3.4192       | 0.0890 | 9.5972 | -7569.6 |
| T-GARCH-Modell mit $GED$ -Verteilung   |            |            |           |              |        |        |         |
| MFE ML                                 | 0.0222     | 0.0311     | 0.9254    | 1.4108       | 0.0865 | 1.4588 | -7604.2 |
| MCMC-MAP                               | 0.0180     | 0.0271     | 0.9265    | 1.8377       | 0.0749 | 1.4579 | -7603.3 |
| MCMC-Mittelwert                        | 0.0190     | 0.0286     | 0.9242    | 3.2117       | 0.0746 | 1.4567 | -7603.4 |
| MCMC-Median                            | 0.0188     | 0.0284     | 0.9246    | 2.6760       | 0.0745 | 1.4562 | -7603.3 |
| GJR-GARCH-Modell mit Normalverteilung  |            |            |           |              |        |        |         |
| MFE ML                                 | 0.0328     | 0.0206     | 0.9115    | 2.2341       | 0.0966 | —      | -7699.8 |
| MCMC-MAP                               | 0.0326     | 0.0200     | 0.9116    | 2.0362       | 0.0972 | —      | -7699.6 |
| MCMC-Mittelwert                        | 0.0339     | 0.0224     | 0.9087    | 3.7738       | 0.0979 | —      | -7699.7 |
| MCMC-Median                            | 0.0336     | 0.0220     | 0.9091    | 3.2559       | 0.0975 | —      | -7699.6 |
| GJR-GARCH-Modell mit $t$ -Verteilung   |            |            |           |              |        |        |         |
| MFE ML                                 | 0.0200     | 0.0294     | 0.9106    | 2.2421       | 0.1000 | 9.6597 | -7587.2 |
| MCMC-MAP                               | 0.0187     | 0.0308     | 0.9119    | 2.2352       | 0.0974 | 9.5894 | -7587.1 |
| MCMC-Mittelwert                        | 0.0205     | 0.0314     | 0.9079    | 3.8772       | 0.1019 | 9.7624 | -7587.1 |
| MCMC-Median                            | 0.0203     | 0.0310     | 0.9083    | 3.3237       | 0.1013 | 9.6525 | -7587.1 |
| GJR-GARCH-Modell mit $GED$ -Verteilung |            |            |           |              |        |        |         |
| MFE ML                                 | 0.0251     | 0.0262     | 0.9112    | 2.2398       | 0.0982 | 1.4602 | -7618.0 |
| MCMC-MAP                               | 0.0177     | 0.0185     | 0.9132    | 1.5611       | 0.0735 | 1.4535 | -7617.8 |
| MCMC-Mittelwert                        | 0.0195     | 0.0214     | 0.9078    | 2.8968       | 0.0747 | 1.4597 | -7617.9 |
| MCMC-Median                            | 0.0193     | 0.0210     | 0.9082    | 2.4413       | 0.0743 | 1.4594 | -7617.9 |

Tabelle 4.32: Schätzergebnisse verschiedener T-GARCH(1,1)- und GJR-GARCH(1,1)-Modelle für um den Mittelwert bereinigte, stetige Renditen des DAX

| Schätzer               | $\alpha_0$ | $\alpha_1$ | $\beta_1$ | $\sigma_0^2$ | $c$     | $\nu$  | Log-L   |
|------------------------|------------|------------|-----------|--------------|---------|--------|---------|
| Normalverteilung       |            |            |           |              |         |        |         |
| MFE ML                 | 0.0116     | 0.1168     | 0.9818    | 2.0441       | -0.0709 | —      | -7699.2 |
| MCMC-MAP               | 0.0118     | 0.1172     | 0.9818    | 3.2057       | -0.0708 | —      | -7698.9 |
| MCMC-Mittelwert        | 0.0120     | 0.1201     | 0.9810    | 5.6700       | -0.0714 | —      | -7699.5 |
| MCMC-Median            | 0.0119     | 0.1196     | 0.9812    | 4.7547       | -0.0713 | —      | -7699.2 |
| <i>t</i> -Verteilung   |            |            |           |              |         |        |         |
| MFE ML                 | 0.0084     | 0.1387     | 0.9869    | 2.0107       | -0.0704 | 9.4181 | -7579.2 |
| MCMC-MAP               | 0.0040     | 0.1364     | 0.9872    | 3.0308       | -0.0708 | 9.3483 | -7579.0 |
| MCMC-Mittelwert        | 0.0051     | 0.1429     | 0.9861    | 5.5993       | -0.0714 | 9.6554 | -7579.5 |
| MCMC-Median            | 0.0050     | 0.1423     | 0.9862    | 4.6639       | -0.0712 | 9.5658 | -7579.2 |
| <i>GED</i> -Verteilung |            |            |           |              |         |        |         |
| MFE ML                 | 0.0102     | 0.1302     | 0.9848    | 2.0244       | -0.0703 | 1.4526 | -7618.0 |
| MCMC-MAP               | -0.0127    | 0.1095     | 0.9853    | 1.9260       | -0.0608 | 1.4506 | -7612.2 |
| MCMC-Mittelwert        | -0.0130    | 0.1163     | 0.9839    | 4.6779       | -0.0614 | 1.4526 | -7612.8 |
| MCMC-Median            | -0.0129    | 0.1159     | 0.9840    | 3.8053       | -0.0612 | 1.4525 | -7612.5 |

Tabelle 4.33: Schätzergebnisse verschiedener EGARCH(1,1)-Modelle für um den Mittelwert bereinigte, stetige Renditen des DAX



| Schätzer               | $\alpha_0$ | $\alpha_1$ | $\beta_1$ | $\sigma_0^2$ | $c$     | $\lambda$ | $\nu$   | Log-L   |
|------------------------|------------|------------|-----------|--------------|---------|-----------|---------|---------|
| Normalverteilung       |            |            |           |              |         |           |         |         |
| MFE ML                 | 0.0273     | 0.0651     | 0.9253    | 1.5928       | -0.5761 | 1.2523    | —       | -7686.9 |
| MCMC-MAP               | 0.0261     | 0.0653     | 0.9262    | 2.6871       | 0.5653  | 1.2381    | —       | -7686.0 |
| MCMC-Mittelwert        | 0.0278     | 0.0652     | 0.9242    | 4.0270       | 0.5746  | 1.2765    | —       | -7686.2 |
| MCMC-Median            | 0.0276     | 0.0650     | 0.9245    | 3.4333       | 0.5722  | 1.2698    | —       | -7686.1 |
| <i>t</i> -Verteilung   |            |            |           |              |         |           |         |         |
| MFE ML                 | 0.0199     | 0.0778     | 0.9220    | 1.5284       | -0.5400 | 1.1640    | 9.7911  | -7570.1 |
| MCMC-MAP               | 0.0193     | 0.0762     | 0.9235    | 1.5690       | 0.5372  | 1.1786    | 9.8167  | -7569.4 |
| MCMC-Mittelwert        | 0.0213     | 0.0768     | 0.9201    | 4.0253       | 0.5429  | 1.2309    | 10.1656 | -7570.2 |
| MCMC-Median            | 0.0212     | 0.0766     | 0.9203    | 3.4129       | 0.5415  | 1.2249    | 10.0512 | -7570.1 |
| <i>GED</i> -Verteilung |            |            |           |              |         |           |         |         |
| MFE ML                 | 0.0218     | 0.0739     | 0.9245    | 1.5182       | -0.5478 | 1.1477    | 1.4604  | -7603.4 |
| MCMC-MAP               | 0.0182     | 0.0613     | 0.9265    | 1.6144       | 0.5699  | 1.1076    | 1.4532  | -7602.7 |
| MCMC-Mittelwert        | 0.0187     | 0.0625     | 0.9229    | 3.0814       | 0.5389  | 1.1949    | 1.4601  | -7602.9 |
| MCMC-Median            | 0.0185     | 0.0623     | 0.9233    | 2.5649       | 0.5363  | 1.1832    | 1.4598  | -7602.8 |

Tabelle 4.34: Schätzergebnisse verschiedener AP-ARCH(1,1)-Modelle für um den Mittelwert bereinigte, stetige Renditen des DAX

| Schätzer                              | $\alpha_0$ | $\alpha_1$ | $\beta_1$ | $\sigma_0^2$ | $\nu$  | Log-L   |
|---------------------------------------|------------|------------|-----------|--------------|--------|---------|
| GARCH-Modell mit Normalverteilung     |            |            |           |              |        |         |
| MFE ML                                | 0.0099     | 0.0688     | 0.9230    | 0.6644       | —      | -6350.9 |
| MCMC-MAP                              | 0.0095     | 0.0680     | 0.9241    | 0.5430       | —      | -6350.9 |
| MCMC-Mittelwert                       | 0.0101     | 0.0701     | 0.9216    | 0.9338       | —      | -6351.0 |
| MCMC-Median                           | 0.0100     | 0.0699     | 0.9218    | 0.8005       | —      | -6350.9 |
| GARCH-Modell mit $t$ -Verteilung      |            |            |           |              |        |         |
| MFE ML                                | 0.0060     | 0.0595     | 0.9362    | 0.6567       | 7.3666 | -6251.3 |
| MCMC-MAP                              | 0.0058     | 0.0589     | 0.9370    | 0.5974       | 7.2561 | -6251.3 |
| MCMC-Mittelwert                       | 0.0064     | 0.0612     | 0.9339    | 0.8924       | 7.5228 | -6251.4 |
| MCMC-Median                           | 0.0063     | 0.0609     | 0.9342    | 0.7828       | 7.4573 | -6251.4 |
| GARCH-Modell mit $GED$ -Verteilung    |            |            |           |              |        |         |
| MFE ML                                | 0.0074     | 0.0626     | 0.9315    | 0.6592       | 1.3922 | -6260.4 |
| MCMC-MAP                              | 0.0050     | 0.0449     | 0.9317    | 0.3604       | 1.3840 | -6260.3 |
| MCMC-Mittelwert                       | 0.0054     | 0.0464     | 0.9296    | 0.6737       | 1.3930 | -6260.4 |
| MCMC-Median                           | 0.0053     | 0.0461     | 0.9300    | 0.5790       | 1.3926 | -6260.4 |
| TS-GARCH-Modell mit Normalverteilung  |            |            |           |              |        |         |
| MFE ML                                | 0.0175     | 0.0792     | 0.9206    | 0.4141       | —      | -6359.3 |
| MCMC-MAP                              | 0.0173     | 0.0788     | 0.9212    | 0.5035       | —      | -6358.7 |
| MCMC-Mittelwert                       | 0.0182     | 0.0798     | 0.9193    | 0.8944       | —      | -6359.9 |
| MCMC-Median                           | 0.0181     | 0.0796     | 0.9196    | 0.7418       | —      | -6359.8 |
| TS-GARCH-Modell mit $t$ -Verteilung   |            |            |           |              |        |         |
| MFE ML                                | 0.0143     | 0.0675     | 0.9323    | 0.4078       | 7.7484 | -6262.4 |
| MCMC-MAP                              | 0.0139     | 0.0663     | 0.9336    | 0.4432       | 7.7920 | -6261.8 |
| MCMC-Mittelwert                       | 0.0151     | 0.0688     | 0.9304    | 0.9050       | 7.8719 | -6263.0 |
| MCMC-Median                           | 0.0150     | 0.0685     | 0.9308    | 0.7554       | 7.8069 | -6262.9 |
| TS-GARCH-Modell mit $GED$ -Verteilung |            |            |           |              |        |         |
| MFE ML                                | 0.0155     | 0.0719     | 0.9279    | 0.4102       | 1.3992 | -6269.2 |
| MCMC-MAP                              | 0.0064     | 0.0601     | 0.9384    | 0.3608       | 1.3899 | -6262.9 |
| MCMC-Mittelwert                       | 0.0078     | 0.0627     | 0.9342    | 0.7020       | 1.3880 | -6263.1 |
| MCMC-Median                           | 0.0076     | 0.0624     | 0.9344    | 0.5855       | 1.3879 | -6263.1 |

Tabelle 4.35: Schätzergebnisse verschiedener GARCH(1,1) und TS-GARCH(1,1)-Modelle für um den Mittelwert bereinigte, stetige Renditen des Dow Jones

| Schätzer                               | $\alpha_0$ | $\alpha_1$ | $\beta_1$ | $\sigma_0^2$ | $c$    | $\nu$  | Log-L   |
|--|------------|------------|-----------|--------------|--------|--------|---------|
| T-GARCH-Modell mit Normalverteilung    |            |            |           |              |        |        |         |
| MFE ML                                 | 0.0157     | 0.0134     | 0.9343    | 0.4083       | 0.1024 | —      | -6275.0 |
| MCMC-MAP                               | 0.0155     | 0.0122     | 0.9352    | 0.5763       | 0.1030 | —      | -6274.5 |
| MCMC-Mittelwert                        | 0.0158     | 0.0145     | 0.9335    | 0.8967       | 0.1022 | —      | -6274.6 |
| MCMC-Median                            | 0.0157     | 0.0142     | 0.9337    | 0.7724       | 0.1019 | —      | -6274.5 |
| T-GARCH-Modell mit $t$ -Verteilung     |            |            |           |              |        |        |         |
| MFE ML                                 | 0.0135     | 0.0113     | 0.9398    | 0.4053       | 0.0973 | 8.5359 | -6202.1 |
| MCMC-MAP                               | 0.0116     | 0.0127     | 0.9420    | 0.6495       | 0.0948 | 8.3235 | -6201.1 |
| MCMC-Mittelwert                        | 0.0123     | 0.0153     | 0.9392    | 0.9403       | 0.0955 | 8.4942 | -6201.2 |
| MCMC-Median                            | 0.0122     | 0.0150     | 0.9395    | 0.8162       | 0.0952 | 8.4096 | -6201.1 |
| T-GARCH-Modell mit $GED$ -Verteilung   |            |            |           |              |        |        |         |
| MFE ML                                 | 0.0141     | 0.0124     | 0.9378    | 0.4065       | 0.0992 | 1.4576 | -6209.2 |
| MCMC-MAP                               | 0.0114     | 0.0104     | 0.9401    | 0.4189       | 0.0834 | 1.4638 | -6208.7 |
| MCMC-Mittelwert                        | 0.0118     | 0.0130     | 0.9368    | 0.7387       | 0.0847 | 1.4555 | -6208.8 |
| MCMC-Median                            | 0.0117     | 0.0126     | 0.9371    | 0.6252       | 0.0846 | 1.4548 | -6208.7 |
| GJR-GARCH-Modell mit Normalverteilung  |            |            |           |              |        |        |         |
| MFE ML                                 | 0.0127     | 0.0089     | 0.9261    | 0.6591       | 0.1041 | —      | -6294.9 |
| MCMC-MAP                               | 0.0123     | 0.0088     | 0.9274    | 0.5838       | 0.1026 | —      | -6294.9 |
| MCMC-Mittelwert                        | 0.0130     | 0.0111     | 0.9236    | 0.9012       | 0.1046 | —      | -6295.0 |
| MCMC-Median                            | 0.0129     | 0.0108     | 0.9239    | 0.7939       | 0.1043 | —      | -6294.9 |
| GJR-GARCH-Modell mit $t$ -Verteilung   |            |            |           |              |        |        |         |
| MFE ML                                 | 0.0090     | 0.0086     | 0.9343    | 0.6554       | 0.0968 | 8.0568 | -6212.6 |
| MCMC-MAP                               | 0.0089     | 0.0077     | 0.9347    | 0.6172       | 0.0976 | 8.0766 | -6212.6 |
| MCMC-Mittelwert                        | 0.0093     | 0.0112     | 0.9315    | 0.9628       | 0.0974 | 8.1512 | -6212.7 |
| MCMC-Median                            | 0.0091     | 0.0107     | 0.9320    | 0.8541       | 0.0969 | 8.0716 | -6212.6 |
| GJR-GARCH-Modell mit $GED$ -Verteilung |            |            |           |              |        |        |         |
| MFE ML                                 | 0.0105     | 0.0086     | 0.9310    | 0.6567       | 0.0997 | 1.4367 | -6221.6 |
| MCMC-MAP                               | 0.0076     | 0.0074     | 0.9307    | 0.3739       | 0.0720 | 1.4374 | -6221.5 |
| MCMC-Mittelwert                        | 0.0080     | 0.0084     | 0.9279    | 0.7116       | 0.0739 | 1.4377 | -6221.7 |
| MCMC-Median                            | 0.0079     | 0.0080     | 0.9283    | 0.6221       | 0.0736 | 1.4369 | -6221.6 |

Tabelle 4.36: Schätzergebnisse verschiedener T-GARCH(1,1)- und GJR-GARCH(1,1)-Modelle für um den Mittelwert bereinigte, stetige Renditen des Dow Jones

| Schätzer               | $\alpha_0$ | $\alpha_1$ | $\beta_1$ | $\sigma_0^2$ | $c$     | $\nu$  | Log-L   |
|------------------------|------------|------------|-----------|--------------|---------|--------|---------|
| Normalverteilung       |            |            |           |              |         |        |         |
| MFE ML                 | 0.0012     | 0.1175     | 0.9849    | 0.5638       | -0.0871 | —      | -6282.3 |
| MCMC-MAP               | 0.0011     | 0.1162     | 0.9851    | 0.7499       | -0.0866 | —      | -6282.2 |
| MCMC-Mittelwert        | 0.0012     | 0.1200     | 0.9845    | 1.1777       | -0.0876 | —      | -6282.6 |
| MCMC-Median            | 0.0011     | 0.1197     | 0.9846    | 1.0049       | -0.0874 | —      | -6282.4 |
| <i>t</i> -Verteilung   |            |            |           |              |         |        |         |
| MFE ML                 | 0.0015     | 0.1116     | 0.9887    | 0.5656       | -0.0823 | 8.2881 | -6204.3 |
| MCMC-MAP               | -0.0021    | 0.1115     | 0.9885    | 0.8109       | -0.0833 | 8.2748 | -6204.1 |
| MCMC-Mittelwert        | -0.0021    | 0.1150     | 0.9882    | 1.2452       | -0.0832 | 8.4672 | -6204.5 |
| MCMC-Median            | -0.0021    | 0.1145     | 0.9883    | 1.0569       | -0.0830 | 8.3822 | -6204.3 |
| <i>GED</i> -Verteilung |            |            |           |              |         |        |         |
| MFE ML                 | 0.0013     | 0.1136     | 0.9871    | 0.5650       | -0.0841 | 1.4482 | -6213.1 |
| MCMC-MAP               | -0.0187    | 0.0968     | 0.9874    | 0.6108       | -0.0732 | 1.4373 | -6213.0 |
| MCMC-Mittelwert        | -0.0192    | 0.1007     | 0.9866    | 1.0153       | -0.0733 | 1.4484 | -6213.5 |
| MCMC-Median            | -0.0191    | 0.1003     | 0.9867    | 0.8620       | -0.0732 | 1.4477 | -6213.2 |

Tabelle 4.37: Schätzergebnisse verschiedener EGARCH(1,1)-Modelle für um den Mittelwert bereinigte, stetige Renditen des Dow Jones

| Schätzer               | $\alpha_0$ | $\alpha_1$ | $\beta_1$ | $\sigma_0^2$ | $c$     | $\lambda$ | $\nu$  | Log-L   |
|------------------------|------------|------------|-----------|--------------|---------|-----------|--------|---------|
| Normalverteilung       |            |            |           |              |         |           |        |         |
| MFE ML                 | 0.0159     | 0.0652     | 0.9344    | 0.4031       | -0.7945 | 0.9680    | —      | -6274.9 |
| MCMC-MAP               | 0.0156     | 0.0631     | 0.9357    | 0.5197       | 0.8128  | 1.0037    | —      | -6274.5 |
| MCMC-Mittelwert        | 0.0155     | 0.0613     | 0.9336    | 0.9257       | 0.7955  | 1.1313    | —      | -6275.4 |
| MCMC-Median            | 0.0155     | 0.0611     | 0.9338    | 0.7901       | 0.7943  | 1.1170    | —      | -6275.4 |
| <i>t</i> -Verteilung   |            |            |           |              |         |           |        |         |
| MFE ML                 | 0.0117     | 0.0588     | 0.9401    | 0.4279       | -0.7692 | 1.1314    | 8.3877 | -6201.3 |
| MCMC-MAP               | 0.0112     | 0.0586     | 0.9410    | 0.5346       | 0.7326  | 1.1313    | 8.3644 | -6200.8 |
| MCMC-Mittelwert        | 0.0117     | 0.0554     | 0.9392    | 0.9624       | 0.7756  | 1.2580    | 8.5950 | -6201.2 |
| MCMC-Median            | 0.0116     | 0.0552     | 0.9395    | 0.8474       | 0.7763  | 1.2493    | 8.5226 | -6201.2 |
| <i>GED</i> -Verteilung |            |            |           |              |         |           |        |         |
| MFE ML                 | 0.0134     | 0.0617     | 0.9379    | 0.4150       | -0.7827 | 1.0507    | 1.4557 | -6209.1 |
| MCMC-MAP               | 0.0119     | 0.0531     | 0.9381    | 0.4225       | 0.8084  | 1.0145    | 1.4506 | -6208.6 |
| MCMC-Mittelwert        | 0.0108     | 0.0499     | 0.9368    | 0.7195       | 0.7630  | 1.1851    | 1.4526 | -6209.0 |
| MCMC-Median            | 0.0107     | 0.0498     | 0.9371    | 0.6149       | 0.7633  | 1.1690    | 1.4525 | -6209.1 |

Tabelle 4.38: Schätzergebnisse verschiedener AP-ARCH(1,1)-Modelle für um den Mittelwert bereinigte, stetige Renditen des DAX

## 4.8 Anwendungsbeispiel zur Modellselektion

Die Entscheidung, welches Modell am besten geeignet ist die bedingten Varianzen zu modellieren, kann mit Hilfe der verschiedenen Methoden aus Abschnitt 3.3 getroffen werden. Exemplarisch wird das Vorgehen hier für die drei GARCH-Modelle mit unterschiedlichen Verteilungsannahmen und dem GJR-GARCH<sub>*t*</sub>- und GJR-GARCH<sub>*GED*</sub>-Modell für die stetigen Renditen des DAX gezeigt.

Für die ML-Schätzung gibt Tabelle 4.39 die Werte des Bayesschen Informationskriteriums (BIC) an.

$$BIC = \log f(Y | \hat{\theta}_i, M_i) - \frac{n_i}{2} \cdot \log T,$$

mit dem ML-Schätzer  $\hat{\theta}_i$  für Modell  $M_i$ , der Dimension  $n_i$  des jeweiligen Modells und  $T$  Beobachtungen.

Außerdem sind die Ergebnisse für Chib's Candidate Estimator (CE) und den Harmonic Mean Estimator (HME) für die MCMC-Schätzung angegeben. Dabei wurde die MCMC-Schätzung wie bereits beschrieben durchgeführt. Um die Anzahl der Parameter zwischen

| Schätzung                | BIC     | CE      | HME     |
|--------------------------|---------|---------|---------|
| GARCH                    | -7753.8 | -7758.3 | -7742.9 |
| GARCH <sub>t</sub>       | -7635.3 | -7635.8 | -7620.6 |
| GARCH <sub>GED</sub>     | -7664.3 | -7675.2 | -7649.8 |
| GJR-GARCH <sub>t</sub>   | -7608.3 | -7612.5 | -7590.3 |
| GJR-GARCH <sub>GED</sub> | -7639.1 | -7646.9 | -7620.8 |

Tabelle 4.39: Verschiedene Kriterien zur Modellselektion

ML-Schätzung und MCMC-Schätzung vergleichbar zu gestalten, wurden die MCMC-Schätzung für die Berechnung in diesem Teil mit einem konstanten Wert für  $\sigma_0^2$  durchgeführt. Miazhynskaia et al. (2006) [75] weisen darauf hin, dass der Candidate Estimator von Chib zu ähnlichen Ergebnissen führt wie bspw. das Bridge Sampling oder das Reciprocal Importance Sampling, daher werden hier nur die Ergebnisse für den CE präsentiert. Miazhynskaia et al. verwenden die Log-Likelihoodwerte anstatt der Likelihoodwerte für ihre Berechnungen. In Abschnitt 3.3 wurde bereits darauf hingewiesen, dass die Verwendung von Likelihoodwerten zu numerischen Problemen führt. Dies trifft insbesondere auf die Berechnung des Harmonic Mean Estimators zu. Bei der Verwendung von Log-Likelihoodwerten besteht dieses Problem nicht. Daher werden auch hier die Log-Likelihoodwerte benutzt.

Mit Hilfe der Ergebnisse aus Tabelle 4.39 können nun die Bayes-Faktoren berechnet werden. Tabelle 4.41 zeigt exemplarisch die Bayes-Faktoren, die sich aus dem CE ergeben, für das GARCH-Modell im Vergleich zu den anderen Varianten und ebenso für das GARCH<sub>t</sub>-Modell.

| Bayes-Faktor          | Interpretation                 |
|-----------------------|--------------------------------|
| $B_{12} < 1/10$       | Starker Hinweis auf Modell 2   |
| $1/10 < B_{12} < 1/3$ | Moderater Hinweis auf Modell 2 |
| $1/3 < B_{12} < 1$    | Schwacher Hinweis auf Modell 2 |
| $1 < B_{12} < 3$      | Schwacher Hinweis auf Modell 1 |
| $3 < B_{12} < 10$     | Moderater Hinweis auf Modell 1 |
| $B_{12} > 10$         | Starker Hinweis auf Modell 1   |

Tabelle 4.40: Jeffrey's Skala zur Interpretation von Bayes-Faktoren

| Modell 1 | Modell 2          | BF          | Modell 1  | Modell 2          | BF          |
|----------|-------------------|-------------|-----------|-------------------|-------------|
| GARCH    | $GARCH_t$         | 6.2940e-054 | $GARCH_t$ | GARCH             | 1.5888e+053 |
| GARCH    | $GARCH_{GED}$     | 8.1307e-037 | $GARCH_t$ | $GARCH_{GED}$     | 1.2918e+017 |
| GARCH    | $GJR-GARCH_t$     | 4.7848e-064 | $GARCH_t$ | $GJR-GARCH_t$     | 7.6022e-011 |
| GARCH    | $GJR-GARCH_{GED}$ | 4.1648e-049 | $GARCH_t$ | $GJR-GARCH_{GED}$ | 6.6171e+004 |

Tabelle 4.41: Bayes-Faktoren

Für die Interpretation der Bayes-Faktoren wird hier Miazhynskaia et al. folgend, die durch Wasserman (1997) [96] modifizierte Skala von Jeffreys (1961) [60] verwendet (Tabelle 4.40).

Es zeigt sich, dass das GARCH-Modell von allen anderen Modell-Varianten dominiert wird. Das  $GARCH_t$  wird allen anderen Modellen vorgezogen mit Ausnahme des  $GJR-GARCH_t$ -Modells. Dieses Modell wird bei dieser Auswahl an Modellen favorisiert.

Die Modellwahrscheinlichkeiten können dazu verwendet werden verschiedene Modellvarianten für die Prognose zukünftiger Werte zu verknüpfen und ein gewichtetes Mittel aus den einzelnen Prognosen zu bestimmen.

## Kapitel 5

# Zusammenfassung und Ausblick

Die weit verbreitete Modellierung von Kapitalmarktrenditen durch GARCH-Modelle wird in der vorliegenden Arbeit aus bayesianischer Sicht demonstriert. Anhand simulierter Datensätze werden Möglichkeiten der Schätzung von GARCH-Modellen durch MCMC-Methoden erläutert und besondere Aspekte der bayesschen Betrachtung diskutiert. Zunächst werden die Ergebnisse der MCMC-Schätzung für große Stichproben anhand von aus der a posteriori Verteilung berechneten Punktschätzern mit den Punktschätzern der ML-Schätzung verglichen.

Die beschriebenen MCMC-Methoden zeichnen sich dadurch aus, dass sie leicht zu implementieren sind und keine hohen Anforderungen an die Spezifikation der Startwerte oder Vorschlagsdichte stellen. Durch Berücksichtigung der Abhängigkeitsstrukturen der Parameter ergibt sich meist problemlos eine gute Akzeptanzrate, so dass i.d.R. keine Anpassung der Kovarianzmatrix notwendig ist. Auch der zeitliche Aufwand der MCMC-Schätzung selbst ist überschaubar.

Das asymptotische Verhalten der Schätzer der GARCH-Parameter mittels ML-Schätzung wird in unterschiedlichen Simulationsstudien untersucht und mit den Ergebnissen der a posteriori Verteilung aus der Bayes-Schätzung verglichen. Es zeigt sich, dass die a posteriori Verteilung der Parameter selbst bei kleinen Stichproben geeignet ist die Verteilung der Parameter abzubilden. Allerdings muss einschränkend bemerkt werden, dass dies nur der Fall ist, wenn die Stichprobe die dem datengenerierenden Prozess zugrunde liegenden Parameter gut widerspiegelt. Durch die Verwendung flacher a priori Verteilungen für alle Parameter erhält man mit dem MCMC-MAP stets Ergebnisse für die Punktschätzer, welche dem ML-Schätzer gleichen. Allerdings wird in kleinen Stichproben hierdurch meist die Varianz überschätzt. In weiteren Simulationsstudien wurde für die Parameter der  $t$ - bzw.  $GED$ -Verteilung das asymptotische Verhalten der ML-Schätzer der Scharparameter



untersucht, dabei zeigt sich in der durchgeführten Simulation, dass der Scharparameter der *GED*-Verteilung bereits in kleinen Stichproben annähernd normalverteilt ist.

Die gewonnen Erkenntnisse werden verwendet, um verschiedene Erweiterungen des GARCH-Modells ebenfalls aus der Sicht eines Bayesianers zu untersuchen. Zu diesem Zweck werden an simulierten Daten aus EGARCH-, GJR-, AP-ARCH-, T-GARCH und TS-GARCH-Modellen unterschiedliche Simulationen zur MCMC-Schätzung durchgeführt. Auch bei diesen Simulationen erfolgt ein Vergleich der jeweiligen a posteriori Verteilungen mit den Verteilungen, die sich aus Simulationsstudien zu den ML-Schätzungen ergeben. Die Ergebnisse entsprechen hierbei den Ergebnissen bei dem GARCH-Modell. Das beschriebene MCMC-Verfahren wird außerdem an zwei einfachen Modellen zur Modellierung von bedingten Korrelationen auch auf multivariate Datensätze angewendet. Ein abschließendes Beispiel demonstriert die Modellselektion an ausgewählten erweiterten GARCH-Modellen und ermöglicht so die Bestimmung der Modell-Likelihoods, welche bspw. verwendet werden kann, um zukünftige Werte der bedingten Varianzen mittels Model Averaging zu prognostizieren.

---

## Literaturverzeichnis

- [1] ASAI, Manabu: Comparison of MCMC Methods for Estimating GARCH Models. In: *J. Japan Statistical Society* 36 (2006), Nr. 2, S. 199–212
- [2] BACHELIER, L.: Theory of Speculation. Übersetzung des französischen Originalartikels in Cootner (Herausgeber). In: *The Random Character of Stock Market Prices*, MIT Press, 1964, 1900, S. 17–78
- [3] BAILLIE, Richard T.: Estimation of GARCH Models from the Autocorrelations of the Squares of a Process. In: *Journal of Time Series Analysis* 22 (2001), November, S. 631–650(20)
- [4] BAILLIE, Richard T. ; BOLLERSLEV, Tim: Intra Day And Inter Market Volatility In Foreign Exchange Rates. 1989. – Forschungsbericht
- [5] BAILLIE, Richard T. ; BOLLERSLEV, Tim ; MIKKELSEN, Hans O.: Fractionally integrated generalized autoregressive conditional heteroskedasticity. In: *Journal of Econometrics* 74 (1996), September, Nr. 1, S. 3–30
- [6] BAUWENS, L. ; LAURENT, S. ; ROMBOUTS, J.V.K.: Multivariate GARCH Models: A Survey. In: *Journal of Applied Econometrics* 21 (2006), S. 79–109
- [7] BAUWENS, Michel Luc und L. Luc und Lubrano: Bayesian inference on GARCH models using the Gibbs sampler. In: *Econometrics Journal* 1 (1998), S. C23–C46
- [8] BAUWENS, Michel Luc und L. Luc und Lubrano: Bayesian option pricing using asymmetric GARCH models. In: *Journal of Empirical Finance* 9 (2002), S. 321–342
- [9] BERA, Anil K. ; HIGGINS, Matthew L.: ARCH Models: Properties, Estimation and Testing. In: *Journal of Economic Surveys* 7 (1993), Dezember, Nr. 4, S. 305–366

- [10] BERA, Anil K. ; KIM, Sangwhan: Testing constancy of correlation and other specifications of the BGARCH model with an application to international equity returns. In: *Journal of Empirical Finance* 9 (2002), März, Nr. 2, S. 171–195
- [11] BERNDT, E. K. ; HALL, B. H. ; HALL, R. E. ; HAUSMAN, Jerry A.: Estimation and Inference in Nonlinear Structural Models. In: *Annals of Economic and Social Measurement* 3 (1974), Nr. 4, S. 653–665
- [12] BILLER, Clemens: *Bayesianische Ansätze zur nonparametrischen Regression*, Ludwig-Maximilians-Universität München, Diss., 1999
- [13] BLACK, Fischer: Studies in Price Volatility Changes, Proceedings of the 1976 Meeting of Business and Economics Statistics Section. In: *American Statistical Association* (1976), S. 177–181
- [14] BOLLERSLEV, Tim: Generalized autoregressive conditional heteroskedasticity. In: *Journal of Econometrics* 31 (1986), April, Nr. 3, S. 307–327
- [15] BOLLERSLEV, Tim: A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. In: *Review of Economics and Statistics* 69 (1987), S. 542–547
- [16] BOLLERSLEV, Tim: Modelling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model. In: *The Review of Economics and Statistics* 72 (1990), August, Nr. 3, S. 498–505
- [17] BOLLERSLEV, Tim: Glossary to ARCH (GARCH). (2007), Oktober
- [18] BOLLERSLEV, Tim ; CHOU, Ray Y. ; KRONER, Kenneth F.: ARCH modeling in finance : A review of the theory and empirical evidence. In: *Journal of Econometrics* 52 (1992), Nr. 1-2, S. 5–59
- [19] BOLLERSLEV, Tim ; ENGLE, Robert F. ; NELSON, Daniel B.: Arch models. In: ENGLE, R. F. (Hrsg.) ; MCFADDEN, D. (Hrsg.): *Handbook of Econometrics*
- [20] BOLLERSLEV, Tim ; WOOLDRIDGE, Jeffrey: Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. In: *Econometric Reviews* 11 (1992), Nr. 2, S. 143–172

- [21] CAI, Jun: A Markov Model of Switching-Regime ARCH. In: *Journal of Business & Economic Statistics* 12 (1994), Juli, Nr. 3, S. 309–16
- [22] CARLIN, B. P. ; LOUIS, T. A.: *Bayes and empirical Bayes methods for data analysis*. 2. Auflage. New York, New York, USA : Chapman and Hall, 2000
- [23] CHIB, Siddhartha ; GREENBERG, Edward: Understanding the Metropolis-Hastings Algorithm. In: *The American Statistician* 49 (1995), Nr. 4, S. 327–335
- [24] CHIB, Siddhartha: Marginal Likelihood from the Gibbs Output. In: *Journal of the American Statistical Association* 90 (1995), Nr. 432, S. 1313–1321
- [25] CHIB, Siddhartha ; JELIAZKOV, Ivan: Marginal Likelihood From the Metropolis-Hastings Output. In: *Journal of the American Statistical Association* 96 (2001), März, S. 270–281
- [26] CONT, Rama: Empirical properties of asset returns: stylized facts and statistical issues. In: *Quantitative Finance* 1 (2001), S. 223–236
- [27] COPELAND, Thomas E. ; WESTON, J. F.: *Financial Theory and Corporate Policy*. 3. Reading (Mass.) : Addison-Wesley, 1992
- [28] COWELS, Mary K. ; CARLIN, P. Bradley: Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review. In: *Journal of the American Statistical Association* 91 (1996), S. 893–904
- [29] DAVID, Ardia: Bayesian Estimation of the GARCH(1,1) Model with Normal Innovations / University Library of Munich, Germany. 2006. – MPRA Paper
- [30] DEMOS, Antonis ; SENTANA, Enrique: Testing for GARCH effects: a one-sided approach. In: *Journal of Econometrics* 86 (1998), Juni, Nr. 1, S. 97–127
- [31] DEUTSCHEBÖRSEAG: *Leitfaden zu den Volatilitätsindizes der Deutschen Börse*. [www.deutsche-boerse.de](http://www.deutsche-boerse.de), Januar 2007
- [32] DING, Zhuanxin ; GRANGER, Clive W. ; ENGLE, Robert: A long memory property of stock market returns and a new model. In: *Journal of Empirical Finance* 1 (1993), S. 83–106

- [33] DUAN, Jin-Chuan: Augmented GARCH(p,q) process and its diffusion limit. In: *Journal of Econometrics* 79 (1997), S. 97–127
- [34] ENGLE, Robert F.: Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. In: *Econometrica* 50 (1982), Juli, Nr. 4, S. 987–1007
- [35] ENGLE, Robert F.: Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. In: *Journal of Business & Economic Statistics* 20 (2002), Juli, Nr. 3, S. 339–50
- [36] ENGLE, Robert F. ; LILIEN, David M. ; ROBINS, Russell P.: Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model. In: *Econometrica* 55 (1987), März, Nr. 2, S. 391–407
- [37] ENGLE, Robert F. ; SHEPPARD, Kevin: Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH. 2001 (2001-15). – University of California at San Diego, Economics Working Paper Series
- [38] EVANS, Michael ; SWARTZ, Tim: Methods for Approximating Integrals in Statistics With special Emphasis on Bayesian Integration Problems. In: *Statistical Science* 3 (1995), S. 254–272
- [39] FIORENTINI, Gabriele ; CALZOLARI, Giorgio ; PANATTONI, Lorenzo: Analytic Derivatives and the Computation of GARCH Estimates. In: *Journal of Applied Econometrics* 11 (1996), Juli-August, Nr. 4, S. 399–417
- [40] GALBRAITH, J. ; ZINDE-WALSH, V.: *Properties of estimates of daily GARCH parameters based on intra-day observations*. 2000
- [41] GAMERMAN, Dani ; LOPES, Hedibert F.: *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman and Hall/CRC, 2006. – ISBN 1584885874
- [42] GELFAND, A. E. ; DEY, D. K.: Bayesian Model Choice: Asymptotics and Exact Calculations. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 56 (1994), Nr. 3, S. 501–514

- [43] GELMAN, Andrew ; CARLIN, John B. ; STERN, Hal S. ; RUBIN, Donald B.: *Bayesian Data Analysis, Second Edition*. Chapman & Hall/CRC, 2003. – ISBN 158488388X
- [44] GELMAN, Andrew ; ROBERTS, Gareth O. ; GILKS, Walter R.: Efficient Metropolis jumping rules. In: BERNADO, J. M. (Hrsg.) u. a.: *Bayesian Statistics* Bd. 5. OUP, 1996, S. 599
- [45] GESSLBAUER, Stefan: *Die Simulation von Verteilungen durch MCMC-Methoden mit adaptiver Verwerfung*. Institut für Statistik, Technische Universität Graz, Diplomarbeit, September 1997
- [46] GEWEKE, John ; MCCAUSLAND, William: Bayesian Specification Analysis in Econometrics. In: *American Journal of Agricultural Economics* 83 (2001), Nr. 5, S. 1181–86
- [47] GEWEKE, John ; TANIZAKI, Hisashi: Note on Sampling Distribution for the Metropolis-Hastings Algorithm. In: *Communications in Statistics* 32 (2003), S. 775–789
- [48] GILKS, W.R. ; RICHARDSON, S. ; SPIEGELHALTER, D.J.: *Markov Chain Monte Carlo in Practice. Interdisciplinary Statistics*. London : Chapman and Hall, 1996
- [49] GLOSTEN, Lawrence R. ; JAGANNATHAN, Ravi ; RUNKLE, David E.: On the relation between the expected value and the volatility of the nominal excess return on stocks. In: *Journal of Finance* 48 (1993), S. 1779–1801
- [50] GOURIÉROUX, Christian: *ARCH Models and Financial Applications*. Springer-Verlag New York, Inc., 1997
- [51] GREEN, Peter J.: Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. In: *Biometrika* 82 (1995), S. 711–732
- [52] GRIMMETT, G.R. ; STIRZAKER, D.R.: *Probability and Random Processes. 2*. Oxford University Press, 1992
- [53] HAAS, Markus ; MITTNIK, Stefan ; PAOLELLA, Marc S.: A New Approach to Markov-Switching GARCH Models. In: *Journal of Financial Econometrics* 2 (2004), Nr. 4, S. 493–530

- [54] HASTINGS, W.K.: Monte Carlo Sampling Methods Using Markov Chains and Their Applications. In: *Biometrika* 57 (1970), S. 97–109
- [55] HENTSCHEL, Ludger: All in the family: Nesting symmetric and asymmetric GARCH models. In: *Journal of Financial Economics* 39 (1995), S. 71–104
- [56] HIGGINS, Matt L. ; BERA, Anil K.: A class of nonlinear ARCH models. In: *International Economic Review* 33 (1989), S. 137–158
- [57] HULL, John C.: *Options, Futures, and Other Derivatives*. 5. London : Prentice-Hall, 2003
- [58] HWANG, Soosung ; PEREIRA, Pedro V.: Small sample properties of GARCH estimates and persistence. In: *European Journal of Finance* 12 (2006), Oktober, Nr. 6-7, S. 473–494
- [59] JEANTHEAU, Thierry: Strong Consistency Of Estimators For Multivariate Arch Models. In: *Econometric Theory* 14 (1998), Februar, Nr. 01, S. 70–86
- [60] JEFFREYS, Harold: *Theory of probability*. 3rd ed.,. Clarendon Press, Oxford, 1961
- [61] JOHNSON, N. L. ; KOTZ, S. ; BALAKRISHNAN, A. W. N.: *Continuous univariate distributions: Vol. 2*. Wiley, New York, 1995
- [62] KIM, Sangjoon ; SHEPHARD, Neil ; CHIB, Siddhartha: Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. In: *The Review of Economic Studies* 65 (1998), Juli, Nr. 3, S. 361–393
- [63] KRISTENSEN, Dennis ; LINTON, Oliver: A Closed-Form Estimator For The Garch(1,1) Model. In: *Econometric Theory* 22 (2006), April, Nr. 02, S. 323–337
- [64] LAMOUREUX, Christopher G. ; LASTRAPES, William D.: Persistence in Variance, Structural Change, and the GARCH Model. In: *Journal of Business & Economic Statistics* 8 (1990), April, Nr. 2, S. 225–34
- [65] LASTRAPES, William D.: Exchange Rate Volatility and U.S. Monetary Policy: An ARCH Application. In: *Journal of Money, Credit and Banking* 21 (1989), Februar, Nr. 1, S. 66–77

- [66] LASTRAPES, William D. ; KORAY, Faik: Exchange rate volatility and U.S. multilateral trade flows. In: *Journal of Macroeconomics* 12 (1990), Nr. 3, S. 341–362
- [67] LEE, Sang-Won ; HANSEN, Bruce E.: Asymptotic Theory for the Garch(1,1) Quasi-Maximum Likelihood Estimator. In: *Econometric Theory* 10 (1994), März, Nr. 01, S. 29–52
- [68] LUMSDAINE, Robin L.: Consistency and Asymptotic Normality of the Quasi-Maximum Likelihood Estimator in IGARCH(1,1) and Covariance Stationary GARCH(1,1) Models. In: *Econometrica* 64, Nr. 3, S. 575–596
- [69] MALMSTEN, Hans: Evaluating exponential GARCH models / Stockholm School of Economics. 2004 (564). – Working Paper Series in Economics and Finance
- [70] MALMSTEN, Hans ; TERÄSVIRTA, Timo: Evaluating Exponential GARCH Models. In: *w* (2002), November
- [71] MALMSTEN, Hans ; TERÄSVIRTA, Timo: Stylized Facts of Financial Time Series and Three Popular Models of Volatility / Stockholm School of Economics. 2004. – Working Paper Series in Economics and Finance
- [72] MANDELBROT, B.: The Variation of Certain Speculative Prices. In: *Journal of Business* 36 (1963), S. 394–419
- [73] MARCUCCI, Juri: Forecasting Stock Market Volatility with Regime-Switching GARCH Models. In: *Studies in Nonlinear Dynamics & Econometrics* 9 (2005), Nr. 4, S. 1145–1145
- [74] METROPOLIS, N. ; ROSENBLUTH, A.W. ; ROSENBLUTH, M.N. ; TELLER, A.H. ; TELLER, E.: Equations of State Calculations by Fast Computing Machines. In: *Journal of Chemical Physics* 21 (1953), S. 1087–1092
- [75] MIAZHYNKAIA, Tatiana: A comparison of Bayesian model selection based on MCMC with an application to GARCH-type models. In: *Statistical Papers* 47 (Oktober 2006), S. 525–549(25)
- [76] MITSUI, H. ; WATANABE, T.: Bayesian analysis of GARCH option pricing models (in Japanese). In: *The Journal of the Japan Statistical Society (Japanese Issue)* 33 (2003), S. 307–324



- [77] NAKATANI, Tomoaki ; TERASVIRTA, Timo: Testing for volatility interactions in the Constant Conditional Correlation GARCH model. In: *Econometrics Journal* 12 (2009), 03, Nr. 1, S. 147–163
- [78] NAKATSUMA, Teruo ; TSURUMI, Hiroki: ARMA-GARCH Models: Bayes Estimation Versus MLE, and Bayes Non-stationarity Test / Rutgers University, Department of Economics. 1996 (199619). – Departmental Working Papers
- [79] NELSON, D. B.: Conditional heteroskedasticity in asset returns: A new approach. In: *Econometrica* 59 (1991), S. 347–370
- [80] NELSON, Daniel B. ; CAO, Charles Q.: Inequality Constraints in the Univariate GARCH Model. In: *Journal of Business & Economic Statistics* 10 (1992), April, Nr. 2, S. 229–35
- [81] ROBERT, Christian P. ; CASELLA, George: *Monte Carlo Statistical Methods*. 2nd. New York: Springer-Verlag, 2004
- [82] SCHLITGEN, Rainer ; STREITBERG, Bernd H. J.: *Zeitreihenanalyse*. 9., unwesentlich veränd. Aufl. Oldenbourg, 2001. – ISBN 3–486–25725–0
- [83] SCHWERT, G. W.: Why does stock market volatility change over time? In: *Journal of Finance* 44 (1989), S. 1115–1153
- [84] STORTI, G.: Minimum distance estimation of GARCH(1,1) models. In: *Comput. Stat. Data Anal.* 51 (2006), Nr. 3, S. 1803–1821. – ISSN 0167–9473
- [85] TAKAISHI, Tetsuya: Bayesian Estimation of GARCH Model by Hybrid Monte Carlo. In: *JCIS*, 2006
- [86] TAKAISHI, Tetsuya: Financial Time Series Analysis of SV Model by Hybrid Monte Carlo. In: *ICIC (1)*, 2008, S. 929–936
- [87] TAKAISHI, Tetsuya: An Adaptive Markov Chain Monte Carlo Method for GARCH Model. In: *ENGINEERING.COMPLEX SCIENCES* (2009)
- [88] TAKAISHI, Tetsuya: Bayesian inference with an adaptive proposal density for GARCH models. (2009)
- [89] TAYLOR, Stephen: *Modelling financial time series*. New York, NY : Wiley, 1986

- [90] TIERNEY, Luke: Markov Chains for exploring posterior distributions. In: *The Annals of Statistics* 22 (1994), Nr. 4, S. 1701–1762
- [91] TSAY, S. R.: *Analysis of Financial Time Series*. New York : Wiley series in probability and statistics, 2002. – ISBN 0–471–41544–8
- [92] TSE, Yiu K.: A test for constant correlations in a multivariate GARCH model. In: *Journal of Econometrics* 98 (2000), September, Nr. 1, S. 107–127
- [93] TSE, Yiu. K. ; TSUI, Albert K. C.: A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-Varying Correlations. In: *Journal of Business & Economic Statistics* 20 (2002), Juli, Nr. 3, S. 351–62
- [94] VRONTOS, D. ; DELLAPORTAS, P. ; POLITIS, D.N.: Full Bayesian Inference for GARCH and EGARCH Models. In: *Journal of Business and Economic Statistics* 18 (2000), April, Nr. 2, S. 187–198
- [95] VRONTOS, D. ; DELLAPORTAS, P. ; POLITIS, D.N.: A full-factor multivariate GARCH model. In: *Econometrics Journal* 6 (2000), S. 312–334
- [96] WASSERMAN, L.: Bayesian Model Selection and Model Averaging / Department of Statistics, Carnegie Mellon University. 1997 (666). – Forschungsbericht
- [97] ZAKOIAN, Jean-Michel: *Threshold heteroskedastic models*. Unpublished paper (Institut National de la Statistique et des Etudes Economiques), Paris, 1991

# Anhang A

## Anhang

### A.1 Grafiken zu den Ergebnissen der robusten Schätzung der Varianzen bei der ML-Simulation

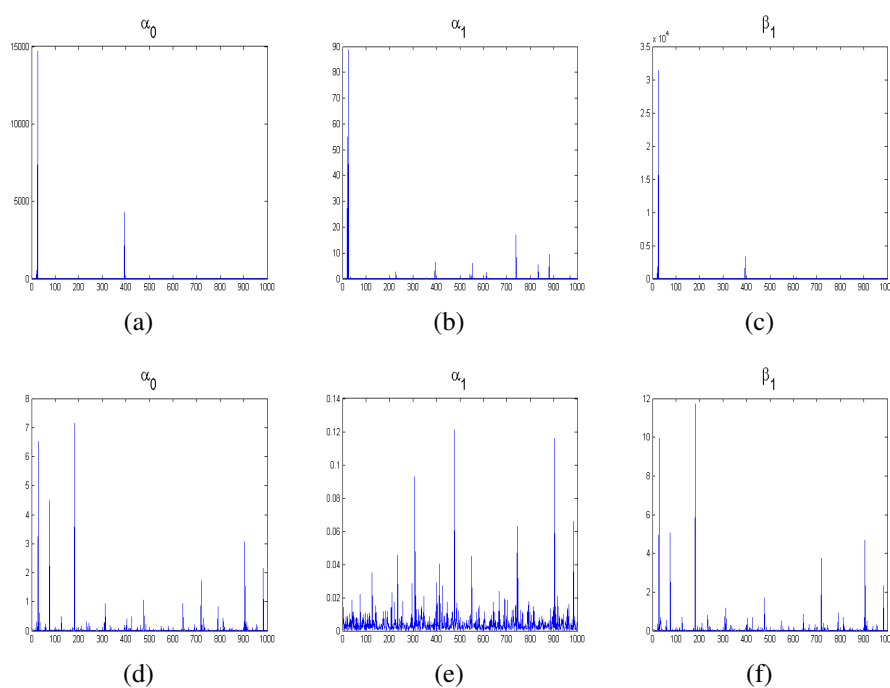


Abbildung A.1: Ergebnisse der robusten Schätzung der Varianzen bei der ML-Simulationsstudie (GARCH) ( $n=100, 250$ )

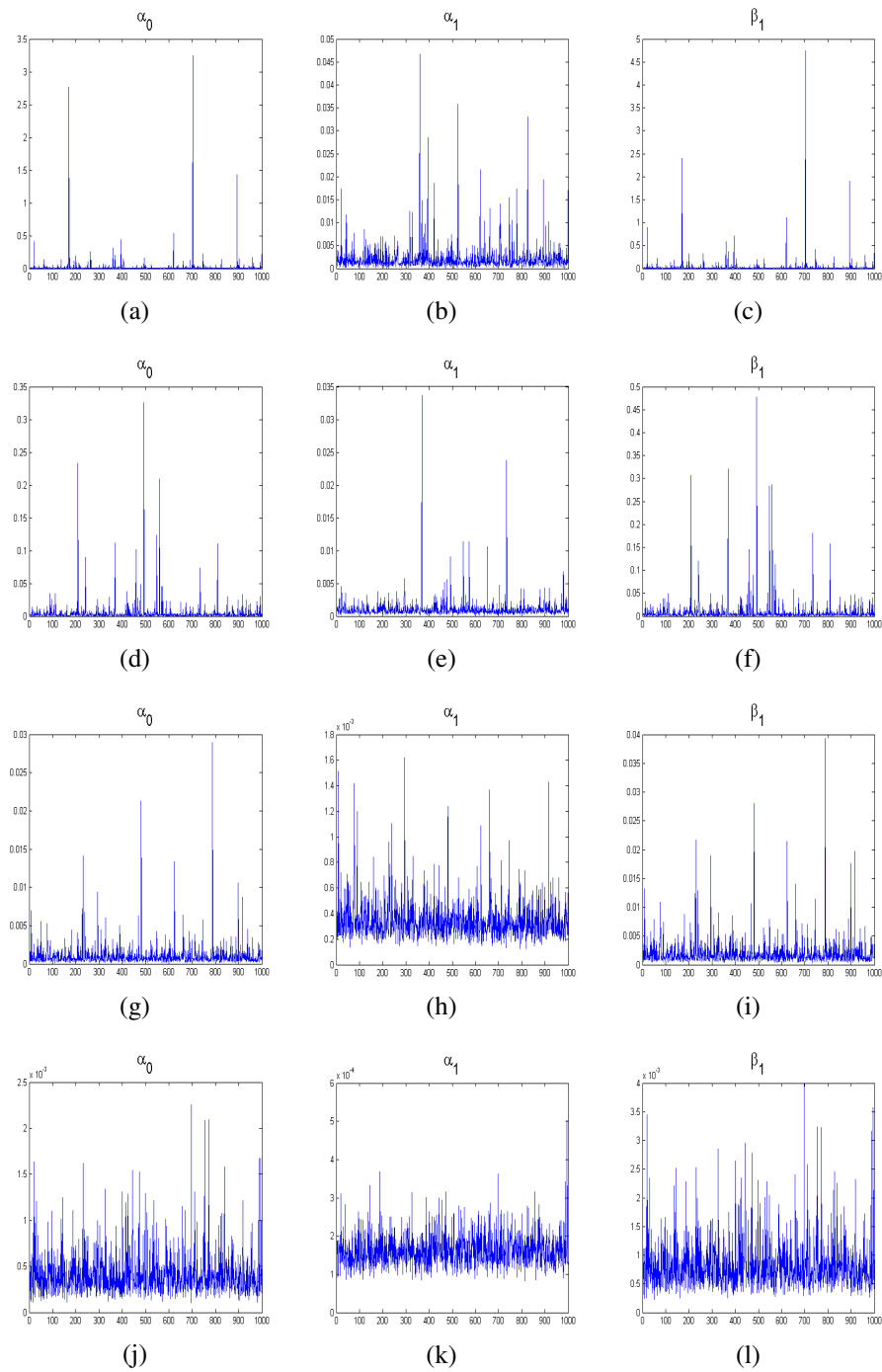


Abbildung A.2: Verteilung der ML-Parameterschätzer bei unterschiedlichen Stichprobenumfängen im Vergleich zu der a posteriori Verteilung der MCMC-Schätzung ( $n=500, 1000, 2500$  und  $5000$ )

## A.2 Grafiken zu den Simulationsstudien zur ML-Schätzung eines GARCH(1,1)-Modells

### A.2.1 Variante 2

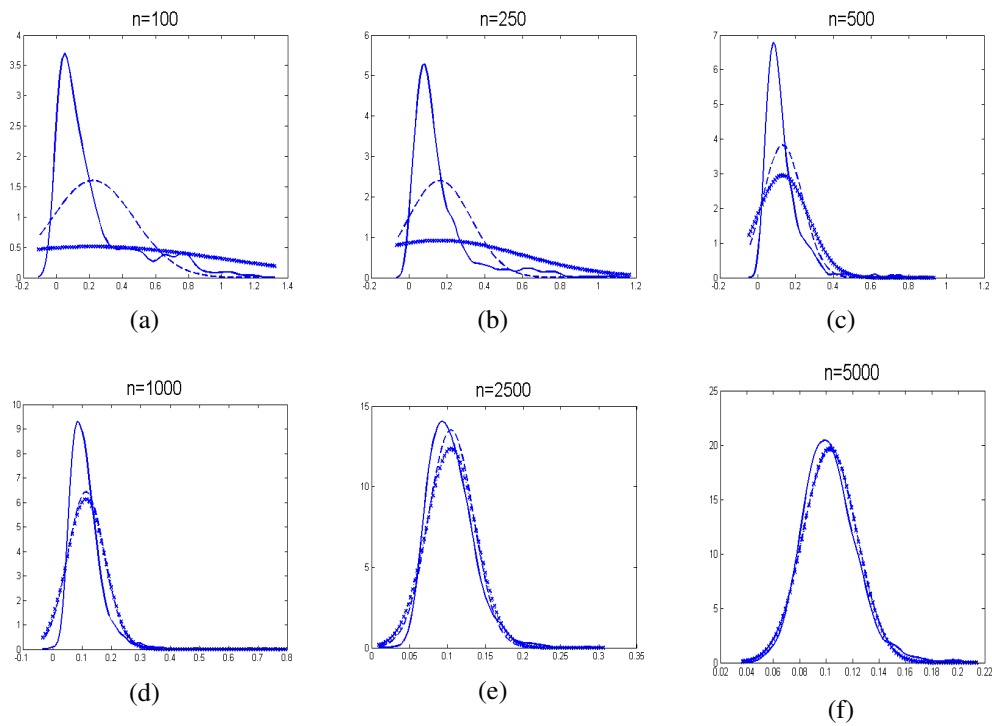


Abbildung A.3: Verteilung des ML-Schätzers des Parameters  $\alpha_0: 0.1$  bei Variante 2, bei unterschiedlichen Stichprobenumfängen

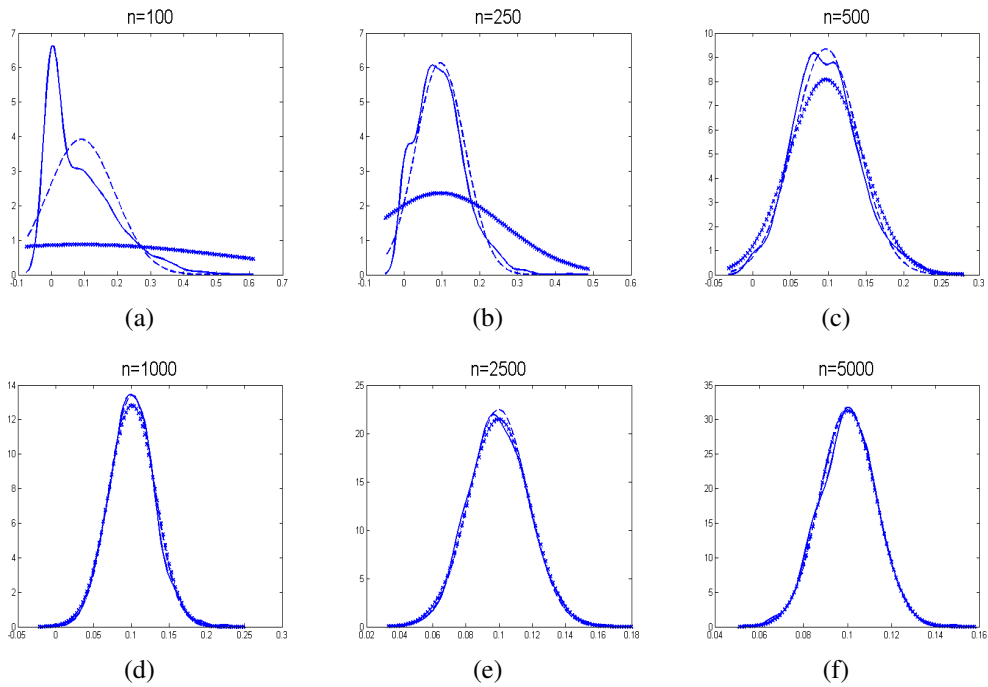


Abbildung A.4: Verteilung des ML-Schätzers des Parameters  $\alpha_1: 0.1$  bei Variante 2, bei unterschiedlichen Stichprobenumfängen

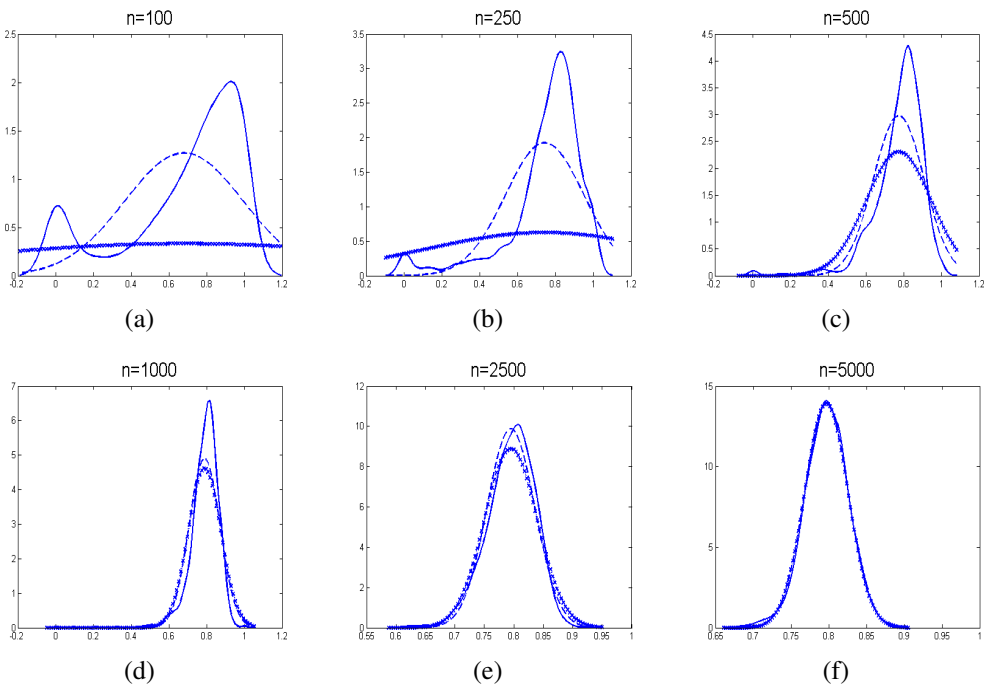


Abbildung A.5: Verteilung des ML-Schätzers des Parameters  $\beta_1: 0.8$  bei Variante 2, bei unterschiedlichen Stichprobenumfängen

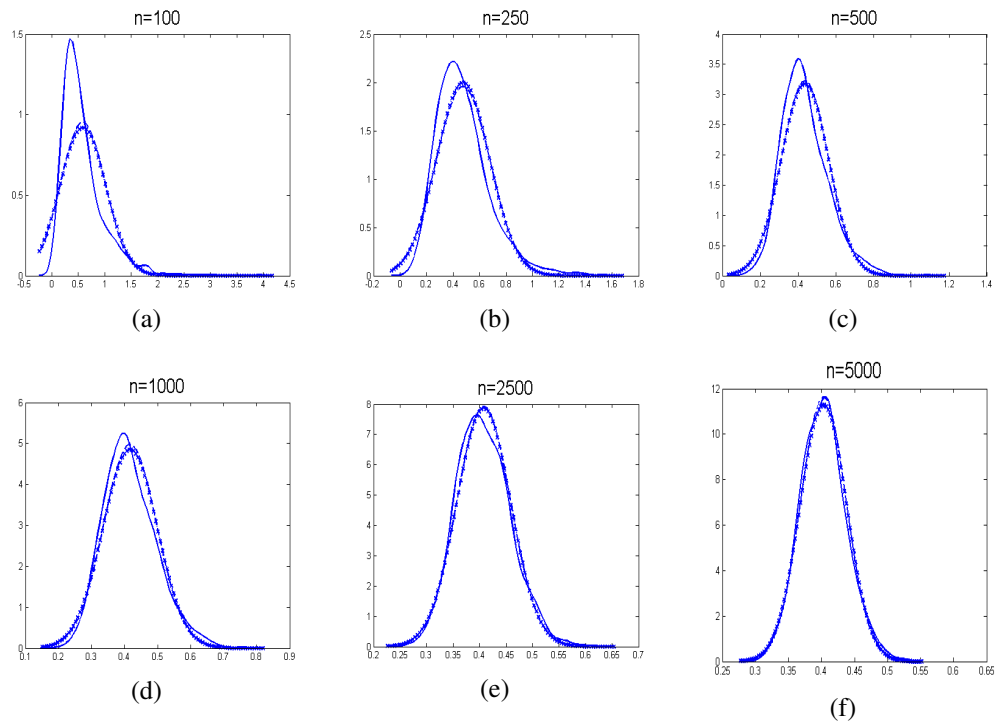
**A.2.2 Variante 4**

Abbildung A.6: Verteilung des ML-Schätzers des Parameters  $\alpha_0$ : 0.4 bei Variante 4, bei unterschiedlichen Stichprobenumfängen

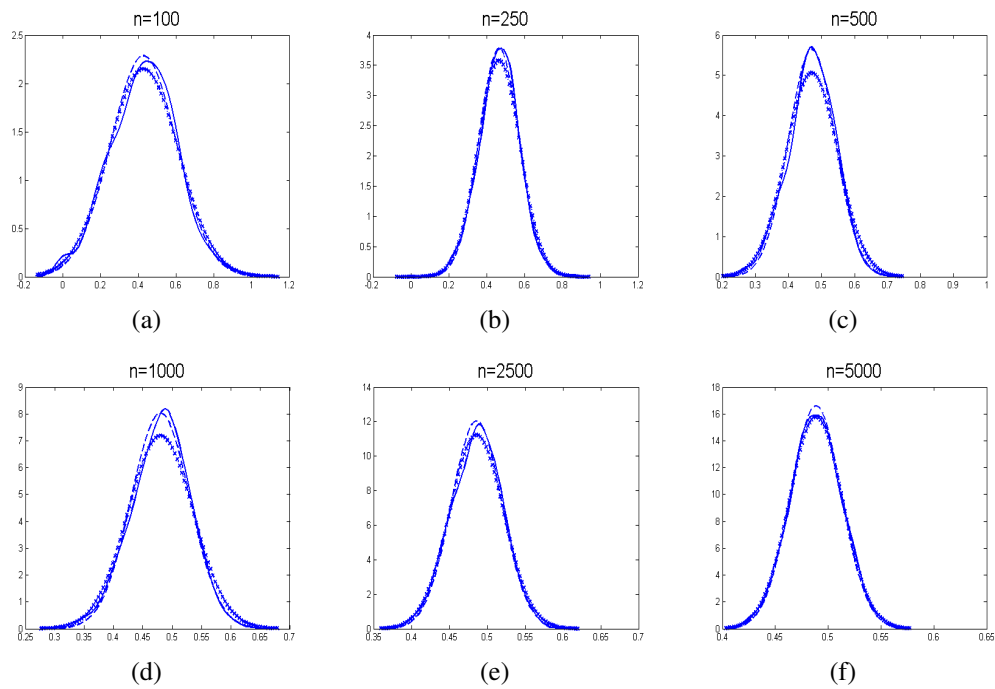


Abbildung A.7: Verteilung des ML-Schätzers des Parameters  $\alpha_1: 0.49$  bei Variante 4, bei unterschiedlichen Stichprobenumfängen

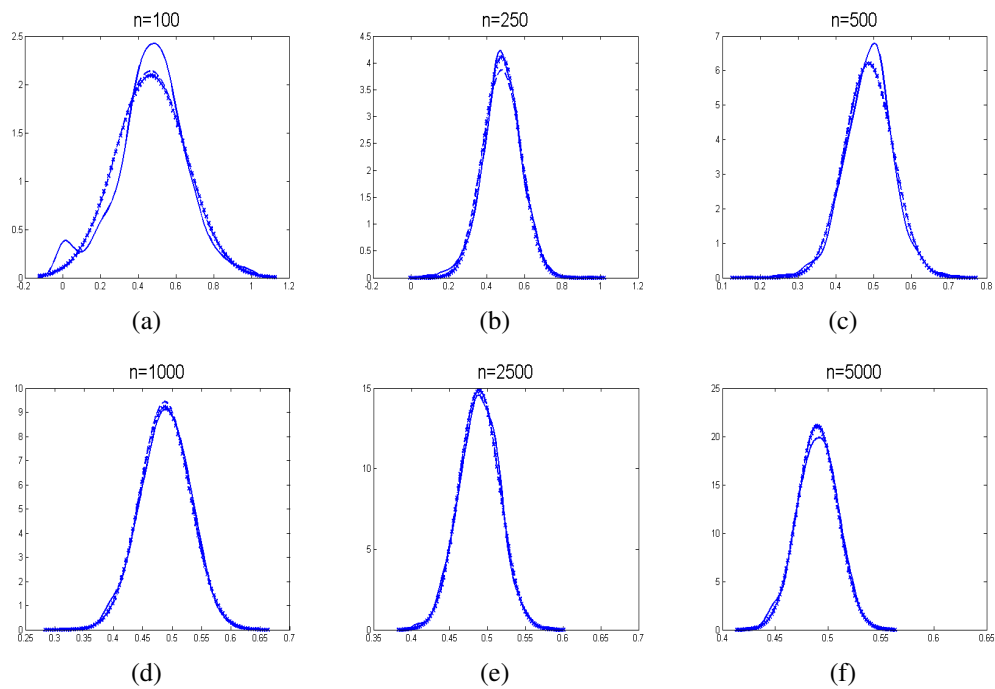


Abbildung A.8: Verteilung des ML-Schätzers des Parameters  $\beta_1: 0.49$  bei Variante 4, bei unterschiedlichen Stichprobenumfängen



| $n$  |                          | $\alpha_0 : 0.1$ | $\alpha_1 : 0.3$ | $\beta_1 : 0.3$ |
|------|--------------------------|------------------|------------------|-----------------|
| 100  | Mittelwert (1)           | 0.0960           | 0.2688           | 0.3454          |
|      | Varianz robust (3)       | 0.0170           | 0.0652           | 0.6802          |
|      | Varianz der Schätzer (2) | 0.0031           | 0.0320           | 0.0914          |
| 250  | Mittelwert (1)           | 0.1015           | 0.2884           | 0.3010          |
|      | Varianz robust (3)       | 0.0040           | 0.0213           | 0.1120          |
|      | Varianz der Schätzer (2) | 0.0017           | 0.0128           | 0.0469          |
| 500  | Mittelwert (1)           | 0.1039           | 0.2919           | 0.2892          |
|      | Varianz robust (3)       | 0.0017           | 0.0063           | 0.0424          |
|      | Varianz der Schätzer (2) | 0.0010           | 0.0056           | 0.0233          |
| 1000 | Mittelwert (1)           | 0.1015           | 0.2973           | 0.2941          |
|      | Varianz robust (3)       | 0.0005           | 0.0030           | 0.0126          |
|      | Varianz der Schätzer (2) | 0.0004           | 0.0028           | 0.0107          |
| 2500 | Mittelwert (1)           | 0.1011           | 0.2988           | 0.2964          |
|      | Varianz robust (3)       | 0.0002           | 0.0012           | 0.0043          |
|      | Varianz der Schätzer (2) | 0.0002           | 0.0011           | 0.0046          |
| 5000 | Mittelwert (1)           | 0.1007           | 0.3003           | 0.2970          |
|      | Varianz robust (3)       | 0.0001           | 0.0006           | 0.0020          |
|      | Varianz der Schätzer (2) | 0.0001           | 0.0006           | 0.0020          |

Tabelle A.1: Ergebnisse für 1000 Replikationen eines GARCH(1,1)-Modells mit  $\alpha_0 = 0.1$ ,  $\alpha_1 = 0.3$  und  $\beta_1 = 0.3$  bei unterschiedlichen Stichprobenumfängen  $n$

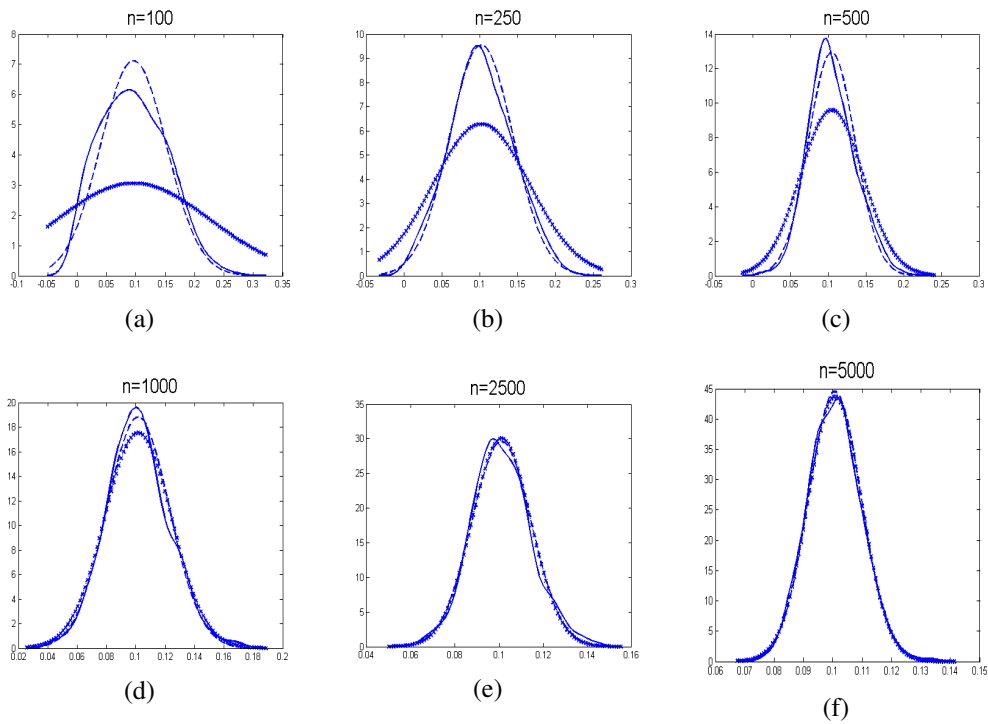
**A.2.3 Variante 5**

Abbildung A.9: Verteilung des ML-Schätzers des Parameters  $\alpha_0: 0.1$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

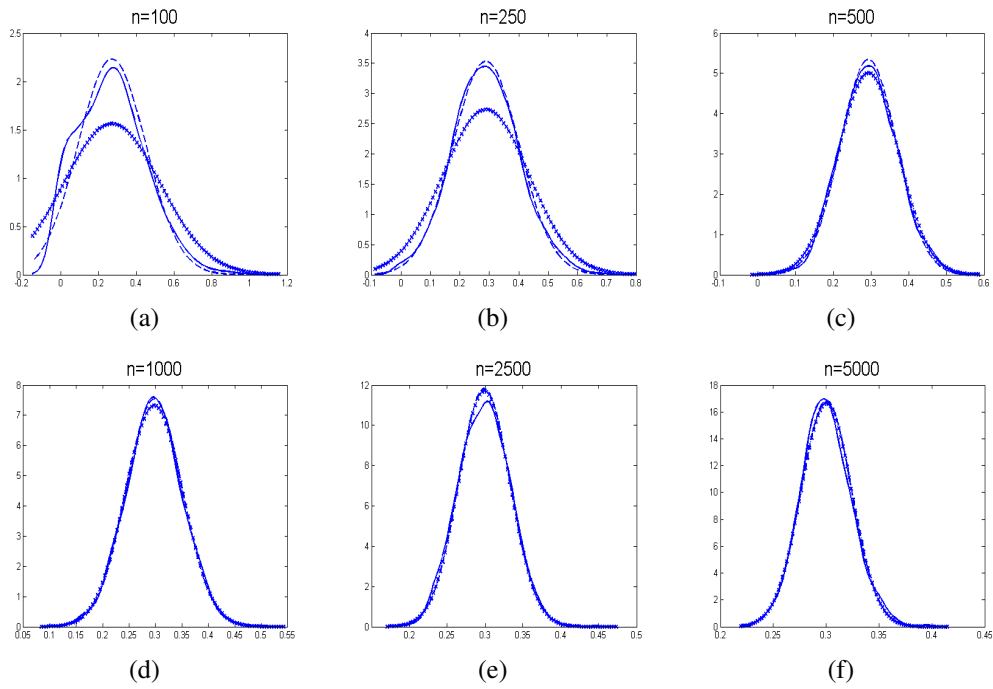


Abbildung A.10: Verteilung des ML-Schätzers des Parameters  $\alpha_1: 0.3$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

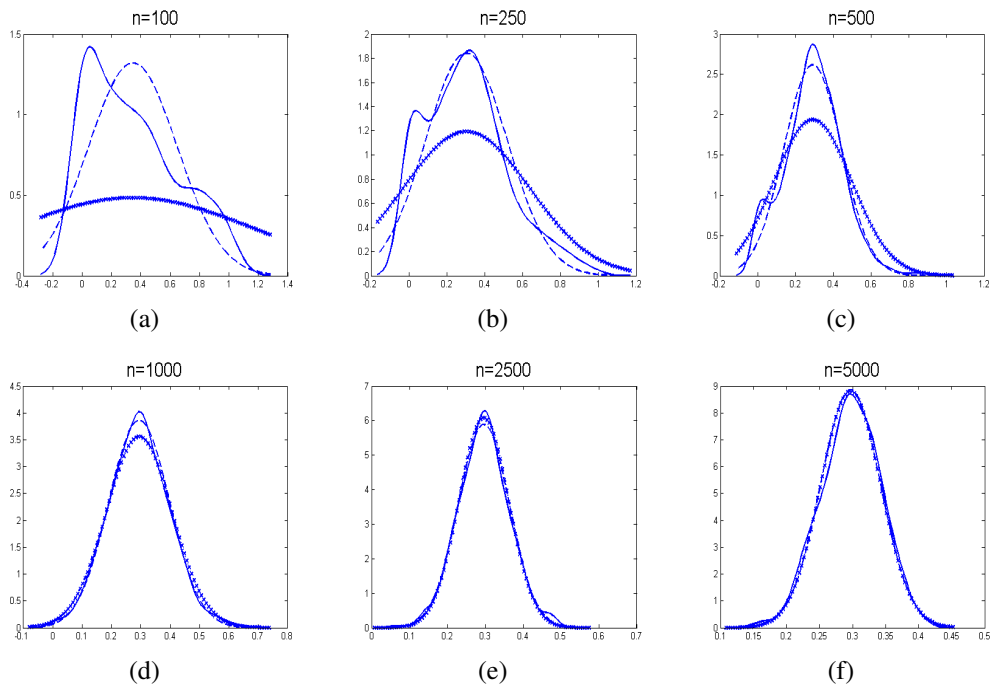


Abbildung A.11: Verteilung des ML-Schätzers des Parameters  $\beta_1: 0.3$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

### A.3 Ergebnisse der Simulationsstudie zum TS-GARCH(1,1)-Modell

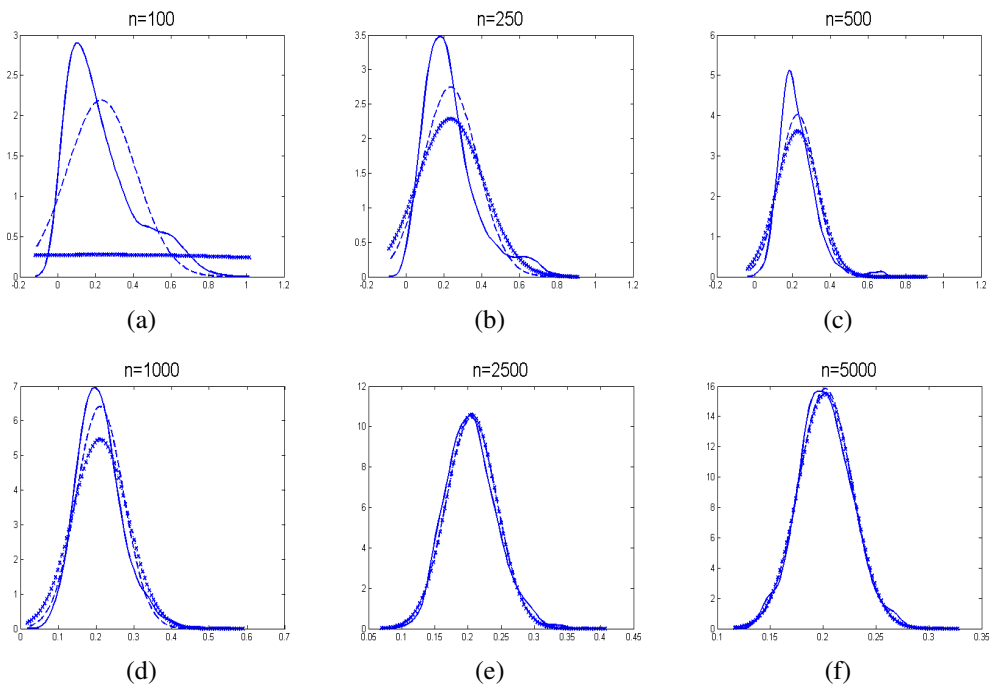


Abbildung A.12: Verteilung des ML-Schätzers des Parameters  $\alpha_0: 0.2$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

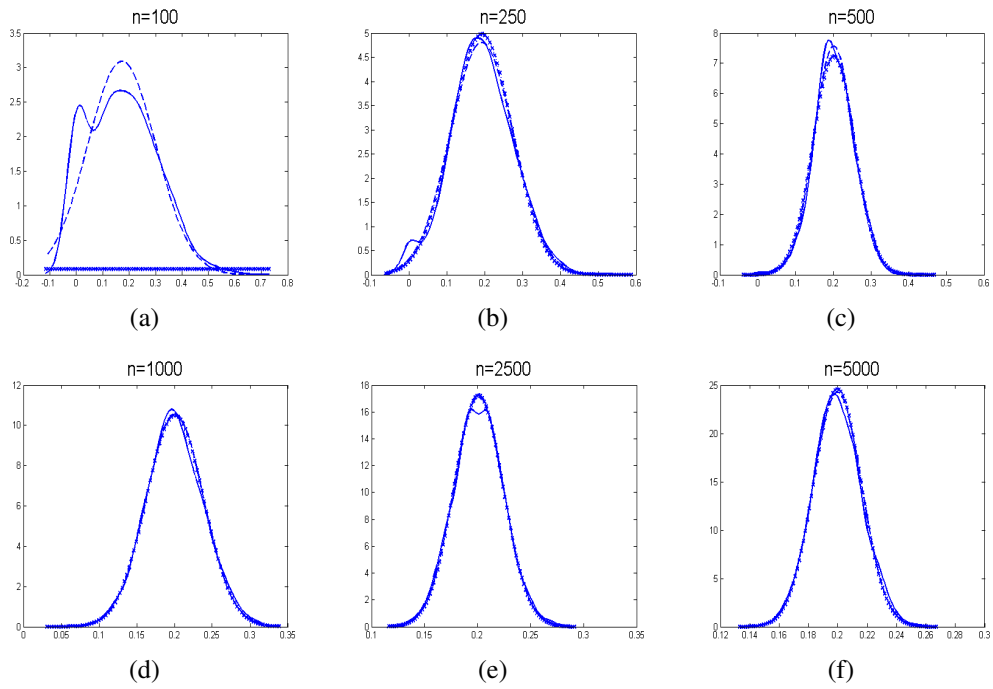


Abbildung A.13: Verteilung des ML-Schätzers des Parameters  $\alpha_1: 0.2$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

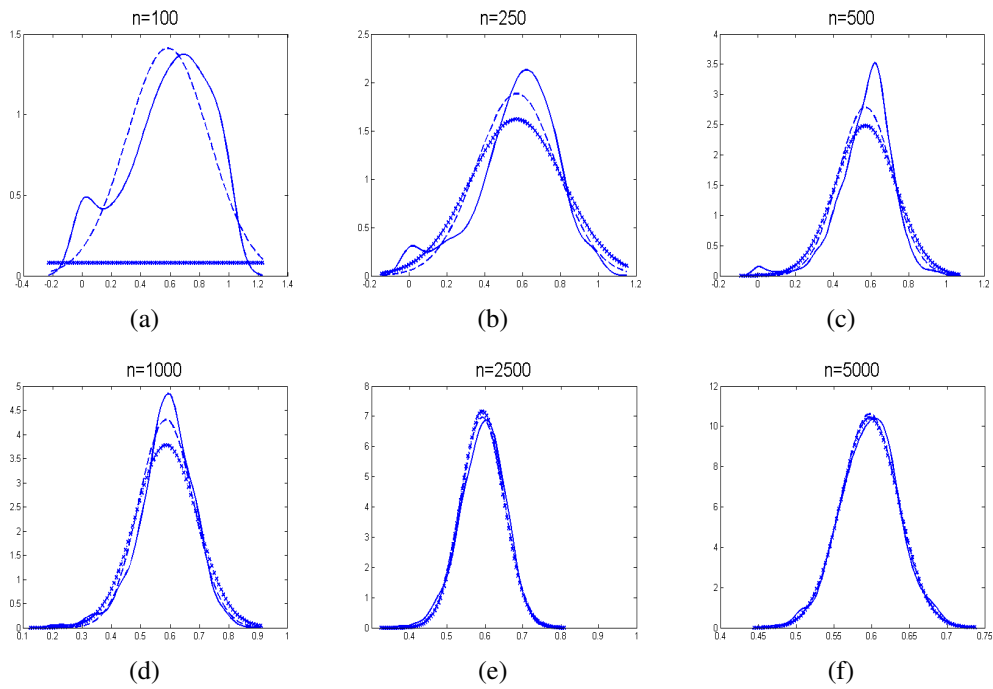


Abbildung A.14: Verteilung des ML-Schätzers des Parameters  $\beta_1: 0.6$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

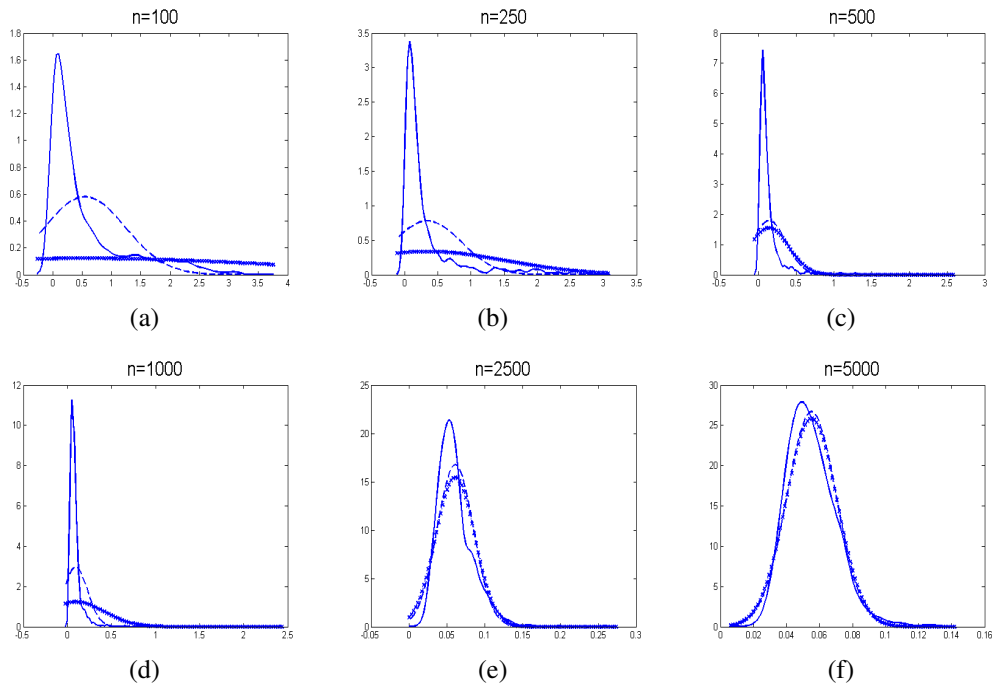


Abbildung A.15: Verteilung des ML-Schätzers des Parameters  $\alpha_0: 0.05$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

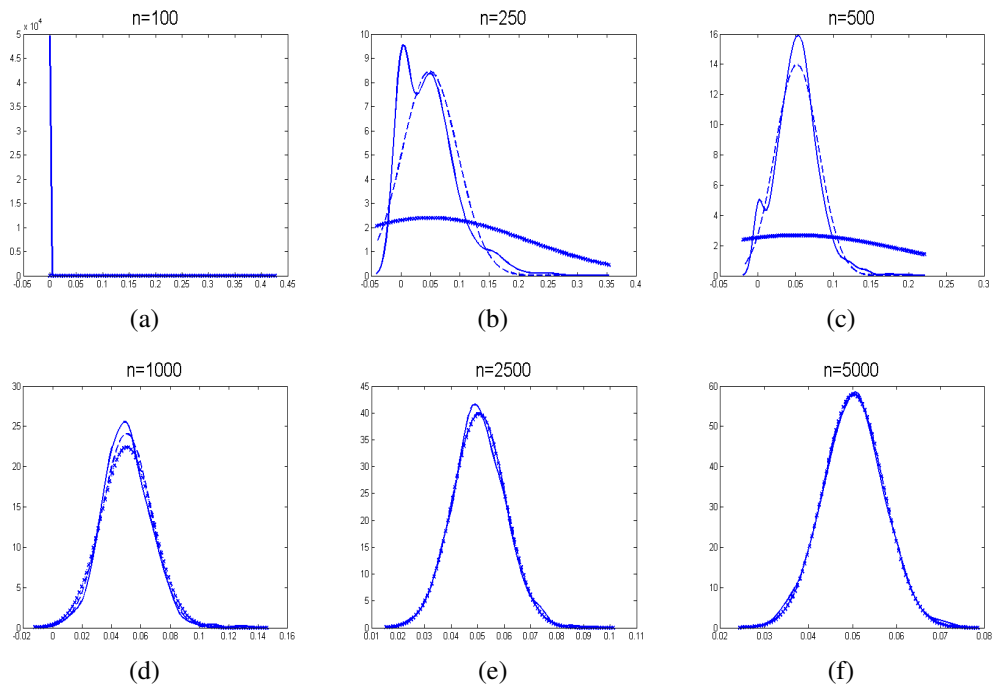


Abbildung A.16: Verteilung des ML-Schätzers des Parameters  $\alpha_1: 0.05$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

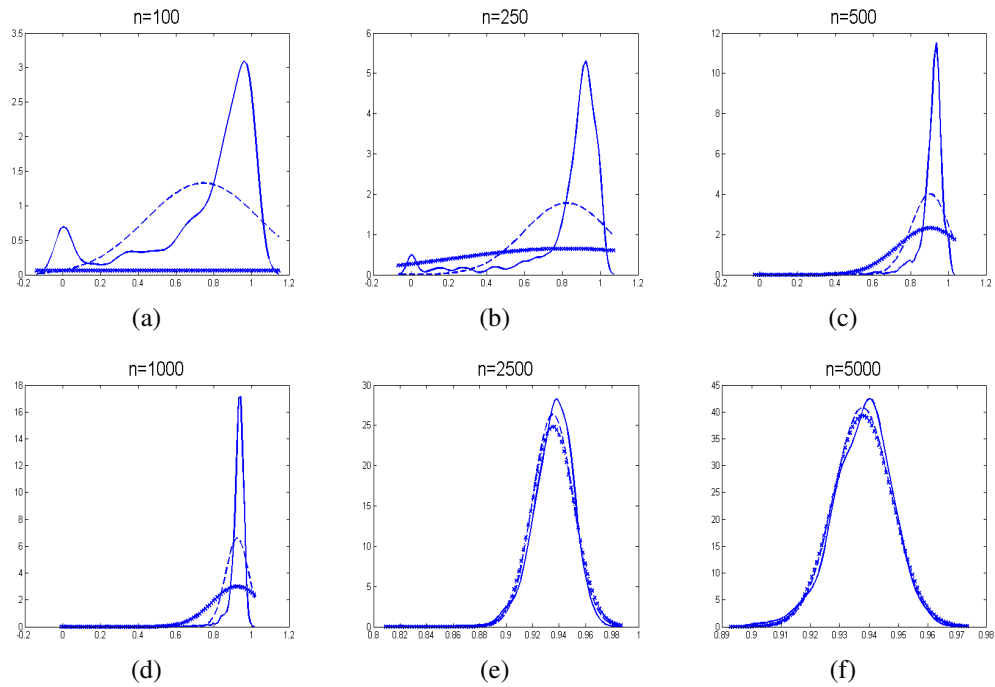


Abbildung A.17: Verteilung des ML-Schätzers des Parameters  $\beta_1: 0.94$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

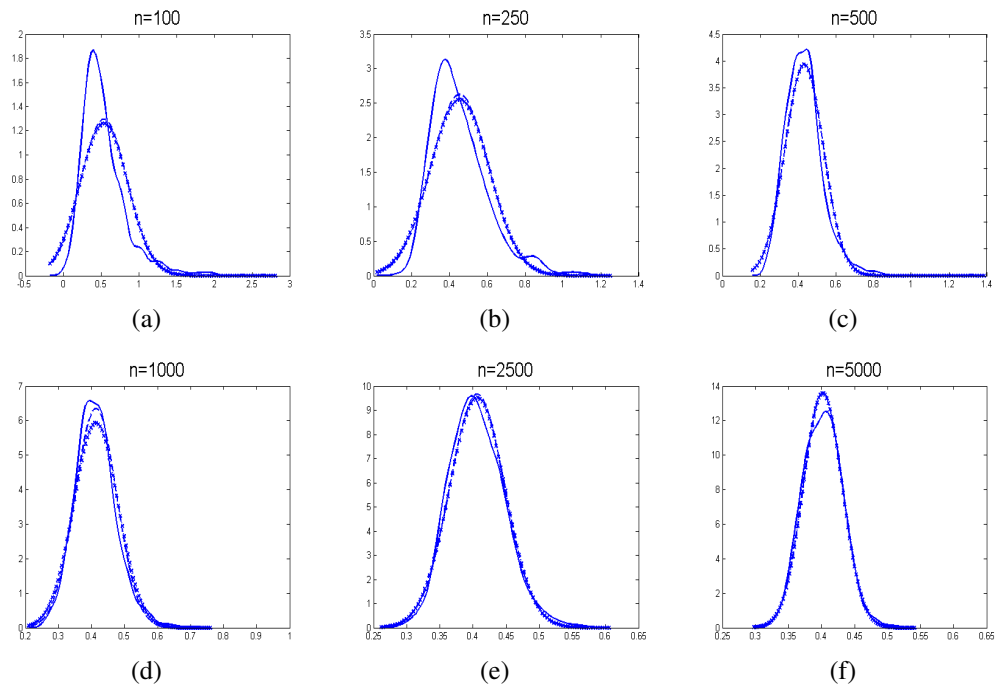


Abbildung A.18: Verteilung des ML-Schätzers des Parameters  $\alpha_0: 0.4$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

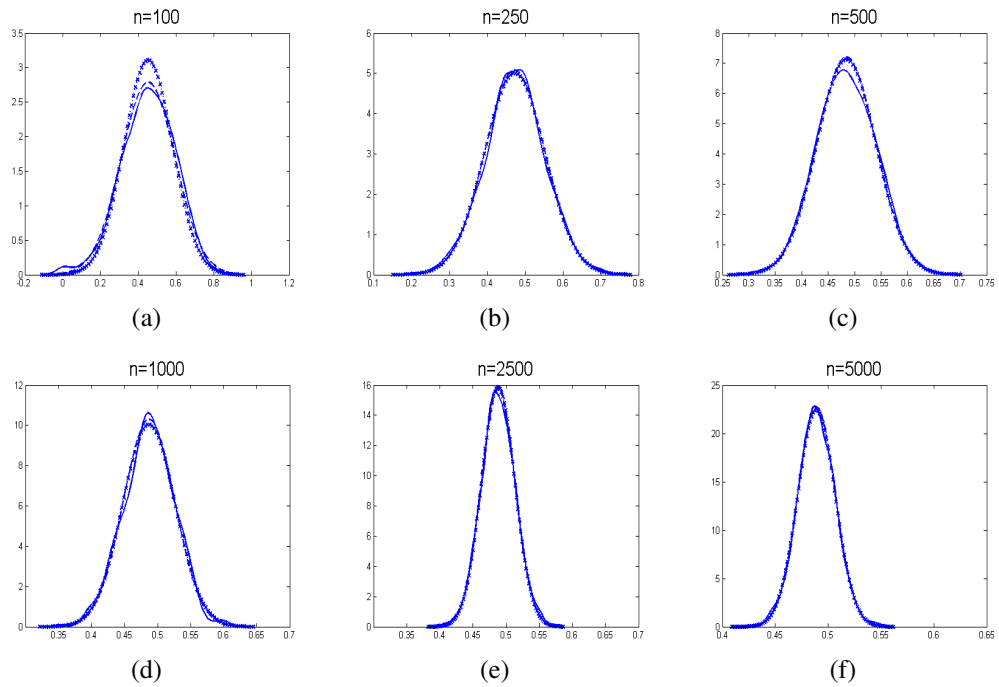


Abbildung A.19: Verteilung des ML-Schätzers des Parameters  $\alpha_1$ : 0.49 bei Variante 1, bei unterschiedlichen Stichprobenumfängen

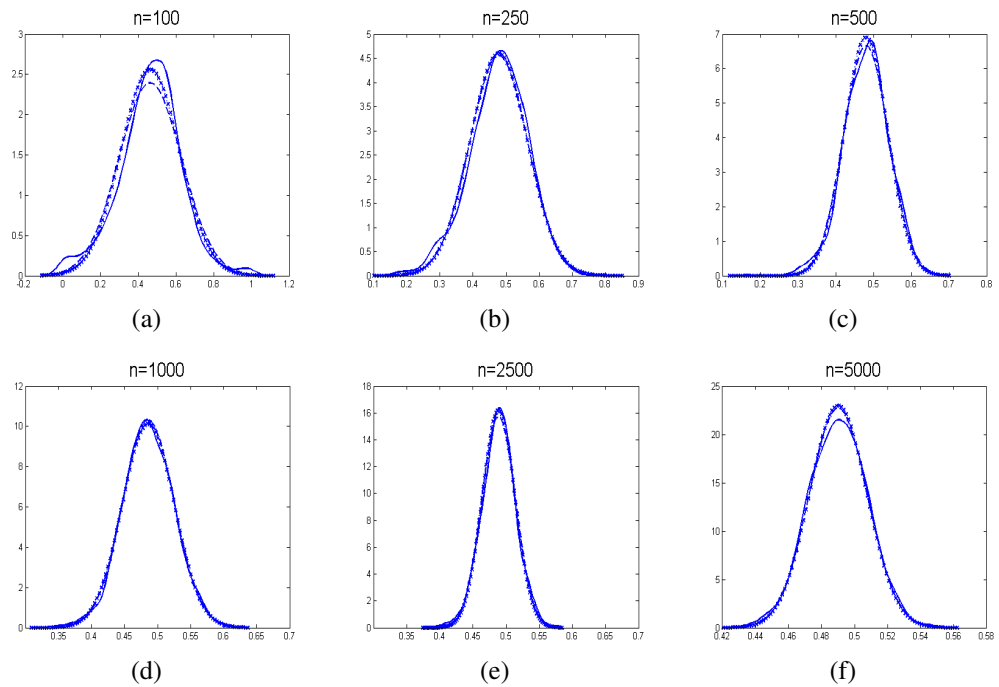


Abbildung A.20: Verteilung des ML-Schätzers des Parameters  $\beta_1$ : 0.49 bei Variante 1, bei unterschiedlichen Stichprobenumfängen



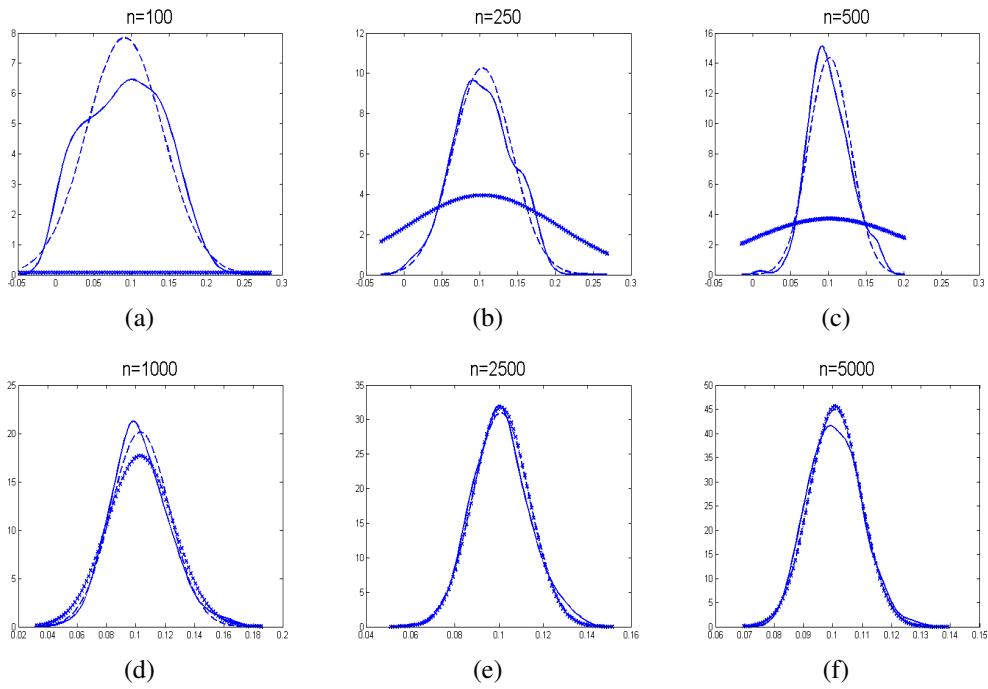


Abbildung A.21: Verteilung des ML-Schätzers des Parameters  $\alpha_0: 0.1$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

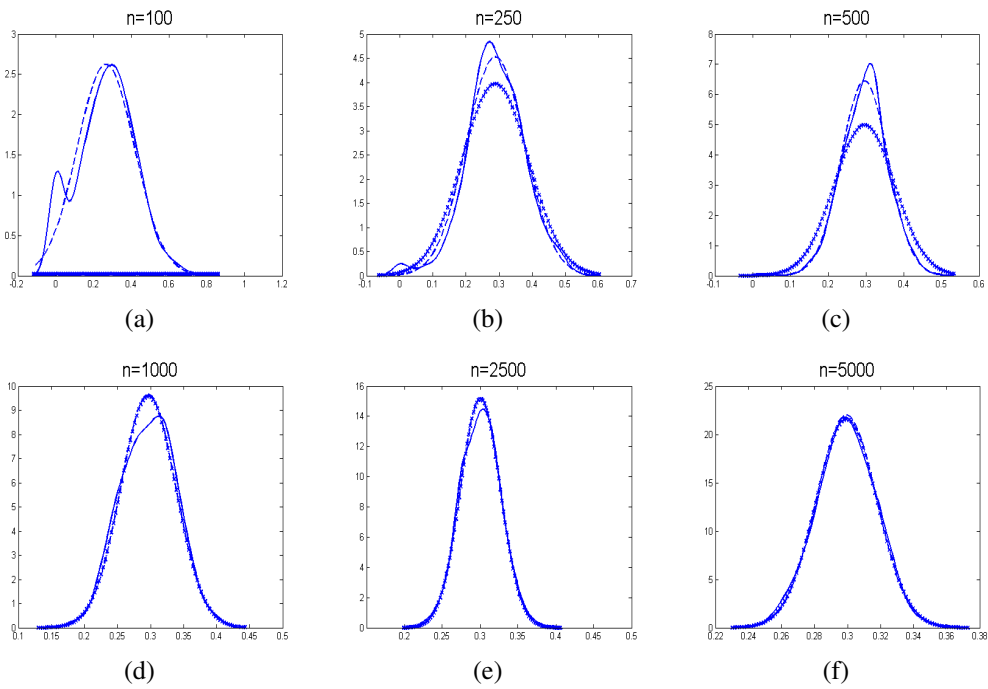


Abbildung A.22: Verteilung des ML-Schätzers des Parameters  $\alpha_1: 0.3$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen

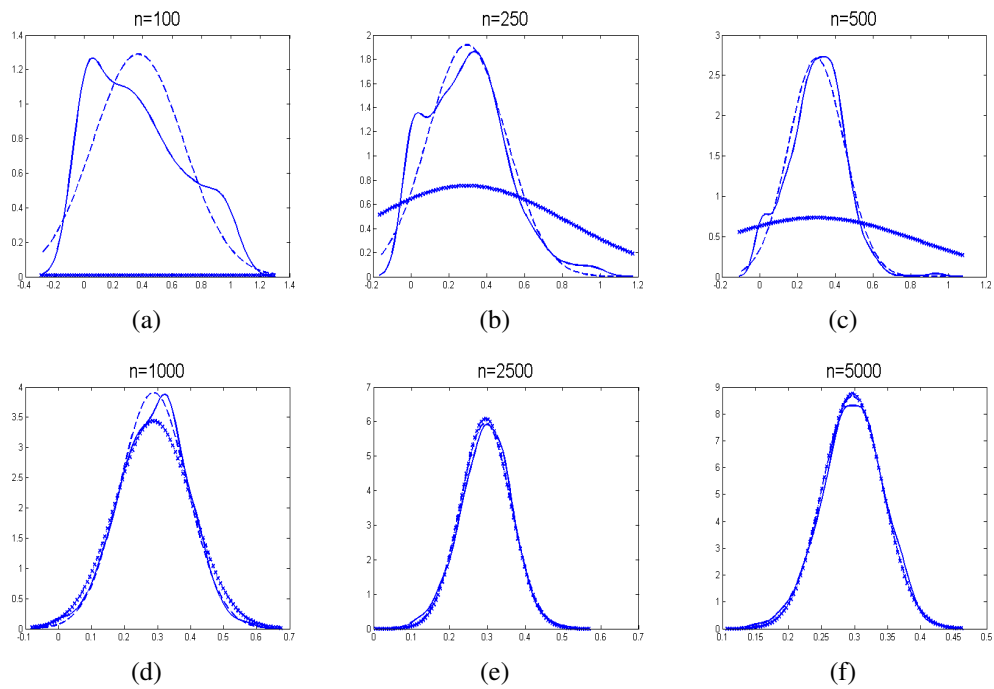


Abbildung A.23: Verteilung des ML-Schätzers des Parameters  $\beta_1: 0.3$  bei Variante 1, bei unterschiedlichen Stichprobenumfängen